

https://dspace.org/

The design of retirement schemes: possibilities and imperatives

Asher, Anthony

2008-02-29T07:20:17Z
http://hdl.handle.net/10539/4497

Downloaded from a DSpace repository. https://dspace.org/

THE DESIGN OF RETIREMENT SCHEMES: POSSIBILITIES AND IMPERATIVES

Anthony Asher

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Sydney, 2007

DECLARATION

I declare that the published and unpublished articles and other elements that make up this thesis are the result of my own work. To the extent usually and reasonably expected, assistance and peer review were received from my supervisor and colleagues, and from editors and referees of the journals and books in which they were published. The thesis is being submitted for the Degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University. I was employed by the Australian Prudential Regulation Authority and Trowbridge Deloitte for the period during which the thesis was written; papers were published with permission.

_____

24th day of October, 2007

ABSTRACT

South Africa has a sophisticated and developed retirement fund industry and an extensive social security system. While the objective of the latter is wider, both are concerned with financial security: particularly in the face of risks of death, disability and old age. It is widely recognised that there are many gaps in coverage. The chapters in this thesis address these gaps and administrative and benefit structures that could be developed to provide a truly comprehensive social security system. In particular, the thesis discusses the retirement and old age recommendations of the Taylor Committee, on which the author served. The vision is of universal coverage for the current state benefits augmented by mandatory employer based group schemes that offer disability, retirement and orphans' pensions. Means tests, the Road Accident Fund and workers' compensation arrangements would be abolished.

The chapters of the thesis are each self-contained, having all been published in – or submitted to – journals, books or conferences. In each, an attempt has been made to review a broader literature than is normally used to discover the impact of some element of the benefit structure, governance or investment policies of retirement schemes on their members. In this context, it is considered to be particularly appropriate to test policies and governance against the standard of justice.

DEDICATION

To Jane, Robert, John and Michael

Companions in co-inherence – and the rough and tumble

ACKNOWLEDGEMENTS

My understanding of retirement fund issues has benefited greatly over the years from interaction with fellow members of the Actuarial Society of South Africa retirement fund and social security committees, the Pensions Advisory Committee, and the board of the Community Growth Fund. I value too the contribution of my colleagues at the University of Witwatersrand, particularly Bill Haslam and Rob Thomson, who deserves particular thanks for his diligence as a supervisor and editor.

In the writing of the chapters, I would also like to thank as editors John Piggot, John Evans and Hazel Batemen of the University of New South Wales Centre for Pensions and Superannuation where I have presented four of the chapters, and Edward Elgar which has published two; the Department of Actuarial Studies at Macquarie University particularly for electronic access to their library; the Australian Prudential Regulation Authority, for which I prepared two of the chapters and which gave me space between other responsibilities to write much of the others; Michael Orszag and the Oxford University Press for commissioning the chapter on Pensions in Africa; and the Actuarial Society of South Africa and the Institute of Actuaries of Australia for opportunities to present and discuss chapters and parts thereof.

My wife Jane deserves particular credit for giving me time and place to write, and much encouragement to finish.

CONTENTS

	Page
DECLARATION	II
ABSTRACT	III
DEDICATION	IV
ACKNOWLEDGEMENTS	V
CONTENTS	VI
LIST OF TABLES	XIII
LIST OF FIGURES	XIV
CHAPTER 1 – INTRODUCTION	1
1 OBJECTIVE	1
2 MY VISION	2
3 JUSTICE	3
4 IS THE TOPIC RELEVANT TO ACTUARIES?	4
5 INSURANCE COMPANIES AND RETIREMENT FUNDS	7
6 LITERATURE AND OTHER SOURCES	8
6.1 The area to cover	8
6.2 Understanding	10
6.3 The approach of this thesis	11
7 THE THESIS	12
REFERENCES	15
CHAPTER 2 - PENSIONS IN AFRICA	19
1 INTRODUCTION	19
2 BACKGROUND	20
2.1 African economies	20
2.2 Poverty	21
2.3 The informal sector	22
3 SOUTH AFRICA	23
3.1 The first pillar	23
3.2 The South African private sector	27
3.3 Regulation	31

3.4 Government employees' schemes 32
4 OTHER AFRICAN COUNTRIES 33
4.1 Civil service schemes 33
4.2 National schemes 33
4.3 Other arrangements 35
4.4 Investments 37
4.5 Governance 37
5 CONCLUSION 38
5.1 Research questions 38
5.2 Policy lessons 39
REFERENCES 40

CHAPTER 3 - MEANS TESTS: AN EVALUATION OF THE JUSTICE OF IMPOSING HIGH RATES OF CLAWBACK ON THOSE OF MODEST MEANS 43
1 INTRODUCTION 43
1.1 Structure 44
1.2 Scope 44
2 JUSTICE 45
2.1 Personal 45
2.2 Relevance 46
2.3 Liberty 47
2.4 Needs 48
2.5 Equality 51
2.6 Redistribution 55
2.7 Efficiency 59
2.8 Just deserts 68
3 REFORM 69
3.1 Bringing it together 69
3.2 Institutional factors 73
3.3 Lobby groups 74
4 CONCLUSION 76
REFERENCES 78

CHAPTER 4 - CONFLICTED SUPER STRUCTURES: ARE AUSTRALIAN INVESTORS BEING SHORT-CHANGED? 83
1 INTRODUCTION 83
2 TYPES OF FUNDS 83
2.1 By trustee 84
2.2 By contractual party 85
2.3 By benefit structure 87
2.4 By manner of investment 88
3 MAIN FORMS OF REGULATORY PROTECTION 88
3.1 Fiduciary responsibility 89
3.2 The main legislation 89
3.3 Equal representation on the board 90
3.4 Fitness of trustees 91
3.5 Investment choice 91
3.6 Accounting procedures 92
3.7 Risk management planning 92
3.8 Waiver of charges on small accounts 92
4 MARKET, REGULATORY AND INSTITUTIONAL FAILURE 93
4.1 Market failures 93
4.2 Regulatory or legal failures 95
4.3 Institutional factors 97
5 TRUSTEES AND POTENTIAL CONFLICTS OF INTEREST 100
5.1 Distribution 101
5.2 Administration 107
5.3 Investment 109
5.4 Trustee discretion over fees 110
6 COSTS AND RETURNS 112
6.1 Costs 112
6.2 Returns 115
6.3 Discussion 115
7 CONCLUSIONS 119
REFERENCES 120
APPENDIX 1: COMMISSIONS: HOW SHOULD THEY BE FUNDED? 123
APPENDIX 2: INVESTMENT RETURNS OF FUNDS (1997–2003) 126
POSTSCRIPT: COMPARISON WITH SOUTH AFRICA 128

CHAPTER 5 - PENSION BENEFIT DESIGN AND THE INTEGRATION OF INSURANCE BENEFITS OVER THE LIFE CYCLE 129
1 INTRODUCTION 129
2 THE LIFECYCLE 130
2.1 Economic research 131
2.2 Income 132
2.3 Expenses 134
2.4 Savings 135
2.5 Nuclear families 136
2.6 Retirement as a social phenomenon 137
3 RISKS DURING THE LIFECYCLE 140
3.1 Death 141
3.2 Disability 144
3.3 Unemployment 150
3.4 Divorce 152
3.5 Retirement 153
4 BENEFIT DESIGN 154
4.1 Cross-subsidies in a typical DB scheme 154
4.2 Integrating death benefits before retirement 156
4.3 Integrating disability benefits 158
4.4 Dealing with divorce 159
4.5 Integrating unemployment benefits 160
4.6 Inappropriate insurance 161
5 THE PROPOSED POLICY FRAMEWORK 162
5.1 Compulsion 162
5.2 Minimum benefits 167
5.3 Charging and reserves 169
5.4 Governance 170
6 CONCLUSION 171
REFERENCES 172
APPENDIX 179

CHAPTER 6 - THE RELATIVE INVESTMENT PERFORMANCE OF THE COMMUNITY GROWTH FUND 181
1 INTRODUCTION 181
2 HISTORY AND STRUCTURE 181
2.1 Unions and provident funds 181
2.2 Formation of the CGF 183
2.3 Investment process 183
2.4 Initial responses 185
2.5 Subsequent developments 186
2.6 New investment criteria 186
3 DETERMINANTS OF SHARE PRICE MOVEMENTS 187
3.1 Overall level of prices 187
3.2 Demand 188
3.3 Supply 188
3.4 Underlying profitability 189
3.5 Individual share prices 190
3.6 Perceptions 192
3.7 Efficient market hypothesis 192
4 SOCIALLY RESPONSIBLE INVESTMENT AND SHARE PRICE 193
4.1 Screening 193
4.2 Active shareholding 194
4.3 Profits and rent 195
4.4 Fiduciary duties 196
4.5 Truth, good governance and creativity 197
5 MEASUREMENT 197
5.1 Earlier studies 197
5.2 Present study 198
5.3 Further analysis 199
5.4 Empowerment 200
5.5 Regression on all data 200
5.6 Regression excluding black empowerment companies 200
5.7 Regression excluding other outliers 201
6 CONCLUSIONS 202
REFERENCES 203
APPENDIX 1: CORRELATIONS BETWEEN VARIABLES 205

APPENDIX 2: REGRESSION EXCLUDING BLACK EMPOWERMENT COMPANIES 206
APPENDIX 3: REGRESSION EXCLUDING BLACK EMPOWERMENT COMPANIES AND OUTLIERS 207
APPENDIX 4: BASIC DATA 208

CHAPTER 7 - MEAN REVERSION IN INVESTMENT MARKETS 210
1 INTRODUCTION 210
2 HOW INVESTMENT MARKETS WORK 211
2.1 Market efficiency 211
2.2 Behavioural finance and bounded rationality 213
3 MEAN REVERSION 216
3.1 Currency markets 216
3.2 Interest rates 218
3.3 Inflation 220
3.4 Equity markets 220
3.5 Possible consequences 225
4 THEORY OF INVESTMENT 226
4.1 Fundamental analysis 226
4.2 Diversification and utility theory 227
4.3 Life cycle applications 227
4.4 The size of the equity risk premium 228
4.5 The growth optimal portfolio 229
4.6 Older superannuation fund members 231
4.7 Exploiting mean reversion and time varying risk premiums 232
4.8 Expenses and taxes 233
5 REGULATORY RESPONSE 234
5.1 Suggestions from behavioural finance 235
5.2 Australian regulations 236
5.3 Recommendations 237
6 CONCLUSION 241
REFERENCES 242

CHAPTER 8 - SMOOTHING INVESTMENT RETURNS 250
1 THE NEED FOR SMOOTHING 250
2 THE MANAGEMENT OF RISK 250
2.1 Longer term risks 250
2.2 Cross subsidies and inequities in DB funds 251
2.3 Managing risks in retirement 253
3 CURRENT METHODS OF SMOOTHING 254
3.1 Lifestyle switching to guaranteed assets 254
3.2 Delaying action 254
3.3 Smoothed bonuses 255
3.4 Alternative algorithms 256
4 A NEW ALGORITHM 257
4.1 A series of forward contracts 257
4.2 Pricing forwards and futures 258
4.3 Smoothed maturity value 259
4.4 Anti-selection 260
4.5 Effects on investment returns 260
5 CREATING ALTERNATIVE INSTRUMENTS 261
5.1 Inflation linked assets 261
5.2 Other linkages 261
5.3 Dividend links 262
6 AN ILLUSTRATION 263
6.1 Smoothing achieved 263
6.2 A problem with actuarial values 264
7 OTHER ISSUES 265
7.1 Guarantees 265
7.2 Application to retirement funds 265
7.3 Investment choice 266
8 CONCLUSION 266
REFERENCES 267

CHAPTER 9 - CONCLUSION: METAPHYSICS AND MOTIVATION 269
1 VOCATION 269
2 PRACTICAL WISDOM 273
3 FUTURE DIRECTIONS 277
REFERENCES 280

LIST OF TABLES

	Page
Chapter 3 - Means tests	
Table 1: Incentives to defer income - South Africa	64
Table 2: Incentives to defer income - South Africa; no means tests	66
Chapter 4 - Conflicted superstructures	
Table 1: Main types of superannuation funds	84
Table 2: Manner of investment of funds' assets	88
Table 3: Asset allocation of funds by type of fund	113
Table 4: Funds' costs as a share of their assets by type of fund	114
Table 5: Total wealth management and insurance industry	116
Table 6: Aggregate gross returns of funds by type of fund	126
Chapter 5 - The integration of insurance benefits over the life cycle	
Table 1: Time allocations by household type	139
Appendix tables: calculations for figure 1	179
Chapter 6 - The relative performance of the CGF	
Table 1: Initial social criteria	184
Table 2: Outperformance of approved shares	198
Table 3: Relative outperformance - 1998 to 2001	199
Correlations between variables	205
Regression excluding black empowerment companies	206
Regression excluding black empowerment companies and outliers	207
Basic data	208
Chapter 7- Mean reversion in investment markets	
Table 1: Accumulation phase	229
Table 2: Drawdown phase	232
Table 3: Costs of trading	234

LIST OF FIGURES

	Page
Chapter 3 - Means tests	
Figure 1: Effects of redistribution	56
Chapter 4 - Conflicted superstructures	
Figure 1 Sources of conflicts	102
Chapter 5 - The integration of insurance benefits over the life	
Figure 1a: Children born in twenties	137
Figure 1b: Children born in thirties	137
Figure 2: Cover and reserves	157
Chapter 6 - The relative performance of the CGF	
Affirmative action residual plot	206
Job creation residual plot	207
Corporate governance residual plot	207
Chapter 8 - Smoothing Investment Returns	
Figure 1: Smoothing (D/Y + inflation) - 10 year monthly JSE	264
Figure 2: Smoothing using D/Y + growth	265

CHAPTER 1 – INTRODUCTION

1 OBJECTIVE

The chapters of this thesis have been written to evaluate the policy proposals in Chapter 9 of Taylor *et al* (2002) (the Taylor recommendations). Chapter 9, Old Age and Retirement, was based on the report of the Retirement Sub-committee[1], which was written by the author of this thesis. That report needed expansion in a number of areas, and the recommendations needed to be justified systematically; conditions which this thesis remedies. It discusses what I see to be the more important questions in the context of an extensive review of the relevant literature.

To be fair to other members of the Committee, it should be recorded that there was not time to discuss many of the recommendations of the Retirement Fund Sub-committee. It should also be recorded that I submitted a letter to the Minister of Social Development dissenting from a number of the Committee's main recommendations. This thesis therefore reflects my views and not necessarily those of other members of the Committee.

The individual chapters of this thesis therefore discuss, in some cases explicitly and in others tangentially, those of the Taylor recommendations that I saw as needing additional research, investigation and possibly explanation. As I have wanted to expose and debate the ideas more widely and to contribute to the development of policy, the chapters have been written to stand alone, and have already been published, discussed at conferences or included in research paper series.

The concerns of this thesis also interact with other issues covered by the Taylor report: unemployment benefits, protection of children, disability, the Road Accident Fund and worker's compensation arrangements. In each of these areas, it was evident that middle-class South Africans obtain their financial security (or at least receive more money) from

---

[1] Although an acceptable version can now be found at: http://www.sarpn.org.za/CountryPovertyPapers/SouthAfrica/taylor/report7.pdf#search=%22Taylor%20%22Road%20Accident%20Fund%22%20abolish%22

insurance companies and retirement funds than from the state, but the authors of the other chapters of the Taylor report focused almost entirely on state provision. Their discussion, and on occasion recommendations, is often consistent with the recommendations of chapter 9. Those reading both the report and this thesis will however see that they are based on different visions of how a comprehensive social security system should serve all South Africans.

2 MY VISION

While the following chapters set out to investigate some of the Taylor recommendations, the thesis effectively outlines a vision of a rather different South African social security system insofar as it affects death, disability and retirement benefits. Little is said about changes to poverty relief, job creation and medical insurance, each of which is arguably more important. Fundamental issues not addressed here are the advantages and disadvantages of pay-as-you-go systems and funded private arrangements. Asher (1997) makes the case for the latter in terms of accountability, the efficiency of investment markets, and greater economic democracy. The social, economic and educational benefits of economic democracy are spelt in more detail in Asher (1992).

The vision consists of two parts that can be briefly described.

First is a universal pension for orphans, the disabled and the aged funded from general tax revenue. This would mean adding an orphans' pension to the current system, and abolishing the means test, which I characterize in chapter 3 as an administrative atrocity.

The second part would be employment based funds very much like current retirement funds, except that membership would be compulsory and they would be required to offer a set of minimum pensions: for orphans, the disabled and the aged. Membership would not be required of workers informally employed or earning below the tax threshold. The funds would all be controlled by elected trustees subject to the fiduciary duties of common law and a government regulator. These benefits would replace those currently offered by the Road Accident Fund and workers' compensation arrangements, which could therefore be abolished.

Such a system would be uniquely South African. The universal state pension and benefits would resemble the New Zealand system, but that country has not much of a contributory scheme. The defined contribution private system would have much in common with

countries that have followed the Chilean model, but would not be imposed on informal workers and would offer a more comprehensive suite of benefits. It would provide an acceptable level of support to the poor, and create a high probability that wealthier families could retain their lifestyle in the event of early death, disability and in retirement. There would be protection for the weak and controls on the power of the strong.

3 JUSTICE

The Taylor recommendations are evaluated against the standard of justice, which provides the imperatives of the title to this thesis. Justice is the standard against which we measure the exercise of power by individuals, and agreements and laws that exercise power over our lives. It is therefore the standard by which we measure government intervention, and is particularly concerned with protecting the powerless against the powerful. In lesser measure, it is also the standard against which we measure the actions of trustees when they exercise discretions given them in terms of their power over the operation of retirement funds.

There is an ancient consensus that justice can be defined as giving to each his due[2]. What is due to people subjected to the power of others? If anything, it must be related to their essential dignity: to their ubuntu (humanness), flowing, for some, from our resemblance to the Creator. This dignity seems to provide sufficient basis for the recognition of objectives or what are sometimes called basic or first-generation human rights: to life, equality, liberty, the provision of basic needs, and recognition of just deserts. The approach taken in this thesis is to recognise the absolute value of these objectives, in the sense that they always have value and that governments are under obligation to attempt to satisfy them as much as practical. Efficiency is added to the list of objectives in to capture the practicalities.

This approach can be distinguished from the "rights-based approach" to social objectives that are occasionally referred to in the Taylor report, which ambitiously attempts to ground minimum social security benefits in law and with the judicial arm of government. The justice-based approach applied here is also concerned with processes and social

[2] See Aquinas (1275) Part II – II Question 58

impacts, less definite in its operation, and looks to both government and private sector for action. Looked at another way, justice can be seen as the application of the golden rule to the relationship between rulers and the ruled.

Justice is, as Lucas (1980) says, the glue that creates social solidarity, the *fraternite* of the French revolution or the flowering of ubuntu: *umuntu ngubuntu ngabantu* – a person is only a person because of others. Its foundation in human dignity gives the lie to claims that such solidarity requires those with authority in a community to be immune from criticism. On the contrary, their position of power means that they are subject to an even higher standard of behaviour.

4 IS THE TOPIC RELEVANT TO ACTUARIES?

Benefit design is not one of the statutory functions required of actuaries, and it may be asked whether it is an issue of direct relevance to actuaries. An answer can be derived from at least three perspectives: from the needs of society, the roles played by actuaries in the retirement fund industry and the self-interest of members of the profession.

The need for the profession (as underlined by their statutory roles) arises from the difficulty faced by the general public in evaluating financial soundness of insurance companies and retirement funds without specialised training and detailed knowledge. There is a need for people with the skills and integrity to monitor the soundness of those institutions that are likely to pay benefits many years after contributions have been made.

The remoteness of the benefits from the contributions creates another need for specialised knowledge. It is difficult for healthy people to envisage and understand the financial needs of the disabled, for parents to know the financial implications to their children of the death of one or both of them, or for young people to project their needs in retirement. Such knowledge cannot emerge spontaneously. It requires a systematic evaluation of the financial needs of the beneficiaries. The arguments for professional discipline in this arena are similar to the arguments for prudence in the management of financial institutions: they both require some deferment of gratification, which can be an unpopular message that may require that those people and organizations that promise premature pleasures should be restrained. There is therefore a need for professional involvement in this sphere.

Given the need, it is not surprising that actuaries, in practice, devote much of their time to the design of retirement fund benefits. Actuaries working with retirement funds are frequently styled, or employed by, "employee benefit consultants". It is true that the actuarial focus is more on the pricing of the benefit than on the design per se, but that is largely because benefit design cannot be often changed and is therefore not often the subject of much thought. When changes are considered, it is inconceivable that actuaries would not have to be consulted and have to consider the contingencies upon which the benefits are paid. While it is argued that they should often know more about these contingencies, their training and experience frequently make them better informed than others involved in these decisions.

This second perspective might appear to be conclusive, were there not a dislocation between the work of so many actuaries and alternative views of the nature of actuarial science. This dislocation is expressed in the difference between the definition of the role of the profession given above, and another commonly used to describe the work of actuaries. The Institute of Actuaries defines actuaries as "providing expert and relevant solutions to financial problems, especially those involving uncertain future events[3]". This definition is so broad that it appears to extend the work of actuaries into a number of investment and commercial areas in which they have no particular experience, and where they would compete with other professionals. As the profession already has a defined social purpose, which includes but is not necessarily limited to its statutory powers, it is suggested that it is not helpful to see actuaries as a group of people with a particular skill looking for anything to do that relates in some way to their training. The profession's statutory powers are limited to the security and pricing of benefits rather than their design but just as actuaries have had to consider the cost of this security, they must surely be concerned that the benefits are worth securing in the first place. It can also be noted that the profession antedates its statutory powers and that members have always been involved in benefit design.

Day (2000) perhaps provides a clue as to why a profession might attempt to avoid a practical definition of their responsibility. He shows that claims to a relatively undirected

---

[3] The Institute of Actuaries Education Committee: "Consultative Document on the proposed new syllabuses prepared by the Education Strategy Working Party", 1991

expertise are not limited to the actuarial profession, and are often tied to claims that the professional discipline is scientific:

"Professions have habitually taken an uncritical stance toward their foundations, and they have often uncritically adopted the rhetoric of dominant social, cultural, and political forces in order to construct and preserve their social power. Upon the rhetoric of "science" and "progress," professions have often harnessed their sleds, sometimes fairly careless of their destination."

Thomson (2004) has persuasively argued that actuarial work involves the application of a scientific discipline, but even if it were not so, the problems to be addressed by this thesis are real, amenable to empirical observation, modeling and application in the work of actuaries. It is however suggested that while actuaries have always had a significant, and often dominant, input into the design of life insurance and retirement fund benefits, they do not always appear to have had the benefit of a full understanding of the real needs of the policyholders and beneficiaries of these funds, nor a coherent framework from which to develop new benefits or improve existing ones. This thesis is also an attempt to address this issue.

It can be argued that actuaries should consider the question in their own self-interest; it is a question of marketing. Some marketing theorists insist that the business of an institution is to provide for the real needs of its clients, and not just to sell more of the products it happens to be producing. In the classic exposition of this view, Levitt (1975) shows the danger of failing to appreciate this distinction. He suggests, for instance, that the American railway companies lost their dominant position in the American economy of the nineteenth century by failing to see they were in the transport business rather than being providers of railways.

The business of a profession can be considered similarly. Long-term survival will depend on developing services that address people's real needs - rather than producing and servicing products that are inadequate or may become so. Phrased more positively, considering people's needs is the only way to generate new ideas of how to meet them.

Rich (1991) quotes the draft mission statement then being prepared by the Actuarial Society of South Africa: "to make a significant contribution towards enabling South Africans to enhance both their individual and institutional financial security using their expertise in the disciplines of ..." The paragraph was dropped as referring to the mission

of actuaries rather than that of the professional body, but it would seem to provide a more fruitful definition of the mission of actuaries.

The actuarial mission can then be outlined as follows:

- People need financial security: their future incomes are both uncertain and limited, while the value of their property is also exposed to various risks.
- The insurance and retirement industry exists to provide for this need.
- It is the responsibility of actuaries to provide the necessary professional skills to ensure that these industries are run soundly with this aim in mind.

As Trowbridge (1989) puts it: "Actuaries are those with a deep understanding of financial security systems, their reasons for being, their complexity, their mathematics, and the way they work." Such an understanding would seem pointless if it were not utilised to provide people with the financial security they really need. It is therefore suggested that the question is therefore not merely actuarial, but perhaps critical to the profession playing its full role in society. The consideration of people's needs, as for instance found in using the financial lifecycle to answer financial security questions, is finding increased exposure in the actuarial literature. Cooper (2002), Cantor and Sefton (2002) and Milevsky (2004a) provide examples.

5 INSURANCE COMPANIES AND RETIREMENT FUNDS

The thesis is limited to considering retirement funds rather than insurance companies in spite of their similarities and frequent overlaps. The major difference between the benefits provided by life insurers to individuals and group retirement funds is that the individual benefits offered by the former have to be sold to individuals, and are primarily subject to considerations related to the distribution channel and are not necessarily related to needs. Retirement schemes are different in that compulsory membership greatly simplifies the question of distribution and they are managed by trustees. Benefit design is far more likely to be governed by what is apparently good for the members rather than designs that are superficially attractive.

Milevsky (2004b), for instance, recounts an attempt to persuade a Canadian company to sell a deferred annuity with no benefit on prior death. The management were persuaded of the benefits to potential policyholders, but ultimately rejected the idea because of administrative difficulties, and the possible public embarrassment that might arise from

the lack of a death benefit. (Insurers can be criticised for paying nothing to heirs on the death of an annuitant in spite of the terms of the contract. The often blatant self-interest of the criticism does not necessarily ameliorate the perceived damage to reputation.)

Another illustration is the annuity portfolios of the two largest South African insurance companies: a large proportion if not most group contracts are with-profit versions that address inflation and longevity risks. The overwhelming majority of individual contracts are however either fixed annuities that offer no protection against inflation risk or "living annuities" that offer no protection against longevity risks[4]. The companies have made minimal effort to sell the more appropriate benefits to individuals apparently because of the difficulties in persuading their sales forces to make the effort to suggest that retirees should sacrifice immediate yield for inflation protection, and make adequate provision for longevity risk.

One cannot therefore discuss insurance policy design without an in-depth consideration of distribution issues, which is beyond the scope of this thesis.

6 LITERATURE AND OTHER SOURCES

Reviewing the relevant academic literature must be the first step of any serious research, but it is suggested that its proliferation over the past few decades means that a literature review is now different.

6.1 The area to cover

Until perhaps 15 years ago, reviewing all the literature on any single topic was a manageable activity for a diligent researcher. There were a limited number of academic journals and a few reputable publishers. Printed citations indices, searches of adjoining library shelves and common room discussions with those from different disciplines could guide one to important ideas in related areas.

In the last decade, this has become more difficult if not impossible. There are frequently too many papers to read – even on the most esoteric subjects. In addition to the formally published research, there are vast numbers of working papers and articles published by

[4] As is evident from the annual reports of the Registrar of Long Term Insurance, available from the FSB.

academic departments, research organizations and policy think-tanks. The internet makes them available almost immediately. The literature review for chapter 7 of this thesis illustrates. An internet search for "mean reversion" and "investment" for section 3 produced almost 100 000 references. Limiting the search to Google Scholar reduced it to 5 000 - still impossibly vast to inform a few pages of discussion. Even writing ten years ago, Campbell *et al* (1997) mentioned the impossibility of covering all the papers that have investigated the efficiency of markets, the subject of the second section of this chapter.

In addition to the likelihood that useful research will be missed, the proliferation of voices makes it more difficult to sift what is important and reliable from the relatively trivial and speculative. While the number of citations and downloads gives a guide to relevance, it is less useful in measuring reliability.

For insight and reliability, one would want to see the extent to which ideas have been subject to review and debate. As to technical competence there is often little to cause concern: academic journals are peer reviewed - many to an extremely high standard. The review process is not however entirely uniform as, for instance, illustrated by Ferson *et al* (2003), who find instances of spurious statistical relationships.

Often more problematic is the interpretation of the statistics. Chapter 6 of this thesis, on the relative investment performance of the Community Growth Fund, makes almost no reference to a fairly considerable literature on the performance of ethical investment. This is because the published papers invariably fail to consider that the relative performance is a function of both demand and profitability[5]. As explained in the chapter, poorer performance may well be explained by higher demand for the shares at the beginning of the period of measurement and have nothing to do with profitability.

It would also seem that even the best journals have become less open to critical debate over time[6], so making it more difficult to understand the areas of controversy, let alone

---

[5] True of the first 5 references out of over 600 on a Google Scholar search: "investment performance" and ethical.

[6] As: Coelho P R P, De Worken-Eley F & Mcclure J E (2005) Decline in Critical Commentary, 1963-2004 econjournalwatch.org/main/intermedia.php?filename=CoelhoetalEconomicsInPracticeAugust2005.pdf

come to a balanced opinion. A particular problem arises when those in the same field appear to ignore each other. The question of the efficiency of markets again provides an illustration. Market inefficiencies arising from the cost of information, and those that arise from transaction costs and institutional structures, are discussed in Stiglitz (2002) and Merton and Bodie (2005) respectively. Joseph Stiglitz and Robert Merton are winners of the Nobel Prize for economics, but these papers summarise their own contributions to the literature about ways to understand market imperfections without cross-references.

For actuaries, there is a further problem that, with minor exceptions, the professional actuarial journals are not indexed in the academic literature. An advantage for papers published in the British Actuarial Journal, which may not be sufficiently appreciated, is that the discussion of the papers at sessional meetings is also published.

There are some shortcuts, such as referring to more recent papers, which will usually summarise at least one string of previous research. This method creates the risk that their importance is overemphasised, but it has often been taken in the chapters in this thesis.

6.2 Understanding

If one cannot cover all the ground, the aim must be to attempt to produce a broad outline with a sample of the more prominent features. The object of research is understanding, and not mere information. This means searching out insights, comparing and contrasting alternative views and approaches, and attempting to synthesise a framework for thinking about the subject.

How this might be done is itself controversial. Pemberton (1999) and Huber and Verall (1999) emphasise the need for both theory and empirical testing in developing actuarial or economic models, and the need for judgment in their application. They (and apparently most philosophers of science) therefore reject the empiricist view that only falsifiable theories can be considered valid.

They do, however, appear to accept the other positivist view that true knowledge ought to be dispassionate, if not disinterested. While superficially attractive, this view is unrealistic: a deep understanding is not likely to come to those for whom the subject has no interest – of a material or other kind. Polanyi (1962) “demonstrates that the scientist's personal participation in his knowledge, in both its discovery and its validation, is an

indispensable part of science itself. Even in the exact sciences, 'knowing' is an art, of which the skill of the knower, guided by his personal commitment and his passionate sense of increasing contact with reality, is a logically necessary part."[7] Strong feelings about a subject may well be a necessary basis from which to launch the pursuit of understanding, although they become counterproductive if they blind us to credible alternatives or inconsistencies in our own views. While it is preferable to reduce the heat in argument, participation in debate with those with whom we disagree is a critical element of learning.

6.3 The approach of this thesis

The process of literature review therefore needs to be both critical and efficient. My input into the Taylor recommendations was based on twenty years of experience in the retirement fund industry, and over a decade of exposure to formal academic literature. Much of this involvement and reading were motivated by indignation at what was regarded as unjust or inadequate. In searching the literature, I have looked principally for justification and criticism of prevailing theory and practice; and for support or otherwise for my own views where they differ from the consensus. In such searches – as is consistent with Simon's theory of bounded rationality – the search was stopped once I was satisfied that I had either found an appropriate justification for the consensus and could therefore accept it, or found support for a credible alternative. There are no doubt many important ideas and research results that were not unearthed, some of which may well have a bearing on the arguments.

In searching for discussion of alternative ideas, the literature search has necessarily ranged considerably further than the standard actuarial and economic literature. Important related fields are sociology, gerontology, the less formal literature of financial advice and publications on social security. The latter is produced by governments and international agencies reflected in government commissions, the publications of the World Bank, International Monetary Fund (IMF) and the International Labour Organization (ILO). This strand however is also politically fractured: the Bank and IMF are more interested in the development of private-sector schemes while the ILO is more interested in extending the coverage of benefits.

---

[7] This pithy summary of the book's theme is taken from the review at Amazon.com.

Whether the thesis does adequately evaluate the supporting and contradictory evidence is for the reader to judge. It should be pointed out that the publication or discussion of the chapters as independent pieces means that each have their own bibliography.

7 THE THESIS

The common thread in the thesis is how different aspects of the retirement system can be adapted to better meet people's needs in a just manner. The different aspects include the:

- interaction between the state and private institutional arrangements discussed particularly in chapters 2 and 3,
- governance of the private arrangements covered in chapter 4,
- benefit structure of private funds in chapter 5;
- investment issues that affect the size of benefits in chapters 6, 7 and 8.

Chapters 2 and 3 discuss the Taylor recommendations for the abolition of the means test for the state old-age pension, and various specific changes to the income tax system. The second chapter (Pensions in Africa) describes the South African retirement fund system and some alternative systems in the context of Africa, with its small and poorly resourced economies. It focuses particularly on the needs of people working informally, suggesting that their retirement needs should be provided by a non-contributory scheme funded by general taxation as is currently the case. While South Africa is richer than other countries on the continent, the same conclusion is reached: governments ought not to attempt to deliver a complete earnings-related retirement system but: "the energies of the state would be better directed towards providing a low level of universal cover and the protection of members in voluntary private sector arrangements." The conclusion is evaluated in terms of the model of justice discussed above that is also used extensively in chapter 3, and finds its echoes in each of the succeeding chapters. The second chapter also raises the question of the governance of retirement schemes, an issue taken up again in chapter 4. The third chapter (Means Tests) makes the argument for the abolition of the means test not only in South Africa, but also in Australia and the UK. It is suggested that means tests create an unnecessary and onerous burden on lower income workers in the formal sector.

Chapter 4 (Conflicted Superstructures) discusses the question of governance. The Committee recommended that all funds be controlled by trustee boards of lay members, 50% elected by members, and that the Financial Services Board report annually on its

efforts to manage conflicts of interest. The chapter describes the Australian superannuation industry, but the governance issues have South African parallels, which are drawn out in a postscript. The main driver of expenses is the distribution system, which also creates a number of fiduciary conflicts. The fiduciary responsibilities of retirement fund trustees are evidently an issue of justice, as those that exercise power should not enrich themselves at the expense of those for whom they have responsibility. Ensuring that trustees conform to their responsibilities is the proper role of the state, whether exercised through the courts or the regulators.

The Committee also recommended that the power pension funds had over the companies in which they invested be increased by changes to company law, but this is not explored further as it was debated and recommended in Asher (1998).

Chapter 5 (Integrating Insurance Benefits) discusses the Taylor recommendations covering compulsory contributions, preservation and the related issue of annuitization. The report recommended that compulsion be limited to those over 40, that preservation requirements be waived only if the members have exhausted their unemployment benefits, and that at least two thirds of the benefit be taken as an inflation-protected annuity. Other benefit recommendations were that funds should be compelled to offer a compulsory disability benefit of 50% of the member's income and orphans' benefits of 20% each, and that there be greater disclosure of the cross-subsidies inherent in a defined benefit (DB) design. The Committee was divided over whether the life and disability cover would make the Road Accident Fund and workers' compensation benefits redundant, but recommended that the question be investigated.

In chapter 5, the extensive economic and social literature on the financial life-cycle is used to expand and clarify a variety of ideas raised in Asher (2000, 2003a and 2003b). This chapter also considers the questions whether formal sector employees should be compelled to contribute to retirement funds, and whether there should be a compulsory suite of benefits. The conclusion is that while there appear to be good reasons for funds to provide death, disability and annuity benefits, compulsory contributions seem to have the effect of creating earlier rather than more satisfying retirements, which provide arguments against high levels of contribution.

Compulsion itself can however be justified. One consequence of the distributional economies of group schemes, and their advantages of integrating insurance benefits, is that they are common goods. As the benefits of common goods cannot be secured

without compelling potential free riders to join the group, this provides good arguments for compulsory membership of group retirement schemes. It also provides further arguments against choice of fund and in some respect investment choice, which Asher (1999) argues, provides more stresses than advantages to members.

The next three chapters consider investment issues. The first of these, chapter 6 (Ethical investment) examines the history of the Community Growth Fund (CGF), set up by trade unions to invest retirement fund moneys. It discusses the theory and some of the practice of using social criteria to choose investments and shows the effects on the performance of the fund. The CGF provides an illustration of ethical innovation that would be extremely unlikely in a government run retirement scheme.

While I was Chairman of the CGF Management Company for 10 years, I did not support the Taylor Committee's investment recommendations that funds invest a proportion of their assets in a defined socially desirable investment universe. Compulsion was not recommended by the Retirement Sub-committee; and I would continue to stand by the points made in Asher and Adler[8]: prescribed assets create hidden and undesirable subsidies. It was however not a point on which I felt sufficiently strongly to include in my dissenting letter.

Chapter 7 (Herding and benchmarking) surveys the current state of investment theory and how it can be best transmitted to retirement fund trustees and members who might otherwise suffer. Recent research suggests that ill-advised investors can lose money not only by inadequate diversification and excessive trading but also by being panicked into buying overpriced or selling under-priced assets. It is suggested that regulators or industry bodies should be active in identifying poorly performing investment managers, and should publish a consensus portfolio to act as a benchmark to reinforce the need for fundamental analysis and appropriate diversification. In the discussion of this chapter[9], it was suggested that it would be better for government to set the benchmark portfolio and restrict members to this allocation. This alternative suggestion however ignores the need to match asset allocation to the particular needs of the members, but also the evidence of

---

[8] Asher, A. & Adler T (October, 1991) "Should Prescribed Assets be Reintroduced?" Trustee Digest, circulated to union elected trustees.

[9] At the 14th Australian Colloquium of Superannuation Researchers, July 2006

Srinivas *et al* (2000) that such restrictions universally reduce returns, apparently by inhibiting market functioning. This chapter thus provides a justification for the Taylor recommendation that the Financial Services Board (FSB) develop performance benchmarks to guide trustees particularly in their choice of international assets.

The last chapter (Chapter 8: Smoothed returns) outlines why retirement schemes should smooth investment returns, and how they can do so in a manner that is fair to all members and consistent with modern financial theory. It argues explicitly that governments should not guarantee the retirement benefits of contributory schemes, but that the beneficiaries should participate in general prosperity through appropriate investment instruments – which mean that retirement arrangements ought to be funded. Merely setting the level and minimum age of a universal flat scheme creates enough political problems. The current European turmoil over excessive, cross-subsidized defined benefit schemes mirrors earlier problems in Latin America and illustrates the point. Chapter 8 does not address any of the recommendations of the Taylor Committee, but is important in that it shows that members do not necessarily suffer if governments do not underwrite pay-as-you-go systems or otherwise provide investment guarantees. Not mentioned in the final version of the chapter is the point that government guarantees of earnings-related pensions involve a subsidy to wealthier members of society that seems difficult to justify.

The thesis concludes with a final chapter that provides a discussion on vocation and wisdom, with a more personal *apologia,* and considers how further research and reform might proceed.

REFERENCES

Asher A (1992) Pension Fund Democracy *Transactions of the Actuarial Society of South Africa* IX.II: 61-91

Asher A (1997) The Actuarially Appropriate Management of Public Pensions: Funding or Pay-as-you-go *Discussion Paper PI-9701* of the Pensions Institute (Birkbeck College, University of London)

Asher A (1998) Effective and Ethical Institutional Investment *British Actuarial Journal* V: 969-1027

Asher A (1999) Loading Members with Investment Choice *Journal of Pensions Management* 5.1: 76-85

Asher A (2000) Designing Benefits for the New Century *Retirement Funding: A Source Book for Trustees* (ed M Dagut) Business Day, Liberty Asset Management, University of the Witwatersrand

Asher A (2003a) Meeting Consumers' Needs in *Understanding Actuarial Management: the actuarial control cycle* ed Clare Bellis, John Shepherd & Richard Lyon, Institute of Actuaries of Australia, Sydney 71-108

Asher A (2003b) Retirement and Old Age (with Marius Olivier) 231-299 and Finance and Tax 595-618 in *Social Security – a Legal Analysis* ed MP Olivier, N Smit & E R Kalula, Butterworths, Durban

Asher A (2005) The Relative Investment Performance of the Community Growth Fund *South African Actuarial Journal* 5:1-26 (Chapter 6 of this thesis)

Asher A (2006) Pensions in Africa in *Oxford Handbook of Pensions and Retirement Income* ed: G L Clark, A Munnell & M Orzsag Oxford University Press (Chapter 2 of this thesis)

Asher A (2007a) Smoothing Investment Returns in *Retirement Funds in Scary Markets* ed H Bateman, Edward Elgar (Chapter 8 of this thesis)

Asher A (2007b) A Christian in Financial Services *Praxis - Journal for Christian Business Management* 14.4: 2-9 (Extracts from Chapters 1 and 9 of this thesis)

Asher A (Forthcoming a) *Conflicted Superstructures: Are Australian Investors shortchanged?* 12th Australian Colloquium of Superannuation Researchers - to be published as a chapter of Selected Papers. Original version can be found at http://wwwdocs.fce.unsw.edu.au/fce/Research/ResearchMicrosites/CPS/Asher.pdf (Chapter 3 of this thesis)

Asher A (Forthcoming b) Mean Reversion in investment markets: the implications for investors and regulators *Australian Actuarial Journal* 13.2 721-770 (Chapter 7 of this thesis)

Asher A (Forthcoming c) Pension Benefit Design and the Integration of Insurance Benefits over the Life Cycle *South African Actuarial Journal* 7:73–115 (An earlier version at http://www.acst.mq.edu.au/docs/research_papers/2007/integrating_insurance.pdf (Chapter 5 of this thesis)

Asher A (Unpublished) *Means Tests: An evaluation of the justice of imposing high rates of claw-back on those of modest means.* Presented to Institute of Actuaries Financial Services Forum http://www.actuaries.asn.au/IAA/upload/public/fsf06_paper_asher_means%20tests.pdf (Chapter 4 of this thesis)

Aquinas Thomas (1275) *Summa Theologica* Translated by Kevin Knight (2006) http://www.newadvent.org/summa/

Campbell J Y, A W Lo & A C MacKinlay (1997) *The Econometrics of Financial Markets*, Princeton University Press

Cantor A C M & J A Sefton (2002) Economic Applications to Actuarial Work: Personal Pensions and Future Rates of Return *British Actuarial Journal* 8.1: 91-150

Cooper D R (2002) Family Fortunes: A guide to saving for retirement *British Actuarial Journal* 8.5: 849-885

Day R E (2000) Tropes, History, and Ethics in Professional Discourse and Information Science *Journal of the American Society for Information Science* 51.5: 469-475

Ferson W E, S Sarkissian & T T Simin (2003) Spurious Regressions in Financial Economics *Journal of Finance* 58.4: 1393-1414

Huber P P & R J Verall (1999) The Need for Theory in Actuarial Economic Models *British Actuarial Journal* 5.2: 377-395

Levitt T (1975) Marketing Myopia *Harvard Business Review* Sept-Oct: 2-14

Lucas J R (1980) *On Justice* Clarendon Press, Oxford

Merton R C & Z Bodie (2005) The Design of Financial Systems: Towards a synthesis of function and structure *Journal of Investment Management* 3.1: 1–23

Milevsky M A (2004a) A Diffusive Wander over the Human Lifecycle: How Quantitative Finance is Transforming the Field of Personal Wealth Management *Working Paper* http://www.ifid.ca/pdf_workingpapers/WP2004FEB_QFIN.pdf

Milevsky M A (2004b) Real Longevity Insurance with a Deductible: Introduction to Advanced-Life Delayed Annuities *Working Paper* http://www.ifid.ca/pdf_workingpapers/WP2004FEB_.pdf

Pemberton J M (1999) The Methodology of Actuarial Science *British Actuarial Journal* 5.1: 115-195

Polanyi M (1962) *Personal Knowledge: Towards a Post-Critical Philosophy* University of Chicago Press

Rich J (1991) Presidential Address *Transactions of the Actuarial Society of South Africa* IX. I: 1-7

Srinivas P S, E Whitehouse & J Yermo (2000) Regulating Private Pension Funds' Structure, Performance and Investments: Cross-country evidence *Social Protection Discussion Paper Series* siteresources.worldbank.org/SOCIALPROTECTION/Resources/SP-Discussion-papers/Pensions-DP/0113.pdf

Stiglitz J E (2002) Information and the Change in the Paradigm in Economics *American Economic Review* 92.3: 460-501

Taylor V *et al* (2002) *Transforming the Present - Protecting the Future. Report of the Committee of Inquiry into a Comprehensive Social Security System for South Africa*, Department of Social Development, Pretoria

Thomson R J (2004) Is Actuarial Science Really a Science? *South African Actuarial Journal* 6:97-103

Trowbridge C L (1989) *Fundamentals of Actuarial Science* Actuarial Education and Research Fund

CHAPTER 2 - PENSIONS IN AFRICA

First published as Chapter 41 "Pensions in Africa" in the Oxford Handbook of Pensions and Retirement Income (2006), ed: G L Clark, A Munnell, M Orzsag, Oxford University Press (816-836)

1 INTRODUCTION

In providing an outline of pension arrangements in Africa, this chapter is particularly concerned with justice. Justice is the traditional criterion for evaluating the performance of governing structures, and this book is intended to describe the theoretical frameworks used to understand the structure and performance of retirement systems. For our purposes, justice can be evaluated by determining whether the policy concerned strives, in a procedurally just manner, to reconcile five sometimes conflicting objectives: equality, liberty, efficiency, provision for people's basic needs, and recognition of their deserts. The main issues raised in this chapter are those where policy appears to fail in achieving these objectives.

Perhaps the thorniest African issues are those of compulsion and governance. Compulsory contributions are an affront to liberty, but are justified on the ground that they provide for basic needs that people would otherwise not provide for themselves, and on the ground of desert, in that people should pay for benefits that the state would otherwise have to provide. Compulsion is also more efficient in that voluntary contributions invariably attract significant marketing costs. In the informal sector however, compulsion is enormously expensive and creates further inequality by acting as an additional tax on the financial and administrative resources of the poor.

On the issue of governance, many would agree with Barbone and Sanchez (1999) that it is the 'first order of business' for retirement reform in sub-Sahara particularly. The inefficiency and corruption of many government structures are clearly unjust. A specific issue in pension arrangements is actuarial fairness that requires the accruing value of future benefits to be closely related to contributions – real or notional. This is a question of desert, and departures ought to be justified as a contribution to greater equality, or as

the provision of basic needs. In Africa, this is a particular problem with civil service schemes.

The chapter is organized as follows. The next section provides some background on the economies of Africa, a short discussion on the causes of poverty and a description of the informal sector. This is followed by a relatively detailed description of the South African pension system. Apart from the author's familiarity with South Africa, its subsistence and informal sectors have much in common with the rest of Africa, while both its public and private sectors provide a number of models for potential development. Its policymakers and entrepreneurs have a growing influence on other African countries, particularly in southern Africa. Section four extends the description to other African countries. A final section also raises some research questions.

2 BACKGROUND

2.1 African economies

Africa makes up almost a quarter of the world's land surface, one eighth of its people but only 4% of its economy even on a purchasing power parity (PPP) basis[10]. For purposes of this chapter, we can divide it into three. South Africa has a diversified modern economy and accounts for some quarter of total economic output, with the balance divided more or less equally between the five Mediterranean countries (including Morocco) and the rest of the continent. In terms of per capita income, on a purchasing power parity basis, South Africa is a medium low income country standing at something under $10,000; the North African countries stand at half of this, while the other Sub-Saharan countries average one tenth.

All African economies are relatively small and concentrated – often dependent on a single commodity. The implication is that all are volatile, and subject to periodic bouts of currency fluctuations and high inflation, except where there are tight foreign exchange controls. This provides one explanation for extensive government intervention in the economy, with its consequential restrictions on private activity and risk of corruption. Economic instability is also invariably linked to political volatility.

---

[10] The statistics in this section are taken from the Nationmaster.com website.

These economic and political risks undermine the security that might be offered by a pension system. It is often not possible to find local investments to provide sufficient diversification. Those that are available will not provide security in times of political and economic uncertainty.

2.2 Poverty

Of the various explanations for poverty, Sachs *et al* (2001) offer little hope in that they suggest that geography is a major cause. The tropics, which include most of the world's poverty and very few prosperous nations, are debilitating environments. In addition, Africa is served by a relatively few harbours and navigable rivers, significantly increasing the cost of transport and trade.

Other analyses relate to institutions and culture, which are not immutable. In economically fragile societies with poorly understood informal institutions, apparently sensible reforms can be counterproductive. Institutional obstacles to development have been highlighted recently by De Soto (2000), in particular that most property cannot be sold, so tying people to their existing land rights and preventing land from being used as capital. He also describes the multiple bureaucratic obstacles to business development in all poorer countries, which often goes hand in hand with corruption. On these points, it can be noted that unmarketable land rights do function as retirement assets and alternatives will need to be found if they can be sold. As already noted, the imposition of formal social security contributions contributes to the bureaucratic burden on enterprise.

Of those who focus on cultural and political issues, Landes (1998) mentions openness to science and trade, trust within society, open government and the value placed on hard work. These are also issues covered by Powelson (1994), who emphasises the benefits of a diffusion of power in society, both as a counter to corruption and a source of personal responsibility and motivation. South African experience described later shows that private retirement funds with democratically elected trustees can contribute to such diffusion.

Somovia (2003), Director General of the International Labour Office (ILO), surveys the immediate causes of poverty: malnutrition and illiteracy, lack of access to markets and capital, and corruption. He calls for the transfer of resources to the poor, debt relief and

the extension of a developed legal and institutional framework. The need for a system of social security to break the life-cycle of poverty is a major goal of the ILO.

2.3 The informal sector

Somavia looks at various elements of the informal economy, which accounts for the overwhelming proportion of employment in Africa, and its interaction with poverty. The informal economy can be defined as that beyond state control: untaxed, unregulated and not using formal legal or accounting systems. It includes subsistence farmers, micro-scale traders and manufacturers, and many domestic workers.

The disadvantages of the informal economy are obvious. Efficiency is low as the division of labour is much more difficult and economies of scale impossible. Access to credit and insurance is limited by the absence of records and formal dispute resolution mechanisms. The latter is also associated with violence. Literacy is of less value, and is therefore encouraged less in the young, which also means that skills are less likely to be passed on. The smaller tax base limits the ability of governments to develop infrastructure and provide social services.

The informal sector, by definition, precludes formal pension arrangements. Outside formal organizations, wage related contributions cannot be collected; money cannot be securely invested, and structured payouts cannot be managed. While formal systems have clear advantages over informal arrangements, it is not clear that the ILO and other agencies understand the significant demands they place on participants. If the demands cannot be met, the costs and the disruption of attempting to create formal systems may outweigh the advantages.

There are also unstable smaller employers on the fringe of the informal sector. Yakoboski *et al* (2001) report the results of their US research which suggests that these employers, and their employees, are reluctant to contribute to formal pension arrangements because of their uncertain future, and the expense of contributing on behalf of lower income individuals. It may be better to use persuasion rather than compulsion to bring them in to formal pension schemes.

3 SOUTH AFRICA

3.1 The first pillar

The first pillar consists of a flat pension, subsidies to old age homes and other support to the needy living alone or with their families. The state's resources are inadequate from many perspectives, but it would seem that the positive social impact does justify the costs.

The main benefit is a monthly old age pension paid out of general revenue, subject to a means test, to men above the age of 65 and women over 60. Recipients must be South African citizens resident in South Africa. The amount, R740 (R3 is approximately U$1 using PPP) monthly for the 2005 fiscal year, although low by OECD standards, is a generous 45% of GDP per capita.

3.1.1 Administration

The Mothers and Fathers Report (2001) contains the following comments that give a flavour of the African setting.

The predominant method of pension receipt, namely queuing on particular days at a specified pay-point, and the problems associated with this over-shadowed all other problems faced by the elderly... Security at pay-points is a great concern of pensioners, particularly when they have to return home after dark… The length of time spent in the queue, lack of shelter, seating, insufficient and filthy toilets and lack of water at pay points are problems across the country. Even where there are halls, pensioners are not necessarily allowed into them. …

People continue to arrive very early and to sleep at pay points. No preference is given to the very frail who have to wait in line along with everyone else. …Many pensioners experience rough and insulting treatment by staff at pay-points. …

Money lenders are active at pay points, take IDs (identity documents) … and harass pensioners. Hawkers and liquor vendors also cause problems… A large number of burial societies, largely unregistered, try to recruit pensioners.

… Security officers accept bribes to let people jump the queues. There are also allegations that pensioners are short-changed by officials….

… Pensioners in hospital have to make their way to pay-points on pension day or face the prospect of losing their pension. Some have died in the queue.

Perhaps the most difficult administration issue arises from the bureaucratic need for rules to determine whether the pensioners are still alive. It seems that such rules often become unreasonable. The South African courts (prompted by non-government organization funded lawyers) are sufficiently reliable to provide some protection, and have recently found certain pension administration to be unjustifiable. One practice that has been outlawed is stopping payment unilaterally in an attempt to counter fraud, and refusing to make back payments in an attempt to save money.

3.1.2 Means tests

An income test phases out the pension at a rate of 50% of other income[11]. Married couples with a combined income over R3,384 do not qualify for any state pension. Those with assets (excluding their homes) of more than R266,000 also do not qualify[12].

Lund (1993) reports that the mechanism is widely misunderstood and inconsistently applied. This should not be surprising. The difficulties of auditing income are immense. Apart from those who are obviously cheating, one just has to think of determining income from casual jobs and the renting of rooms, and of irregular interest, pensions and transfer payments, let alone translating income in kind from subsistence farming[13]. The means test is not, and probably cannot be, enforced. Unenforceable policy is bad policy.

If this were not enough, it is also inefficient. The distribution of income and the failure to implement the test mean that the pension is paid to almost 85% of old age pensioners. The Smith Committee (1995) estimated that the costs of abolishing the test would be some R1.5 bn of which two thirds could be recovered through the tax system. This is

---

[11] The formula for the means tests are set out in the South African Government Gazette 18771 of 31/3/1998

[12] The asset test has not been phased in since changes published in the Government Gazette 22852 of 23/11/2001.

[13] The author has some experience attempting to quantify earnings for tort claims. On one occasion he was able to gather evidence that the in-kind income was ten times higher than that initially claimed.

entirely dwarfed by tax concessions to retirement funds that are justified by the putative savings created by the means tests. The test also offends vertical equity in that the maximum rate of tax levied on high incomes is currently set at 42% while poorer pensioners lose 50% of the state old age pension for every rand of private income. Willmore (2001) expresses surprise that it has not been abolished. International agencies suggesting the introduction of means tests may well do more harm than good.

3.1.3 Non-citizens

Rules to determine the benefits to be enjoyed by immigrants require some balance between equality, need, desert and affordability. This is of particular concern in South Africa where the state old age pension is significantly larger than anything offered by its immediate neighbours. As in many similar situations, the rules governing the acquisition of citizenship can be contested at times.

3.1.4 Social impact

The Smith Committee (1995) records that the amount payable to whites was some eight times that payable to blacks in the mid-sixties, but subsequent reforms led to final equalization in 1993. Even before equalization, social pensions provided an important source of income particularly in poor rural areas. Ardington and Lund's (1994) study in Kwazulu Natal found that the pension income made up more than half the income of a third of rural households. Not only does the pension relieve poverty among the elderly, it turns old people into economic assets rather than liabilities to their families, and it seemed unlikely to create incentives that might distort the labour market or fertility behaviour. Case (2001) confirmed that families with the pension, if it was shared within the family, enjoyed measurably better nutrition and health.

Bertrand *et al* (2003) introduce an element of scepticism with their findings that labour market participation of working age adults in extended families is lower in families with the pension, which decline from an already low 24% and to 21%. The effect was largely explained by the participation rates of older men. From one perspective, this is a matter for concern as the pension appears to make subsistence more palatable and reduce the incentives to move into the formal sector. On the other hand, facilitating dignified early retirement to a rural area can be defended as providing openings for younger workers. The negative labour market incentives do however need to be made clear as they

undermine the arguments currently made in South Africa by supporters of a basic monthly income grant for all.

3.1.5 Institutions for the aged

As in the developed world, long term care is becoming an increasing problem. It is widely accepted that it is desirable for people to remain outside formal institutions for as long as possible, but increasing numbers will inevitably require admittance. Many who require admittance are unlikely to be able to afford the R4 000 monthly cost[14]. The costs in OECD countries of frail care are much higher. Rappaport (2001) reports them as being as much as $6 000 monthly in the USA. South African costs are lower mainly because the wages of caregivers are lower.

Involvement by government or private charities in these homes is inevitable. Old age homes for whites were previously given generous subsidies by the Department of Social Development, but these have been significantly reduced in the last decade, so undermining the viability of many organisations. Subsidies can be justified to support particularly needy individuals, but whites are disproportionately represented amongst those in frail care not least because of their longer life expectancy and smaller families. Subsidies that largely benefit whites are difficult to justify politically so soon after the removal of apartheid.

The Mothers and Fathers Report also gives a graphic description of these institutions:

… some homes are dirty and pervaded by a strong smell of urine. The quality and quantity of food is often below standard …In some homes residents don't have their own clothing, laundry is not marked and face-cloths and toothbrushes are shared. … Many residents seem to spend their days seated in rows around a room or along a veranda - waiting for the next meal.

…Staff are generally low-paid and untrained and few qualified nurses are employed to care for elderly residents. …

The high level of theft in homes could be described as endemic and nobody seems to take responsibility for addressing it. Some homes have been targeted by criminal elements….

---

[14] Personal communication from Peter Asher, board of The Association for the Aged, Durban.

3.1.6 Other social services

The alternative, 'care in the community' may however be little better, as the report continues:

(this) has become increasingly difficult for many old people due to the absence of community services in most areas, inadequate housing and unaffordable service charges. …Many elderly persons are being abused by their children and grand-children. … Elderly people living alone are dying from malnutrition and neglect… Official agencies do not make home visits to elderly people. … A meals on wheels service is run by one national organization but it reaches fewer than 10,000 elderly persons …Specialist geriatric clinics and nurses have been totally phased out. Clinic staff tends to say home care is 'not our business'. Equipment banks and lending depots for wheel-chairs and other aids are not supported by clinics and hospitals and tend to run out of stock. …

3.2 The South African private sector

The private sector consists of employer or occupationally based schemes and voluntary products largely sold by commercial life insurers which provide for top ups, the needs of the self-employed and perhaps some of those informally employed. The Financial Services Board (FSB) (2003) shows that group retirement funds for which they were able to provide statistics have almost 9 million members (but over 1 million are retired and there are a number of duplications). Contributions amounted to R60 billion. To these occupational scheme contributions, another R8 billion in contributions to annuity products with life insurers can be added. A large proportion of the R27 billion of regular premium life assurance is also written as endowment policies intended to mature at retirement. Total assets amount to over R900 billion, and appears to account for some half (or more) of the country's total assets.

3.2.1 Coverage in the organized formal sector

The term 'organized formal sector' is used here to describe businesses registered for tax purposes. These businesses usually belong to one or other labour or employer federations, while self-employed people are likely to be organized in professional associations. While noting that the statistics are invariably unreliable, the South African Institute of Race Relations Survey (2000, 356) reports estimates that this sector accounted for some 75% of remunerated employment in 1996.

Those employed in this sector are invariably covered by private pension arrangements, whether as part of their employment contract or voluntarily. The actual extent of retirement fund coverage is debatable. Some surveys - such as AMPS[15] - suggest that only 50% of employed people are contributing to a retirement fund. This does however not reconcile with the level of contributions to retirement funds reported to the Financial Services Board, which amount to almost 20% of personal remuneration. The 9 million active members reported would be more than the number of formal jobs. The latter figures tend to corroborate the Smith Committee estimates that some 80% of formally employed workers are covered by retirement funds. The AMPS survey recipients may have misunderstood the questions asked. Evidence from around the world, such as from Ferris (2000) and Glass (2001) is that the average person finds the design of retirement arrangements too complicated to understand.

3.2.2 Explaining the high coverage

The South African second pillar is more efficient and has a greater coverage than that of Chile in spite not being compulsory. The unusually wide coverage of the organized formal sector can be explained by a number of factors. Not least of these is the absence of an earnings related state scheme. Of the other factors, and the one perhaps with most lessons outside South Africa, is the focus on group schemes. These have significant advantages over arrangements where individuals have more choice. Relatively generous tax concessions are only available to group schemes where employees have no choice but to join. Such compulsion is an infringement of freedom, but the benefits appear to outweigh this loss.

Foremost of the benefits, and now widely recognised, are cheaper administration and minimal marketing costs. Almost all the information required for administration can be obtained from the employer's personnel records. The contributions are collected in one monthly amount from the employer. Large superannuation funds are also able to negotiate lower investment management fees. The costs of selling group schemes are proportionately much lower than selling to individuals. Keeping charges low can make a difference of 20% or more to final pension payouts. Compulsory group schemes also

---

[15] The All Media and Product Survey conducted by the South African Advertising Research Foundation – private communication.

routinely offer life, disability and health cover without medical questions. Those in poor health, who might not be offered cover at all under an individual contract, are thus able to obtain significant cover at ordinary rates. This generates further cost savings.

The contribution that group schemes can make to industrial democracy is not often considered. Members can participate in the election and monitoring of trustees, who in turn are able to vote the shares owned by the fund. This is not possible with individual retirement contracts. The South African focus on group schemes arises partly from the early start, with the Pension Funds Act (1956) providing focused regulation. There is a widespread perception amongst employers that offering retirement benefits helps retain staff, while trade unions have seen the intrinsic benefits and the opportunities to use the funds as a focus for organization. This has facilitated a number of industry wide determinations, in terms of the Labour Relations Act, requiring employers to contribute to an industry fund or set up a company specific fund. The marketing effort of life insurance companies has often been directed at the employers because of the life companies' influence as major shareholders. This, in turn, has arisen because of their management of retirement fund moneys.

3.2.3 The informal sector

The income over which contributions to retirement or insurance funds become economical is debatable. Participants in the South African industry suggest that monthly administrative costs are unlikely to be kept at below R30, while distribution costs (marketing in the private sector and enforcement in the public sector) may double this. The costs for someone earning even R1 500 monthly (twice the monthly pension) and contributing 15% of their income for retirement would be at least 3% of income (and 15% of contributions).

Requiring people who earn much below these levels to contribute to retirement savings is equivalent to an additional tax. The South African experience makes this clear. Legislative reforms in the nearly nineties freed micro-lenders from the constraints of the Usury Act and led to an explosion of loans at interest rates that vary from 20% p.a. to 80% p.a. or more. The statistics suggest that most employed South Africans have such loans. While the morality of such rates is debatable, they reflect the shortage of funds in the low income and informal sector. At the same time, retirement funds earn net returns

of some 10% p.a. Requiring poor people to borrow at 50% and invest at 10% is clearly exploitative.

Most people employed in this sector rely on the government disability and old age grants for insurance and retirement cover. The current level of these grants is sufficiently high relative to the earnings of the majority of recipients for them not to need further savings. Funding these benefits through consumption taxes is much more efficient and contributes more to equality than collecting contributions. James (1999) also makes many of these points, suggesting that they may be applicable outside South Africa. She also points out that universal pensions provide support for women who have not been active labour market participants and who gain nothing from contribution based systems.

Compulsory contributions appear to me to be only in the interest of people employed in the formal financial sector, and here I would include policy makers and government regulators. Their narrow focus on the formal sector and their self interest appear to combine to blind them to the cruelty of compulsory requirements.

3.2.4 Benefit coverage

In spite of these high levels of coverage, which have been all but universal for the white population for forty years, over a third of the white population still draws the means tested old age pension, with women being disproportionately represented. Two interrelated reasons can be identified. The first is the lack of preservation of benefits on changing employers. The second is that elderly widows suffer because their husbands have spent their withdrawal and lump sum retirement benefits. This is an international problem - as described in Auerbach and Kotlikoff (1991) in a US study. It is an injustice that deserves urgent consideration in the reform of family law.

3.2.5 Types of fund

The earliest South African funds were DC in that they were funded by endowment policies. They were replaced by DB funds during the sixties because of the greater predictability of the benefit in DB funds. A swing back to DC provident funds (that provided lump sums and not pensions) began in the early eighties largely in response to pressure from black trade unions. Black South Africans were first allowed to join trade unions in 1979, and in the early eighties organised a series of strikes over plans by the

government to limit pre-retirement withdrawals. As in many countries, trade unions continue to be involved in fund administration.

Kerrigan (1991) describes the unions' objectives: fairer withdrawal benefits, lump sums on retirement, greater influence on investment policy and the power to elect trustees. The DC funds did away with the complicated cross subsidies of DB, and the lump sums offered greater ease in avoiding the means test that at that stage reduced the state pension by 100% of other income.

Employers encouraged the shift to DC benefits. Firstly, they were reluctant to permit newly elected trustees to make decisions that could lead to investment losses. Secondly, they were not averse to reducing the investment risks inherent in DB design, and towards the end of the eighties, they saw that AIDS threatened a significant increase in the cost of insurance benefits. The swing gained momentum when the high returns on equity investment were seen to lead to better benefits for members.

3.3 Regulation

The standard regulatory structure is based on the Pension Funds Act, which sets up the Registrar of Pensions. The Financial Services Board Act (1990) sets up the FSB, whose chief executive officer assumes the functions of the Registrar. Funds regulated by the FSB are subject to proper governance procedures. They must have a set of rules, a board of management of which 50% are elected by the members[16], report regularly to members and produce audited accounts and actuarial valuations (if they are self-administered). The board of management or trustees owe their primary fiduciary duties, which derive from the Common Law, to members.

A South African innovation is the Pension Funds Adjudicator set up to 'dispose of complaints …in a procedurally fair, economical and expeditious manner'[17]. The intention was to empower members who could not use expensive legal procedures to establish their rights. The first adjudicator was flooded with complaints, not least because of his personal energy and ability to address critical issues in legally creative ways if necessary.

---

[16] Section 7A of the Pension Funds Act (1956), in force since 15 December 1998.

[17] Section 30D.

Some of the complaints were resolved by mediation, but hundreds have required written rulings[18]. In shifting the balance of power towards members, it seems that the innovation has been resoundingly successful. While few will agree with all his determinations, the office and the approach deserve consideration and perhaps imitation in other countries.

The Pension Funds Act has always provided protection to DB members to ensure that the funds are fully funded. Withdrawal benefits were however considerably less than the actuarial reserve so penalising early leavers, and inflation was often allowed to erode the real value of pensions significantly. Members transferring to DC funds also suffered losses. Employers taking contribution holidays benefited from the surpluses thus generated. In a relatively high inflationary environment, these flaws in benefit design accelerated the move to DC. Unprecedented legislation in 2001 requires funds to use their surpluses (where they exist) to go back to 1980 and compensate pensioners and those who have withdrawn.

3.4 Government employees' schemes

South African government sector employees are generally covered by DB schemes. There is one for central government employees, others for local government, and various para-statal organizations. Unusually, these are actuarially valued and fully funded, although the assets of the DB funds are largely held in government stock.

During the eighties, however, the funds were removed from actuarial oversight and quickly reduced to some half their previous funding levels. The reasons partly arose from the inflationary erosion of fixed interest assets, but also from excessive benefits, particularly the notorious 'buyback' provisions that gave enormous discounts to senior members of the funds. Wassenaar's (1989) exposé which led to the re-imposition of proper controls reports that this bounty was limited to about 45 000 more senior public servants. He estimated the cost at R5 billion in 1986 terms, but the total deficits required perhaps ten times that number in additional contributions (and tax concessions) before they were extinguished in the mid-nineties. Another abuse of the system was to artificially inflate final salaries by promotions just before retirement - as the rules based the pension on the last day's salary.

---

[18] Some 162 are recorded in 2001 on the website: http://www.fsb.co.za/pfa/deter2001.htm

4 OTHER AFRICAN COUNTRIES

Many of the South African issues have parallels in other African countries. Detailed information on the benefits available is however not as readily available. One source is 'Trends in Social Security' published by the International Social Security Association (ISSA) and available on their website to members. It gives some indication of the nature of retirement arrangements in different countries by reporting on planned changes. Barbone and Sanchez (1999) give some helpful tables.

4.1 Civil service schemes

Separate DB schemes for the civil service are the norm internationally. They are frequently overly generous: the problem is not limited to developing countries such as South Africa and Brazil, with the Actuary (2004, 14) reporting unfunded liabilities in the UK of over half that country's GNP. Benefits for military are often particularly favourable with very young retirement ages. In poor African countries, the pensions of government employees do not have to be particularly generous to absorb a large proportion of the state's resources. ISSA data reports retirement ages for civil servants that are being raised to 60 in some countries, and that the contributions required for the Egyptian scheme amount to some 30% of remuneration. Both suggest generous scheme benefits.

This is not to object to DB arrangements *per se*, but accruing benefits should be clearly defined and bear a reasonable relationship to current income. Actuarial management is necessary for this, but insufficient to prevent abuse. Reform is particularly difficult as those in the government departments responsible for drafting legislation are frequently the main beneficiaries. They are well placed to defeat the intentions of any reforms – not least by last minute surreptitious insertions.

4.2 National schemes

Most African countries have set up national schemes for all those of working age, that, as discussed, necessarily exclude those employed in the informal sector. In some cases, this is recognised in the regulations of the scheme. In other cases the informal sector is legally included, in which case the law is often an empty letter or another source of confusion and corruption. Coverage is higher in North Africa where it is reported to

exceed 80% of the workforce in Tunisia and Egypt, but Barbone and Sanchez show most countries have coverage of less that 10%. This may, of course, represent a high proportion of the formal sector outside of government employment.

4.2.1 Provident funds

A number of ex-British colonies, as elsewhere in the world, introduced national provident funds around the time of independence. These initially provided lump sums at retirement and some ancillary insurance benefits in return for contributions from private sector employees. Lump sums rather than pensions are particularly attractive to members who do not have access to the banking system in retirement. This is the case for much of Africa, where pensioners in rural areas can be charged exorbitant fees by local shopkeepers to cash cheques - if they arrive in an unreliable post. Lump sums can be used to upgrade housing and purchase cattle, traditionally the source of wealth in Sub-Sahara.

A number of African provident funds failed because of inadequate investment returns that reduced payouts to trivial amounts in some cases. Governments make it difficult to make international investments, and there are not enough local opportunities. The funds thus proved vulnerable to all the problems listed in section 2.1 above: political interference, economic volatility and inflation. Mounbaga (1995) illustrates from Cameroon. Some of the African funds are now being converted into national pension schemes. Nigeria is an exception having just chosen to follow the Chilean model. The different influences of the ILO and the World Bank have presumably contributed to this difference of approach.

4.2.2 Earnings related schemes

The ex-French colonies introduced national earnings-related pension schemes for private sector employees. These protect the members from direct investment losses, folding the problem into general government finances, which are thereby placed under further pressure.

One apparent exception, which is described by Chaabane (2002) and receives favourable mention by van Ginneken (2003), is that of Tunisia. Imaginative and energetic action appears to have been successful in extending coverage to the self-employed and informal sector. Chaabane does not however report on the efficiency of the approach taken and the

exercise is too recent, and the government too authoritarian, to be confident of ongoing success. Chourouk (2003), in a rather confusing paper, suggests that the Tunisian and other North African schemes are facing financial difficulties.

4.2.3 Non-contributory schemes

Willmore (2001) describes the four countries he says have universal systems. Three (Mauritius, Botswana and Namibia) are in Africa and have been influenced by the South African experience. He reports that the schemes are popular in New Zealand and Mauritius where they have been established for some time, without being too expensive (4% and 2% of GDP respectively). In spite of the administrative difficulties mentioned in 3.1.1 above and acknowledged by Willmore, universal pension schemes are probably the most cost effective way of getting money to poor and rural areas.

4.3 Other arrangements

4.3.1 Large private sector employers

The employees of multi-national companies are invariably covered by pension funds as part of the companies' world wide human resources policies. These funds do not always offer the same protection to locals as to ex-patriates. In some cases, there is discrimination in eligibility for membership; in others different practices on withdrawal.

In Anglophone countries with a less extensive first pillar, particularly Kenya and Zimbabwe, larger local companies also have corporate schemes. As in South Africa, both employers and employees appear to have encouraged the move to DC arrangements – in spite of investment restrictions on international investments. Members have not however benefited from all the legislation of the type mentioned in section 3.3 above.

Anecdotal evidence is that the mortality experience of group schemes does not reflect the increases expected from AIDS deaths because local employees return to their rural homes to die, and so do not claim their entitlements. The author's experience of South African funds makes this credible: large numbers of retirement fund beneficiaries do not claim. Illiteracy plays a role, as does the cultural reluctance of men to let their wives know of potential life cover benefits.

4.3.2 Voluntary arrangements

Van Ginneken (2003), from the ILO, suggests the development of voluntary special schemes to give the self employed access to pension benefits. He envisages community based non-profit organizations such as some of those involved in micro-financing either offering pensions or acting as agents for large formal institutions. Burial and rotating credit societies provide other examples that clearly flourish in Africa. He suggests that they should be subsidized in order to remain attractive to all members - because they are not actuarially fair.

There are a number of models that have historically provided something of this type of benefit. 'Industrial' life insurance involves the weekly collection of premiums and pays out lump sums. At its peak in the UK, more than half the households contributed. It also had some success in South Africa. It has however been phased out largely because it is impossible to give decent value for money. The costs of collection invariably make up more than 30% of the premiums.

Friendly societies provide an alternative, community based, and therefore cheaper, model that goes back to the European middle ages, and has proved itself in a number of social and economic environments. The British version pays disability benefits predominantly, which – if there is no retirement age – is precisely what is required for poorer communities where voluntary retirement is a luxury.

Subsidies to voluntary schemes are likely to be abused, and are regressive in that only the wealthier members of society are likely to be able to save. James (1999) warns specifically against them. The first step rather is to remove obstacles to the formation of voluntary arrangements in the form of inappropriate regulations. Subsequent work is to provide financial and technical advice, and appropriate consumer protections. The need for the latter may not always be evident. Informal arrangements are effectively beyond the reach of formal law and can be unreliable. The detailed interviews with a number of burial society members, reported in Thomson and Posel (2002), illustrate the great difficulty of building trust and enforcing agreements in informal arrangements.

4.3.3 The extended family

Kaseke (1999) writes of the extended African family and its role in the provision of financial and social support to the aged and other needy members. The large families that

have characterised the rapid population growth of the last century – and which thereby can be called 'traditional' – do provide a level of financial security through the sharing of risk. Certainly, his view that many children provide security in old age appears to be held widely. The sharing also appears to go beyond the extended family to village communities.

Pensioners too old to work do receive support from their families, but Peil (1992) shows that many do not seem to have contact with adult children. Somavia (2003) reports that more than 40% of Africans over 64 are still obliged to work. AIDS has greatly aggravated this problem, which is often made worse by the need to care for orphaned grandchildren. If it can be afforded, even a small old age grant seems desirable.

4.4 Investments

Nineteen African countries have local stock exchanges[19], although most are small. Plans to rationalise some of them will not only improve liquidity and efficiency, but also provide potentially more diversity for retirement fund investments. Real diversification will require investment in economies removed geographically, as both economic and political risks can be concentrated regionally. Except in South Africa, the capitalization of these is too small to absorb more than a fraction of funded pension liabilities.

Insulation from political risk would mean that the retirement funds themselves would have to be outside the control of their national government. Anything less would be of limited value, but it is not entirely clear whether governments can be persuaded to lose control of such a large pool of assets.

4.5 Governance

In suggesting reforms that will aid governance, Barbone and Sanchez (1999) advocate simplification of institutions, reducing the role of government and formalising the management of the national schemes. This should go along with a reduction in regressive subsidies particularly, and with increased competition. Simplification clearly contributes

[19] Found, October 2004 at http://allafrica.com/business/

to efficiency, but illiteracy should not be confused with the inability to cope with complexity.

Few would disagree on the need to address the accompanying and more difficult issue of corruption, which might well serve as a heading for this section. Addressing corruption will be easier in simpler systems where accountability is easier to enforce and power is diffused from government to private but accountable institutions. Diop (2003) lauds reforms in Senegal, which have created a more autonomous management of the social security system, and appear to offer greater stability as well as improvements in service. The way in which South African unions have both used pensions as an organising force and contributed to pension fund democracy provides an interesting model of power diffusion

Given their resources and temptations, it is doubtful whether the public sector in most sub-Saharan countries has the capacity to administer even simple national retirement funds. If it is accepted that the overwhelming majority of the population in the informal sector would be better off without compulsory schemes, then there is a strong argument for their abolition. The energies of the state would be better directed towards providing a low level of universal cover and the protection of members in voluntary private sector arrangements. The South African experience shows that these can cover a high proportion of the formal sector.

5 CONCLUSION

5.1 Research questions

We know surprisingly little about the actual and potential effects of pensions and social security benefits on the lives of people, especially those living in the third world. What are the costs of levying contributions on informal sector workers? Are means tests more effective than suggested here in targeting the poor and ameliorating their deprivation? How do families react to unexpected losses of income and of pension benefits? In order to answer these questions, I believe one has to have panel data as suggested in Asher (2001). Cross sectional studies record the memories and impressions of those interviewed, which are often unreliable.

Research also has a role in exposing and addressing corrupt practices. Careful analyses of institutional structures and the interests that they serve is a necessary first step to reform. Three South African examples can be provided. Wassenaar's (1989) book provided much of the impetus for the subsequent withdrawal of the over-generous state benefits. The recording and publicity given to administrative corruption over many years has clearly brought some measure of relief to state pensioners. The success of the pensions fund adjudicator has depended partly on adequate research, and on the development of the legal protections available.

The role of retirement funds in the development of local capital markets also requires some exploration. This relates to the management of the investments of national schemes and to exchange control and other pressures placed on private funds to invest locally. Africa would not appear to provide many role models but experience elsewhere may prove helpful.

5.2 Policy lessons

Justice means, *inter alia*, defending the powerless against the powerful. In Africa, the poor and powerless are to be found mainly in the informal sector. Pension arrangements for them need careful consideration; a number of commentators have come to the conclusion that a universal non-contributory pension is the solution.

African civil services are particularly powerful, and need to be constrained. Actuarial management of their own pension schemes may help to reduce their ability to extract excessive benefits from this source. Limiting their involvement in national schemes and encouraging second pillar group arrangements may help develop countervailing powers.

The interests of women also need consideration. Non-contributory schemes provide more equitably for those who have not worked in the formal sector. Family law reform is particularly necessary to protect married women whose husbands consume more than a fair share of lump sum benefits.

Finally it should be said that addressing these issues from outside the continent is fraught with dangers. Injustices provoke indignation, and foreigners might do best to support the indignant who have been wronged by providing the fruits of research and technical if not material support. As the New International Version has it: 'speak up for those who cannot speak for themselves'... Particular examples worth considering are the efforts of

the trade unions, the Pensions Fund Adjudicator and the private support for Court actions in South Africa.

REFERENCES

The Actuary (2004) Danger – unexploded tax bombshell *The Actuary, September.* Staple Inn Actuarial Society, London

Ardington E & F Lund (1994) *Pensions and Development: The Social Security System as a Complementary track to Programmes of Reconstruction and Development.* Centre for Social and Development Studies. University of Natal

Asher A (2001) *South African Panel Study in Income Dynamics* Annual convention of the Actuarial Society of South Africa http://www.assa.org.za

Auerbach A J & L Kotlikoff (1991) Life Insurance Inadequacy - Evidence from a Sample of Older Widows *NBER Working Paper No. w3765*

Barbone L & L-A Sanchez (1999) Pensions and Social Security in Sub-Saharan Africa: Issues and options *Africa Region Working Paper Series No. 4* (The World Bank)

Bertrand M, S Mullainathan & D Miller (2003) Public Policy and Extended Families: Evidence from pensions in South Africa *World Bank Economic Review*17.1: 27-50

Case A (2001) Does Money Protect Health Status? Evidence from South African Pensions *National Bureau of Economic Research Working Paper* 8495

Chaabane M (2002) Towards the Universalization of Social Security: The experience of Tunisia *Extension of Social Security Paper No.4.* International Labour Office, Geneva

Chourouk H (2003) Pensions in North Africa: The need for reform *The Geneva Papers on Risk and Insurance* 28.4: 712-726

De Soto H (2000) *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else* Bantam, London

Diop A Y (2003) Governance of Social Security Regimes: Trends in Senegal *International Social Security Review* 56.3-4: 17-23

Ferris S (2000) 'Women Making Choices' *Macquarie Research Paper* (www.actuary.mq.edu.au/research_papers/200001_women_making_choices.pdf)

Financial Services Board (2003) *Annual Report* (http://www.fsb.co.za/)

Glass R D (2001) Investment Education or Advice? That's the wrong question to ask! *Pensions World* March 28-33

James E (1999) Coverage under Old Age Security Programs and Protection for the Uninsured - What are the issues? *Policy Research Working Paper 2163* The World Bank

Kaseke E (1999) Social security and the elderly; the African experience *The ACP-EU Courier* 176: 50-52
http://europa.eu.int/comm/development/body/publications/publications_courier_en.htm

Kerrigan G K (1991) The Role of COSATU Affiliated Trade Unions in Retirement Provisions in SA *Transactions of the Actuarial Society of South Africa* IX.I: 177-198

Landes D S (1998) *The Wealth and Poverty of Nations: Why are some so rich and others so poor?* (W.W. Norton, New York: ISBN: 0393040178).

Lund F (1993). Social benefits in South Africa. *International Social Security Review* 46.1.

Mothers and Fathers (2001). *Mothers and Fathers of the Nation: The forgotten people*? Ministerial Committee on Abuse, Neglect and Ill- Treatment of Older Persons: Executive summary. http://www.polity.org.za/govdocs/reports/welfare/2001/elder.html#content

Mounbaga E (1995) The Investment of Social Security Reserves during Periods of Crisis: The experience of Cameroon *International Social Security Review* 48.2: 49-59

Peil M (1992) Family Help for the Elderly: A comparative study on Nigeria, Sierra Leone and Zimbabwe *Bold*, *Quarterly Journal of the International Institute on Ageing* 2.3: 2-4

Powelson J P (1994) *Centuries of Economic Endeavour: Parallel paths in Japan and Europe and their contrast with the Third World* (University of Michigan Press, Ann Arbor)

Rappaport A M (2001) 'Setting the Stage: An overview of the Retirement 2000 issues' *North American Actuarial Journal* 5.1: 1-11

Sachs J D, A D Mellinger & J L Gallup (2001) The Geography of Poverty and Wealth *Scientific American* 284: 70-76

Smith Committee (1995) *Report to the Minister of Finance of the Committee on Strategy and Policy Review of Retirement Provision in South Africa* (Government Printer, Cape Town)

Somavia J (2003) *Report of the Director-General: Working out of Poverty* International Labour Conference (International Labour Office, Geneva)

South African Institute of Race Relations Survey (2000) *Annual Survey 1999/2000*, Johannesburg

Thomson R J and D B Posel (2002) The Management of Risk by Burial Societies in South Africa *South African Actuarial Journal* 2: 83-128

van Ginneken W (2003) Extending Social Security: Policies for developing countries *Extension of Social Security Paper No. 13* International Labour Office, Geneva

Wassenaar A D (1989) *Squandered Assets* Tafelberg, Cape Town

Willmore L (2001) Universal Pensions in Low Income Countries *Discussion Paper* Initiative for Policy Dialogue, Pensions and Social Insurance Section, Oxford Institute of Ageing

Yakoboski P, P Ostuw & B Pierron (2001) The 1999 Small Employer Retirement Survey: Building a better mousetrap is not enough *North American Actuarial Journal* 5.1: 131-132

CHAPTER 3 - MEANS TESTS: AN EVALUATION OF THE JUSTICE OF IMPOSING HIGH RATES OF CLAWBACK ON THOSE OF MODEST MEANS

Presented to the Institute of Actuaries of Australia Financial Services Forum
11 and 12 May, 2006

1 INTRODUCTION

Integration between flat state pensions, which make up the first pillar of a pension system, and other sources of retirement income is achieved in many countries[20] by means tests that reduce the first-pillar pension. It is currently the subject of some debate in the UK and South Africa, and was also subject to detailed investigation in Australia during the mid-nineties.

This paper is written to evaluate the recommendations of the Taylor Committee of Inquiry into a Comprehensive Social Security System for South Africa: means tests should be abolished and the taxation of retirees and retirement funds increased in order to compensate. It attempts to focus on these two interrelated issues.

The conclusion is that in South Africa, Australia and the UK means tests cannot be justified in their current form. The standard justification for means tests is that they make state pensions more affordable by targeting them towards those in greatest need. Means tests can however be seen as a tax on other income (and assets) in which case they are unrelated to questions of expenditure and affordability. A more careful evaluation of means tests against the different criteria of social justice suggests that the major problems are those of an unwarranted interference in the lives of pensioners and the efficiency of administration and incentives. It is not so much about the money but the mechanisms of the means test.

---

[20] Willmore (2001) reports that only 4 countries have universal non-means-tested schemes: New Zealand, Mauritius, Botswana and Namibia. Hancock *et al* (2005) say that 18 out of 30 OECD countries include a means tested component in their state systems. The rest offer a minimum pension from an otherwise earnings-related mandatory pension system.

1.1 Structure

Part 2 is the main part of the paper. In it, the policy issues governing means tests and the taxation of pension systems are systematically evaluated against a range of criteria required by justice. The criteria used are need, equality, efficiency (covering administration and incentive impacts), liberty (including freedom from unnecessary administrative interventions) and just deserts. As is appropriate when considering political questions, the third part considers the institutional environment and the interests of the most important participants in the institutional structure – private and public. The paper then goes on to suggest possible reasons for the absence of debate and the perpetuation of means tests, and makes recommendations as to how the policy issues can be more systematically addressed.

1.2 Scope

The first-pillar pension under consideration is that paid from general revenue to those over some relatively arbitrary retirement age. The needs of those below this age are not considered. I am concerned here with arrangements in South Africa, Australia and the United Kingdom (UK), being most familiar with their systems. The arrangements of interest are the state pension and means tests, how they are determined and administered, and to a lesser extent the tax treatment of contributions, investment income and benefits paid by pension or superannuation arrangements.

The paper includes no detailed analysis of the variety of different allowances and variations, although there is some discussion of the principles of grandfathering arrangements. There is also no specific costing of any proposals, as such specifics relate only to a particular year's government budget, and irrelevant to the debate about general principles.

Means tests and pension and superannuation fund taxation are not always considered together. From a policy perspective, however, it is difficult to consider one without the other. The question is discussed from different perspectives at other points in this paper. At this point it is suggested that there must be a *prima facie* case that the fungibility of money applies to taxation and the clawback of social security benefits required by tests of income.

2 JUSTICE

This section first looks at the relevance of justice to this topic, and then applies the five criteria mentioned above to the question of means testing and taxation. The criteria used find common justification in our recognition of the dignity of every person, which has a symmetry with our own self-respect. While they are debated, the criteria are not novel and would have to be mentioned in any serious application of moral philosophy to financial or social questions, as they are in chapters 7, 9, 10 and 11 of Hausman and McPherson (1996) for instance. Their application to the question at hand, however, does not appear to have been done systematically.

Similar analyses of the related subject of social security and the taxation of investment income have been performed. Diamond (1977) evaluates the US social security system explicitly in terms of redistribution, market failure, paternalism and efficiency. McNulty (2000) examines taxation issues using the criteria of fairness, economic allocative efficiency, welfare, and simplicity, with fairness incorporating notions of equality, reward and retribution. Fairness is also used in this paper to describe this particular subset of the justice criteria.

Discussions on the relevance of justice frequently bring forward the arguments that the concept is personal and that it is irrelevant to the way people behave. Neither is entirely true.

2.1 Personal

People certainly place a different personal value on the different elements of justice. Europeans appear to place a much greater value on equality and a lower value on liberty than North Americans (Alesina *et al*, 2004). Union members usually place more emphasis on equality than managers - and less on desert. MacIntyre's (1988) different intellectual traditions of rationality and justice provide different rationalizations for this variety, but it would be a mistake to see the variety as proving that our views of justice are entirely personal with no wider application. Without the idea of justice, government could be based only on the whims of the powerful. There is widespread agreement on the procedures required to attain it, and much on its constitution. There is near universal agreement of what constitutes gross injustices: exploitation, corruption, etc.

Tyler and Smith (1995), in a review of the social science literature on justice, show that people are often more concerned with just procedures than just outcomes. The argument being made here is that that each of the criteria has an absolute value in the process: *ceteris paribus* it is never wrong to value people's needs, create greater equality, efficiency and liberty or give people their just deserts. The subjective personal differences arise from the compromises that practicalities force on us. In another paradigm, justice gives to each person a set of minimum human rights. Our indignation is justified if the criteria or the rights are not considered in the process. To think otherwise is confused, deviant or malicious. The critical issue is the process rather than result: one might justifiably consider a criterion or right and decide to override it in particular circumstances.

2.2 Relevance

The view that justice is irrelevant to people's behaviour can only be sustained by a cynicism that not only holds that people exclusively pursue their own self-interest but that they are not prepared to compromise for the common good.

While this view may have some traction in the economics literature as justified for instance in Stigler (1982), it finds minimal acceptance in any other discipline. In politics and philosophy justice provides the rules that allow for the production of the common goods of security and the other benefits of good government. As Lucas (1980) puts it: "Justice is the bond of society ... the condition under which I and every man can identify with society, feel at one with it, and accept its rulings as my own." Even the most self-interested must take note of it if they are not to be excluded from society.

From the other social sciences, Tyler and Smith (1995) confirm that:

> "People's actual behavior is also strongly linked to views about justice and injustice. A wide variety of studies link justice judgments to positive behaviors such as willingness to accept third-party decisions; willingness to help the group; and willingness to empower group authorities. Conversely, other studies link the lack of justice to sabotage, theft, and on a collective level, to the willingness to

rebel or protest. In other words, how people feel and behave in social settings is strongly shaped by judgments about justice and injustice."[21]

2.3 Liberty

This is perhaps the easiest criterion to deal with. In general, people do not like to be told what to do. Intervention of government in the lives of people is undesirable *per se.* Where it is necessary, it should be minimized by number, extent and impact:

- One intervention is better than two.
- Intervention in one life is better than intervention in two.
- An intervention that takes one hour or costs one dollar is better than one that takes or costs two.

The implications in this context follow:

- A single system to allocate money to individuals and to recoup from others is preferable.
- It should have as few rules as possible.
- It should require the least possible effort from those involved.

With means tests, the work of the tax authorities in making income assessments is duplicated by the welfare authorities. The tests are applied at much lower levels of income than those considered for tax purposes and also apply to assets ignored for tax purposes.

It can be noted that apparently respectable discussions of social security issues, such as Feldstein (1985), assume that a large proportion of the population are myopic and can be justifiably compelled to act in a particular way. In the case of pensions, they can be forced to contribute, or be subject to the indignity of detailed investigation of their means. This line of thinking appears to fall short of justice in that it would require a careful analysis before deciding people were myopic. Even if they were found to be so, procedural justice would require there to be a careful balancing of the advantages of intervention against the loss of liberty.

---

[21] Page 1 with references to other studies removed.

Hancock *et al* (2005) suggest that the British pension-credit may provide a test of extent of the indignities created by an intrusive evaluation of means. The UK pension credit is a means-tested top-up of the non-means-tested basic state pension. The take up rates are relatively low: the explanation appears to be that the costs of claiming – in terms of effort or social stigma – outweigh the benefits. A means-tested pension is devalued by the cost of obtaining it.

2.4 Needs

Meant by needs is the wherewithal for a minimum acceptable standard of living. If the retirement age is set so high that those above the age are not readily able to work for a living, need provides a ready justification for the payment of a pension to those without another source of support. It does not however provide a justification for means tests.

2.4.1 The minimum standard

Open to debate is what constitutes a minimum acceptable standard of living. It must clearly include a place to stay, enough for basic water, food and clothing and basic health care. This paper is not intended to debate what constitutes an adequate level as it is clearly highly context specific. Minimum pensions, in the countries discussed here, also appear to be somewhat higher - for most recipients - than that required for bare necessity. This is contentious, but the evidence is that even the relatively small South African pension raises whole families out of poverty - as, for instance, suggested by Case (2001).

Contention about this issue arises from a common argument that state pensions are required to deliver pensioners from poverty. A recent UK government publication[22], for instance, boasts that millions of pensioners have been lifted out of "absolute poverty" which is defined in a footnote as 60% of the national median disposable income. This is however a measure of relative not absolute poverty: it relates to equality rather than need. The argument from need suggests necessity and urgency. A view that the large proportion of the population of Australia or the UK living on their respective state pensions are in real economic need does not seem tenable.

---

[22] Principles for reform: the national pensions debate published by the Department of Pensions and Works (dwp.gov.uk/publications/dwp/2005/pensions_debate/)

In South Africa, 40% of children are reported as being clinically malnourished. Justice would apply the same standard of need to all members of society – not just pensioners. In the face of the needs of the families that do not include a pensioner, the argument from need - to increase the minimum income of pensioners – appears impossible to sustain.

2.4.2 Perceived needs

To check this conclusion, one can ask whether poorer people actually consider their retirement income to be sufficient. In an international survey[23], AXA finds that significant numbers of retirees do regard their income as insufficient. This varies from 30% in Canada to 60% in Japan, with Australia at 36% and the UK at 40%. The US percentage is 31%, which varies from 15% for individuals with incomes in the top third of their survey to 64% for the low income third. This latter breakdown is unfortunately not available for other countries. One conclusion that may be drawn is that the absolute level of income is not the only issue that leads to dissatisfaction .This again suggests that equality and not need is more likely to be important when setting the level of a state pension.

In their evaluation of public attitudes to pension policy in Australia, Evans and Kelley (2005) report weak evidence that older people believe that pensions should be lower than younger people think the level should be. This again suggests that need is perhaps not an issue.

2.4.3 Targeting

There are grounds for believing that there is no practical way of targeting benefits precisely at those that need them. The nature of the needs is complex and fluctuating; the measurement is inevitably somewhat simplistic, limited to a particular time and prone to error. Where attempts have been made at triangulation (measuring needs using different methods including surveys of the views of the poor and their neighbours), significantly different answers are obtained[24]. If it is true that pensions are invariably large enough to

---

[23] AXA (2006) AXA Equitable Retirement Scope: Retirement — a new life after work? Report axaonline.com/axa/pressroom/2006/2006_01_24_Retirement_Survey_International.pdf

[24] Two working papers from the Centre for International Studies, University of Toronto illustrate this point. (www.q-squared.ca/papers.html) (See overleaf)

meet all essentials, there is then a strong argument for a pragmatic approach to measuring income and needs. Getting into too much detail represents a spurious level of care, and is an unwarranted interference in the lives of the poor. Income – as defined by the tax collecting agencies - is probably an adequate measure of lifestyle for most purposes.

2.4.4 Allowances

Cases of exceptional needs can be addressed by specifically targeted social security allowances: housing and care particularly. In South Africa, the departments of health, housing and education all provide means-tested benefits. Local governments and state owned enterprises provide other allowances or price reductions to those with special needs, of whom pensioners figure prominently. Their design and implementation however requires considerably more effort than has been allocated.

These allowances appear to offer little in the way of a systematic approach to the provision of needs, but are often a further offence against liberty with their intrusive questions and their artificial incentives to distort spending patterns. Cheaper prices for some items in a consumption basket often come with a higher cost in others. Rural areas are a case in point with cheaper housing and perhaps food but much higher costs of transport and health care[25]. How can a local government tell that subsidies for urban housing are targeted at the neediest? There is a vast array of different needs and expenses, and the enormous difficulty in calibrating benefits to meet them undermines the rationale for having specific allowances.

Social security departments do not normally appear to make any systematic evaluation of the standards of living of different groups of recipients. Travers (1995) makes this point in the Australian context. Perhaps as a result of a recommendation in the Taylor Committee report, the South African Department of Social Development has recently

---

Appleton S & D Booth (2005) Combining Participatory and Survey-based Approaches to Poverty Monitoring and Analysis

Barahona, C & S Levy (2005) The Best of Both Worlds: Producing National Statistics using Participatory Methods.

[25] The detailed Financial Diaries survey of the expenditure habits of 150 poor families in South Africa found rural families spend 2.5% of their incomes on medical care, as against the 1% spent by wealthier urban families. See www.financialdiaries.com/files/Financial%20Diaries%20Medical%20Focus%20Note.pdf

commissioned a Human Science Research Council's survey of social grant beneficiaries: two waves of a panel survey were performed in 2005. This is to be welcomed and perhaps imitated. If the welfare departments cannot measure needs, however, it would not seem that other arms of government are likely to be able to do so.

These other means tests are not the particular concern of this paper as they impinge on the needy other than the aged. To the extent that they are important to pensioners, two points can be made. The first is that any analysis of needs should be holistic and take all means and expenditure into account. This may be appropriate in managing long term care for the aged and possibly in offering subsidized housing.

The other point, deriving as much from the principle of efficiency as from need, is that some of the basic needs or rights can profitably be offered in kind. This is normally the case for medical care, which is effectively free in all three countries for those prepared to queue at public hospitals. Many South African municipalities have introduced a "life-line" tariff for utilities that provide a minimum supply of water and electricity free to every household. These are to be preferred to means-tested subsidies.

There are other types of self-targeting benefits that may not be appropriate. Hancock *et al* (2005) were quoted earlier in suggesting that means tests introduce various costs and create a social stigma that helps target pensions at those that need it most. There is some attraction in this type of self-targeting, but it should be noted that the neediest people may be excluded from benefits. It is far from clear that social stigma will select the neediest; it will certainly select the shameless, but the impoverished prim may prefer to starve.

2.5 Equality

If pensions are greater than the amounts required by urgent need, equality provides justification for their current size. It is however difficult to find a justification for means tests arising from this criterion. One possible exception is raised in subsection 2.6.2.

2.5.1 The principle

Equality refers in this case to equality of welfare, interpreted broadly, of income and of status. For Christian, socialist and economist RH Tawney: "What is repulsive…is that some classes should be excluded from the heritage of civilization which others enjoy, and

that the fact of human fellowship, which is ultimate and profound, should be obscured by economic contrasts, which are trivial and superficial." [26]

Equality can be variously justified as desirable for its own sake, as maximizing society's total utility from the widely observed fact that marginal utility tends to decline with income, for its impact on improved health (Wilkinson, 2005), for its contribution to the development of mutual trust and social capital so discouraging crime (Uslaner, 2002), or because it leads to greater happiness (Layard, 2005), although Alesina *et al* (2004) suggest that this does not apply universally. Tyler and Smith (1995) refer to extensive documentation of how feelings of relative rather than absolute deprivation underlie much social unrest: it is the inequality that seems to provide the spark.

In order to clarify a point for the intended audience of this paper, it should be emphasized that the argument here is that – if there were no other considerations – everyone should, morally, have an identical amount of money to spend. This is clearly a fundamental assumption of those who would incline to the political left. It is also, just as clearly, not recognized by some on the political right. The consequence of the model of justice used here is that failing to recognize equality as a principle is wrong. (Having considered the desirability of equality, one might however make a just decision that does not promote equality.)

As this may appear a rather bold claim, some further discussion may be warranted. The argument for equality of spending power is analogous to that of equal votes. They both depend on the essential equality of persons (of adult age and perhaps excluding a limited number of others for reasons of their personal competence or as punishment.) As an example of its application which may be easier to accept, one might consider the shares given to each adult in the privatization of some state-owned firms in Eastern Europe. The principle may be more acceptable if it is couched in such terms: if there is largesse to distribute, it can be fair to distribute it equally.

Equality should incorporate both horizontal equity, which means like people are treated in the same way, and vertical equity, which means that differences in treatment are proportional to the differences between the people concerned. In terms of taxation,

---

[26] Hausman and McPherson (1996, 138) quoting Tawney (1931, 139) *Equality* New York: Harcourt, Brace and Co.

horizontal equity requires those of equal income to pay equal amounts of tax. Vertical equity requires that the rate of tax applied to people of different income increases smoothly with their income and requires increasing marginal rates of tax.

2.5.2 The measures

A case can be made for a system of benefits and taxation that is more cognizant of purchasing power, but for the sake of practicality, this section only considers net income available for spending. Equality would be fully served if income tax and social security benefits left everyone with the same spending power.

Mehran (1976) considers some of the single measures of equality that provide some measure of the dispersion from complete equality, the best known being the Gini co-efficient. Single figure indices however often hide as much as they reveal, and there are considerations not captured by any one of them.

Rawls (1971) suggests that – rather than equality – people would choose a maximin principle – that maximized the income of the poorest - if they did not know where they would fit in the income distribution. While this may be extreme, the income of the poorest people may be considered of special importance. This is related to the object of equality of status to the extent that it is related to income inequality. If this is so, significant gaps between the incomes of neighbouring classes might also be considered as undesirable.

Another question is the period over which to measure equality. In practice, this reduces to whether people should be permitted to defer their income when they are being taxed at higher marginal rates to years when the marginal rates are lower. As the tax year is essentially arbitrary, it seems acceptable for people to be able to make such a deferral so long as the income deferred is eventually taxed (even if the person has died). Emerson (2005, 4) says it has been argued that tax deferral is unfair, presumably because it largely benefits the wealthy. In fact, it identifies and benefits wealthy people who will have a lower marginal tax rate in retirement, and so can be seen to discriminate more accurately between degrees of wealth.

2.5.3 The taxation of income

Perhaps the most hotly contested of taxation debates is whether tax should be imposed on income or consumption, or more particularly whether it is fair and efficient to tax

investment income. McNulty (2000) provides a detailed evaluation of the debate, finding no obvious grounds to prefer one over the other. Changing from one to the other does however create significant transitional issues and is clearly not desirable. The current system in most countries of a balance between income and consumption taxes seems worth retaining.

In this context, it seems clear that the taxation of investment income contributes to equality. Positive real investment income (after adjusting for inflation) undoubtedly gives the beneficiary greater spending power. It is therefore reasonable that it should be taxed. There is a view that this represents double taxation, but this – as McNulty (2000) and others have shown - is nonsense. If I earn 100 and pay tax of 20 I have been taxed. If I earn 100 the next year and pay tax of 20 again, I have been taxed twice but this is not double taxation in the sense of being unfair – it is new income. If I invest 80 of my earnings and earn a real return of 5, taxing the 5 is not double tax, but a tax of new income. (Taxing income annually does mean that the rate of tax that applies to the investment income earned over a number of years is higher than that earned over one year.[27] This is another consequence of the arbitrary annual taxation of income, and is not necessarily unfair, although might be characterised as double taxation.)

This misunderstanding takes a different form in the context of pension taxation, in that it is argued that double taxation occurs when both investment income and benefits are taxed. There is, in this case, double taxation, but it is not unfair as the taxation of benefits is matched by the exemption of contributions as shown by Brown (1993).

Those with investment income are, by virtue of their assets, wealthier than those with similar earned incomes. Both vertical and horizontal equality would thus allow for a higher rate of marginal tax on investment income. The argument is sometimes made that taxing the investment income of retirement funds is unfair to the poor. This is straining at a gnat and swallowing a camel. The two important points not made are that the poor face higher rates of clawback from means tests, and that most of the investment income of retirement funds is being earned on behalf of the wealthy. The poor with retirement savings are, probably, better off than those without, so using a higher rate of tax than their direct marginal rate is not an offence against horizontal equality.

---

[27] $(1+(1-t)i)^n < 1+(1-t)\{(1+i)^n -1\}$ for $n>1$, tax at rate t

2.5.4 The implications

From the perspective of equality of spending power, it is clear that tax and state pensions (and thus the means tests that determine them) should be considered together. They are the two complementary parts of any system of creating equality: amounts taken from the wealthy and given to the poor.

It also becomes clear that means tests that apply a clawback of benefits at a higher rate at lower incomes than rates of tax at higher incomes are an offence against vertical equity. In South Africa, the means test on private income reduces the pension by 50% against marginal rates that currently vary from 17% to 40%. In Australia, the rate of clawback is 40% against tax rates that vary from 16% to 48.5%, but the taxation of superannuation contributions can increase the loss to as much as 49% of the proceeds. The UK basic rate of clawback is 40%, which is the same as the highest marginal rate of tax.[28]

2.6 Redistribution

Arguments for means tests, as for instance in Hancock et al (2005) and the Australian debate discussed in 3.1.2 below, refer invariably to the affordability of the state pension. This is however a limited perspective, seeing the issue as that of allocating the government's welfare budget and failing to see means tests as revenue raising items. The welfare budget should also be seen as part of government's role in the redistribution of income. To characterize this as a question of how much the rich can afford to give to the poor is a caricature. Questions of redistribution involve the division of a common pie; some of the slices may appear too big or too small but talk of affordability is meaningless.

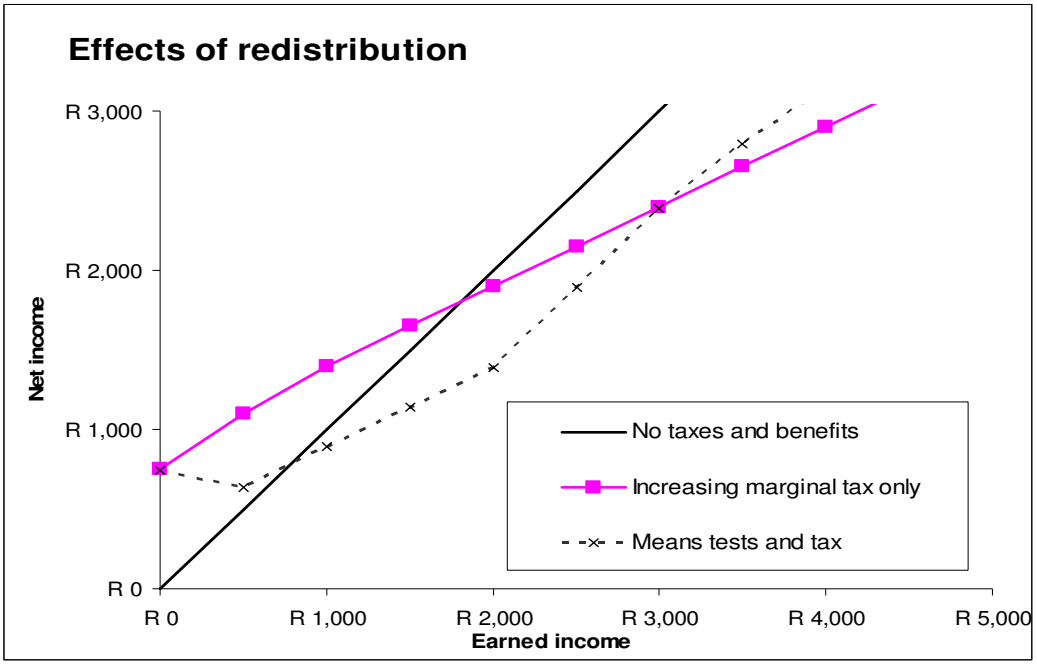
The major issues are illustrated by figure 1. The straight unmarked line shows net income equal to earned income – if there were no taxes and benefits. The dashed line shows – more or less - the effects of the South African means tests on net income after adding in the state pension and deducting the clawback and tax. The other marked line shows the impact of abolishing the means tests and only applying higher tax rates. What is made clear is that means tests would target the poorest if it were possible to identify them

[28] Information on these can be found at websites of the relevant government departments: sassa.gov.za, centrelink.gov.au; thepensionservice.gov.uk. The first is however not yet fully operational.

accurately, but that there are significant penalties applying to those just above the lowest income band.

The graph shows the reality in South Africa that, at some income levels, the marginal rate of clawback exceeds 100%. The higher marginal rates arise from the use of means tests to clawback some of the allowances mentioned in subsection 2.4.4. These include allowances for housing, rates and taxes, school fees (which are relevant in multi-generational homes where pensioners support grandchildren) and health care. In Australia, higher marginal rates apply at the point where the pensioner loses rights to a pension, which also mean loss of access to the Pharmaceutical Benefit Scheme and various subsidies on transport, rates and most importantly, residential aged care. At some incomes, the marginal rate therefore also exceeds 100%. Hancock *et al* (2005) report that it only reaches 91% in the UK.

Figure 1



Before moving on to the other criteria, it is useful to consider the balance of high and low incomes in the countries being considered. The UK and South African budgets[29] show social security payments at some two thirds of personal tax (including social security

[29] finance.gov.za/ and budget2005.treasury.gov.uk/page_09.html

contributions in the UK). In Australia, official figures[30] show personal tax is more or less balanced by social security spending. Abolishing the means tests would represent additional expenditure of some 5% of the respective national budgets. Some half would be recouped from simplifying the tax system by removing tax concessions to the elderly. The rest would have to be funded by an increase in tax or a reduction in the level of benefit. How taxation expenditures on retirement provision can be reduced for this purpose is discussed in section 3.1.3 below.

The arguments from equality therefore seem clear. The state pensions do contribute to equality, but greater equality would arise if the combined rates of clawback and tax increased with income. In terms of figure 1, this paper is intended to address only the kink (or dip) in the line created by the means test, and to suggest that it is unnecessary and unfair. The degree of redistribution between pensioners changes the slope of the line, the degree of redistribution between generations its intercept. A range of different lines would be compatible with justice, but only a smooth curve, with a declining slope, satisfies the requirements of equality.

2.6.1 Lifetime considerations

Spread over a lifetime and looked at from the impact on average rates of tax, the means tests on pensions are not particularly significant. Given that contributions to retirement funding represent a relatively small part of any individual's wages, the effects of inequalities in the clawback and taxation of retirement benefits is considerably diluted. A 50% clawback of retirement benefits that have a present value of 10% of salary over the lifetime increases the marginal – and average – rate of tax on income by 5%. As a large proportion of the benefits can be taken as tax-favoured lump sums, the actual impact is reduced to perhaps half in current South African circumstances[31]. This means that it is probably less than 20% of the income and expenditure tax paid by South Africans

---

[30] budget.gov.au/2005-06/bp1/html/index.htm

[31] The present value of the state old age pension is more or less equal to the tax free lump sum. If a new pensioner's savings were three times the tax free lump sum then the means test would apply to two thirds of their retirement savings. This is where the affect of the means test is at its maximum, so the average must be less.

earning an average income. While significant at any one time, this is less than some of the changes that can be observed in average tax rates in the past.[32]

2.6.2 Comparison with unemployment and disability benefits

There is another argument from equality that needs consideration.

Pensions could conceivably be seen as unemployment or disability benefits based on a presumption that people over the retirement age are unable to work. Certainly, state pensions, unemployment and disability benefits are often set at more or less the same level. If one were just to consider the recipients' material needs, it would be fair to treat the beneficiaries equally and apply similar means tests.

There are however wider economic and social reasons to encourage people of working age to return to the workforce, as well as considerations related to their own welfare. Disabled people need rehabilitation for their own sake. Some return to the workforce is usually a possibility, and is generally recognised as a cure in its own right. Unemployment is similar: both from their own perspective and that of the national economy, people are better off if some way can be found for them to return to work. The tests that apply to unemployment and disability might better be called capacity tests rather than means tests. They should also include additional requirements to seek the necessary medical attention, or to be actively searching for a job.

These considerations barely apply to the retired. In many cases, they might well prefer to continue to work, but the personal, social, economic and budget imperatives are not as pressing. Even if the costs of re-entering the workforce do not rise with age, there is a shorter period over which to recover the human capital invested and it becomes increasingly uneconomic to attempt to do so. If the retirement age is set high enough, it is a reasonable presumption that the pensioners cannot economically return to the workforce.

---

[32] For 30 years of Australian rates see http://www.bendzulla.com/ref8.html. South African changes will be found to be of the same order. The last 10 years have however been relatively stable and can be found at: http://www.sars.gov.za/it/Income%20Tax%20Tables/Income%20Tax%20-%20Tax%20Rates-Duties-Levies%20-%202004.pdf

This suggests that abolishing the means test has an increasingly strong case for those at more advanced ages. It is interesting that Travers (1995) reports that, for a period during the seventies, the means test was abolished in Australia for those over 70.

For younger retirement ages, it becomes likely that retirement, unemployment and disability benefits are confounded. Means tests on their own will discourage people from returning to the workforce, but this brings us to the incentives discussed in the next section.

2.7 Efficiency

Efficiency is taken to include lower costs of administration and the provision of productive incentives. Efficiency is not normally explicitly considered as an issue of justice, but it must often be considered in order to make a balanced evaluation of policy options. It is argued here that it should function as a criterion of justice. The question to ask is whether it would be unjust to fail to meet one of the other criteria because of cost considerations? If not, then efficiency would appear to act as one of the components of justice.

In some of the social security literature[33], efficiency refers to the proportion of social security payments that are paid to those under some poverty level of income. As concluded in section 2.6, this artificially characterizes the question as a budgeting one or as an artificial caricature: how the state (on behalf of wealthy taxpayers) can best afford to target a needy group?

This section considers deadweight losses that arise from means tests. It concludes that the criterion of efficiency produces strong arguments against means tests. The first two sub-sections discuss administrative difficulties and costs; the next four discuss incentives to save.

2.7.1 Administration

Means tests in the three countries being discussed are controlled by government departments responsible for social security and not those responsible for finance, which control the collection of taxes.

---

[33] As, for instance, in Lundberg (2005)

The challenges of auditing assets and income for means tests are difficult to overstate. Apart from catching those that cheat, one should review records of income from casual jobs and the renting of rooms, interest on formal and informal loans, pensions and transfer payments, and – particularly in South Africa - estimate income in kind from subsistence farming. This is impossible given the budgets of these agencies. The UK website mentioned in footnote 28 has a calculator that asks 48 questions in order to provide a first estimate of the means-tested benefits available. Centrelink is the agency responsible for implementing the tests in Australia. Its website lists seven different kinds of income and six different types of assets that have to be considered twice a year. In many instances, there is no way in which the answers can be independently confirmed, while the level of complexity makes the assessment difficult even for diligent applicants and competent administrators with goodwill and plenty of time.

The difficulties are illustrated by a recent audit of Centrelink – reported in Pender (2004). The audit found the agency made some 250,000 errors a month, with an "actionable" error in 52% of the cases investigated, and that almost all folders contained administrative errors. Another audit of Centrelink by the Australian National Audit Office[34] reviewed the files for proof of identity and found only 24% completely correct, although most of the errors were of a minor administrative nature. This is not to suggest that Centrelink is any worse than other agencies. It is to suggest that the proper administration of means tests is not possible. There is far too much to do.

Any attempt to quantify the additional costs of means tests is necessarily suspect given that they are almost certainly not performed properly. The National Association of Pension Funds (2005; 35) however quotes costs reported to the Parliamentary Select Committee of £54 annually for the UK means-tested pensions credit against £5.40 for the non-means-tested basic pension. The Select Committee found that the Pensions Service (the UK social security agency) was to be commended for its accessibility, but expressed concerns with accuracy rates purported to be at 90%.

Abolishing means tests would make administration possible. Three quarters of the questions in the UK calculator would be irrelevant (assuming that many of the social

---

[34] See particularly table 2.1 in the report, which can be found at: anao.gov.au/WebSite.nsf/Publications/4A256AE90015F69BCA256B22000BE66B

security allowances would be reduced to take the universality of the pension into account) and much the same would apply in other countries. All the difficult questions relating to assets and private income would be abolished. Pensioners could then be subject to the same tax rules as other people.

There would – of course - be the additional administrative work of paying pension benefits to more people. This would however not appear to be a major burden as almost all the wealthier people concerned would have bank accounts and the exercise would be largely automated. For those paying more tax than the pension, it could be avoided altogether if there were the necessary electronic linkages.

2.7.2 Unenforceable law is unjust

Laws that cannot be enforced are unjust laws as they are likely to fail each of the criteria at some point. The administrative difficulties referred to in the previous paragraphs indicate that means tests cannot be enforced uniformly. Anecdotes from South Africa and Australia suggest that cheating is widespread. This would be reason alone to attempt to find an alternative to means tests.

2.7.3 Work incentives and redistribution

Redistribution of income requires higher tax rates on higher incomes, and creates a risk that this will act as a disincentive for higher earners. It does however reduce taxes on lower income earners which is likely to have compensating effects. We are here concerned with changing the incidence of tax and clawbacks on retirement income, which is likely to have an attenuated impact on motivation during the working years.

The studies reported by Tuomala (1990) suggest that "most labour supply studies of men seem to indicate backward-sloping supply curves." Higher income taxes can lead to men taking less leisure because they need to work harder for the same standard of living (described as an income effect), but leisure is relatively cheaper because of the loss of the lower net wage, which may mean that more leisure is taken (described as a substitution effect). It appears that the income effect is often more powerful for men. Studies of women on the other hand suggest normal supply curves: lower tax rates increase work effort.

This suggests that it would be very difficult to make the case that small increases in the marginal tax rate of the wealthiest are likely to have a significant impact on their

incentive to work hard. The other suggestion relevant to South Africa is whether higher tax rates would create an incentive for the wealthy to emigrate. There appears to have been no local research into either of these incentive questions, and the ambivalence of other research suggests that they are unlikely to be of great importance. They are therefore not considered further.

There is the additional question of whether people understand the size of the incentives anyway. Hills[35] reports that in the UK "only 17% of basic rate taxpayers and only 28% of higher rate taxpayers can identify their own rate of tax relief."

2.7.4 Deferral incentives and life time smoothing

It is frequently argued that tax incentives are required to get people to save enough for their retirement through private retirement funds. This again assumes a level of myopia that is difficult to test because almost all countries offer some level of incentive, compulsion or minimal provision. The AXA survey[23] suggests that pensioners are, as a whole, happier with their income than those still working, which may mean that the balance is about right.

While the case for myopia may be overstated, there is the alternative argument, mentioned in 2.5.2 above, that people whose income is "bunched" ought to be given an opportunity to lower their average rates of tax by deferring income.

Whether seen as incentives or opportunities for tax rate smoothing, it would seem that the rewards for deferral should be applied more or less equally to different earning patterns. The current system however creates significant disparities for three different earning patterns:

- In some cases the marginal rate of deduction on contributions will be the same as the marginal rate on pensions. This is likely to apply to those at the extremes of the income distribution. They see no gain from deferment.
- Those in the middle of the income distribution will be able to claim tax deductions on contributions (technically described as employer contributions in Australia) at a higher rate than the rate of tax that they pay on pensions. As people

[35] Presentation by John Hills to the 2005 AGM of the Pensions Policy Institute pensionspolicyinstitute.org.uk/uploadeddocuments/John_Hills_AGM_Slides_7_June05.pdf

may be in the position of paying no tax on their pensions, this difference can be as high as 40% in South Africa and the UK and 33.5% in Australia (where the contribution tax is not refundable).

- In many cases however, the rate of tax saved on contributions is lower than the rate of clawback imposed on the pension. For those on low wages, they might obtain no deduction for contributions and yet suffer loss through the means test at 50% in South Africa, 40% in the UK and 49% in Australia (where the 40% rate of clawback has to be increased to allow for the 15% tax on superannuation contributions, which is not refunded).

Concessions on investment income change this ratio. The longer the concessions are in force, the greater the benefit of contributing. The combined effect is shown in table 1 for a variety of South Africa rates. An Australian table would follow the same pattern. Hills produces a range of between 80% and 255% for the UK equivalent of the last column.[36]

There are indeed significant rewards for upper-middle-income people, but for the quarter of the population (in all three countries) that is likely to be drawing a reduced state pension as a result of means tests, there are significant disincentives to deferring income. People in this category will be misled by generic financial advice that retirement savings are tax advantaged.

As already noted, lump sums enjoy significant tax concessions in every country. In South Africa, members of provident funds can take their entire benefit as a lump sum, which may often be too small to be taxed so the members face a marginal rate of zero. All superannuation benefits in Australia can be taken as lump sums with a significant tax free allowance, but the un-refunded 15% tax on contributions means that the effective rate is at least 15%. In both countries higher marginal rates apply to larger payouts. In the new UK system introduced in April 2006, 25% of the value of the pension at retirement will be tax free.

The need for an incentive is reduced in Australia where contributions are compulsory. Additional voluntary contributions (whether they are made in or out of the superannuation system) are however also subject to means tests.

---

[36] As for footnote 35.

2.7.5 Work incentives after retirement age

As discussed in 2.6.2, retirement, disability and unemployment benefit issues are difficult to distinguish if entitlement to pension is set at a fairly young age. Remaining employed becomes less economical in any event, and high rates of clawback create significant disincentives to work. With South Africa's high levels of unemployment, encouraging early retirement can be seen as opening up jobs for younger people. The flat population pyramid makes it a relatively inexpensive way of providing unemployment benefits. For countries facing population aging that have labour shortages, higher retirement ages appear to be a necessity, and people ought to be encouraged to remain in the workplace as long as possible. Means tests ought then to be supplemented by other measures to encourage people to return to work.

Table 1: Incentives to defer income - South Africa with means tests

	Marginal rate of deduction for contributions	Years of investment income tax concessions	Marginal tax or clawback rate on benefits	Benefit: contribution ratio
Same tax rates	18.0%	10	18.0%	104%
	18.0%	40	18.0%	128%
	40.0%	10	40.0%	116%
	40.0%	40	40.0%	227%
Reducing tax rates	18.0%	10	0.0%	127%
	18.0%	40	0.0%	156%
	25.0%	10	0.0%	144%
	25.0%	40	0.0%	205%
	30.0%	10	0.0%	158%
	30.0%	40	0.0%	251%
	30.0%	10	25.0%	118%
	30.0%	40	25.0%	188%
	35.0%	10	18.0%	143%
	35.0%	40	18.0%	252%
	35.0%	10	25.0%	131%
	35.0%	40	25.0%	231%
	40.0%	10	0.0%	193%
	40.0%	40	0.0%	379%
	40.0%	10	30.0%	135%
	40.0%	40	30.0%	265%
Increasing tax rates	0.0%	10	50.0%	48%
	0.0%	40	50.0%	39%
	18.0%	10	50.0%	64%
	18.0%	40	50.0%	78%
	30.0%	10	50.0%	79%
	30.0%	40	50.0%	125%
Ratio of greatest to least		9.70		

Assuming 10% nominal investment return, 30% tax on investment

2.7.6 Tax on investment income

It can be seen from table 1 that the effect of the tax concessions on investment income (which can be seen by comparing the effects of 40 years vs. 10 years of contributions) is small relative to the differences between the rate at which contributions are deducted and the rate at which they are taxed. The concessions do however have a deadweight loss as they provide a distorting effect on investment decisions, allowing for a variety of tax-planning strategies.

It is however frequently argued that tax concessions for savings have positive externalities in that they promote investment and increase economic growth. It is perhaps necessary to rehearse, briefly, the refutations of each of the three mistakes in this argument.

Saving in tax-favoured superannuation barely increases savings. In a review of the literature on factors affecting saving, Smith (1990) finds that savings through private-sector pension funds are likely to reduce other types of savings, although they do contribute to an increase in overall savings. Knox (1992), in a typical study, has found that the reduction in other forms of savings means that total personal saving appears to be increased by only 30% to 40% of contributions to pension funds. Half, or more, of this increase is contributed by government in the form of tax concessions. Thus pension and superannuation funds make a much smaller contribution to increasing national savings than is sometimes assumed - probably less than 20% of the amounts saved.

Saving does not necessarily increase investment. The classical view is that an increase in savings will reduce the cost of investible funds, and so encourage investment. The Keynesian view is that increased savings represents a postponement of consumption, and that this will discourage investment in plant that would otherwise have provided for this consumption. While there is a positive relationship between savings and investment, it is not at all clear which comes first.

Investment does not necessarily increase economic growth. The debate about the relationship between investment and growth is unresolved. Aghevli *et al* (1990) point out that the positive correlation that clearly exists between savings, investment and growth is consistent with several different hypotheses about causes. It is just as likely that growth causes savings and investment as for savings and investment to cause growth.

If private capital accumulation through retirement savings made a difference to prosperity, one would expect a significant difference between otherwise similar economies such as those of the UK and France, and the Netherlands and Belgium, as they have taken very different paths in this respect. In fact, the differences are barely evident. It is, therefore, not true that there are proven macro-economic benefits for giving special tax consideration to retirement funds.

Even if it were demonstrated that there were significant advantages in encouraging superannuation savings, that would not by itself justify tax concessions. There are few economic activities that do not have similar advantages; as has been one of the major concerns of economists since Adam Smith, and which Olson (1965) particularly shows, the better policy is to avoid all concessions to special interests.

Table 2: Incentives to defer income - South Africa no means tests

	Marginal rate of deduction for contributions	Years of investment income tax concessions	Marginal tax or clawback rate on benefits	Benefit: contribution ratio
	18.0%	10	18.0%	94%
Same tax	18.0%	40	18.0%	72%
rates	40.0%	10	40.0%	105%
	40.0%	40	40.0%	131%
	18.0%	10	0.0%	115%
	18.0%	40	0.0%	88%
	25.0%	10	0.0%	130%
	25.0%	40	0.0%	116%
	30.0%	10	0.0%	143%
	30.0%	40	0.0%	143%
	30.0%	10	25.0%	107%
Reducing	30.0%	40	25.0%	107%
tax rates	35.0%	10	18.0%	129%
	35.0%	40	18.0%	144%
	35.0%	10	25.0%	118%
	35.0%	40	25.0%	132%
	40.0%	10	0.0%	175%
	40.0%	40	0.0%	218%
	40.0%	10	30.0%	122%
	40.0%	40	30.0%	153%
	0.0%	10	0.0%	87%
	0.0%	40	0.0%	44%
Identical	18.0%	10	18.0%	94%
tax rates	18.0%	40	18.0%	72%
	30.0%	10	30.0%	100%
	30.0%	40	30.0%	100%
Ratio of greatest to least		5.00		

Assuming 10% nominal investment return, 30% tax on investment

Table 2 provides an indication of what the incentives to defer income were to become if investment income was taxed at 30% and means tests abolished. It is assumed that no-one would save for a higher income in retirement than while they were working, and so would never be faced by a higher marginal tax rate. The range of the ratio of benefits to contributions is halved, but it remains less than 1 for those with a marginal rate of less than 30%, which may be seen as unfair. Apart from the arguments made in the fourth paragraph of 2.5.3 above that this can be reconciled with both horizontal and vertical equity, the after tax return on retirement fund investments is likely to be higher than the alternative instruments used by the poor[37], or they can negotiate membership of alternative savings mechanisms that better suit their circumstances. Mandatory superannuation makes the latter impossible in Australia unless people were given the option of joining non-complying unit trusts where they paid their own tax on the investment income.

2.7.7 Alternative investments

Paying inflation protected state pensions to all pensioners would provide them with a useful additional investment instrument. It would also allow for retirement funds to adopt a more adventurous investment policy with their investments.

The introduction of a universal government guaranteed pension would reduce – even eliminate – calls for the state to artificially create instruments that match pension payments or otherwise subsidize pensions for the wealthy. Blake (2000) notes suggestions that the state should provide not only more long-term inflation-linked assets, but also instruments that would absorb some of the longevity risk inherent in annuities. Vitas (1998) says that Chile and other South American governments provide minimum guarantee of investment returns, which has a similar effect although with moral hazards.

---

[37] Of the 2 out of 150 Financial Diaries families with unit trusts, which can be assumed to offer comparable returns to retirement funds, one was a taxpayer (facing a marginal rate of 25%) the other not. Most families had bank accounts which offer lower yields. Although informal loans to friends and family may have offered higher returns, this would have been at some risk. See http://www.financialdiaries.com/files/Financial%20Diaries%20Financial%20Styles%20smaller.pdf

2.8 Just deserts

Recognizing just deserts is equivalent to saying that people's actions matter and must have their proper consequences. In the context of benefits and taxation, this suggests a correspondence between the two. It is possible to find some arguments for means tests from this perspective, but they largely vanish under closer scrutiny.

2.8.1 State pensions

State pensions, unlike private income, might be said to be un-deserved and thus it could be argued that they can be treated differently when it comes to clawing them back.

The amounts of clawback are determined by reference to private income and not the pension. In order to justify means tests using an argument from desert, one would have to make a far more difficult case. It would be necessary to prove that those on a partial pension (subject to the means test at 50% in South Africa) were less deserving than those with the full state pension.

It would appear to be an impossible task to justify a blanket clawback without a much more careful analysis of just deserts. Needy pensioners may have been significant contributors to the fiscus over their lifetime, and their current low income may be the result of economic bad luck, poor health, or a commendable generosity - rather than idleness or imprudence. From the perspective of desert, such distinctions would have to be made. That they have not been made, suggests that this idea of deserts is not regarded as an important justification for the high rate of clawback applied in means tests.

One could imagine practical ways of implementing such a policy: for instance, applying it to those who had not been taxpayers for 20 years or more (as is the case in some countries), and also to those who could not explain why they were no longer in possession of their lump sum retirement benefits. Alternatively, it may be regarded as just too hard to achieve, or desirable to forgive undeserving recipients for their failure to pay adequate taxes.

Another response would be to justify the payment of a universal (non-means tested) pension on the grounds of desert rather than need or equality. In the first place, most pensioners are likely to have worked since they were children, and even if they have not paid income tax or contributed greatly to indirect taxes, they are all likely to have made some contribution to the economic wealth of society. It can also be argued that some of

the state's taxation income arises from economic rents (e.g. the value of land, mineral wealth and band-width) that should be given equally in some form to all residents. It appears, from research such as Evans and Kelly (2005) and St John and Willmore (2001) that there is majority public support for universal pensions in a number of countries including Australia and New Zealand, which has had one for some time. It can be argued that a universal pension would make a real contribution to social solidarity.

2.8.2 Grandfathering

The taxation arrangements of all three countries are greatly complicated by the protection of historic privileges, or grandfathering. This applies in some cases to allowances and income that are not means-tested, and to income and benefits that are not taxed. The resultant complexity is widely regarded as expensive and a deadweight loss on the economy.

One argument made to defend grandfathering is the rule against retrospective legislation. The rule relates to the procedural principle of certainty. It relates to desert because, in this context, it is necessary to prevent people from being unfairly penalized for making decisions that subsequently prove to be inappropriate because of changes in the law, and provides justification for transitional arrangements. It cannot however be used as a legitimate argument to protect tax privileges, to prevent government from making relatively minor changes to rates of tax, or when the individuals concerned are able to rearrange their circumstances to avoid harm that they would otherwise have suffered.

The UK is in the process of eliminating all grandfathering in its pension tax regime. South African and Australia might consider doing the same.

3 REFORM

3.1 Bringing it together

This part of the paper first summarizes the arguments for reform, discusses the results of the Australian debate of the nineties to understand why it failed to lead to the abolishment of the means tests, and then suggests how reform might be achieved.

3.1.1 The debate

Means tests are justified as making them more affordable by targeting the needy. It has been shown, however, that the issues are not so much about saving expenses but about distributing tax and clawback charges equitably. Means tests are a poor mechanism for this because:

- They involve a greater intrusion into the lives of most pensioners, applying an unfair social stigma.
- They are not well targeted at needs.
- They present declining levels of clawback and tax, and so are offences against the vertical equity element of equality.
- They are expensive to administer, probably impossible to do so accurately.
- If intended as incentives to save more or work in retirement, they do not fulfil this function for a large proportion of the target population.
- They might be justified in terms of just deserts, but this would require the introduction of a proportional residence requirement or a measure of how long tax was paid.
- They might be justified as treating the recipients of pensions, disability benefits and unemployment benefits equally, but differentiation can be justified in terms of the desirability (from an individual and societal perspective) to get people back into the workforce.

3.1.2 The Australian debate

The means-test debate in Australia in the mid-nineties may be instructive in understanding how other debates may proceed.

The Institute of Actuaries of Australia (1994) recommended the introduction of a universal non-means-tested state pension in a package of other superannuation suggestions. They placed a particular emphasis on the inefficiencies arising from attempts to avoid the impact of the means test – which is a real problem in Australia, but appears to be less so in South Africa. The submission emphasised that the Institute was advocating an alternative structure rather than a particular package of changes to taxation but, after some debate, modelling of the cost implications was included in this and a subsequent submission. This proved unfortunate, as Gallagher (1994) showed that the costs had been underestimated and that the proposals appeared to significantly benefit

wealthier pensioners. This led Barber *et al* (1995), in their strategic review of the pensions' means tests for the Government, to reject the Institute's recommendations. The arguments that the costs were too high appear to have been accepted by subsequent papers to the Institute, such as those of Somogyi *et al* (1995) and Rice (1998).

The outcome of this debate is unfortunate. If a much better administrative arrangement happens to unfairly favour wealthier pensioners, the obvious solution is to increase the taxes that apply to this group. Such an increase would also justify the removal of the administratively cumbersome reasonable benefit limits that currently apply to the wealthiest.

3.1.3 Reducing tax expenditures

Section 2.6 argues that the redistribution of wealth can be made fairer by abolishing means tests and current tax concessions to pensioners, and probably some increase in other forms of taxation. Increasing the tax on the investment income of retirement funds would be an obvious source for this additional requirement. (As suggested above, another alternative would be to tax pensioners at a higher rate than other people. This would not have to be any more complicated than the current position in all three countries.)

The question this section addresses is whether a tax on the investment income of retirement funds would currently be sufficient to abolish the means tests. It could be made to be so in future by adjusting the level of the state pension and other taxes at the time.

Abolishing the means test in South Africa would currently cost some R10bn, which could also be recouped by taxing payments and the reduction or abolishing of tax concessions for older people. The exact amounts are not readily determinable, but it would probably not be necessary to entirely remove tax concessions on investment income (R6bn).[38]

Abolishing the means test in Australia would cost some $10bn, but some 30% of this would be immediately recouped by taxing the resulting pensions, and the balance could

[38] This is extrapolated from Smith (1995) and the South African budget reports on retirement fund tax.

be recouped by removing tax concessions to superannuation investment income ($5.7bn) and removing the senior tax offset ($1.8bn).[39]

The National Association of Pension Funds (2005) says the UK system has been called the "most complex in the world". The report suggests the introduction of a citizen's pension that would be set at a level that required no means tests. Their estimate of the net cost in 2010 is of the order of £13bn, of which £2bn would be recovered by the tax system. They suggest that the balance of the cost should be recovered from redirecting the national insurance contributions for the earnings related state scheme and raising the retirement age. UK pension funds currently pay limited tax on their investment income so there would also be scope to recover amounts from investment income.

Reducing tax expenditures on retirement funding would significantly simplify the system of retirement fund taxation, creating obvious efficiency gains to government and people who would not have to devote as much time to understanding the system (and avoiding its negative consequences):

- It abolishes means tests.
- It eliminates tax concessions currently enjoyed by the elderly.
- It allows for the elimination of the tax concessions currently enjoyed by retirement funds, which also reduces opportunities for tax arbitrage.
- It removes the need for caps on benefits for the wealthy.

Administration would also be simplified. Extrapolating UK costs of the savings credit to the other countries suggests savings in the hundreds of millions as the social welfare agencies no longer perform means tests. State pensions could be paid without deduction. Employers and financial institutions paying benefits would determine PAYG deductions on the assumption that a pension was being paid to those over the retirement age. The tax collecting agencies would be responsible for ensuring that those with multiple income sources were assessed correctly. It is not clear how much additional work this would create; for most institutions it would represent a change in the work that they do but not an increase. The Australian Tax Office, for instance, already has data on some 50% of pensioners. The same is likely to be true in the UK and in South Africa in respect of those pensioners who have private incomes.

---

[39] See Tax Expenditure Statement at treasury.gov.au/documents/950/PDF/TES_2005.pdf

3.2 Institutional factors

If the arguments against means tests are so strong, why do they persist? If we can find no good reasons, perhaps there are bad ones?

Lundberg (2005) discusses some of the institutional and political reasons for the structure of pensions systems in different countries. He raises the role of history, political interests and "policy entrepreneurs". It may well be that the answer mainly lies in historical institutional accidents.

3.2.1 Fault lines

In all countries, recommendations as to the level of the state pensions and clawbacks are made by the same people who look at the levels of unemployment and disability benefits. They are given a constrained budget, and effectively prevented from considering rates of taxation.

There appears to be a wide gap between the people in the departments responsible for social security and those for taxation. The best of the former are motivated mainly by compassion, the latter by efficiency; they come from different intellectual disciplines, read different newspapers and journals, appear likely to vote for different parties and are advised internationally by different agencies. The International Labour Office (ILO) advises the former and the World Bank and International Monetary Fund (IMF) interact with the latter. Ervik (2003) and Asher (2002) comment on some of these differences.

Policy development in social security departments such as the South African Department of Social Development operates within a given budget that has to be applied to meet urgent needs and reduce inequality. The department officials are required to treat different groups of their clients fairly. There is no question of making a trade-off with the tax on investment income of retirement funds or tax concessions given to pensioners. In both cases, concessions are likely to be regarded as being favourable to pensioners in general, and thus worth supporting. Given the limited budget and the range of needs that present themselves, a steep means test meets a number of the criteria of justice listed above.

From the perspective of tax policy, social security means tests are regarded as a reduction in expenditure and not as taxation income. They do not fall within the purview of those responsible for determining marginal tax rates and the question of the taxation of

retirement fund income. People developing policy face ongoing arguments that taxation should be reduced, and counting the clawback of pensions as a reduction in expenditure is consistent with doing so. Within the confines of their brief, they also meet many of the criteria of justice. The combination however does not.

3.2.2 A South African experience

My experience in South Africa was of enormously difficulties in obtaining an interdepartmental view on the matter. I was a member of the steering committee set up to produce recommendations that arose from the National Retirement Consultative Forum in 1997. This was set up by the Department of Finance (since renamed the Treasury), as a result of the recommendations of the Smith Committee (1995), to attempt to gain a national consensus on a variety of pension issues – including all aspects of compulsory membership and taxation. The steering committee suggested to the Department, in early 1998, that the matters could not be determined independently of the Department of Welfare (since renamed Social Development). Whether as a result, or independently, an inter-departmental task team was convened by the Department of Social Development in 1999 to review the social security system. It recommended the investigation of a comprehensive and integrated social security structure. The output of this investigation was the report of the Taylor Committee (2002) of which I was also a member. The report recommended that means tests (for pensions and other categorical grants, such as children's allowances) be abolished and the lost revenue recaptured through the taxation system. The recommendation was however overshadowed by the Committee's major recommendation (from which I dissented) of a new, non-means-tested, basic income grant payable to all residents. The controversy over the latter may well overshadow the important, unanimous, recommendation that means tests should, wherever possible, be avoided.

3.3 Lobby groups

The institutional arrangements would probably not be sufficient to maintain means tests in their current form if it were not for a few powerful interests.

3.3.1 Potential losers

Potential losers may be expected to resist reform.

Major losers will be wealthier individuals who are currently benefiting from tax concessions offered to pension funds. Those currently drawing benefits will be barely affected if the majority of the additional revenue comes from tax on investment income. They could however be effected by an increase in the tax on pensions.

It may be that, in the long run, moving away from means tests would put some pressure on the level of the state pension, and so potentially also disadvantage those with the lowest income, who are currently not affected by the means test. This is to be regretted, although the argument of 2.4.2 above is that there is no way of accurately measuring who are the poorest, and that targeting is invariably spurious.

The winners are those of lower middle income who are, or would be, subject to means tests.

3.3.2 The gerontocracy

Mulligan and Sala-i-Martin (1999) describe how the old in every country, with plenty of time to lobby, extract a greater share of resources, and shape the pension system to maximise their power. Their thesis would explain tax concessions that favour the old, and how these are seldom questioned in political debates. All tax expenditure on retirement funds may fall into this category, particularly concessions such as the exemption of the investment income of pensioners within retirement funds - in South Africa and Australia.

3.3.3 The industry

The long term savings industry, which includes retirement funds, insurers, and service providers, including actuaries, has an interest in a larger, more tax-favoured, industry. In South Africa and Australia, it can be relied upon to call for the reduction of taxes on pension and superannuation fund investment income particularly.

3.3.4 Public interest groups

Non-government organizations can speak effectively for the rights of pensioners. They understandably interpret their role as to speak for the poorest, and so are unlikely to be immediate supporters of a change that benefits those slightly wealthier. As suggested above however, the losses to those on the lowest income will be small, and are abolishing means tests is ultimately likely to benefit everyone in the lower income brackets.

In South Africa, the position is clouded by the fact that whites will benefit more proportionately from the abolition of the means test. Lost tax concessions for wealthier pensioners will however fall almost entirely on whites.

In all countries, public interest groups may baulk at paying pensions to the wealthy. They may also not be easily persuaded by proposals that the wealthy will be worse off once their tax concessions are removed. One way of addressing this objection is to increase the level of income and assets at which the means test applies so that very few people are affected. As long as the level is inflation adjusted, this removes most of the objections to the means test. It is however administratively untidy.

3.3.5 The beneficiaries

It can be noted that the beneficiaries are not likely to be represented by anyone, nor are they likely to speak up for themselves. There is evidence that they feel stigmatized.

3.3.6 Recommendations

If this analysis of the problem is true, the next step is to recommend that responsibility for determining means tests be transferred to the same government office as that responsible for deciding the incidence of taxation. This responsibility should extend to the evaluation of all means tests offered by central, state and local government. The clawback of means should be recorded clearly as a charge on income rather than a reduction in pensions and other allowances. If the logic of this paper is ineluctable, then one could expect such an office to recommend a move to a more just approach to means tests.

The administration of the evaluation of means should also be centralised with the administration of taxation. It should then become more efficient and lead to lower levels of intrusion in people's lives.

4 CONCLUSION

This paper has shown that there are good reasons to abolish means tests and replace them with additional taxes in South Africa, Australia and the UK. The reasons relate to money and mechanisms.

On the question of money, means tests for pensioners should not be seen as one of targeting of needs and affordability. Targeting that effects over half the pensioner population does not deserve the title; needs are invariably measured relatively and arise more from notions of equality than personal necessity; seeing the problem as one of affordability misses the fact that the means tests can be seen as a source of state revenue rather than a reduction in expenditure. The financial question is not one of affordability but of redistribution. The high rates of clawback - identified by the inequitable kink in figure 1 – create vertical inequality and perverse incentives - as demonstrated in table 1. Relatively simple tax changes can easily recover the loss of income arising from abolishing the means tests.

It is the mechanisms that are used rather than the redistribution question that create the most egregious problems. Three major offences against justice are the unwarranted interference in the lives of pensioners, the injustice of laws too complex to be correctly administered, and the deadweight costs to both the affected pensioners and the government of attempting to do so.

The internal structure of government and an alignment of established interests make it difficult to achieve reform. Reform is not however impossible.

- Government needs to be persuaded that means tests and taxation should be considered together as a single policy issue. Means tests operated by different government departments and local governments, which can push marginal rates of clawback to over 100%, should be included.
- Those in the retirement industry who lobby for tax preferential treatment for contributions, without considering the needs of the means-tested, need to be corrected.
- Public interest groups need to be persuaded that focusing on the poorest may be too narrow.

There is a moral imperative for those with power. As the New International Version has it: ‘speak up for those who cannot speak for themselves’.

REFERENCES

Aghevli B B, J M Houghton, P J Montiel, D Villaneuva & G Woglom (1990) *The Role of National Saving in the World Economy: Recent Trends and Prospects* Occasional Paper 67 International Monetary Fund, Washington, DC

Alesina A, R Di Tella & R McCulloch (2004) Inequality and Happiness: Are Americans and Europeans Different? *Journal of Public Economics* 88: 2009–2042

Asher A (2002) Book review: New ideas about Old Age Security, ed R Holzman & JE Stiglitz *South African Actuarial Journal* 2:161-163

Barber J, G Moon and S Doolan (1994) Targeting for Equity, Final Report of the Strategic Review of the Income and Assets Tests Australian Government Publication Service

Blake D (2000) Does it Matter What Type of Pension Scheme You Have? *The Economic Journal* 110.2: 46-81

Brown R (1993) Tax Assistance to Qualified Retirement Savings Plans: Deferral of Waiver *Journal of Actuarial Practice* 2.1: 159-166

Case A (2001) Does Money Protect Health Status? Evidence from South African Pensions *National Bureau of Economic Research Working Paper* 8495

Clare R (2001) *Equity and retirement income provision in Australia* Association of Superannuation Funds of Australia (superannuation.asn.au/policy/equity.pdf)

Connolly E & M Kohler (2004) The Impact of Superannuation on Household Saving *Research Discussion Paper* Reserve Bank of Australia

Diamond P A (1977) A framework for social security analysis *Journal of Public Economics* 8: 275-298

Emerson C (2005) *Tax, benefits and retirement income incentives* Pensions Policy Institute Seminar www.pensionspolicyinstitute.org.uk/news.asp?p=128&s=2&a=0

Ervik R (2003) *Global Normative Standards and National Solutions for Pension Provision: The World Bank ILO, Norway and South Africa in Comparative Perspective* Stein Rokkan Centre for Social Studies www.ub.uib.no/elpub/rokkan/N/N08-03.pdf

Evans M D R & J Kelley (2005) Assessing Age Pension Options: Public opinion in Australia 1994-2001 with comparisons to Finland and Poland *Working Paper Melbourne Institute of Applied Economic and Social Research* melbourneinstitute.com/wp/wp2004n21.pdf

Feldstein M S (1985) The Optimal Level of Social Security Benefits *Quarterly Journal of Economics* 100: 303-320

Gallagher P (1994) *Analysis of the Institute of Actuaries' Proposals for Reform of the Retirement Income System* Submission to the Strategic Review of the Pensions' Income and Assets Tests, rim.treasury.gov.au/content/rtf/CP94_5.rtf

Hausman D M & M S McPherson (1996) *Economic analysis and moral philosophy* Cambridge University Press

Hancock R, S Pudney, G Barker & M Hernandez (2005) *What Should be the R ole of Means-testing in State Pensions?* Pensions Policy Institute Seminar pensionspolicyinstitute.org.uk/news.asp?p=155&s=2&a=0

Institute of Actuaries of Australia (1994) Submission to Senate Select Committee on Retirement Income Policy *Transactions of the Institute of Actuaries of Australia:* 115-185

ILO (1998) *Social Security Financing* Geneva published jointly by the International Social Security Association

Knox D M (1992) An Analysis of the Option to Tax Investment Income Generated by Occupational Pension Funds *XXIV International Congress of Actuaries* 4: 465

Lundberg U (2005) *Social Democracy Lost - The Social Democratic Party in Sweden and the Politics of Pension Reform (1978-1998* Arbetsrapport/Institutet för Framtidsstudier

National Association of Pension Funds (2005) *Towards a Citizen's Pension Final Repor*t pensionspolicyinstitute.org.uk/uploadeddocuments/TCP_Final_Report_sep_05.pdf

Layard R (2005) Happiness and Public Policy: A challenge to the profession http://www.iew.unizh.ch/study/courses/downloads/happy.pdf

Lucas J R (1980) *On Justice* Clarendon Press, Oxford

MacIntyre A (1988) *Whose Justice? Whose Rationality?* Duckworth, London

McNulty J K (2000) Flat Tax, Consumption Tax, Consumption-Type Income Tax Proposals in the United States: a tax policy discussion of fundamental tax reform *California Law Review* 2095-2184 law.washington.edu/courses/hjorth/T507_Win05/507documents/mcnulty.htm

Mehran F (1976) Linear Measures of Income Inequality *Econometrica* 44.4: 805-809

Mitchell D, A Harding & F Gruen (1994) Targeting Welfare *Economic Record* 70 (210): 310-335

Mulligan C B & X Sala-i-Martin (1999) Gerontocracy, Retirement, and Social Security *NBER working paper w7117*

Olson M (1965) *The Logic of Collective Action: public goods and the theory of groups* Harvard University Press, Cambridge, Mass

Pender H (2004) *Public policy, complexity and rulebase technology* The Australia Institute Discussion Paper, tai.org.au/Publications_Files/DP_Files/DP67sum.pdf

Rawls J (1971) *A Theory of Justice* Belkup, Cambridge, Mass

Rice M (1998) *The Age Pension: Our unstable pillar* Institute Of Actuaries Of Australia Sessional Meeting

Skott P (1988) Finance, Saving and Accumulation *Cambridge Journal of Economics* 12.1

Smith R S (1990) Factors Affecting Saving, Policy Tools and Tax Reform: A review *IMF Staff Papers* 37.1: 1-70

St John S & L Willmore (2001) Two Legs are Better than Three: New Zealand as a model for old age pensions *World Development* 29.8: 1291-1305

Select Committee (2005) *Work and Pensions - Third Report* UK House of Commons Select Committee on Work and Pensions publications.parliament.uk/pa/cm200405/cmselect/cmworpen/43/4315.ht

Smith Committee (1995) *Report to the Minister of Finance of the Committee on Strategy and Policy Review of Retirement Provision in South Africa* Government Printer, Cape Town

Somogyi S, A P Gale, C O Helenius, D M Knox, H B Sarjeant & R J Stevens (1995) Financing The Ageing: Now is the time to act *A Discussion Paper prepared by the Financing the Ageing Committee of the Institute of Actuaries of Australia*

Stigler G J (1982) *The Economist as Preacher* Basil Blackwell, Oxford

Taylor V *et al* (2002) *Transforming the Present - Protecting the Future* Report of the Committee of Inquiry into a Comprehensive Social Security System for South Africa, Department of Social Development, Pretoria

Travers P (1995) The Delivery of Welfare Services: What level of access? *Proceedings of the Industry Commission conference on Equity, Efficiency and Welfare* www.pc.gov.au/ic/research/confproc/equefwel/equefwel.pdf

Tuomala M (1990) *Optimal Income Tax and Redistribution* Clarendon Press

Tyler T R & H J Smith (1995) Social Justice and Social Movements Institute of industrial Relations *Working Paper* University of California, Berkeley, http://ideas.repec.org/p/cdl/indrel/1088.html

Uslaner EM (2002) The Moral Foundations of Trust Cambridge University Press

Vittas D (1998) Regulatory Controversies of Private Pension Funds *World Bank Policy Research Working Paper* 1893 papers.ssrn.com/sol3/papers.cfm?abstract_id=620656#PaperDownload

Vito T & H Lee (2000) Tax policy for Emerging Markets: Developing Countries *IMF Staff Papers 00/35*

Wilkinson R (2005) *The Impact of Inequality - how to make sick societies healthier* The New Press

Willmore L (2001) Universal Pensions in Low Income Countries *Discussion Paper* Initiative for Policy Dialogue, Pensions and Social Insurance Section, Oxford Institute of Ageing

Yoo K-Y & M de Serres (2004) Tax Treatment of Private Pension Savings in OECD Countries and the Net Tax Cost Per Unit of Contribution to Tax-favoured Schemes *OECD Working Paper No. 406*
papers.ssrn.com/sol3/papers.cfm?abstract_id=607185#PaperDownload

CHAPTER 4 - CONFLICTED SUPER STRUCTURES: ARE AUSTRALIAN INVESTORS BEING SHORT-CHANGED?

Presented at the 12th Australian Colloquium of Superannuation Researchers - to be published by Edward Elgar as a chapter of the Selected Papers

1 INTRODUCTION

Are members of superannuation funds in Australia being short-changed? The different types of superannuation funds do charge significantly different fees. Some of the difference clearly arises from the quality and scope of the service offered, but there are other, less innocent, explanations for the high level of fees. Of particular concern are the conflicts of interest in some organizational structures.

The next section of this chapter identifies the various types of funds. The third describes the main kinds of regulatory protection provided to members, some of which give rise to additional expenses. The fourth section examines the role of market or regulatory failures in explaining the apparently excessive level of charges of some funds. The fifth section looks at the functions performed by fund trustees, and identifies potential conflicts of interest as an alternative explanation of additional charges. The sixth section looks at the charges associated with the different types of funds and identifies the main drivers of expense. The final section suggests some topics for additional research and addresses some policy questions.

2 TYPES OF FUNDS

Australian superannuation funds are classified in a variety of ways in legislation and in the industry statistics published by the Australian Prudential Regulation Authority (APRA)[40]. There are inconsistencies in the definitions depending on the purpose for which they are used, but the main classifications are summarized in Table 1.

---

[40] These can be found on the superannuation section of their website (apra.gov.au). They currently include the "Quarterly Superannuation Performance" and the "Annual Superannuation Bulletin", which replaced "Superannuation Trends" in 2005. The latter has been used in this chapter.

Table 1: Australia - Main types of superannuation funds

APRA CLASSIFICATION	Is employer sponsored	Operates in the public sector	Operates in the private sector
Funds with 1–4 members			
Small APRA funds	Possibly	No	Yes
Self-managed superannuation funds	Possibly	No	Yes
Approved deposit funds with 1-4 members	No	No	Yes
Funds with 5+ members			
Corporate funds	Yes	No	Yes
Public sector funds	Yes	Yes	No
Industry funds	Possibly	Possibly	Possibly
Retail funds	Possibly	No	Yes
Special purpose funds			
Pooled superannuation trusts	No	Possibly	Possibly
Approved deposit funds with more than 5 members	No	Possibly	Possibly
Eligible rollover funds	Possibly	Possibly	Possibly

Note: 'Possibly' means that the fund may in some cases have the characteristic in the column heading
Source: Author's mapping of the APRA statistical publications and SIS legislation

In this chapter, we are concerned only with complying funds.[41] Complying funds qualify for favourable tax treatment of investment income, and contributions to such funds may be tax deductible. These funds can be classified in four ways: by trustee, by contractual party, by benefit structure and by manner of investment.

2.1 By trustee

With the exception of retirement savings accounts, all superannuation funds must be managed by a trustee. The trustee is generally a corporate body with its own board of directors but in some circumstances may refer to a group of natural persons. I mean both

[41] Defined in part 5, division 2 of the *Superannuation Industry (Supervision) Act 1993* (SIS Act)

types when referring to trustees in this chapter. There are currently three classes of trustee licence, the non-public offer entity licence, the public offer entity licence, and an extended public offer entity licence[42] for trustees of a public offer fund that also operate other funds.

2.2 By contractual party

Superannuation funds can also be classified according to the party that initiates the contract of membership.

2.2.1 Employer-sponsored funds

A fund is a 'standard' employer-sponsored fund[43] if all contributions to the fund are made as a result of an arrangement between the employer and the trustees, and just an employer-sponsored fund if the contributions are made as a result of an arrangement between the members and their employer. APRA's statistical collections sort employer-sponsored funds into three types: corporate funds, public sector funds and industry funds. Corporate funds are sponsored by a single employer or group of employers in the private sector. Public sector funds are sponsored by the government or its agencies.[44] Industry funds are sponsored by a diverse group of employers in the same or a related industry, in either the public or the private sector, where the common link and driving force is often a representative union. Legislation in 2005 entitled most members to choose their funds.[45]

2.2.2 Public offer funds

Public offer funds[46] are open to individuals, self-employed or otherwise, who then have a direct arrangement with the trustees to which the employer may or may not be party. Some are employer-sponsored funds, especially industry funds that have opened their doors to the public. But most are commercial companies that initially catered to

---

[42] Defined in section 29B of the SIS Act and SIS regulation 3A.03.

[43] Defined in section 16 of the SIS Act.

[44] This definition is not to be confused with that given in section 10 of the SIS Act, which refers to funds set up by some law or under the control of a state body.

[45] See part 3A of the Superannuation Guarantee (Administration) Act

[46] Defined in section 10 of the SIS Act.

individuals, but later also to groups of employers. They are the retail funds of the APRA statistics.

Approved deposit funds[47] are public offer funds whose operations are restricted to receiving a transfer of benefits from other funds or certain payments from the Australian Taxation Office (ATO).

A pooled superannuation trust[48] is a unit trust whose only investors are other superannuation entities. Life insurance companies may invest in a pooled superannuation trust; designate a portion of their assets as a 'virtual' pooled superannuation trust; or place the assets backing annuity products in a pooled superannuation trust of their own. The assets backing life insurers' annuity products are shown in the APRA statistics as the 'balance of statutory funds'. Pooled superannuation trusts are not counted in the statistical returns, as this would result in double counting of superannuation assets. No statistics are collected for virtual pooled superannuation trusts.

Retirement savings accounts[49] are open to applications from individuals and operate like bank deposits. They are intended for people with small balances in their superannuation funds, and for those who wish, for short periods, to make use of the capital guarantee they offer. They do not have registered trustees and are not defined as public offer funds. Members have a contractual relationship with the bank or life insurance company offering the account, which may then be seen as having a fiduciary responsibility towards its account holders.

2.2.3 Member-initiated funds

There are two types of small, member-initiated funds. The first are the self-managed funds, in which each trustee is also a member of the fund.[50] They are regulated by the ATO rather than APRA because no prudential risk is involved. The second are the small

---

[47] Defined in section 10 of the SIS Act.

[48] Defined in section 10 of the SIS Act and regulation 1.04 of the SIS Regulations.

[49] Defined in, and administered in accordance with, the Retirement Savings Accounts Act 1997.

[50] Defined in section 17A of the SIS Act.

APRA funds.[51] Established by members, their operations are delegated to a separate trustee. They are regulated by APRA because of the prudential risk involved.

2.2.4 Eligible rollover funds

Trustees face various restrictions[52] in the amounts they can charge in administration fees. If they believe that a member's account is too small to be managed economically, they are allowed to transfer that person's benefit into an eligible rollover fund (without first obtaining the permission of the member).[53] Eligible rollover funds are intended to provide low-cost services for small accounts, and so are restricted in what they can charge.

2.3 By benefit structure

Defined benefit funds are defined variously in the SIS Act as funds that have at least one member with a prospective benefit related to salary, or taking a fixed pension.[54] Accumulation funds[55] are defined as those funds that are not defined benefit funds. The APRA statistical returns add an intermediate category of hybrid funds that have both defined benefit and accumulation members. However, because trustees are allowed to categorize funds subjectively, some retail funds with defined benefit members have defined themselves to be accumulation funds for purposes of the statistics.

---

[51] These are defined as 'certain small funds' in section 121A of the SIS Act.

[52] See part 5 of the SIS Regulations.

[53] Defined in part 24 of the SIS Act.

[54] Under section 83A of the SIS Act, a defined benefit fund is a public sector fund that has at least one defined benefit member (that is, a member who is entitled at some point to be paid a benefit related to salary); or a private sector fund that has at least one defined benefit member, and arrangements whereby contributions are not hypothecated to individual members. Section 228 dispenses with the latter limitation but widens the definition of a defined benefit member to include members receiving a fixed pension. Regulation 1.03 of the SIS Regulations uses the section 228 definition, while the Superannuation Guarantee (Administration) Act 1992, which sets out the compulsory minimum contributions to superannuation funds, uses the definition given in section 83A.

[55] Defined in regulation 1.03 of the SIS Regulations.

APRA statistics show that over two-thirds of the assets in Australian funds are in accumulation funds, with over 90 per cent of the balance residing in funds defining themselves as hybrid funds. The accumulation funds can be divided further into those that offer choice of investment type and those that go further to offer choice of investment manager. The latter are often called master trusts, although there appears to be no legal definition of this term.

Table 2: Manner of investment of funds' assets

APRA CLASSIFICATION	Assets held by the fund	Tax on investment earnings
Directly invested	Shares, property, fixed interest etc. managed internally	Paid by fund on direct income
	Shares etc. managed under a specialist mandate	Paid by fund on direct income
Placed with investment manager	Wholesale or retail unit and property trusts	Paid by fund on income that is passed on by the manager
	Pooled superannuation trusts	Paid by pooled superannuation trust
Held through life insurance companies	Life insurance policies	Paid by life insurance company

Source: Author's mapping of APRA(2003) and taxation legislation

2.4 By manner of investment

The underlying assets of a fund can be held directly, or through a range of alternative vehicles. Table 2 shows how the manner in which APRA classifies fund investment is related to the type of investment held and the tax consequences of each type.

3 MAIN FORMS OF REGULATORY PROTECTION

This section discusses the main forms of regulatory protection provided to fund members. Some regulations can be expected to limit the costs associated with superannuation; others may increase them. The discussion focuses mainly on the APRA-regulated funds (that is, those managed by people other than the members), and on accumulation funds.

3.1 Fiduciary responsibility

The commercial law framework for Australian superannuation funds is based on the law of trusts. It is set out, inter alia, by Glover (2002):

> Investors, and more particularly members of superannuation funds, enjoy significant protection from the common (or general) law. Corporate officers and advisers with whom we are concerned are disciplined at general law as 'fiduciaries'. The term refers to the law's code for the maintenance of the honesty and integrity of persons in positions of ascendancy and trust. Its centrepiece is an 'inflexible rule' which prohibits fiduciaries, such as corporate officers and advisers, from putting themselves in positions where their interest and duty conflict.[56]

'Interest' refers to the personal financial interests of fiduciaries, 'duty' to their duty to members or perhaps shareholders. Glover goes on to identify conflicts of interest and duty where a director may owe different duties to the shareholders of different companies. He points out that the traditional formulation of the rule is strict: 'Possibility of a conflict is enough to attract the rule. … Exposure of one's self to temptation is made wrongful. Directors, as it were, are not given the chance to be dishonest' (Glover 2002). He describes a failure in this area as fraud.

The prohibition on exposing oneself to conflicts of interest is common in the rest of the world, where it is also sometimes included in the basic law of agency. Civil law jurisdictions, which do not recognize trusts, incorporate the provision in other law governing fiduciaries.

The responsibilities of a fiduciary also include a requirement to be prudent, and a duty of care.

3.2 The main legislation

Superannuation funds are subject to various, largely concessionary, tax provisions set out in the Income Tax (Assessment) Act 1936 and the Income Tax (Assessment) Act 1997.

---

[56] Glover is here quoting from *Bray v Ford* [1896] House of Lords Appeal Case 44, 51; see also *Phipps v Boardman* [1967] 2 House of Lords Appeal Case 46, 123.

The confusingly named Superannuation Guarantee (Administration) Act 1992 requires employers to make superannuation contributions on behalf of employees, currently set at 9 per cent of remuneration. The 'guarantee' arises from the obligation placed on employers to guarantee that they have made the required contributions. The government takes on no financial obligations and there are no investment guarantees.

The Superannuation Industry (Supervision) Act of 1993 and its accompanying set of regulations—the Superannuation Industry (Supervision) Regulations 1994 (henceforth, the SIS Act and Regulations)—provide for compliance with tax, guarantee and other policy issues, and for prudential supervision. They can be seen as codifying, adapting and extending the provisions in the general law.[57]

3.3 Equal representation on the board

The board of trustees of retail funds will normally consist of individuals with varying degrees of commercial connection with the shareholders. Other boards of trustees may consist of member representatives, employer representatives and independent trustees. The SIS Act requires half the trustees of a standard employer-sponsored fund to be member representatives,[58] who may be elected directly or appointed by a representative union. Employer representatives may also be appointed by an organization representative of the employer's interests. While employer-appointed trustees can be expected to have greater technical knowledge of issues related to superannuation, member-appointed representatives are likely to be better at communicating with members. The trustees' duty to place members' interests first is reinforced by this diversity in the right to elect or appoint trustees.

Public offer funds that have standard employer-sponsored members are only required to have policy committees on which members must be represented. The latter have no rights beyond the right to be consulted.

Pensioners are not entitled to representation on fund boards.

---

[57] In particular, part 6 of the SIS Act sets out provisions that require trustees to be prudent and free from inappropriate influences (direction).

[58] Part 9 of the SIS Act.

3.4 Fitness of trustees

APRA has the power to disqualify inappropriate persons from functioning as trustees, actuaries or auditors.[59] The skill level required of a trustee in trust law is that of 'an ordinary prudent person'[60] without any particular skills in the management of superannuation. Recent regulations[61] however require the trustees of a fund to have 'skills relevant to the duties of a trustee', and APRA's fit and proper guidance note suggests that the trustees as a whole should have 'sufficient knowledge' to make 'informed decisions'. APRA therefore appears to require somewhat higher standards of knowledge and skill than the law.

3.5 Investment choice

The protection of the general trust law against conflicts of interest in the area of investment choice is overridden in at least two sections of the SIS Act. First, the SIS Act makes specific provision to permit investments made at less than arm's length.[62] The protection given is that the price must be determined on an arm's-length basis. APRA says of investments made on an arm's-length basis: 'The test to apply is to consider whether a prudent person acting with due regard to their own commercial interests would have made such an investment'.[63] Second, the SIS Act contains complicated provisions that expressly permit in-house investments in related parties, including the employer sponsor, under some circumstances.[64] The requirement in trust law to invest prudently is made explicit, and the trustee is required to develop a strategy that considers investment risk, diversification, liquidity and the ability to meet liabilities.[65]

---

[59] Parts 15, 16 and 17 of the SIS Act.

[60] Section 52(b) of the SIS Act.

[61] SIS Regulation 4.14 (4)

[62] Section 109 of the SIS Act.

[63] Superannuation Circular No. II.D.5.

[64] Part 8 of the SIS Act.

[65] These are set out in section 52(8) of the SIS Act.

3.6 Accounting procedures

Trustees are required to keep the assets of the fund separated from their own; prepare minutes, records and annual accounts; and make them and any other relevant information available to members.

3.7 Risk management planning

The SIS Regulations require trustees to submit a formal risk management strategy and business plan to APRA. The business plan must specifically address all the risks faced by the trustees and the fund, including:

- governance-related risk (definition, delegation and segregation of responsibilities, outsourcing, fraud);
- operational risk (information systems and records, strategy);
- investment-related risk (market, credit or liquidity); and
- risk mitigation through insurance.

3.8 Waiver of charges on small accounts

Some accounts of less than A$1,000 are given member protection, to ensure that the administrative charges on the accounts do not exceed their investment earnings.[66] This is a rather wasteful element of the system given the miniscule value of the long-term benefits; after all, A$1,000 represents less than a year's contributions for those on the minimum wage. In the meantime, uneconomic small accounts are proliferating—there are over 25 million accounts but fewer than 10 million workers. A high proportion of the uneconomic accounts belong to young, part-time and casual workers with low wages and high turnover. They would be much better off if they had the money to repay their

[66] Defined in part 5 of the SIS regulations in terms of the general power to make regulations set out in section 31 of the SIS Act.

debts.[67] In other instances, members should be encouraged to consolidate their accounts rather than be given this trivial protection. The member protection rules do, however, establish the principle that charges can be capped.

4 MARKET, REGULATORY AND INSTITUTIONAL FAILURE

The puzzle that this chapter addresses is the difference in costs between retail and industry public-offer funds. Retail funds charge members 1–4 per cent of the account balance per annum in fees, in addition to upfront deductions of up to 5 per cent of the contributions. Industry funds offer a similar if not identical service for charges of not much more than 1 per cent per annum, normally with no entry charge, despite having much smaller account balances on average. It is sometimes held that the apparently excessive costs of the retail funds are due to some form of market or regulatory failure. This section briefly analyses this possibility.

4.1 Market failures

The standard economic explanations for market failures of this sort are monopolistic practices and information asymmetries.

4.1.1 Monopolistic practices

There are a dozen or more large institutions competing in the public offer market, as well as many smaller players. Given the relative ease of entry and the numerous participants, one might expect a high level of competition and a low level of fees in the industry. The Productivity Commission (2001), which specifically investigated the effect of legislation on competition and barriers to entry, found no evidence of a lack of competition in the industry. Partly as a consequence, it recommended additional restrictions to entry in the form of compulsory licensing of all trustees, with a view to giving members additional regulatory protection. In studies undertaken for industry bodies, Clare (2001) and Rice and McEwin (2002) both concluded that the charges of funds largely reflected a competitive industry. The Australian Competition and Consumer Commission (ACCC),

---

[67] The notion that people should begin to save early for retirement is demonstrably wrong—certainly if they want to own their own homes. This is because the net yield on superannuation assets is likely to be much lower than the interest on a mortgage, let alone other consumer debt.

in a more detailed 1998 study of the consequences of a proposed merger between two large players in the funds industry, also found no evidence of a lack of competition in the industry.[68] Thus, if fees are too high, it would appear to be necessary to look elsewhere for an explanation.

4.1.2 Information asymmetries

It has been argued that funds are able to charge excessively high fees because members, or their employers, are kept poorly informed about the full extent and level of the fees being charged. This possibility has led to considerable debate about the adequacy of product disclosure, especially the need for improved comparability of fee structures across funds, and for the disclosure of commission earnings by intermediaries.[69]

The Australian Securities and Investment Commission (ASIC) is responsible for regulating the behaviour of all financial market participants, including the trustees of public offer funds. ASIC also regulates financial advisers. Legislative changes in the past 5 years [70] have set new standards for the disclosure of information, including requiring the licensing of financial advisers by ASIC.

There can be no doubt that the fee structures of superannuation funds are complex but not insuperably so for a reasonable person prepared to devote time to the subject. A report by ASIC (2004a) provides a basis for determining a standard approach to disclosure. It lists eight different types of fees and makes some suggestions for uniformity. Further regulatory developments in this area seem inevitable.

---

[68] Of the takeover by the National Australia Bank group of Lend Lease in 1998, it says:

"The merger does not breach the concentration thresholds in any possible market. A number of possible markets were examined: broad 'funds management' market; 'wholesale' and 'retail' funds management markets; separate markets for life insurance, superannuation, retail investment products; and narrow markets based on particular products (the latter are the least likely market definitions due to considerable demand- and supply-side substitutability). E.g., wholesale funds management—merged entity 7.4%, post-merger CR4 [concentration ratio of the top four firms] 35.2%; life insurance 22% and 60%; super 13.8% and 46.8%; retail investment products 20.3% and 50.3%" http://www.accc.gov.au/content/index.phtml/itemId/538590 accessed 23 March 2007.

[69] Some of the issues related to disclosure are discussed at greater length in Bateman (2001).

[70] The Financial Sector Reform Act (2001) and subsequent changes to the Corporations Act (2001).

The investment management component of superannuation charges is particularly transparent and subject to competition in master trusts, which offer a wide choice of investment manager. As discussed below, however, the fees of master trusts do not appear to be particularly competitive.

Again, it seems that information asymmetry does not offer a sufficient explanation for the high fees charged by some funds, suggesting that other factors must be at work.

4.2 Regulatory or legal failures

4.2.1 Fiduciary responsibility

The inflexible prohibition of trust law against conflicts of interest and duty is not always obeyed, or enforced. It is often argued that it is inconsistent with modern commercial practices—mainly those of directors with many interests or of large commercial firms providing advice to a large number of clients.[71]

'Everybody is doing it' is, however, a poor argument. It is easily rebutted by 'but that does not make it right'. The requirement to avoid conflicts of interest and duty is an ancient legal doctrine that provides one of the foundations of the division of labour that allows for our current economic prosperity. It follows from the recognition of human frailty in the face of temptation. Undermining it may risk undermining all contracts of agency, all delegation, all trusteeship, all professionalism.

The concept of fiduciary responsibility is particularly important in long-term contracts such as superannuation. It is difficult if not impossible for some employees to change funds. In such circumstances, competition and symmetrical information at the point of sale are of little value; the essential point is to ensure that the fund continues to be managed on behalf of the members. Of course, the need to attract new members and retain those that do have the option to move provides an ongoing competitive discipline, but this is not necessarily enough.

Eliot Spitzer, the New York attorney-general from 1998 to 2006, has stated that conflicts of interest have become institutionalized in many parts of the financial services industry

---

[71] Teele (1992) provides an example, and gives instances of legal judgments and sections of the Corporations Act that have diluted the effects of the law.

(O'Brien 2004). His efforts to prosecute a number of financial firms for succumbing to conflicts of interest led to apparently reputable organizations in the United States paying billions of dollars in fines and compensation. This scandal was similar in many respects to the mis-selling problems of the UK life insurance industry in the 1990s.

It appears that, prima facie, conflicts of interest cannot be ruled out as an explanation for excessive fees in the Australian superannuation industry.

4.2.2 Agency risks

Agency risks have been investigated in the economics literature over the past three decades, although, perhaps surprisingly, they are not often mentioned in the context of market failure. The cure suggested in the literature is greater monitoring of agents and better alignment of the incentives of agents and principals.[72] The increased disclosure that has thus far been the major regulatory tool in Australia perhaps arises from this economists' perspective. But an alternative more legal view, expressed below, is that the key to the management of agency risks lies in a more appropriate application of the law of agency.

4.2.3 Specific prohibitions

The essential prohibitions in the law of agency are prohibitions of conflicts of interest, which underline the fiduciary's duty of loyalty:

- An agent or trustee is entitled to remuneration but should not make a secret profit. The level of remuneration should not only be disclosed, but should be agreed— in this case—by the member or sponsoring employer.
- An individual should not enter into any contract both as a principal and as an agent or trustee of another person.
- The agent or trustee should not make a profit, disclosed or otherwise, that might be seen as influencing the agent's decision to the possible detriment of the principals—or in this case of the members.

---

[72] Debatable perhaps, but see the definition of agency costs in *The Economist*'s useful economics dictionary at <http://www.economist.com/research/Economics>.

This is not to say that it is always clear whether conflicts exist or not. Even if there are no conflicts of interest, principals cannot be sure that their agents perform their duty of care, being diligent and earning their remuneration. These problems are not however addressed here.

The three prohibitions set out above are examined in more detail in section 5 of this chapter, where I look at the functions of trustees in greater detail. Before doing so, however, it may be helpful to look at some of the insights of institutional economics.

4.3 Institutional factors

Keneley (2004) makes reference to some of the relevant literature in the field of institutional economics. This literature focuses on issues relating to the structure of firms and of markets, and the way that these have evolved to deal with transaction and information costs. Both within firms and in the market, unwritten 'rules of the game' can govern behaviour—with both positive and negative effects.

4.3.1 Regulatory capture

Of some interest in this context is the theory of collective action, which predicts that small organized groups can influence regulation and capture benefits from the rest of society. One example of this is the recognition that large organizations may, intentionally or otherwise, influence regulation to make it more difficult for small organizations to flourish.

The tax, disclosure and prudential regulations surrounding Australian superannuation have reached what is perhaps an incomparable level of complexity. The main superannuation acts and regulations run to over 800 pages; the 2001 financial sector reforms added 150,000 words to this legislation. Advisers must understand not only the difficult provisions of the two income tax acts but also the rules for social security benefit payments. These complexities cannot but entrench the position of large firms that can afford the necessary specialists, thus making it more difficult for smaller firms to compete and driving up the price of administration. This point was recognized but dismissed by the Productivity Commission (2001) without giving reasons. It should perhaps be revisited.

4.3.2 Rules of the game: investment 'cults'

Given the common and not entirely unfair references to the 'cult of the equity', some consideration may be given here to the belief system surrounding investment markets. The belief that it is possible to beat the market appears to be more widespread than is justified by the research. Investment managers, operating on behalf of fund members, fulfil the important additional economic function of allocating savings to where they are most productive. If the investment managers are successful in their task, the expected return on all assets, after adjustment for risk, will be the same for all managers. Put differently, no manager will be able to outperform any other manager when considered *ex ante*. The intervention of random events does mean, however, that the returns produced by different managers will differ *ex post*—even in an efficient market.

It is a difficult problem to discern whether the market is efficient and whether the *ex post* difference results from skill or is random. Behavioural economists have found that people are prone to recognise patterns in random events where none exist. Ferson *et al* (2002) found that this problem extended even to experts using sophisticated statistical tools. Blake and Timmermann (2004) reported findings showing that UK investors were influenced excessively by good past performance in spite of its possible lack of meaning.

The question of whether the investment performance of managers persists over time has been researched intensively. In a finding that remains largely unchallenged, Carhart (1997) showed that it does—but only in the case of underperforming funds. He found weak evidence of persistent outperformance among funds that were able in some way to quickly capture new evidence in the market. On the other hand, he found strong evidence of persistent underperformance by other funds. Part of the underperformance can be explained by excessive trading; the rest must come from the purchase or sale of investments at uneconomic prices. One possibility is that managers are trading at an uneconomic price for some personal benefit, at the expense of fund members. Blake and Timmermann (2004) conclude that information on past performance should be made available to members so that they are able to avoid underperforming funds.

Coleman, Esho and Wong (2003) found some evidence of the persistent underperformance of retail funds in Australia. This requires further investigation; a starting point would be to establish whether conflicts of interest can explain the transactions concerned.

In spite of the difficulties outlined above, there is some assistance available to help retail investors choose the best managers. This includes the specialist media, including the financial sections of most newspapers; financial advisers to individuals including the apparently increasing availability of fee-for-service financial advisers; investment consultancies that collect and analyse data; asset consultants, who largely advise companies; and managers of managers, who work for master trusts. At this stage there appears to be no evidence one way or the other that any of these are successful in distinguishing between the investment performance of two apparently competent managers. They may, however, be able to help investors avoid the less reputable managers.

Why does this industry not focus more intently on what appear to be excessive expenses? In spite of the advances in computers that should have reduced costs over the past 30 years, charges for investment management can be two or three times what they were in the 1960s. Moreover, they are two or three times higher than the costs actually reported.[73] This represents a real puzzle given the competition within master trusts referred to earlier.

Taleb (2001), who provides an insider's view of the funds management industry, repeatedly makes the point that completely random results will lead to wide differences in the performance of investment managers. Random differences are the result of luck, not good judgment, but people (and he includes himself in this) are not really able to distinguish between the two. The consequence appears to be that we are too easily persuaded to pay too much for investment management.

This position may be overstated. The excessive investment charges (if they exist) are a small percentage of assets, and make little difference when accounts are small. It is justifiable, in some instances, to ignore them. But as funds mature and their average account balances increase, the amounts involved become considerable. It may be that members do not adequately appreciate the difference, which is why greater disclosure and more investor education are needed. This education should focus mainly on teaching

[73] Investment expenses, as reported by life insurance companies, amount to some 0.2 per cent of assets. Even allowing for some outsourcing not included in this figure, this is considerably lower than the 0.5–1.5 per cent charged under typical circumstances.

investors that past performance does not govern future returns—except perhaps in the case of underperforming managers, who are likely to continue to perform poorly—and that fees make more of a difference to returns in the long run.

4.3.3 Rules of the game: commissions

The early 1990s saw growing dissatisfaction with traditional front-end-loaded insurance and superannuation savings contracts. In these, as much as the entire first-year premiums on a new policy were often allocated to commission and other selling expenses. The apparent excesses of the front-end-loaded commission system led to a move to a new system involving much smaller front-end charges on savings contracts but the addition of trail commissions. The effect of this change was to reduce the initial costs of taking out a policy, but to increase it over the term of the policy, as explained in Appendix 1. Even a 0.5 per cent annual trail will erode the value of a person's contributions by 30 per cent after 30 years. It may be that the long-term effects of the new system were not initially evident, and that changes will be made as the larger trails begin to bite. Many people are evidently making such changes already, as can be seen from the growth in self-managed funds, which are more cost effective once the value of an account exceeds about A$100,000.

4.3.4 Social relationships

Another perspective is given by two sociologists. In a careful two-year study of the culture and behaviour of nine large pension funds in the United States, O'Barr and Conley (1992, page 70) confirm the author's experience of 'surprising and sometimes disturbing' evidence of 'an unsystematic approach' to investment decisions. They observe that 'relationships are often more important than managing the bottom line in evaluating and deciding whether to retain managers'.

5 TRUSTEES AND POTENTIAL CONFLICTS OF INTEREST

We now turn to the ways in which conflicts of interest may lead to excessive costs. Figure 1 shows the main cashflows in the superannuation relationships, and highlights flows that may arise from potential conflicts of interest. The discussion that follows is grouped around the main functions of trustees: distribution, administration and investment.

5.1 Distribution

Something must bring members to a fund initially. Trustees must then find ways of persuading members or their employers to continue to contribute to the fund, and of attracting new members. This could also be called the marketing function.

5.1.1 Non-public-offer employer-sponsored funds

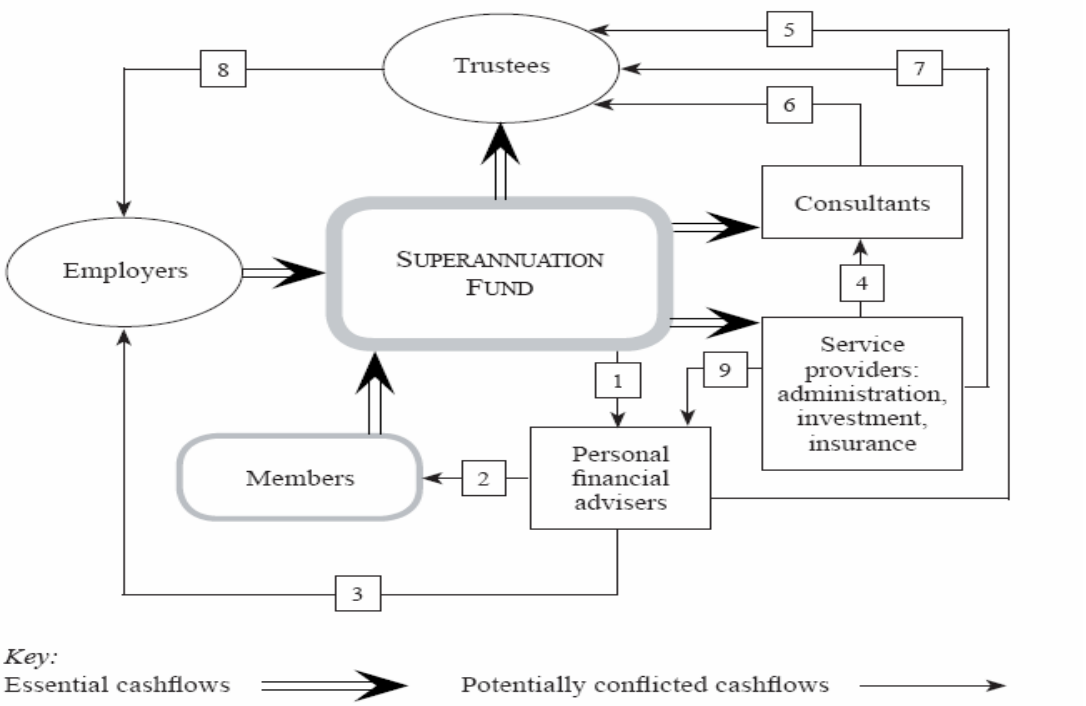
The trustees of corporate and public sector funds deal directly with employers and members, thus avoiding direct distribution costs. However, they still need to justify the fund's ongoing existence, and to market its benefits. Trustees wear two hats if they are also directors or managers of the employer. This creates a potential conflict of interest when discretionary benefits are discussed: as trustees, managers are supposed to put the interests of members first, but as managers, they have an interest in reducing the employer's contributions.

The equal representation rules provide some balance in this respect. They can also be expected to create a certain competitive tension between employer representatives and member representatives on the board of trustees, although there is some debate about the productivity of this tension. It has been suggested that union directors are valuable members of the board because they are able to provide information that might not otherwise be available, and because they often have more time to devote to the board than other outside directors (Anonymous 1982). But Williamson (1984, page 1206) worries about the appointment of union directors 'deflecting strategic decision makers from their main purposes … by forcing them to address operating level complaints …' and, more seriously, about 'the problem of opportunism that inclusion of partisan constituencies on the board invites.'

5.1.2 Public-offer industry funds

Industry funds with direct links to sponsoring employers are also able to avoid direct distribution costs and enjoy the 'benefits' of any tension between member and employer-appointed trustees. In cases where trustees are not members of the fund, there is an agency risk that they will pursue their own interests at the cost of members.

Figure 1: Sources of Conflicts



Flows causing potential conflict of interest and duty:

1	Adviser owes principal duty to member but is paid by fund, trustee or life company—sometimes in soft dollars that are dependent on volumes.
2	Adviser passes on cost of additional commission and rebates to member, so reducing the member's superannuation account.
3, 8	Adviser, trustee or related-group company rebates commission or provides other benefits to employer.
4	Service provider makes discretionary or soft dollar payments to consultant.
5, 6, 7	Trustee or group makes profit from holding in or payments from adviser group, service provider or consultant.
9	Financial adviser receives payment from service provider.

The non-profit nature of industry funds also sets up a tension between them and the retail funds that is identical to the tension between mutual and for-profit insurance companies. The non-profit industry funds and mutuals do not face the same conflicts of interest between members and shareholders as the for-profit retail funds and insurance companies. It is however widely held that their trustees or directors are not subject to the same level of discipline by members or policy holders as that imposed on the elected

directors of retail funds and insurance companies by company shareholders (see, for example, HM Treasury 2004).

In developing a distribution channel, industry funds rely on direct contact with employers; union pressure on employers to persuade them to join; and, more recently, advertising to attract individual members.

5.1.3 Retail funds

Retail funds do not have the inherent links to members of the funds just discussed, and so must market themselves actively if they are to attract and retain members. The retail funds are mainly invested in life insurance policies,[74] and have in large measure inherited their distribution systems from the life insurers that initiated them. While superannuation may be distinguished from life insurance, it is probably unhelpful to make the distinction in this context. Many funds pay for life and disability insurance for members. The remuneration of the sales channel will include commissions from insurance sold to the same clients for insurance that is both inside and outside the superannuation fund.

The life insurance sales system can be classified into four channels. The first is the traditional high-pressure channel, paid purely on commission. It is peopled by agents, normally in some 'dealer group', who sell only one company's products, and by brokers who sell the products of more than one company. These people have been described as the archetypal salespersons. A scholarly analysis of their social status, and the extreme moral and financial pressure placed on them, is given by Guy Oakes in a book reviewed by Smith (1991: page 287):

> The general thesis of this book is that these salesmen … are subject to two inherently conflicting principles: a commercial idiom that stresses sales at all costs and a service idiom that stresses concern and sensitivity for the needs of clients. … [T]he author reviews the five essential steps of selling insurance: prospecting, approaching, interviewing, closing (making the sale) and providing service to clients. … After presenting both idioms, he offers a brief discussion of the major antimonies of the profession, namely, sacralization and manipulation, opportunism

[74] Retail funds held 57 per cent of their assets in life insurance companies in 2003, as can be seen in Table 3 discussed in section 6.1.

> and professionalism, toughness and sensitivity, sincerity and dissimulation. The book ends with an analysis of the ways in which most salesmen succeed in remaining ignorant of their own conflictual situation.

Given the unpleasantness and costs of this system, it may be asked why it perseveres. Zultowski (1979) surveys research on the widespread belief that life insurance 'is sold and not bought' that provides some justification for the high-pressure approach. The agents reported that less than 20 per cent of sales were initiated by the policy holders, but the policy holders felt the statistic was closer to 50 per cent. There appears to be no formal research into the matter since then but there is also little evidence that the position has changed.

In Australia, where superannuation contributions are compulsory, this argument cannot be said to hold. It may be argued, however, that high-pressure sales forces actively seek out small businesses and the self-employed and persuade them to contribute to superannuation, thus contributing towards greater compliance with the law.

Berger, Cummins and Weiss (1997), who investigated the coexistence of insurance distribution systems with widely differing fee structures, found that higher charges did not translate into higher profits, although they did seem to be related to a better-quality service. It would be an exercise beyond the scope of this chapter to determine whether any of the participants in the industry earned monopoly-type profits. What is clear is that the additional costs of the distribution system of retail funds accounts for a significant proportion of the difference in charges between them and industry funds. An open and perhaps more pressing question is whether the extra costs are justified.

A conflict of interest occurs when commission-remunerated salespeople purport to be 'financial advisers'—a conflict that is aggravated if they claim to be independent. A strict interpretation of the law on conflicts would prohibit salespeople from purporting to give 'advice', and then receiving a commission from a third party that depended on the nature of that advice. However, the compensation given to those who had been mis-sold life insurance in the United Kingdom was not based on this strict interpretation of the law, but rather on legislation that required advisers to demonstrate they had given the 'best advice'.

In the Australian context, it would seem that members of corporate and industry funds receive less expensive advice than those in retail funds. On either a best advice rule or a

strict interpretation of the law, this might seem to expose some retail funds to similar demands for compensation. However, the SIS Act and the Corporations Act expressly permit the payment of commissions to advisers, thus undermining a strict interpretation of the general law.

The second main channel for life insurance sales is less pressured but also commission based. In this channel, the incomes of the salespeople are not entirely dependent on life insurance or superannuation sales; they may also broker general insurance, occasionally recommend a product in their role as professional accountants or lawyers, or be based in bank branches where they also serve customers' other needs. Their business can be said to come to them; they are described as the 'farmers' of the life insurance business, against the 'hunters' of the first channel who have to go out and find their customers. For brokers, bank branches that generate leads can be twice as efficient as having to prospect for clients. To the extent that they are commission remunerated, they face the same conflicts as the salespeople in the first channel. But to the extent that they cannot justify their commissions in terms of the need to find clients, they may face a greater burden in justifying high commissions. Some members of this channel charge fees for their advice and deduct their commissions from these. In such cases, there is no conflict of interest.

A problem, which applies to both channels and was investigated by ASIC (2004b), develops when a company pays overriding (additional) commissions to apparently independent brokers if they place larger volumes of business with it. While widely accepted previously, this practice is now discouraged.

The third channel for life insurance sales is the bank branches where employees are paid regular salaries and bonuses rather than by commission. Here it would seem to be the bank that faces the conflict of interest. Glover (1995) suggests that banks may in some circumstances owe a fiduciary duty to their customers. To the extent that a bank, and not its employees, receives the commissions paid by a retail superannuation fund, and thus makes a secret profit, there is a conflict of interest between it and its customers. Determining this profit is not easy. Australian banks report relatively small profits from superannuation business, even though the estimates given in the next section suggest that superannuation and investment management charges make up perhaps half of their consolidated fee income from all sources, or 25 per cent of their total revenue. The actual contribution to profit depends on a relatively subjective allocation of overheads, so cannot easily be determined.

The final channel is direct distribution of life insurance products by mail or internet. There is unlikely to be a conflict of interest if the charges are adequately disclosed, although a problem may arise if a bank or other apparently reputable organization endorses an overpriced product. A bank that endorses a product priced at several times the going market rate may well be exposing itself to subsequent claims for damages.

5.1.4 The role of employers

Employers are compelled by law to contribute to superannuation on behalf of their employees. Employers who are not governed by an award would appear to owe a common law duty of care and good faith to their employees when selecting appropriate superannuation fund.

It is clear from the decline in the number of corporate superannuation schemes over the past few years that there are advantages for employers in merging their funds into larger industry and retail funds. APRA data on fund transfers show that it is the retail funds that have benefited most from this trend, even though they charge considerably more than the industry funds, and often because of the commissions that are paid to the intermediaries who initiate such a move.

A significant factor in this trend has been the willingness of retail funds to tackle the costs and other difficulties associated with administering residual defined benefits. Although defined benefit funds do not need to do as much data processing as accumulation funds, and so ought to be cheaper to run, their greater need for actuarial services creates a significant overhead. As a result, the smaller defined benefit funds, which often suffer from poor systems and controls, can be considerably more expensive to run than accumulation funds. While the retail funds define themselves as accumulation funds in the APRA statistics, some have been active in acquiring hybrid schemes—a market that appears of less interest to the industry funds. Retail funds have apparently also been more flexible in accommodating other idiosyncrasies in the fund rules.

These factors may explain the success of retail funds in acquiring defined benefit and hybrid funds, but they do not explain why they have been able to acquire so many accumulation funds. Is it because employers are reacting against union domination of the industry funds? Are the retail funds so very much more efficient administratively? Or is it that the industry funds do not give enough support and advice?

Or one could be more suspicious. Do employers receive benefits from dealing with commissioned advisers in addition to advice on the superannuation fund? Is it possible that employers' duties to employees are being compromised by pressure from bank managers? The latter possibility is suggested by the market share gained by the life insurance companies attached to the four large banks. The growth in their assets over and above that of other life insurers, during the years 1998–2003, suggests that the four large banking groups captured over 70 per cent of the flows out of corporate funds in 1998–2003, despite having an overall market share of only 40 per cent.

5.2 Administration

Trustees have to collect contributions and keep records of all member details. They must pay benefits to members, or arrange for members' assets to be transferred into other funds. Other administrative tasks include the payment of taxes and levies, ensuring fund compliance with legislation and regulations, and amending the trust deed where necessary. In this context, the provision of life and disability insurance can also be regarded as an administrative service. Trustees need to be aware that some of their duties may give rise to a conflict of interest.

5.2.1 Outsourcing

Outsourcing may give rise to conflicts of interest. If trustees pay for the costs of outsourcing out of their own fees, paid to them by members, a conflict of interest does not arise. But if instead they use the fund's money to pay for an outsourced service, then a conflict may arise if the supplier is linked in some way to the trustee. Trustees face a conflict between their duty to members and their interest in, or duty to, suppliers, if they exercise their own discretion when choosing and paying an outsourced supplier. This problem is particularly relevant in the retail fund sector, where trustees and service suppliers frequently form part of the same group of companies, and in industry funds, where union-appointed directors often have links with union-owned suppliers.

A second area of conflict arises when the consultants to a fund recommend an outsourced service provider, and in consequence receive a commission from that service provider. An egregious example would be for consultants to recommend unnecessary or expensive insurance on which they are paid a commission.

5.2.2 Delegation

Not many trustees are experts on superannuation; they are therefore reliant on the help of experts in the administration of their funds and the investment of assets. Some trustees may be tempted to abdicate and yield power to the professionals. An argument can be made that members are best represented by lay (non-expert) trustees who control the professional service providers. It needs to be emphasized that this system requires some diligence on the part of lay trustees if they are to preserve the balance of power. Professional service providers will always face the temptation to usurp power, for reasons of efficiency as well as commercial interest.

Two of the trends currently observable in the retirement fund industry illustrate the dangers of giving greater power to the professionals. The first is the tendency to devolve the power of choice of investment manager from trustees to members, even though the members are in a weaker position than the trustees when it comes to monitoring the value of investment management services. This surely explains some of the increase in costs that has inexorably accompanied the introduction of choice. The second is the suggestion that paid professional trustees sit on the board, as recommended, for instance, by Myners in his review of institutional investment in the United Kingdom (HM Treasury 2001). This again has apparent advantages, but clearly weakens the position of lay trustees on the board. If trustees want independent advice, they would be better advised to contract for it rather than surrender some of their powers.

5.2.3 Trustee discretion over benefits

Another potential area of conflict arises when trustees who are also members of the fund amend the trust deed or exercise their discretion in the governing rules to increase the benefits payable to one or more class of beneficiary. This often involves the application of reserve funds.[75] It seems impractical to prevent members from being trustees. Nevertheless, they do need to be conscious of when their interests are affected; ensure that those interests are appropriately disclosed; and absent themselves from votes on such matters.

---

[75] Maclean (2000) discusses how this problem can be managed and how it has been dealt with by the courts.

5.3 Investment

The money collected from members has to be invested with a view to providing the best possible mix of risk and return. Trustees normally delegate the choice of assets to a professional investment manager. The assets themselves are normally held by a custodian who buys and sells on instructions from the investment manager. Outsourcing this function to organizations linked to the trustees potentially creates the same kinds of conflicts described above for the outsourcing of administration. Giving members a choice of investment manager removes this problem so long as the choice is unfettered and members are given unbiased advice.

5.3.1 Stockbroking and shelf fees

The investment industry has its own peculiar, institutionalized conflicts in the soft commissions paid by stockbrokers to investment managers and, occasionally, other investment advisers. These normally take the form of research services and terminals that are provided free if the stockbroker earns sufficient commissions from the investment manager. That the commissions paid by stockbrokers are seldom disclosed separately in the accounts of superannuation funds means that the profits are entirely secret, thus aggravating the problem. This is not to recommend a narrow focus on the direct costs of dealing, as this could be counterproductive. Every market dealer knows that a large buyer or seller can move the price of the more liquid stocks, and competent dealers can easily justify their charges. These charges should, however, be disclosed.

Another conflict arises when investment managers are required to pay shelf fees for the right to be included on the investment menu of a master trust. If not disclosed, these would constitute a secret profit. But they represent a conflict even if disclosed, as fund administrators may be biased towards those managers that offer to make higher payments.

5.3.2 Business opportunities

A more difficult element of the conflict-of-interest prohibition is that which arises from business opportunities. Investment managers who find unusual investments that they regard as good value may face a conflict between investing in them on their own behalf and including them in their clients' portfolios. Buying on one's own account before one's clients is clearly prohibited, but failing to buy the asset for clients at all is not as clear cut

as it may be justified on the grounds that it is inappropriate (for example, too speculative a use of trust money).

This argument has dubious provenance, as modern investment theory would suggest that risk is related more to general exposure to the market than to exposure to individual assets. It is perhaps even more dubious when the assets are held in a superannuation fund for internal staff—as might happen in a financial institution—which then appears to outperform the funds for clients. The best way for investment managers to manage this particular agency risk is to hold all their personal assets, especially superannuation, in pooled funds that are available to their clients. This is required by some institutions, although the idea has not found its way into the ethical codes.

On another front, the right to participate in potentially lucrative initial public offerings involves a conflict of interest if the investment manager is allocated shares or underwriting commissions that are passed on selectively to clients.

5.3.3 Self-investment and adviser-linked investments

The exemptions in the SIS Act for in-house assets include significant accommodation for corporate superannuation funds to make investments in employers—if these investments were made before the relevant part of the legislation was passed. Public sector funds also face pressure to invest in the debt of the sponsor. In both cases, investment returns may suffer.

Also worth mentioning, because it has led to a number of actions by APRA, is the practice of financial advisers recommending investment choices in which they have a personal interest. This is more prevalent among fringe operators, but major groups with private equity or other more exotic investment options are also compromised if the profits they make are not entirely disclosed.

5.4 Trustee discretion over fees

Trustees are entitled to remuneration for their services. However, the level of this remuneration should be disclosed and it should be predetermined. A contract that permits trustees to increase the fees paid to themselves places them in an impossible position of conflict. Any increase clearly benefits them and harms the members; thus, this cannot be permitted.

The problem seldom arises in corporate funds where trustees normally fulfil their duties as part of their general work obligations. The same is true of industry funds, although they are more likely to pay attendance fees or, in some cases, fees determined by an external body such as the Remuneration Commission. If these allowances can be determined objectively to compensate trustees for expenses or losses incurred as a result of their duties, they do not create a conflict. If the fees are significant, then mechanisms need to be found to ensure that they are not determined by the trustees themselves.

It is common for retail funds to charge a fee that may be increased at the discretion of the trustee, normally with some period of notice. If members are entitled to move their accounts to another fund without penalty, then this may be an acceptable practice. But if their accounts are locked in—by the decision of the employer, for instance—then this discretion should not be exercised by the trustee without giving members a chance to change funds.

5.4.1 Conflicts Faced by Corporate and Life Insurance Directors

It would seem that the fiduciary duties of the directors of a corporate trustee are indirect, through their responsibility to the corporate trustee and its shareholders. In terms of the SIS Act,[76] they have a duty to exercise 'a reasonable degree of care and diligence' in ensuring that the trustee fulfils its duties.

It would not be surprising if the remuneration of directors was linked to the profitability of the trustee, or if their performance objectives included increasing revenue or reducing costs. Although the expense of their activities might have to be borne by members, as long as the activities themselves were legal, this would be unremarkable. It can be argued that a remuneration package that rewards directors for increasing business is in the interest of members. Directors do obviously face conflicts of duty when making decisions that might involve the corporate trustee in a conflict of interest. An inflexible application of the law would prevent them from making such decisions.

The same comments apply to life insurance directors, in the general law and in the Life Insurance Act 1995, because 85 per cent of their business is related to superannuation. Under the Life Insurance Act, they are required to put the interests of policy holders first

[76] Section 52(8).

where these might conflict with those of shareholders. This would seem to prohibit them from exercising discretion to increase policy charges unless perhaps policyholders had a right to surrender their contracts without loss.

6 COSTS AND RETURNS

This section looks in greater detail at the differences in costs between the different types of funds, and discusses whether they might be a consequence of some market or regulatory failure.

6.1 Costs

The typical costs of the different types of funds are reported in Clare (2001) and Rice and McEwin (2002), in their survey of charges in the superannuation industry. This section confirms most of their findings.

The costs of funds are reported to APRA, but not in a form that allows for easy comparison. The main reason is that expenses are not reported unless the assets are directly invested (see Table 2 above for the different types of indirect investment). Thus, with some minor exceptions, the charges of the investment managers of the indirect investments, and the tax on investment income earned by pooled superannuation trusts and life insurers, are not reported. Funds invested entirely or almost entirely in life insurance policies may report no expenses at all to APRA; the life insurer pays them all.

The accounts of pooled superannuation trusts and life companies, but not of wholesale trusts, are reported to APRA. Using this information, it is possible to estimate the asset allocation – as is done in table 3 below.

Table 3: Asset allocation of funds by type of fund

	Total (A$ million)	Direct & Special Mandate	Life Insurance Companies	Pooled: Whole-sale	Pooled: Retail	Proportion in Quarterly Survey
Corporate	55,104	45%	7%	44%	5%	45%
Industry	62,579	44%	3%	41%	12%	44%
Public sector	115,767	87%	0%	10%	3%	87%
Retail funds	192,316	10%	57%	26%	7%	10%
Small funds	127,504	66%	12%	22%	0%	66%
Post-retirement	12,679	0%	100%	0%	0%	
Total	565,949					

Source: APRA (2003), APRA's unpublished Quarterly Survey

It is therefore possible to estimate the costs incurred in the underlying vehicles by examining their APRA returns.[77] The costs of life insurance companies appear to average some 1.9 per cent of assets per annum; those of pooled superannuation trusts average around 0.7 per cent of directly managed assets. Using this information and the asset allocations shown in Table 3,[78] it is possible to estimate the total costs incurred by the different types of funds—before any deduction for tax, and without any allowance for the profit margins of life companies and pooled superannuation trusts (see Table 4). As funds pay tax at 15 per cent, members would probably be charged 15 per cent less than

[77] Life company expenses were estimated from APRA (2001), which remains the latest reported. The information on pooled superannuation trusts was obtained from a specially requested APRA statistical report.

[78] APRA's quarterly *Superannuation Trends* reports the allocation of assets across the various investment channels, but does not break this down by type of fund. Its quarterly survey of over 350 funds with assets of over A$60 million does not cover the small funds, but accounts for some 94 per cent of the assets of the remainder. It therefore seems likely that its figures on asset allocations can be used as a reasonable approximation for the whole. The allocation of small funds' assets to the various investment channels is then estimated as a residual.

the figures shown in the final column. However, the profit margins of retail funds may more than make up for this.[79]

These costs do not include stock brokerage, but probably include most other commissions. There may also be other costs that are netted off income. The totals are more or less consistent with the published charges of the larger industry and retail funds, although the incompleteness of the data mean that the results should be treated as mere approximations.

Table 4: Funds' costs as a share of their assets by type of fund (%)

	Direct Charges	Specialist Mandates	Indirect charges by Life Offices	Indirect charges by Wholesale Managers	Indirect charges by Retail Managers	Total Charges
Corporate funds	0.48	0.24	0.07	0.32	0.07	1.17
Industry funds	0.65	0.19	0.03	0.30	0.18	1.35
Public sector funds	0.40	0.39	0.00	0.07	0.05	0.91
Retail funds	0.84	0.03	1.09	0.19	0.11	2.26
Small funds	1.22	0.00	0.24	0.16	0.00	1.62
Post-retirement streams			1.91			1.91

Note: It is assumed that specialist mandate fees are 60 basis points per annum; that the wholesale holders of life insurance policies (corporate, industry and public sector funds) are charged only 100 basis points; and that retail unit trusts charge 150 basis points. These are no more than informed guesses based on discussions with industry participants and a sample of product disclosure statements.

Source: Author's calculations based on APRA (2001, 2003).

The total costs of the entire industry can now be estimated (Table 5 below). It is assumed here that life company profit margins are about 25 per cent of costs,[80] and that banks account for some 60 per cent of the retail funds. While these assumptions are

[79] It would not be possible to determine objectively the profit margins of the different types of funds, even if all the data were available. This is because the determination and allocation of overhead expenses is necessarily subjective.

[80] This figure can be justified by reference to the planned profit margins reported to APRA and found in APRA (2001).

approximate, it does appear that fund management charges make up perhaps as much as 50 per cent of the non-interest, fee, commission and other income reported by Australian banks and perhaps 25% of total revenue.[81] This would explain, in large measure, the importance attached to funds management by the banking industry. It would be interesting, although difficult, to investigate whether this plays a role in the high return on equity enjoyed by the big Australian banks, and whether this could be ascribed to a lack of competition.

6.2 Returns

The table in Appendix 2 calculates the relative investment performance of the different types of funds after taking hidden expense differentials into account. Once this is done, it would appear that there is very little difference between the gross investment performance of corporate, industry, public sector and retail funds. This is to be expected and so tends to confirm the overall level of charges as determined here. While these results are inconsistent with the findings of Coleman, Esho and Wong (2003), the differences can be explained by the incorporation of hidden expenses and taxes.

6.3 Discussion

Australians are paying some 3 per cent of their total personal incomes in charges to the fund management industry, broadly defined. These can be broken down into three types of charges: distribution charges, administration charges and investment management charges.

6.3.1 Distribution costs (retail funds)

The cost of distributing retail products amounts to some 10 per cent of new contributions, or as much as 1.5 per cent of the income of the 30 per cent of Australians who have superannuation in retail funds. This figure of 10 per cent is higher than the

---

[81] There is currently (March 2007) a gap between 2001 and 2005 in the statistics available on the APRA website. Table 6 of *Australian Banking Statistics* gave a breakdown of fees and net interest income as can be seen at http://www.apra.gov.au/Statistics/Australian-Banking-Statistics-June-2001.cfm. These statistics were derived from the June 2003 report.

charges actually disclosed by the funds, but includes charges on rollovers from one fund to another in the numerator while excluding the rollovers themselves from the denominator. The justification for this is that rollovers are not new money, making marketing payments difficult to justify from a national or policy perspective. Including gross rollovers would halve the distribution costs of retail funds to 5 per cent, bringing them in line with disclosed costs.

Table 5: Total wealth management and insurance industry

(A$ million)	**Direct Charges**	**Total Charges**
Superannuation funds		
Corporate funds	260	700
Industry funds	410	900
Public sector funds	460	1,200
Retail funds	1,600	5,000
Small funds	1,600	2,200
Post-retirement streams	0	300
Superannuation total	4,330	10,300
Other providers		
Life companies		2,200
Public unit trusts		2,400
Total wealth management		**14,900**
Type of charge		
Distribution (retail funds)		2,100
Administration		7,200
Investment management		5,600
Total		**14,900**
Major institutions		
Life insurers		5,700
Banks		8,700
Other		500
Total		**14,900**

Source: Author's calculations based on APRA (2001, 2003) and other tables in text

Distribution costs appear to account for over 40 per cent of the total charges of retail funds. In return, investors receive some assistance in complying with the superannuation 1guarantee legislation, advice on financial planning and budgeting, help in negotiating Australia's complex tax and social security regulations, and advice on the choice of investment manager and asset class. Some of this is necessary but much of it is not. In particular, the investment advice provided by retail funds is of doubtful value (especially where advisers receive commissions related to their advice).

Mandatory disclosure of distribution costs does not appear to have reduced expenses elsewhere in the world, and seems unlikely to do so in Australia. A more rigorous interpretation of the common law conflict-of-interest prohibitions may be necessary to reduce distribution costs significantly. Prohibiting the payment of commissions entirely would be draconian. An alternative might be to restrict the payment of commissions to circumstances in which the intermediary does not purport to give advice. This might be easier to manage if salespeople were to sell a limited range of products with specified terms and conditions. This is already the case in the United Kingdom, where 'stakeholder pensions' are 'CAT marked'.[82]

6.3.2 Administration charges

Administration costs account for about half of funds' total charges. However, the costs of administering self-managed funds—averaging A$7,300 per fund per annum, or A$3,100 per member —are disproportionately high, amounting to about one-quarter of total charges. It is difficult to explain the costs of self-managed funds as being driven by their greater flexibility, given the enormous range of investment choice available in public offer funds. Indeed, direct property is the only type of investment *not* available in public offer funds, and, as Roberts (2002) shows, self-managed funds make relatively little investment in this area.

Other possible explanations for the higher costs of self-managed funds are their outlays on expensive retail funds, or the opportunities self-management provides for reducing tax and the means test. Given that self-managed funds provide almost entirely for relatively well-informed, wealthy individuals who could use public-offer industry funds if they chose to, and who might in any case be expected to be able to look after themselves in

[82] CAT' refers to (limits on) charges, (wide) access and (standard) terms.

negotiations with administrators, the latter seems the most likely driver of the higher costs. This is wasteful; reducing tax is a negative sum game as the tax must be collected elsewhere and the costs are not recoverable.

Allocating administration charges (before tax) to the larger funds produces a cost of A$56 per member per annum for industry funds; A$136 for retail funds; A$168 for public sector funds; and A$328 for corporate funds. These figures appear to be in line with the published charges of public offer funds. The costs of public sector and corporate funds may be understated to the extent that their administration is subsidized by the employer.

The differences in cost appear to be related to economies of scale, the difficulties of administering residual defined benefit schemes and the profit margins of the commercial enterprises (guessed as 25 per cent of costs). There may also be a difference in the level of service provided. A large part of the administrative costs of retail funds may be related to their more active quest for members; the high charges of corporate and public sector funds may be related to the higher average account balances of members, who are likely to take a more active interest in their accounts.

Given that the retail funds appear to make an adequate profit at an annual average of A$136 per account, we can make the heroic assumption that this provides the benchmark for a quality service. This would mean that unnecessary administrative charges in the superannuation system amount to at over A$1.5 billion—or about half the total cost of superannuation. (This makes no judgment about the administrative cost of insurance and other non-superannuation investment contracts.) Of this, around 40 per cent can be ascribed to the small self-managed funds and appears to be related to tax planning opportunities. The balance would appear to come from residual defined benefit expenses. It may well be worth investigating ways of simplifying these, for example, by switching members to accumulation funds while providing investment guarantees.

6.3.3 Investment management charges

Investment costs seem to average 60–80 basis points annually. This is shy of the charges displayed on the websites of some of the large retail funds, and of the charges claimed for unit trusts and about equal to the investment costs claimed by industry funds. Members of superannuation funds would therefore seem to be getting a relatively good deal, even if the profit margin in these charges appears to be at least 50 per cent.

While the investment management industry is clearly a competitive one, the high profit margins for this function do merit further investigation, perhaps along the lines suggested by institutional economics (see section 4 above). Investors should be educated to understand the random nature of investment markets and the significance of costs for long-term returns. It should be emphasized that past performance is likely to predict future performance only in the case of underperforming funds.

There ought to be scope for a reduction in superannuation investment charges, but it is however difficult to see how significant savings could arise.

7 CONCLUSIONS[83]

The higher costs of retail funds are largely explained by the costs of their distribution system. While a purported benefit of the system is that it includes advice to members, much of this advice may be biased (because advisers are paid by commission) or is unnecessary. The system also contains a number of other conflicts of interest that may lead instances of excessive charging. Part of the distribution costs of retail funds, and what looks like the higher administration costs of self-managed funds, arise from the complexity of the tax and social security codes and - in the case of the self-managed funds - the many opportunities for tax planning.

The investment management charges of all funds appear to be high relative to the underlying costs. This is difficult to explain, as the costs are clearly disclosed and there is significant choice and competition in this area. However, one factor appears to be investors' lack of knowledge about investment markets, and in particular their failure to understand the impact of charges on long-term returns.

The main policy conclusions would seem to be as follows. First, where conflicts of interest exist, further actions should be taken to eliminate them or mitigate their consequences. Second, the government should simplify the country's tax and social security regulations. Third, there is a need for investment education that emphasizes the significance of costs for long-term returns, the random nature of investment markets and the meaninglessness of differences in past performance—except in the case of

---

[83] This is the original conclusion. A postscript at the end of the chapter makes some applications to South Africa.

underperforming funds. Finally, trustees should expressly be permitted to pay for financial advice for members without incurring any liability for wrongful advice.

Acknowledgements

Particular thanks are due to Merrie Hennessy and Neil Esho for their considerable contribution to the process of writing this chapter. The author was employed by APRA when the paper was first written, but the views expressed here are his alone and cannot be ascribed to his employer.

REFERENCES

Anonymous (1982) An Economic and Legal Analysis of Union Representation on Corporate Boards of Directors, *University of Pennsylvania Law Review*, 130: 919–956

APRA (Australian Prudential Regulation Authority) (2001) *Half Yearly Financial Bulletin on Life Insurance*, APRA, Sydney, December

APRA (Australian Prudential Regulation Authority) (2003) *Quarterly Superannuation Trends*, Sydney, December

ASIC (Australian Securities and Investment Commission) (2004a) *Revised Good Practice Model for Fee Disclosure* Sydney http://www.asic.gov.au/asic/pdflib.nsf/LookupByFileName/revised_fees_model.pdf/$file/revised_fees_model.pdf (accessed 23 March 2007)

ASIC (Australian Securities and Investment Commission) (2004b) Disclosure of Soft Dollar Benefits *Research report* Sydney, June

Bateman H (2001) Disclosure of Superannuation Fees and Charges *Discussion Paper 03/04* Centre for Pensions and Superannuation, University of New South Wales, Sydney

Berger A N, J D Cummins and M A Weiss (1997) The Coexistence of Multiple Distribution Systems for Financial Services: The Case for Property–Liability Insurance *Journal of Business* 70.4: 515–546

Blake D and A Timmermann (2004) *Performance Persistence in Mutual Funds: An Independent Assessment of the Studies* Prepared by Charles River Associated for the Investment Management Association, Financial Services Authority, London

Carhart M M (1997) On Persistence in Mutual Fund Performance *Journal of Finance* 52.1: 57–82

Clare R (2001) *Are Administration and Investment Costs in the Australian Superannuation Industry Too High?* Association of Superannuation Funds of Australia (ASFA) Research Centre, Sydney

Coleman A D F, N Esho and M Wong (2003) The Investment Performance of Australian Superannuation Funds APRA *Working Paper* Australian Prudential Regulation Authority, Sydney

Ferson W E, S Saskissian and T T Simin (2002) Spurious Regressions in Financial Economics NBER *Working Paper No. 9143* National Bureau of Economic Research, Washington DC

Glover J (1995) Banks and Fiduciary Relationships *Bond Law Review* 7.1: 50–66

Glover J (2002) Conflicts of Interest in a Corporate Context in *Corporate Crime Workshop* Acquaah-Gaisie G (ed) (Monash University, Dept of Business Law and Taxation, Melbourne, 2002) 39-57

Keneley M (2004) Adaptation and Change in the Australian Life Insurance Industry: An historical perspective *Accounting, Business and Financial History* 14.1: 91–109

Maclean D (2000) Superannuation Trustees and Conflicts of Interest *Australian Law Journal* 74.1: 47–57

HM Treasury (2001) *Institutional Investment in the United Kingdom: A Review* Her Majesty's Treasury, London, March, <http://www.hm-treasury.gov.uk/media/2F9/02/31.pdf>

HM Treasury (2004) *Myner's Review of the Governance of Life Mutuals* Her Majesty's Treasury, London, December, available at <http://www.hm-treasury.gov.uk/independent_reviews/myners_review/review_myners_index.cfm>

O'Barr W M and J M Conley (1992) *Fortune and Folly: The wealth and power of Institutional investing* Business One Irwin, Homewood IL

O'Brien L (2004) How to Restore the Fiduciary Relationship: A conversation with Eliot Spitzer *Harvard Business Review* 82.5: 70–77

Productivity Commission (2001) *Review of the Superannuation Industry (Supervision) Act 1993 and Certain Other Superannuation Legislation Report No 18* Canberra, 10 December, available at **<**http://www.pc.gov.au/inquiry/super/finalreport**/>**

Rice M and I McEwin (2002) *Superannuation Fees and Competition* Phillips Fox Actuaries and Consultants report to the Investment and Financial Services Association, Sydney

Roberts M (2002) *Self-managed Superannuation Funds: Overview*. Presented to the 10th Annual Colloquium of Superannuation Research, University of New South Wales, Sydney

Smith C W (1991) Book Review: The Soul of the Salesman: The moral ethos of personal sales *Social Forces* 70.1: 287–289

Taleb N N (2001) *Fooled by Randomness: The hidden role of chance in the markets and in life* Texere, New York NY

Teele R (1992) The Necessary Reformulation of the Classic Fiduciary Duty to Avoid Conflicts of Interest or Duties *Australian Business Law Review* 22.2: 99–113

Williamson O E (1984) Corporate Governance *Yale Law Journal* 93.7: 1197–1230

Zultowski W H (1979) The Extent of Buyer-initiated Life Insurance Sales *Journal of Risk and Insurance* 46.4: 707–714

APPENDIX 1: COMMISSIONS: HOW SHOULD THEY BE FUNDED?

There is clearly a need for a distribution network for life insurance and superannuation products, and for advice on income tax and means test management. If advice is desirable, how should it be funded?

The Previous System

When most advisers lived off front-end-loaded commissions on insurance policies, the business model would have been structured along the following lines.

- The agent would be paid, on average, 50 per cent of the first year's premium on each new policy in commission and very little if any thereafter.

- If the average policy premium was equal to 2.5 per cent of the client's salary, then an agent's income per client would average: Client's salary x 2.5% x 50% = 1.25% of the client's salary.

- The agent would need to sell about 80 policies a year to earn as much as an average client. Usually it did not take long for agents to attain this level of income. In the meantime, the life insurance company employing the agent would provide some bridging finance to cover the first few months of training and establishment.

- If the agent was able to persuade each client to take out a new policy every four years, the agent would need to build up a portfolio of about 300 clients to maintain the same level of income.

In practice, very few agents proved able to find this number of clients, resulting in a high turnover of agents, and many lapsed policies. While policy holders lost money on lapsed policies because they received little in the way of refunds, and life companies bore heavy costs in building up agency forces, the agents themselves took significant financial risks—up to 80 per cent dropped out in the first two years, with many earning only minimal income during this period.

The Present System

The new rules of the game make front-end-loaded contracts unacceptable. The new commission structure for superannuation is as follows.

- Agents are paid up to 4.5 per cent of the client's contributions (including rollovers) and up to 0.4 per cent per annum of the client's assets.

The consequences of this system for agents and clients are as follows:

- If clients pay, on average, 12 per cent of their salaries into superannuation, on which the agent collects 3 per cent, and if they have accumulated one year's income in assets, on which the agent collects 0.3 per cent per annum, then the agent will take: Client's salary x 12% x 3% + client's salary x 0.3% = 0.66% of the client's salary.

- This means that the agent needs only (100/0.66) or about 150 clients to earn as much as a client does on average, but that the agent is not viable until she has accumulated all 150. In the interim, the agent will need to make up the shortfall in income elsewhere. This is very difficult for most people, suggesting that there will be a decline in the number of new agents (as in fact appears to have been the case over the last few years). It should be noted, however, that one way for new agents to address the problem is to persuade clients to transfer their benefits from one fund to another, and earn a commission on the transfer.

- At the same time, the costs to clients have risen. Total charges under the new system are more than twice those of the old front-end-loaded system, reducing the benefit of a lifetime's investment by 20 per cent or more.

This system is obviously very much in the interests of existing financial advisers. As they reach retirement, they are apparently able to sell their renewal commissions to the life companies and banks at multiples of 2.5–4 times their annual value.

Towards a New, Fee-based System?

The most desirable system would be one in which payment for advice was separated from the decision on how much to invest, and where. At the moment, product providers do not want to offend financial advisers—their major source of business—by recommending fee-for-service brokers (although, interestingly, the Corporations

Regulations[84] exempt fee-for-service advisers who give tax advice from registration under the Corporations Act).

It is also difficult even for trustees with the best intentions to change the system. Members would rather pay for advice indirectly through their superannuation fund than pay for it directly, because of the tax advantages of doing so and because they do not have to take the cash out of their current household budget. On the other hand, it would be a brave trustee who paid for advice for members out of the fund. Such a trustee would rightly be concerned about falling foul of the sole purpose test contained in section 62 of the SIS Act, which does not permit superannuation moneys to be used for any purpose other than those listed in the act: being retirement savings, life insurance and the payment of commissions. Also, members might sue the trustee if the advice provided should prove to be inadequate. This is an area in which some clarification from the regulator would be helpful.

[84] Corporations Regulations 2001, 7.1.29(4).

APPENDIX 2: INVESTMENT RETURNS OF FUNDS (1997–2003)

The investment returns of funds for 1997–2003 are shown in Table 6. They are included here to show that the apparent difference in investment performance between retail and industry funds is largely explained by the difference in hidden expenses. The returns provided to APRA are not intended to measure investment performance, and are not sufficiently detailed or accurate to be reliable for this purpose.

Table 6: Aggregate gross returns of funds by type of fund (%)

	Corporate Funds	Industry Funds	Public Sector Funds	Retail Funds
Returns to end June[a]				
1997	20.07	15.85	17.93	12.02
1998	10.13	9.70	9.72	7.47
1999	8.45	9.35	9.17	8.83
2000	13.21	12.35	14.63	9.76
2001	5.87	5.43	4.57	4.94
2002	-3.35	-0.85	-4.86	-3.48
2003	-0.53	0.90	-0.89	-0.53
Adjustments for indirect charges				
Expenses[b]	0.48	0.53	0.13	1.48
Income tax[c]	0.12	0.04	0.00	0.41
Capital gains tax[d]				
1997	0.32	0.08	0.01	0.60
1998	0.09	0.04	0.00	0.21
1999	0.05	0.03	0.00	0.33
2000	0.16	0.06	0.01	0.40
2001	-0.01	0.00	0.00	-0.01
2002	-0.10	-0.02	0.00	-0.36
2003	-0.09	-0.02	0.00	-0.23
Foundation units				
1997				0.70
1998				0.53
1999				0.39
2000				0.30
2001				0.22
2002				0.17
2003				0.12

Estimated gross performance				
1997	20.99	16.50	18.07	15.20
1998	10.79	10.31	9.86	10.09
1999	9.07	9.96	9.31	11.44
2000	13.96	12.97	14.77	12.34
2001	6.43	6.00	4.70	7.04
2002	-2.92	-0.30	-4.74	-1.79
2003	-0.07	1.46	-0.76	1.24
Arithmetic average	8.32	8.13	7.32	7.94

Notes:

a The investment performance of all except corporate funds is based on quarterly averages as reported in APRA's *Superannuation Trends*; estimates of the performance of corporate funds are extrapolated from the performance of the larger funds participating in APRA's unpublished quarterly surveys.

b The indirect charges of life companies and pooled superannuation trusts are estimated from the expenses reported for their underlying investments; the indirect charges of unit trusts are rough estimates based on a sample of Product Disclosure Documents.

c Assumes an income tax rate of 15 per cent on 4 per cent of the underlying assets.

d Assumes a capital gains tax rate of 12.5 per cent on gross returns less 500 basis points, and that 50 per cent of future benefits are counted.

e Foundation units arise from front-end-loaded contracts sold in the past, the costs of which are still being deducted from members' returns. The amounts shown here are informal estimates after discussion with industry participants.

Source: Author's calculations

POSTSCRIPT: COMPARISON WITH SOUTH AFRICA

While the Australian and South African retirement industries share a number of characteristics, the contrasts are interesting.

The major legal difference is that retirement funds in South Africa are not assets held in trust by a trustee but have been given their own legal persona by the legislation. The Pension Funds Adjudicator has expressed "the view that a pension fund at common law is a universitas personarum - a voluntary association with corporate personality"[85]. This does not appear to make much of a difference to the fiduciary duties of the managers of the funds or to the structure of the industry.

It may however partly explain the greater involvement of South African members in fund governance and a lower level of choice of fund. Few South Africans can choose to move the fund balances without changing employers, and there is no equivalent to the self-managed superannuation schemes.

The next chapter mentions Rusconi's (2004) analysis of South African charges, which is similar to that given in section 6 of this chapter. Investment and administration charges appear to be of the same order. Distribution costs are also higher for individual rather than group contracts. The major difference is that most of the commission is payable in advance in South Africa rather than as a trail. The discussion in appendix 2 above is relevant. It should however be noted that South Africa's higher inflation increases the costs of front end commissions. This is because higher rates of inflation and nominal investment returns reduce the value of the nominally level premiums paid in later years.

Given the absence of compulsion in South Africa, one would expect the costs of distribution to be higher. Rusconi's analysis shows that they are very much higher for shorter term policies. Providing group arrangements for people would appear to provide one way of avoiding such high charges.

---

[85] "A universitas is distinguished from a mere association of individuals by the fact that it is an entity distinct from the individuals forming it, that its capacity to acquire rights or incur obligations is distinct from that of its members, which are required or incurred for the body as a whole, and not for the individual members." Quoted in *Venter v Protektor Pension Fund* (PFA/KZN/20/98) from *Webb v Northern Rifles 1908 TS* 462

CHAPTER 5 - PENSION BENEFIT DESIGN AND THE INTEGRATION OF INSURANCE BENEFITS OVER THE LIFE CYCLE

Forthcoming in the South African Actuarial Journal

1 INTRODUCTION

Investment risks aside, retirement schemes can significantly aid their members navigate pre-retirement risks: death, disability, dismissal and divorce, as well as post-retirement longevity risks. This paper discusses the possibilities - and the imperatives - in the context of the South African retirement fund industry. It is written particularly to evaluate the relevant recommendations of the Taylor Committee (Taylor *et al,* 2002), on which the author served.

Providing retirement, death and disability benefits within a single scheme allows the charges for each to be offset against each other, creating direct cost savings and allowing for less underwriting, and therefore wider coverage. The reason for the offset is that members of groups subject to higher mortality and disability risks will need less retirement income. Integrating payments made from retirement funds during periods of voluntary and involuntary unemployment can reduce the moral hazards associated with unemployment insurance. The value of savings for retirement can also provide alternative funding for periods of involuntary unemployment, which can later be made up by delaying retirement. Divorce is probably not insurable, but retirement schemes should not provide perverse incentives and could be made more flexible to deal with a greater variety of life histories.

This paper places these risks in the context of what we know about the financial lifecycle, identifying strands of research that provide insight into the stages and the risks. Section two of the paper describes the lifecycle; section three the associated risks. These sections expand material from Asher (2003). Some effort is also expended in attempting to understand the underlying reasons for the reduction in retirement age experienced over the past century.

Section four then discusses how retirement funds can provide for non-investment risks before retirement. Defined benefit (DB) funds do offer some of the integration required, but are now recognised as too opaque, unfair and often excessively risky for employers. The challenge is therefore to adapt the design of defined contribution (DC) retirement funds to provide the necessary integration. Some of the suggested changes would be easier after tax and other regulatory reform.

Section five then proposes a regulatory structure for private pension funds - in particular the questions of compulsory membership and minimum requisite benefits. It is concluded that all benefits should be paid as pensions. South Africa law does not currently require compulsory contributions to retirement funds, but the idea is frequently discussed and was recommended by the Taylor Committee. This section evaluates the justice of regulations requiring compulsion and minimum benefits using also the theory of collective action.

Section six concludes that there is much work ahead in developing pension funds that efficiently meet the needs of all their members.

2 THE LIFECYCLE

The financial life cycle provides a framework for considering the financial security needs of the middle class members of retirement funds and their immediate families. In outline, it divides a person's adult life into two main stages: of accumulation while in employment, and de-cumulation in retirement. Within the first stage, there is normally a period before children, a period with children in the household and a final period when they have left.

Financial security is defined here as a level of certainty that the members of the middle class family can maintain their accustomed standard of living in the face of idiosyncratic risks. (These are risks that affect only some households, not those that would lead to a general decline in standards of living.) The effects go beyond the financial: continuity in housing and leisure consumption are required for people to continue to interact easily with their family and friends.

The financial lifecycle does not really apply to casual workers and subsistence farmers who cannot save and so are obliged to work while they are able, and rely on *ex gratia* state or family support when they are not. Those who inherit wealth that they pass on to

their children may also display less of a change of financial circumstance over their life course, and have limited needs for insurance or retirement funds.

2.1 Economic research

The interest of economic research in the financial life cycle is spelt out in Modigliani's (1986) Nobel lecture. Deaton (2005) places it in wider perspective, which includes principally the determinants of savings. The idea is hardly novel: the need for the young to accumulate wealth, and of the old to be supported, are present in the three thousand year old account of Jacob's youthful accumulation of wives and herds under Laban, and reliance on his sons in his later years[86]. Urbanization, and its attended mobility, has however made it necessary to develop new institutions to replace the extended family and rural community as a source of financial support[87]. Beginning with Germany, a number of countries introduced national pension arrangements that clearly recognised the lifecycle. Rowntree (1901) provided an early systematic evaluation of the variations of poverty over the life cycle: affecting families with young children and the elderly particularly.

In spite of its macro economic origins, and appropriately because the etymology of economics derives from the management of a household, the elements of the cycle as they affect individuals have also been intensively investigated in the economics literature. Polachek and Siebert (1993) provide an excellent reference for the factors affecting income, which is used in the following sections, as is Hadjimatheou's (1987) useful treatment of consumption and savings questions. In the actuarial literature, Cooper (2002) also describes the results of a model of savings and consumption over the working life.

More recent research is particularly sophisticated. Gourinchas and Parker (2002) calibrate a model that incorporates utility of consumption over the lifetime, and stochastic investment returns, with US household savings data. They show the

---

[86] Jacob's adult life is covered in Genesis 29 to 48.

[87] The Canadian Museum of Civilization website has a useful summary of an extensive literature review of the economic history of the late nineteenth and earl twentieth centuries. (civilization.ca/hist/pensions/1867-1914_e.pdf) for the summary and (civilization.ca/hist/pensions/cpp-biblio_e.html) for the bibliography.

precautionary motive supplements lifecycle savings. Gomes and Michaelides (2005a,b) return to Modigliani's original problem by showing how the savings pattern generated by the life-cycle can be calibrated to the equity premium and the level of interest rates for countries with different demographic structures.

2.2 Income

We know that average full time incomes generally rise initially - until at least the early thirties - with people's increasing skills and commitment to the workforce. The process has been explained by "human capital theory", of which Becker (1983) is the originator. Human capital incorporates knowledge and skills that arise from education, training and practice. Murphy and Welch (2002) model the relationship with US wages over a 30 year period.

Human capital depreciates over time, more rapidly in an environment of rapid technical and institutional change. The rate of depreciation is however difficult to determine, with Groot (1998) finding rates of 15% per annum, and Arrazola and De Hevia (2004) finding them much lower. The capacity to work also eventually declines with age – beginning at an earlier age for manual workers. Once people begin to think of retirement, they are also likely to reduce their investment in their human capital because it will be of use for a shorter period. Incomes tend to decline from this point on.

Noonan (2005) reports on difficulties experienced by some older workers in keeping their jobs. Stock and Beegle (2004) find that this can arise from discrimination against older workers and retirement rules and incentives that encourage earlier retirement. These both appear to contribute to the observed fall in average incomes as retirement approaches.

2.2.1 Education

The rise in incomes is quicker and less protracted for those with less education, who are more likely to rely on physical strength. Increasing education delays entry into the workforce, but income growth often persists into the fifties. Income growth perseveres for longer in larger organizations as a result of promotions to senior positions. This might be interpreted as returns to firm-specific capital, but Neal (1995) however finds evidence only of industry-specific returns to experience. The high income of senior people in large organizations therefore requires other explanations.

There is considerable data on the investment returns to education. Psacharopoulos (1994) summarises his review of the literature:

> "...primary education continues to be the number one investment priority in developing countries; the returns decline by the level of schooling and the country's per capita income; investment in women's education is in general more profitable than that for men; returns in the private competitive sector of the economy are higher than among those working in the public sector; and that the public financing of higher education is regressive."

The advent of computerization appeared to have had the effect of increasing the wages of those able to use them, but Allen (2001) finds that it has not reduced the wages of older workers relative to younger ones.

2.2.2 Sex

The pattern of income by age and sex is likely to differ in different societies, but the research is largely limited to developed countries, and heavily weighted to the USA. The observed patterns do however have intuitive appeal:

- Single women's salaries are similar to those single men's salaries by age.
- Married women's income drops at the birth of the first child, and they seldom catch up. Kristen and Malone (2005) explain a large portion of the loss by the additional housework associated with children, which reduces the opportunity to earn.
- Married men have the highest average income. On the one hand, they have partners who take a disproportionate share of domestic chores so enabling them to invest more in their earning power, and on the other they are expected to meet more of the family's financial demands.

Discrimination, most often on grounds of race, sex and age can create deviations from these generalizations. Hellerstein *et al* (1996) in a study of 60,000 US plants finds evidence that older males are relatively overpaid, while less qualified women are underpaid relative to productivity. Such statistics are obviously situation specific, but are consistent with general impression of relative social power. The discrimination against older workers reported in above relates to the ability to obtain a job, so is consistent with older males earning higher wages on average.

Waldfogel (1998) finds evidence that more generous maternity leave reduces the differences between women with and without children, suggesting that the cost of re-establishing oneself in the workforce (rather than the efforts spent on housework or discrimination) explains some of the lower wages earned by married women.

The previous two paragraphs are significant for people deciding to retire: a return to work at a later age may well be at a lower income level, which means that the decision cannot be taken lightly, and that it represents a loss of future wage earning power.

2.3 Expenses

It is particularly difficult to measure the costs of living precisely, given the vast number of potential purchases and the ability to substitute when prices change. Nelson (1988) discusses the difficulties in measuring economies of scale, and Deaton (1998) provides an overview of some of the difficulties in measuring prices over time and quality changes.

Expenses do however have a different trajectory to income over the lifecycle:

- Setting up house - and buying cars - is expensive. To obtain an idea of the costs, one can use the websites of general insurance companies to calculate the value of household contents, and it can amount to years of income. This obviously has significance for those who need to set up another household after a divorce.
- The cost of children generally rises with their age and drops when they leave home. A range of methods have been used to adjust household income for the number of adults and children. There must be some overhead costs, and marginal costs are likely to reduce as the number of people increase: more expenses can be shared, bulk purchases can reduce unit costs, and there will be more opportunities to provide goods and services for other members of the household. It is doubtful whether any households reach the size where diseconomies of scale arise from the need to achieve co-ordination.

Family living allows for economies in expenses compared with living singly. The benefits of living together are not however only financial and continue into retirement. Brantervik *et al* (2005) report that Swedish aged living alone or with spouses in poor health are more likely to be malnourished. Ogg (2005) looks at European data and describes the particular problems faced by the separated and divorced.

- Medical costs appear more or less proportional to the number of people in the household until retirement, after which they rise rapidly. Evidence however is that they (especially if they include the cost of care) are much heavier in the years preceding death rather than related to age *per se,* especially if one includes the cost of care. Van Weel and Michels (1997) suggest that medical costs in the last year of life account for 40% of lifetime costs.
- Other expenses reduce after retirement, more so as health deteriorates as found, for instance, by Banks *et al* (1998).

2.4 Savings

Savings depend on the interaction of income and expenses. The following crude stages can be outlined:

- Twenties and increasingly early thirties: people start work, establish a separate household, marry and have children. They save for deposit on buying a home. Young people are likely to be "liquidity constrained" when they start work: consumption is constrained by an inability to borrow.
- Late twenties to late forties: most middle class families will borrow to buy a home. Paying off debt is prudent before saving for retirement can begin.
- Forties and fifties: at this stage, pension provision becomes important. Pension assets of at least fifteen times the required pension are probably necessary. This converts to perhaps 7 years of income or 30% of earnings over twenty years.
- Healthy over sixties: Depending on finances, inclination and opportunity, people can choose to work or enjoy the leisure of retirement. There is likely to be some withdrawal of savings for leisure activities.
- Unhealthy over sixties: People are increasingly unable to work and in need of help in their activities of daily living. Significant dissaving may occur for medical expenses and formal care.

The cycle is effectively normative: deviations can be costly. Delaying children increases the risk of congenital defects and infertility. Too little saving may lead to financial hardship in old age; excessive savings may create an unnecessary reduction in consumption and more than enough assets later.

It is clear that there are many working people who think that they should save more. Many of those already retired also think that they should have saved more while they

were working. Hurst (2003), analysing the Michigan Panel Study on Income Dynamics (PSID) data, finds that there are a significant number of families that fail to plan for retirement, and whose expenditure myopically follows changes to income. They predictably end up with less assets in retirement, although appear to have adequate pension benefits. Laibson *et al* (1998) provide a detailed analysis of motivation for saving, and document that a majority of their US survey are in favour of compulsory savings and limitations on pre-retirement withdrawal - seeing them as helpful restrictions on their own behaviour. This research helps explain the dispersion of wealth at retirement that Venti and Wise (1998) found to be less dependent on lifetime income than expected.

On the other hand, in an international survey[88], AXA finds that the majority of retirees in most countries regard their income as sufficient. As many report that their post-retirement living standards have increased as have reduced. It is not entirely clear therefore whether the view that people do not save enough arises from a clear picture of the life-cycle or from an un-informed, but established, myth that more saving for retirement is always good.

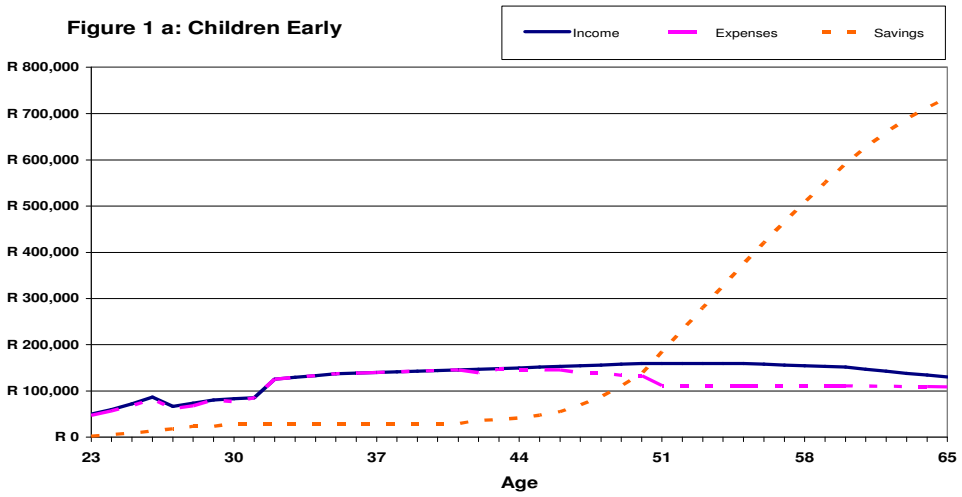
The models shown in figures 1a and 1b suggest that the popular view that one should save lots and save early is misguided. (The underlying data is given in the appendix.) The numbers are stylised and chosen to produce level consumption over the lifetime. They illustrate that couples will usually be liquidity constrained when they are attempting both to buy a house and look after children. If they want to enjoy consumption over their lifetime that varies less, they should save less at that time. Only if children are delayed should savings become significant earlier in adult life.

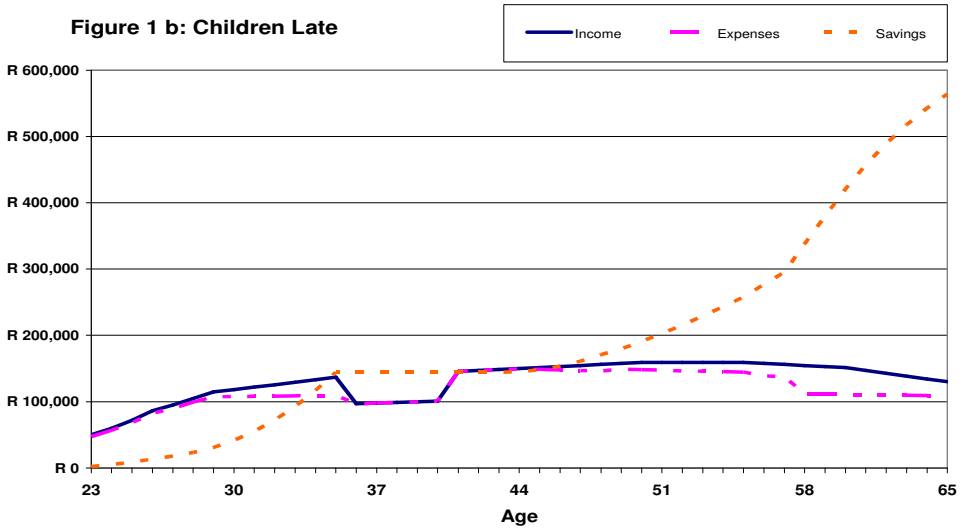
2.5 Nuclear families

The lifecycle is not only a middle class, but a nuclear family construct. It assumes that retired people are not supported by their children within an extended family, and – in the form described in figure 1 – assumes that parents support their young children. The

[88] AXA (2006) AXA Equitable Retirement Scope: Retirement — a new life after work? Report axaonline.com/axa/pressroom/2006/2006_01_24_Retirement_Survey_International.pdf

lifecycle would be less pronounced and retirement funds less necessary in societies characterised by extended families living together.





2.6 Retirement as a social phenomenon

As a middle class, nuclear family construct, the life-cycle as outlined above is largely of modern origin. Costa (1998) shows how retirement ages have declined in all the countries for which statistics are available for most of the twentieth century. People can work until older ages, but institutionalised retirement ages, public pensions funded out of both earmarked and general contributions, and tax concessions, have encouraged them to retire earlier. Buetler *et al* (2004) confirm, with individual data, that compulsory savings does create an incentive for earlier retirement.

Schulz (2002) writes of the emergence of retirement as an active phase in the middle class over the past century. He refers to "irrefutable evidence" that a major reason has been the need to ease unemployment at younger ages. While unemployment at younger ages has been high at times and may have been a contributing factor to lower retirement ages, table 1 below shows that leisure is not currently distributed evenly over the life-cycle – if Australian data is representative. Earlier retirement has accompanied a significant increase in the employment of women, including the mothers of young children who work almost three times more hours than men in the last of the family phases shown. Jager *et al* (2003) confirm that this extra work creates stresses in their report on European countries. They suggest that while the Dutch and Swedes have similar views on the equality of the sexes, the Dutch approach of specialization within the household (women are more likely to work part-time) leads to lower levels of stress.

Erlich and Kim (2005) find a link between the generosity of national retirement benefits, family disruption and lower levels of fertility. They explain this by a reduced need to have children for support in old age, and the greater economic freedom enjoyed by women when they have their own pensions. Other explanations are possible. Generous pension arrangements are however characterised as much by higher contributions and earlier retirement, as higher pensions with less need for support from children. Higher contributions for younger people have themselves created a greater need for younger women to work, and earlier retirement has created a greater demand for their labour. These stresses would also cause family disruption and lower fertility.

This issue needs to be explored as it relates to the debate on compulsion that is discussed in section 5 below. The shifting of working hours from richer and older, but active, men to working mothers - as a result of institutional changes - suggests an intersection of class, patriarchy and gerontocracy power interests rather than needs. The idea of the gerontocracy is described by Mulligan and Sala-i-Martin (1999). They suggest that active pensioners make up a dominant political bloc in all democratic countries when it comes to the defence of social structures that benefit and build its constituency. Such structures would include discrimination against older workers that pressure them into an earlier retirement.

Table 1: Time allocations* by household type (Table 5 in Apps and Rees, 2003)

House-hold type	Life-cycle phase	Male hours of work			Female hours of work		
		Market	Domestic	Total	Market	Domestic	Total
I & II	1	2314	554	2972	1811	1014	2285
	2	2359	1369	3775	15	4103	4118
	3	2394	1358	3859	111	3664	3762
I	4	2415	1185	3600	355	3170	3525
	5	2362	817	3179	504	2393	2897
	6	2367	815	3182	670	1781	2451
	7	1862	1093	2955	0	2026	2026
	2	2432	1464	3943	1378	2908	4286
	3	2478	1306	3784	1615	2679	4291
II	4	2464	1095	3559	1915	2285	4200
	5	2527	856	3383	2120	1753	3873
	6	2369	958	3327	2217	1568	3785
	7	2062	1083	3145	1202	1458	2660
I & II	8	0	1458	1458	0	1750	1750

*Weighted data means, hours pa.

Type I	Traditional – female partner has minimal workforce participation
Type II	Non-traditional – female partner works long hours or full-time
Phase 1	No dependent children and a female partner under 40.
Phase 2	Children with an average age under 5.
Phase 3	At least one child aged 5 to 9 years.
Phase 4	Children are predominantly in the 12 to 14 year age group.
Phase 5	Older dependent children still living at home.
Phase 6	No children, female under 55 or male under 60; one working
Phase 7	Between phases 6 and 8
Phase 8	Both partners retired.

A Marxist analysis of this apparent distortion of the social structure would lead to a search for the "ideological superstructure" or "false consciousness" that would justify it. In South Africa and other Anglophone countries, where funded private retirement funds make up a core element of the financial system, there is a set of myths around the benefits to individuals and society of saving for retirement. For individual benefits, there is common misperception addressed in 2.4 above. On economics, there is the unsupported view that retirement saving contributes to investment and to economic growth discussed *inter alia* in section 2.7.6 of chapter 3 above.

There is also the commonly held view that a successful career involves hard work when one is young followed by an early retirement. These myths are institutionalised by compulsory pension fund membership with their consequential earlier retirement ages, subsidies for mothers to be in the workforce, and reinforced by an ideology of equality that leads younger mothers to accept fewer hours of leisure.

While issues of aging have been the subject of investigation in the sociological literature, the assumption is that discrimination based on age is due to stereotyping and prejudice that should be addressed by "information, contradiction and confrontation" (Bytheway, 2005). To the extent that social structures have advantages for particular groups and are embedded in the legal and institutional structure, however, there is the further need for political reform and institutional development.

The existence of an ideological superstructure is not intended by this writer to suggest that its proponents are hypocritical, selfish or stupid. The economic arrangements that generate it and the theoretical constructs that support it are likely to emerge slowly as by-products of other developments. Recognising the ideology, and initiating the education, reforms and developments necessary to overcome its negative impacts are also likely to take time.

3 RISKS DURING THE LIFECYCLE

Much intriguing data on the volatility of income, expenses and savings has come from the PSID, and equivalents such as the British Household Panel Survey and the European Community Household Panel. Duncan (1988) reports from the PSID that relatively few US families remain persistently in poverty, but that a third or more of the population suffer significant drops in per capita household consumption at some point in a decade.

The most important cause of the reduction is reduced wages (normally reduced overtime), or total unemployment. Changes to family composition (particularly leaving home and divorce) are also major contributors to financial instability. Also of interest is the spread and volatility of income progression in the light of changes to inflation and other economic variables.

This section considers the size and nature of these risks and the potential for retirement schemes to provide insurance cover.

3.1 Death

Death before retirement can create financial stress for remaining family members. This section describes the needs and suggests how the benefits should be determined.

3.1.1 Orphans

The most obvious need is for the support of orphans. The presence of children in a household not only increases expenses but also diverts time from parents' opportunities to earn. Most poverty, worldwide, is felt by single parent families (most often headed by mothers) with young children. The problem is greatly exacerbated if both parents are dead, which is likely when AIDS was the cause.

It would appear highly desirable for all retirement funds to offer life cover for orphans. A problem arises if the funds do not know of the existence of the children of members. This can be avoided if the premiums are charged to all members, but this may be seen to be unfair. The reintroduction of a dependent's tax rebate would provide greater incentives for parents to report the existence of dependent children to their employers, and thus to their fund.

3.1.2 Spouses

To the extent that wives have sacrificed a significant portion of their earned income in order to look after their children, they will be more dependent on the earnings of their husbands, and will need cover in the event of his death. Cover for children's expenses may not compensate for this loss. Cover on the wife's life is not normally necessary as household arrangements usually appear to enhance the husband's earnings.

In those rare circumstances where husbands reduce investment in their own human capital to take up household duties, the needs will be reversed.

Offering cover to men but not women is likely to be seen as unfair discrimination, so spouses' benefits should be made optional in retirement funds. This allows the cover to be adapted to fit all circumstances.

Requirements for cover go beyond retirement. Society's poorest members include elderly widows and divorcees - as described in Auerbach and Kotlikoff (1991). It appears that their husbands spend most of their retirement benefits on themselves.

3.1.3 Debt repayment

Banks and other moneylenders frequently require cover in order to ensure quick and easy repayment on death. The cover may well be necessary to protect the family, but there is frequently no attempt to determine the borrower's total need for the insurance. Credit life cover for smaller loans particularly appears to be sold purely to earn additional commissions for the supplier.

If retirement funds offer life insurance, then additional cover may well represent over-insurance. There would therefore be an argument for laws governing retail loans to require moneylenders to accept the cover from retirement funds as security in the event of death. This in turn would require the funds to be able to pay some of the proceeds of life cover to the lenders, and to notify them before cover ceases.

The amounts recoverable would need to be limited to a portion of the cover in order to ensure that families were not impoverished in order to pay excessive debts. If benefits were only payable as pensions, this would be easier to ensure.

3.1.4 Funeral expenses

Funeral expenses are important for people where the cost of a culturally appropriate funeral is significant relative to income. This cover is widely offered by informal burial societies and a variety of group and individual life policies that normally cover the whole family. While not permitted in a retirement fund, over half the employers responding to the Sanlam Survey (2006) offer funeral benefits to the employees; most of these cover children.

South African law[89] limits the insurance that can be taken out on the life of children. The limits may be circumvented - mainly inadvertently - if people are members of a number of different groups offering funeral cover. This presents a perverse incentive. Fisher and Young (1965) report that restrictions on the amount payable on the death of a child were introduced in 1875 in the UK for fear of encouraging infanticide. In 1948, this was taken further and life cover on children under 10 was prohibited altogether. While infanticide is seldom reported in South Africa, there are over 6 000 unnatural deaths[90] of children under 14 annually, which provide grounds for doubt. Given the ubiquity of funeral cover, it would appear undesirable to offer more.

3.1.5 Other death related expenses

Life insurance is also bought to provide a bequest to children, or if there are other assets, to pay the estate duty. This type of cover does not represent a need and should at most be optional in a retirement fund.

3.1.6 Determining the requirements

The amount of life cover required by the dependants of an individual in any future year can be derived as either:

a) net expected income in the year times the proportion not spent on oneself, or

b) one's contribution to household expenses and upkeep of others.

Lilly (2006) provides a detailed analysis as how the required cover might be determined by a financial advisor. His analysis is relatively complicated and includes relatively minor items such as specific provision for funerals. Given the difficulty of determining future income, needs and investment returns precisely, it would not seem that such detail can be justified. Rough calculations based on one of the two formulae suggested above are likely to be adequate.

---

[89] Section 55 of the Long Term Insurance Act 1998 restricts the maximum cover that can be given for children to R10 000 for those under 6 and R30 000 for those under 14.

[90] This number can be calculated from "Causes of death in South Africa 1997-2001: Advance release of recorded causes of death" (P0309.2) published by Statistics South Africa. Actual deaths are however believed to exceed reported deaths by a significant margin.

The cover required for each year is likely to be relatively level, when expressed as a percentage of income. Allen (1970) reports common US benefit formula were for spouse's pensions of between 20% and 50% of income with orphans' pensions of between 5% and 15% each. The greater participation of women in the workforce since that time would suggest that proportions should be more heavily weighted toward orphans.

A formula for orphan's cover could be 15% of income for each child until they become financially independent, with maximum cover of 40% or 50% of income. These percentages would approximate a division of the income of a household into one part for each child and one for each adult.

The need is clearly for a monthly pension, but South African income tax rules on annuities make lump sum cover more tax efficient because the entire annuity and not just the interest portion is taxable. The same applies to spouse's pensions. Antiquated tax rules of this type are designed to capture investment annuities, but have been amended to allow structured settlements in many other jurisdictions[91]. Tax regulations should also be amended in South Africa and extended to cover insurance based annuities. In the interim, even retirement funds that offer monthly cover should offer a commutation option that allows the beneficiary to purchase a tax efficient annuity with the commutation.

3.2 Disability

Where disability leads to an unpreventable loss of income, there is a need for income replacement – whether the disability is temporary or permanent, total or partial. Disability also leads to additional costs of living and often of working. Zaida and Burchardt (2005) confirm this from their survey of a sample of disabled people in the UK.

It appears that insufficient attention is given to disability insurance in South Africa, where fewer people have private disability cover than life cover, although more people would have cause to claim during their working life. The position can be quite different if governments offer cover. In developed countries particularly, disability income

---

[91] For discussion see Chapter 22 of the report of the Road Accident Commission Report 2002 (Satchwell J) http://www.parliament.gov.za/pls/portal30/docs/FOLDER/PARLIAMENTARY_INFORMATION/PUBLICATIONS/ROAD/s10-22.pdf

payments can be very significant. An OECD report[92] shows a variation from 1.3% of GDP in Canada to 5.6% in Norway. This variation suggests that it is possible to make expensive errors in the design of a system of insurance, either denying needy claimants, or overpaying those who could be working.

This section attempts to outline a sensible and compassionate approach to the issues.

3.2.1 Rehabilitation

Disabled people need rehabilitation. Tate (1992) confirms that return to the workforce and "normal" life is often a possibility, and often the best cure. Rehabilitation may require special equipment and other provision, and needs special effort from family, employer and insurance provider. It is difficult, painful and sometimes humiliating for the disabled person. It is however superior from a human point of view and appears to be less costly. These points are illustrated by De Jong *et al* (2006) who find, in their Dutch study, that the cost of disability benefits can be significantly reduced by more rigorous management of the payment of benefits – without a consequential knock-on in higher unemployment benefits.

Goodman and Stapleton (2006) describe what they call a "new paradigm" of attempting to re-integrate the disabled in society. Their point is that current US government expenditures are not appropriately geared to this objective. It is suggested here that this new paradigm can be overplayed. It is not new, having been, *inter alia,* the inspiration for the phasing out of hospitals for those with mental illnesses in many countries. This Tucker (1990) reports the consequences in the growth of the homeless: over a third of those in American cities had been patients at institutions for the mentally handicapped. Rehabilitation is not easily achieved.

3.2.2 Heterogeneity

It was clear from representations to the Taylor Committee that there is a wide range of differing needs, and that the disabled were not uniformly represented. It seems to this author that four main groups can be identified. If there is to be a new paradigm for the

---

[92] From Marin B & Prinz C (2003) *Facts and Figures on Disability Welfare: A Pictorial Portrait of an OECD Report* Vienna: European Centre for Social Welfare Policy and Research, 36 quoted in Social Security Administration (2006)

disabled, it is suggested that it should be based on recognition of the heterogeneity of disabilities, which this classification attempts to capture in some way.

3.2.3 Physical disabilities needing rehabilitation

In this category are those who are intellectually able but blind, deaf or have lost the use of their limbs, as well as many with epilepsy, multiple sclerosis and similar diseases. It is characterised largely by people who want to return to work, although they may need encouragement at times. The category has, practically, to be divided between those who would otherwise be capable only of unskilled work, and those with the ability to perform skilled work. Tate (1992) confirms the former find more obstacles in returning to work – even in his US study, where unemployment is less of a problem than South Africa.

Representations to the Committee from people in this category[93] confirmed that they want to participate as fully as possible in society. To the extent that they need help, they would like training and equipment that gives access to jobs. With help, they expect to be able to make a full economic contribution. It was argued by some that the disabled should not be paid income replacement benefits at all, as the benefits would tend to marginalize them.

For disabilities with these characteristics, it does seem appropriate that insurance pays the costs of rehabilitation rather than income replacement. Given the demography and high unemployment rates of the unskilled in South Africa, and the absence of unemployment benefits, some income replacement benefit would however appear to be necessary for those with fewer skills.

3.2.4 Backache and depression

This group includes people suffering from non-specific backache and some types of depression: painful conditions that make many aspects of life and work more difficult, but do not prevent working. Claims in this category fluctuate dramatically with economic

[93] The Committee received representation from the SA Federal Council on Disabilities, the National Council for Persons with Physical Disabilities in SA, the SA National Council for the Blind, the Deaf Federation of SA and the SA National Epilepsy League

and social change, and can represent more than half of all disability expenditures[94]. The Taylor Committee however received no representations from organizations representing people in this category.

There is evidence that insurance cover for these conditions is largely misconceived. That on backache is strongest. Nachemson (1999) records that backache alone accounts for one third of claims in North America and in 1987 accounted for claims made by 8% of the Swedish workforce. His review of the medical research shows that the diagnosis rather than the condition is the cause of most claims, and that "compensation prolongs symptoms and disability". He continues: "In our efforts to be kind to people, politicians and physicians perhaps have been doing the wrong thing. We are perhaps making people sicker by some insurance schemes." He is particularly scathing of some workers' compensation arrangements where there is a lax management of claims. He reports on widespread success in reducing claims for backache by reducing the benefit payable, and limiting claims for work related injuries to the 6 weeks it takes for physical healing to occur.

Many claims for depression may fall into the same category, but medical research does not appear to have investigated this in sufficient depth to be sure. In a recent in-depth study of 19 individuals, Millward *et al* (2005) however report on "the unwitting role that can be played by the health care system in reinforcing the 'sick role' and in so doing providing a continued justification for an 'off-work' identity."

There is evidence that modern medicine can cure many forms of depression cost effectively[95], which should suggest that the number of claims should be falling rather than increasing as is the case in many countries.

It would seem from a financial and social perspective that benefit payments for these conditions should be circumscribed by strict claims management and, where necessary,

---

[94] As reported in Berkowitz M (2002) Designing an Early Intervention Demonstration to Return Applicants for Social Security Disability Benefits to Work www.dri.uiuc.edu/research/p01-01c/final_report_p01-01c.doc and by Conti et al (2006)

[95] The US National Alliance on Mental Health claims $7 saved for every $1 spent. www.nami.org/Content/NavigationMenu/Inform_Yourself/About_Public_Policy/Policy_Research_Institute/Policymakers_Toolkit/Facts_for_Policymakers_Treatable_Causes_of_Disability_-_Major_Depression.htm

reduced benefits. As an alternative to paying a fixed lower benefit for backache and depression, the total benefits available for this category might be limited to a stated pool of money, which would be shared among those claiming at a particular time. The benefit would therefore be self-adjusting, its size declining as more people were claiming in this category. By reducing claims and making them more stable, it would significantly reduce the premiums for disability insurance so making it more attractive for those that really need it.

3.2.5 Intellectual and mental conditions

This grouping includes people of working age with intellectual shortcomings, the severely mentally-ill, and some people addicted to drugs and alcohol. The Committee received submission from organizations serving rather than representing people in this category[96]; the focus was on sheltered employment and hospital services.

Noble (1998) confirms that there are people, with these conditions, who will not be able to participate in workplace without ongoing expert, or at least informed, assistance, which sometimes should be delivered in an institutional setting. Their families are often not able to cope on their own. The idea that such people should be looked after in the community is meaningless unless there are appropriate social structures to embody the concept. If structured care is not forthcoming, people in this category can be disruptive: occupying and defacing public spaces, and possibly being a physical danger. The US National Association of Mental Illness (NAMI) report[97] that 16% of the US jail population suffers from serious mental illness, while 30% of those in public psychiatric hospitals are forensic patients.

NAMI is an active lobbyist for more appropriate spending on the mentally ill, believing that it could save money in the USA in the relatively short run. They are advocates of supportive housing (within communities), adapting jobs to be more suitable and earlier

---

[96] Cheshire Homes and the Department of Health

[97] See www.nami.org/update/unitedcriminal.html (no longer active) and http://www.nami.org/Template.cfm?Section=Press_September_2006&Template=/ContentManagement/ContentDisplay.cfm&ContentID=38175

and more specialised treatment.[98] Some of these ideas are likely to be applicable to South Africa.

If permanent rehabilitation is not possible and income replacement monies likely to be wasted, benefits for people in this category need to include structured support that perhaps be deducted from compensation for lost income. Benefits should in many of these cases be paid to third parties and subject to appropriate controls. Retirement funds and insurers should investigate collaboration with organizations supporting such people.

3.2.6 Other conditions

Those in this category require income replacement. Included are people who are bedridden or housebound. The Committee received a number of representations from organizations representing people living with HIV/AIDS, but all from organizations concerned with state support for orphans rather than for adults who had previously been members of pension funds.

If people have a condition that is temporary they need support until they are able to return to work. If the condition is not temporary, they will need income support until death. One concern is that people in this category may leave employment because they have exhausted their sick leave, and then lose their life insurance. This is a particular problem where there is no disability cover or if it ceases after a fixed period.

3.2.7 Long term care

Broe (2005), arguing against the "traditional 'age structure' that homogenises the old", uses three age related categories for the retired. He describes the middle group between approximately 75 and 85 as "generally mobile independent and cognitively together, but in 50% brain function is at risk if stressed and then they need some assistance – and 16% have a dementia." For the very old over 85, "70% have difficulties with cognition, executive tasks and/or with balance, gait, mobility and activities of daily living." Broe is concerned with the need for medical care for the older and oldest old, and also with preventing declining brain function. Tilse *et al* (2005) reviews the literature on ageing

---

[98] See the fact sheets at www.nami.org/Template.cfm?Section=policy_research_institute&Template=/ContentManagement/ContentDisplay.cfm&ContentID=14596

individuals and the help they need in managing their finances, and report on a survey that finds that almost one in three Australians has given help to an aged or disabled person in the previous year. It can be noted that most of the assistance is informal and appears to be inadequately monitored.

Institutional long term or frail care is likely to be required for those in Broe's older old categories. In an Australian context, which is similar to that of middle class South Africans, Madge (2000) considers the costs to the Australian state, which amount to some 1% of GDP for care and 2% for medical costs. The costs borne by government are dependent on family and social conditions; women in some age categories are three times more likely to need institutional care than men, not least because couples are likely to care for each other and men are more likely to pre-decease their wives.

Provision for frail care might be incorporated into retirement fund benefits. As an insurance benefit, however, it would be subject to moral hazard: family members will be given less incentive to provide the necessary care. The incentives would be better aligned if there were a savings component that allowed for some money to be left over for bequests. The amount required would however be large, particularly when compared to smaller pensions. It is difficult to envisage retirement funds withholding such amounts from relatively poor pensioners for what may be decades between retirement and entry into frail care. The current method of state-provided means-tested care may well therefore be the most appropriate funding method. Means tests in this context represent relatively little additional intrusion into the lives of the people concerned, and are efficient in providing incentive to family members to limit state-subsidised care. This is in contrast to the means tests for the age pension which are intrusive and inefficient - as described in Chapter 3 of this thesis. Both types of means tests do fall more heavily on those with relatively modest incomes, but the former have at least some merits.

3.3 Unemployment

It is often difficult to disentangle periods of unemployment caused by a disability from those caused by a lack of jobs for which an individual is qualified. Those rendered unemployed by economic and technological change can in some ways be described as disabled. Such unemployment may be insurable to the extent that it arises as a result of retrenchment. Benefits could be paid for the period required to retrain but would be

subject to increasing moral hazards if claimants were not sufficiently motivated to return to work.

As with disability, the payment of unemployment benefits should also be accompanied with active retraining and pressure to return to work. Recent OECD[99] summaries of the international research confirm this. They describe the policies needed to successfully combat unemployment. Three are relevant to this section:

- "make it more attractive for people to work than to stay on welfare benefits;
- "make sure that benefit recipients receive high-quality services, particularly in relation to their quest for jobs, and that this is monitored closely, backed up by the threat of benefit sanctions;
- "provide people of working age with the education and training opportunities that they need to get jobs and raise their incomes."

The South African position before recent changes to the South African Unemployment Insurance Fund (UIF) rules was that individuals were allowed to resign, claim on the UIF and simultaneously spend their retirement fund withdrawal benefits – so making significant additional money available to them if they were previously liquidity constrained, and providing a perverse incentive to claim. UIF benefits are no longer available to those that resign, but claimants can still access their retirement fund withdrawal benefit while still receiving UIF benefits. This is a problem that ought to be addressed.

Retirement schemes cannot help cover youth unemployment that occurs before people have had the opportunity to become members. Stiglitz and Yun (2002) suggest that national schemes could be designed to allow people with little or no retirement savings to borrow against future contributions. The suggestion should perhaps be explored in the context of the funding of tertiary education and training, but is not taken further in this paper because of the moral hazards.

---

[99] From the summary of the "Employment Outlook and OECD Jobs Strategy" found at: http://www.oecd.org/document/31/0,2340,en_2649_34731_36899679_1_1_1_1,00.html

3.4 Divorce

Divorce frequently involves financial distress as a separate household has to be formed and run. A drop in consumption is almost inevitable. The causes of the other stresses that arise from divorce are controversial. Amato (2000) surveys the literature and characterises the debate as polemical:

> "Some scholars see the two-parent family as the fundamental institution of society—the setting in which adults achieve a sense of meaning, stability, and security and the setting in which children develop into healthy, competent, and productive citizens. According to this view, the spread of single-parent families contributes to many social problems, including poverty, crime, substance abuse, declining academic standards, and the erosion of neighbourhoods and communities.... In contrast, other scholars argue that adults find fulfilment, and children develop successfully, in a variety of family structures. According to this view, divorce, although temporarily stressful, represents a second chance for happiness for adults and an escape from a dysfunctional home environment for children."

This rather balanced conclusion seems to stand in contrast to the findings of his survey, which overwhelmingly confirms the negative impact of divorce on living standards and all aspects of well-being of both parents and children, with women suffering the economic effects more. While the studies do not find that divorce is always a destructive experience, they do find that divorce, apart from any pre-existing characteristics, is itself a cause of misery.

More recent research confirms all of these findings. Sigle-Rushton (2005) examines the hypothesis that early fatherhood, like early motherhood, disrupts the lifecycle accumulation of human capital, and thus leads to ongoing financial disadvantage. Avellar and Smock (2005) look at co-habiting couples who break up their relationship and find, as with married couples, that the brunt of the loss of household economies of scale are born by the women, with the presence of children leading to a higher probability of descent into poverty.

Particularly relevant to South Africa are the clear results of Sampson and Groves (1989) that family breakdown is associated with crime. More specifically, but not as widely investigated, Comanor and Phillips (2002) show that the absence of the father is itself a

major contributor to crime. Available data[100] show that 20% of South African children do not live with their natural mothers, while fewer than 50% live with their natural fathers.

It is thus important that retirement funds do not create perverse incentives and aggravate destructive behaviour, while on the other hand attempt to provide flexibility to ease the financial distresses of divorce.

3.5 Retirement

Retirement itself can be stressful.

3.5.1 Reduced expenditure for some

Banks *et al* (1998) look at the consumption of UK households at retirement and try to explain a small decline. They conclude that at least some arises from negative information that becomes available at retirement, which they interpret to mean that many pensioners have saved insufficiently. This would tend to be confirmed by Hurst's (2003) results. Other possible reasons would be the loss of human capital arising from the act of retirement itself, or the reduction in prices arising from pensioner discounts. Hurd and Rohwedder (2006) in perhaps a more careful consideration, find that the decline in spending is volitional: more time is spent on inexpensive pleasures.

3.5.2 Difficult transitions

Larson and Pedersen (2005) look at a Danish sample of pathways to early retirement. They identify 3 major paths: direct; after a period of drawing unemployment or disability benefits; and other paths that incorporate a variety of transitions including a long period of unemployment without a benefit. In their sample, more than one worker in five appears to have been driven to retirement by unemployment; even more appear to retire for health reasons. These difficult transitions are more likely to affect lower income people. Hayward *et al* (1989) find that better educated and paid people have lower rates of ill-health retirement and disability.

[100] 1995 October Household Survey, reported in Case et al. p14

3.5.3 Other disadvantages

Even those who retire voluntarily appear to decline in health. Dave *et al* (2006) carefully analyse a rich US panel data set, and find evidence that reduced opportunities for socialising, and a lower requirement to exercise sufficiently, appear to cause lower levels of health. This appears to be true even for those who found retirement removed them from a stressful job.

4 BENEFIT DESIGN

This section discusses how retirement fund benefits can be adapted to meet the contingencies discussed in section 2. Stiglitz and Yun (2002) make the formal arguments for such integration and mention the Provident Funds of Singapore and Malaysia as exemplars. While these national funds[101] do provide a degree of integration, it is mainly administrative. They provide what large South African companies provide as employee benefits to Provident Fund members: a choice of partial insurance schemes and mainly lump sum benefits[102]. The more complete integration suggested in this section would offer greater benefits.

4.1 Cross-subsidies in a typical DB scheme

DB schemes typically included both risk and retirement benefits: widows' and orphan's benefits and early retirement on grounds of disability. The payment of a joint pension after retirement also offered more protection to divorced spouses and widows.

The reasons for the switch from DB to DC benefits were normally not related to risk benefits. Fore and Hammond (2005) list some of the reasons in the USA as: “higher worker mobility; the maturing of DB plans in older and highly competitive manufacturing sectors with few new entrants, workers’ desire to control their own saving and the high costs of DB plan management, along (in more recent times) with unexpectedly poor stock market performance and low interest rates.” They suggest that

---

[101] See Asher M G (2000) *Social security reform imperatives: the southeast Asian case* http://www.outlookmoney.com/scripts/IIH021C1.asp?sectionid=6&categoryid=33&articleid=1610&NoCache=5%2F1%2F2006+2%3A05%3A15+PM

[102] See the Sanlam Survey (2006) for more detail.

the main reason has been the "unintended consequence of the risk-management mechanisms adopted to reduce risk, correct agency problems, and address pension administrative concerns."

Kerrigan (1991) describes the South African reasons. The trade unions were particularly supportive of the need for change because they obtained fairer withdrawal benefits, lump-sums on retirement, greater influence on investment policy and the power to elect trustees. The DC funds did away with the complicated cross-subsidies of DB, and the lump sums offered greater ease in avoiding the means test that, at that time, reduced the state pension by 100 percent of other income. Employers encouraged the shift to DC benefits. First, they were reluctant to permit newly elected trustees to make decisions that could lead to investment losses. Second, they were not averse to reducing the investment risks inherent in DB design, and toward the end of the 1980s, they saw that AIDS threatened a significant increase in the cost of risk benefits.

Many of these disadvantages can be addressed, but it is suggested that DB schemes will not again find favour because of the unfair cross-subsidies intrinsic to their design. Actuarially fair pension benefits are such that the present value of the contributions is equal to the present value of the benefits: every member gets what they deserve. Actuarial fairness can be determined *ex ante*: making reasonable assumptions about the future and locking people into predetermined benefit formulae, or *ex post*, using actual experience to determine benefits.

Pure DC funds are actuarially fair *ex post*. This is one of their main attractions. DB funds could be made fair *ex ante,* but they normally pay higher benefits to older entrants and the married, and often to those with children. Those where the benefit is based on final average salary give better benefits to those whose salaries increase the most in their last years of work. Skilled people are paid faster growing incomes than manual workers, and so the poor and unlucky subsidize the rich and lucky.

The cross-subsidies can be significant. Consider the effect of an increase in salary on the value of the expected pension. The effect grows larger as retirement approaches and is proportional to years of service. Ignoring survival factors, and assuming a real rate of interest of 2%, the effect will be more than twice as much for a 64 year old as a 24 year old. If the older person has 40 years of service and the other just one, the effect on the older person's pension will be almost 100 times greater.

The problem arises because a lower than expected salary increase gives rise to a reduction in the value of the pension. The 64 year old with 40 years service may well have a pension worth ten times annual income. A salary increase 5% less than expected (not impossible in times of high inflation when the person is only a year from retirement) will mean that the pension declines in value by an amount equal to 50% of annual income. The reverse applies to those with increases higher than expected. Such dramatic side effects of small salary increase differentials are impossible to justify.

Career average DB designs do not suffer from this problem, and can be made fair if income is revalued to adjust for inflation. A revalued career average scheme is very similar to a DC fund with the crediting rate having the same effect as the rate of revaluation. The difference is that a career average scheme with a uniform rate of pension accrual requires increasing contributions with age[103]. There would not appear to be much to choose between such a DB fund and a DC fund with investment guarantees.

4.2 Integrating death benefits before retirement

Thompson (1998; chapter 10) points out that the pension payable can be enhanced if the death benefit payable to survivors before retirement is limited to their real needs and does not include a return on contributions. The improvement is of some significance, depending on mortality rates. Allowing for no payment on death before retirement, the savings for 30 years of level contributions might vary from some 10% to 40% - the former based on US social security mortality[104] and the latter on Dorrington's (1989) estimate of black mortality in South Africa.

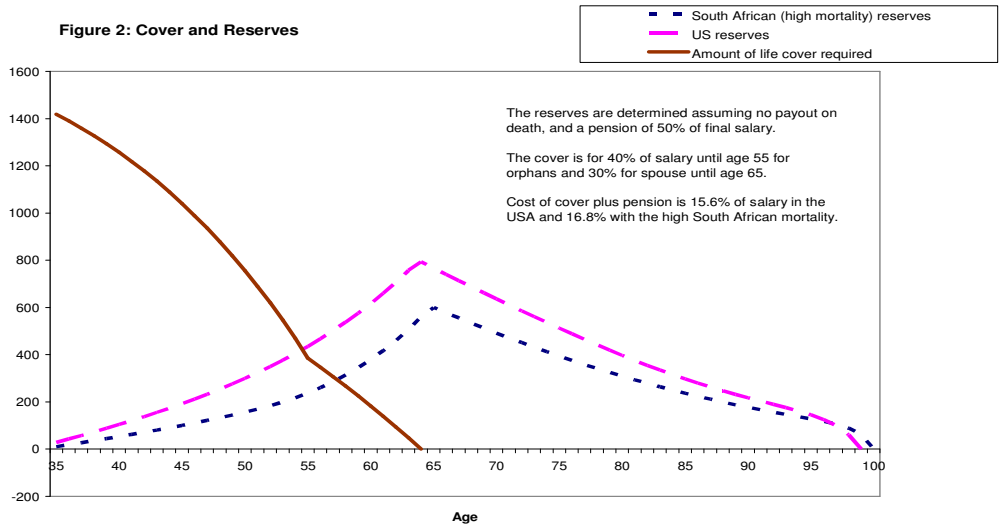
Figure 2 below shows how a pension fund would accumulate and de-cumulate reserves for members experiencing these mortality rates. It also shows the life cover that might be required for a family with three children where there was also a need for a spouse's pension (40% of salary for the children and 30% for the spouse). While the cost of this cover would be 1.9% and 10.7% of salary respectively, the total cost of both retirement

---

[103] An accrual rate of n% requires a contribution of n% $v^{65-x} a_{65}$ - perhaps with an adjustment for mortality. Discounting at a positive rate will mean that the contributions are increasing.

[104] From www.ssa.gov/OACT/STATS/table4c6.html

and mortality benefits is 15.6% and 16.8% - not greatly different for vastly different mortality rates.



The mortality difference between members of most retirement funds is likely to be much smaller than these extremes. A good case can be made for charging the same rate for this package of benefits on the grounds of equality, administrative simplicity and the elimination of underwriting. In contemporary South Africa, it would, in any event, be difficult to apply differentiation that disadvantaged any particular racial group.

In times before AIDS became a problem, being able to pass recruitment interviews would have been sufficient to qualify for life insurance cover. Given the recent improvement in AIDS treatments, it is quite possible that this position can be restored if the employer has a good record of ensuring treatment protocols are followed. If this is not possible in contemporary South Africa, one might be able to subsidise the cover of HIV positive members, or choose to exclude them from cover rather than to reduce or eliminate group insurance benefits for everyone.

The most obvious way of achieving this integration is to separately account to members for retirement and death benefits, with the former leading to a credit being added to the retirement account from the death release, and the latter leading to a charge. Members are already accustomed to the charge; the credit may well come as a pleasant surprise and will have the advantage of educating members about the benefits of annuitization. The

credit will exceed the charge once the retirement account exceeds the death benefit. As can be seen from figure 2, the lines may cross some years before retirement.

4.3 Integrating disability benefits

Disability can be seen as analogous to retirement as both involve a withdrawal from the workforce. Indeed, the discussion of both retirement and disability above suggests that there are limited reasons for the state to provide encouragement for retirement of the able – except if it can be seen to reduce unemployment amongst the young. It is thus ironic that South African retirement funds are not permitted to provide for disability benefits. Allowing retirement funds to pay disability benefits would be administratively convenient, and allow for better integration of benefits. In the interim many employers offer group benefits through separate insurance schemes.

4.3.1 Offset with post-retirement mortality

Of particular importance is the offset - between disability and mortality impairment after retirement - that would be readily available to a largish retirement fund offering disability insurance and compulsory annuitization, but difficult and probably inappropriate for an individual.

The difficulty arises from the need to evaluate the degree of impairment before an impaired life annuity can be issued (by a fund or an insurer) where annuitization is not compulsory. In addition, even if the disabled member is unable to obtain an impaired life annuity, the providers of regular annuities will need to take anti-selection into account in the pricing of their annuities, which increases the price of annuities even further.

The likelihood that an impaired life annuity would be inappropriate arises from the wide variety of conditions that lead to disability and impaired mortality, as well as the fact that different rules will invariably be applied by different funds. A taskforce of the Institute of Actuaries of Australia[105] reported: "Analysis of total and permanent disablement statistics from an insurer indicates that two-thirds of those claiming total and permanent disablement benefits do not have a substantially reduced life expectancy." On the other hand, it is clear that those taking early retirement can experience a significant reduction

[105] www.actuaries.asn.au/IAA/upload/public/2006_0728_SuperTax_IAA_Sub5_Disability_final.pdf

in life expectancy. The Continuous Mortality Investigation Reports[106] gives the additional mortality of UK life office early retirement pensioners in CMI 16. Table Pen1.2a provides a smooth additional mortality loading of some 30% under 70, 20% under 75 and 5% under 80. This produces annuity rates some 25% higher than normal – equivalent to an addition of 10 years at ages between 55 and 65, and more for younger ages.

These annuity enhancements would be significant relative to disability costs. Duggan *et al* (2005) calculate that the US decision in the early eighties to increase the age of full entitlement to an age pension from 65 to 67 will lead to an increase in disability claims equal to just 4% of the savings in pension payments.

4.3.2 Temporary disability

Disabling conditions can persist for some time without being obviously permanent. In many instances, there will be a possibility of recovery. For such cases and all disabilities at earlier ages, it is clear that people require insurance and retirement fund trustees should endeavour to ensure that they are offered as employee benefits outside the retirement fund – until it becomes possible to do so internally.

It is suggested that all disability cover should include provisions for rehabilitation benefits in addition to income replacement, and that the terms governing the payment of benefits should make the distinction suggested in section 3.2.2 above. Funds and insurers should be resolutely committed to restoring people to the workforce. Contracts should be drafted to that they allow for additional support for living expenses where necessary, and pay benefits to third parties if appropriate.

If regulation makes this difficult for a retirement or benefit (disability) fund, arrangements should be made with a medical aid to fill the gaps.

4.4 Dealing with divorce

Divorced spouses and their children frequently have difficulty in accessing their fair share of retirement fund benefits.

---

[106] Published in the UK by the Institute and Faculty of Actuaries

One obstacle is that the law on divorce[107] effectively includes only the withdrawal benefit from the retirement fund. Another is that lump sums paid on retirement are spent by the member on himself with the consequences of poverty for divorced spouses and widows recorded by Auerbach and Kotlikoff (1991).

Ensuring that there are no financial incentives for one party to initiate divorce should clearly be an objective of law reform. In its absence, trustees of retirement funds should ensure that their funds do not contribute to such incentives. Given that men often have a financial incentive to initiate a divorce, this means that the benefits of wives and widows should be protected as much as is possible.

This could take two forms:

a) The payment of benefits as annuities that are likely to be shared with wives, or can be attached by Court orders if partners are separated.

b) Ensuring that lump sum payments are appropriately divided between spouses in terms of their marriage contractual arrangements. If, for instance, marriage is by South Africa's default accrual system, it would seem that lump sum benefits accruing during the marriage belong equally to both spouses – even if they are not separated. It seems that trustees not taking this into account when paying benefits may find themselves exposed to claims for negligence.

4.5 Integrating unemployment benefits

The withdrawal benefits of retirement funds are available to fund periods of unemployment. This has the effect of reducing retirement funding, which at least in some cases, will lead to later retirement. Stiglitz and Yun (2002) show this is theoretically desirable, and Larson and Pedersen (2005) show that it appears to work well in Denmark where those who have not claimed from the unemployment fund are able to draw down the equivalent of a no claim bonus and go on early retirement. It is suggested however that retirement fund rules should withhold the withdrawal benefit until UIF benefits have been exhausted. It would also meet the need of unemployed members – and their spouses - if the withdrawal benefit was paid in monthly instalments until it is extinguished, or the

[107] See discussion in the South African Law Commission discussion paper 77 at ww.law.wits.ac.za/salc/discussn/dp77.wp

member is re-employed. Such an arrangement might be made more palatable to members who fear unemployment if it were accompanied by an insurance benefit that extended the benefit paid by the UIF.

4.6 Inappropriate insurance

Some of the insurance currently offered to South Africans resembles an unlucky dip. Benefits payable only because of accident are of particularly dubious value. If cover is required, then it is required regardless of the cause. The attraction of accidental benefits is that they apparently have a lower price, are easier to sell as no medical questions have to be asked, and they can be offered by short-term insurers as well as long-term assurers.

Disability contracts may also offer benefits on a fixed scale for loss of the use of various parts of the body. These bear no necessary relationship to the financial losses involved but may give the impression that they can replace disability insurance that offers more complete coverage.

The position is aggravated by legislated schemes that only provide cover for accidents. Workers' compensation regulations require employers to offer cover to employees injured while on duty. Accident cover is the cheapest way to do this. It is also offered by the Road Accident Fund (RAF) for death and disability if the driver can be shown to have caused the accident. Furthermore, the Assessment of Damages Act 1969 prevents the deduction of the victim's own life assurance benefits in assessing a claim for damages. This means that the victim who successfully claims may be over-insured. Disability benefits are likely to be reduced to ensure that members are not able to claim more than 75% of their income in benefits, but Nachemson's (1999) criticisms of workers' compensation arrangements mentioned in section 3.2.4 above are likely to apply.

Also problematic are lump sum benefits, which are common but invariably inappropriate as disability cover.

a) The need for a monthly amount to pay expenses is the same as on retirement or bereavement. Disability, however, is likely to be more stressful, making it more difficult to manage a large lump sum especially if the condition involves intellectual dysfunction.

b) Lump sums are normally paid on total and permanent incapacity after a waiting period of at least six months. They cannot easily be adapted for temporary or partial disability. While they do fund costs of rehabilitation, the cover is at best belated.

c) If the amount of the lump sum is not reduced as the members' accumulated retirement savings increase, lump sum cover may also be excessive at ages close to retirement.

Trustees of retirement and benefit funds should obviously avoid inappropriate lump sum and accident covers. Given that statutory accident cover is likely to apply for some time, they could consider offering a benefit that was reduced in the case of the member receiving workers' or RAF compensation. While the amounts concerned may not appear significant, one by-product would be to make members conscious of their total need for cover. An alternative approach would be for the employer to offer full life cover and not merely accidental death cover as part of its workers' compensation obligations.

5 THE PROPOSED POLICY FRAMEWORK

These considerations provide the basis for the design of pension funds that meet people's needs. Legislation however requires greater justification than this. Justice, the traditional criterion for evaluating policy recommendations, can be evaluated by determining whether the legislation is intended, in a procedurally just manner, to maximise the realization of five sometimes conflicting objectives: equality, liberty, efficiency and recognition of people's just deserts, as well as provision for their basic needs. The proposals developed below consider each of these objectives. This section looks first at the compulsory membership of funds, the minimum benefits that should be provided with a brief consideration of consistent charging and governance requirements.

5.1 Compulsion

On the question of compulsion *per se,* perhaps the most common argument is that people are myopic. While bolstered by Hurst's (2003) findings reported in 2.4.3 that many people actually want to be compelled to save, justice would require there to be a careful balancing of the advantages of compulsion against the resultant loss of liberty. One cannot argue that a section of the population is myopic without wondering whether it is the putative experts that have missed something. The discussions in section 2.6, 3.2 and

3.5 certainly suggest that it is disability benefits that are really required, and that experts might be requiring people to save too much for retirement.

5.1.1 Retirement funds as common goods

A more satisfying argument can be developed if the group or national provision of retirement and risk benefits is considered as a common or public good. Olson (1965) writes about the "anarchistic fallacy", which is the "illusion that mutually useful relationships will spontaneously evolve in a free society". Assuming self interest only, individuals will not contribute to public goods - where personal costs exceed benefits. Group members must be persuaded that there will be limited free riding (other individuals not paying their fair share).

Pension funds (with the other benefits recommended here) provide a common good at two levels. At a community level, if people who are able to provide financial security for themselves fail to do so, they become free riders and they or their families a burden on others. At a national level, the amount spent on social welfare for those who could have looked after themselves is not great. Tax concessions for retirement funds exceed any savings in social pensions, as shown, for instance, in Chapter 3 above. The free riding would be more likely to be felt in local communities, where people may feel obliged to support their relations and friends that have fallen on hard times – or feel guilty if they do not. Such support is likely to be unsatisfactory if entirely *ad hoc*; while as soon as thought is given to an appropriate structure, it becomes clear (in terms of just deserts) that people should make their own contributions.

At the employer level, more efficient group schemes might not come into existence because of the likelihood that members will free ride on the efforts of the few that might set up the scheme. Once schemes were in existence, without compulsion, the prevarication of some members would constitute free riding in that it would require the additional costs to sell benefits to each individual. Rusconi (2004) shows that these can more than double the costs. These benefits extend to the reduced need for underwriting for active members, and the ability to support disability benefits for older members if they are required to annuitize their retirement benefits.

There is a further advantage in that better informed members of a group can protect the more naïve. Campbell (2006) describes various research which suggests that a large proportion of financial service consumers (buyers of housing finance in his examples)

fail to make obviously advantageous decisions. In group schemes, sophisticated members of the group will act as buyers for all, while communal pressure or regulatory intervention can prevent exploitation of the weak or naïve.

It is true that members of group retirement funds have to surrender some of their power to make individual choices. It is therefore consistent with maximising people's freedom that retirement funds should limit the benefits offered to a minimum level likely to be needed by all (or at least the overwhelming majority) of members. Members with additional requirements can make voluntary provision for these. Trustees may however provide for optional additional benefits to be offered by the retirement fund.

5.1.2 Contributions

Accepting the advantages for those who can afford to contribute, it does not follow that compulsory membership should be extended to those in the informal sector for whom contributions cannot be collected efficiently. In Chapter 2 above, it is argued that this represents an unwarranted interference and an inefficient tax on people struggling to make a living in this way. The current state pension provides an adequate base for retirement, and would be even more suitable if the means test was abolished as suggested in Chapter 3.

It also does not follow that significant savings contributions are necessary early in the lifecycle. It can be seen from figure 1 in section 2.4 above that retirement saving should often not begin until the age of 40 or later. Before that time, people are better off repaying the mortgage on their homes. The costs of borrowing on a home loan and simultaneously investing through a retirement fund are significant. The margin of home loan rates over matching fixed interest investments are invariably more than 3% per annum in South Africa, to which must be added marginal retirement fund charges and the 9% tax on retirement fund income. If individuals invest in shares, the costs are balanced by the equity premium, but one would have to take an optimistic view of the latter to produce a positive expected benefit. Members are therefore unlikely to gain, except exposure to unnecessary risk.

This is the reason that the Taylor Committee recommended compulsion, but that those under 40 be permitted to redirect the contributions into repaying their home loans. Forty is probably too early for compulsory contributions, especially for those who have their children late or are re-establishing their household assets after a divorce. A provocative

alternative would be to suggest 60 as the age after which compulsion would apply. Most people are capable of working until 70 and the annuity rate for a man of 70 is about 10 times the annual annuity. This means that a single male could live on 50% of his income for the rest of his life, if he contributed 50% for ten years from age 60. From the calculations in 4.3.1, this could probably include disability insurance if he were pre-committed to annuitization.

Compulsory contributions with possible withdrawals for the purchase of a first house (for both the first and subsequent families) and for periods of unemployment appear reasonable compromises. Both are currently permitted by the Pension Funds Act 1953.

5.1.3 Insurance cover

The Committee also recommended compulsory orphans' and disability cover. The arguments have been made above: those who do not pay for cover will leave orphans that are provided for by others – or themselves be provided with disability benefits. There is a greatly reduced need for underwriting, so allowing more people to obtain necessary cover at cheaper prices. Supporting the hypothesis that more sophisticated members protect the more naïve, members of group schemes would also appear to get less of the inappropriate lucky dip cover described in section 4.6 above than individual policyholders. In South Africa, they are more likely to have inflation protected pensions. It can be noted that there is already compulsory cover for work and some road accidents.

5.1.4 Annuitization

The Committee recommended that all retirement funds offer pensions on retirement, but did not recommend whether these should be compulsory or be a default option allowing for commutation.

If some people are myopic before retirement, Broe's (2005) findings about intellectual impairments increasing with age, suggest it is much more likely after retirement. This makes it surprising that when it comes to compulsory annuitization and compulsory contributions, most people appear to be persuaded of the need for the latter, but not of the former. One reason is the disincentive of the means test, but this does not apply to all. The papers in Fornero and Luciano (2004) suggest a number of reasons for the unpopularity of annuities: poor value of money, bequest motives, underestimation of life expectancy, crowding out by compulsory annuitization of state benefits, or that a

sufficient level of wealth or family support makes longevity insurance unnecessary. Suggested by the sociological arguments of section 2.6 are the interests that have allowed lump sums to be spent by older men, who leave their ex-wives and widows in relative poverty. This latter suggestion also provides a possible explanation for the otherwise inexplicable tax advantages enjoyed by lump sums over annuities.

Weighing heavily in favour of some compulsion is the arguments made in 4.3.1 about funding ill-health retirement benefits by pre-commitment to annuitization. Another is the likelihood that people over 85 at least need assistance in managing their finances and that annuities provide the necessary budgeting and administrative services. The iterative argument is that because people cannot know whether they will become incapable of managing their finances next year, they should make provision at all times. If people should be compelled early in life to save for when they cannot look after themselves, and prevented late in life from withdrawing savings for similar reasons, permitting the proceeds to be spent between retirement and senescence appears to defeat both purposes.

5.1.5 Occupational or national arrangements?

Choice of retirement fund therefore undermines the protection offered by sophisticated buyers, economies of scale, universal insurance cover and annuitization. The fact that compulsory fund membership is almost universal suggests that the benefits from efficiency, extended coverage of needs and a closer match to people's desert are generally seen to outweigh the restrictions on people's freedom. This does not address the question of whether membership should be of a mandatory national scheme or the occupationally based schemes that cover most of the formally employed in South Africa.

The possible benefits of a national scheme would be for greater efficiencies, and greater equality in charging for risk benefits - as the well off could subsidise poorer members. There is no doubt that economies of scale are possible; Rusconi suggesting that larger funds can achieve cost savings of 50%. His statistics however suggest that there will be few further economies to be gained from funds over R1 billion in assets and perhaps over 1,000 members. The Taylor Committee suggested that ways should be found of rationalising small retirement funds to gain these advantages.

The calculations in section 4.2 suggest that over a lifetime, the difference in contribution between those with the lowest and those with the highest mortality would be some 2% of salary. The maximum redistributive impact would therefore be of the same order as a 1%

increase in the top marginal rate of tax. This seems insignificant. Although this will be higher with AIDS, the potential for proper treatment suggests that this is a temporary feature around which it would be inappropriate to redesign a retirement system.

A national scheme can also offer investment guarantees. Investment risks and funding issues are not considered in this paper. It can however be argued the state's resources ought not be spent on guaranteeing the investments of the wealthier.

Employment based group schemes therefore appear to achieve almost all the advantages of a national scheme, but allow for greater freedom, more democratic control and may well foster efficiency and innovation through competition of service providers. State run schemes are also more subject to political pressures to create excessive benefits for some classes of members. If these advantages are not already captured by the concept of subsidiarity, it would seem an appropriate principle to apply to this question: government should be as close as practical to the people governed.

5.2 Minimum benefits

This section sets out the minimum benefits that it appears would be required by almost all members. The pensions should, obviously, all be subject to mechanisms that would be expected to provide compensation for inflation.

5.2.1 Contributions

Contributions, whether expressed as employers or employees, should be at least sufficient to cover insurance benefits and administration costs and leave at least 7% for retirement savings. The minimum contribution rate is clearly debatable. This suggestion is based on what would be required for a pension of 50% of average earnings for a single male at 65, who contributes for 30 years and earns an average of 2.5% of income. The amount in the savings account should be available until the age of 45 to repay a loan on the member's house. Contributions should be payable to normal retirement age, which should not be less than 65.

5.2.2 Insurance benefits

The fund should provide pensions to orphans on death. Orphan's pensions should be at least 15% of taxable income each, with a maximum per family of at least 40%, payable

until 18, or 23 if the beneficiary is in full time study, or life if the beneficiary is severely physically or intellectually disabled.

Spouses' pensions should be optional, but if they are offered they should be payable until normal retirement age, which should be defined as the current age for access to the state pension.

The funds should also provide disability pensions of at least 50% of income plus normal contributions to the pension fund. Cover should continue until normal retirement age. The cover should include income replacement and the costs of rehabilitation and permit partial benefits. The benefit should be payable until retirement. It should be permissible to reduce the cover for non-specific back injuries and untreatable depression, and to pay benefits to a third party if this appears to be in the interests of the member.

5.2.3 Unemployment

Members losing their jobs (for whatever reason) should be entitled to withdraw the savings element of their pension, subject to the following limitations:

- No money should be paid while the member is entitled to UIF benefits.
- All benefits should be paid as regular pensions unless a court determines that part of the pension should be made available as a lump sum in order to settle a debt or other obligation. Some part of the pension should be protected against debts unless the debt was incurred fraudulently. The regular payment should be no greater than the maximum monthly disability benefit.
- Provision should be made for divorced and estranged spouses to have part of the benefit transferred to their pension fund.
- Once re-employed, any remaining retirement account should be transferred to the fund of the new employer.

The fund should be permitted to take out unemployment insurance so that members do not draw down in their savings accounts.

5.2.4 Retirement

Early retirement should be covered by disability insurance if it is caused by ill-health or by the unemployment rules otherwise.

A lump sum commutation of part of the annuity could be permitted at retirement age. The simplest approach to this would be to retain the current limitation of 33%.

The balance of the benefits should be paid out in a lifetime annuity with a spouse's reversion of at least 50%. If no residual benefit is permitted, there is no need for a minimum withdrawal level. A maximum withdrawal limit could be set with reference to an appropriate joint annuity.

Trustees should be required to ensure that divorced and estranged spouses receive their fair share of both lump sums and annuities.

5.2.5 Discontinuities

This structure would not necessarily prevent financial distress, and might have to be supplemented by additional insurance or savings for any of the following.

- Insurance cover for orphans and spouses.
- Disability income.
- A shortfall in the retirement benefit when compared with the family's current standard of living whether funded by earnings or a disability or survivors' pensions.

Funds should provide members with appropriate inflation adjusted projections at normal retirement date to alert them to any further needs for savings. They might give their members the opportunity to pre-commit to a particular target at normal retirement date, and annually calculate the contributions necessary to meet it. This would be in the spirit of the observation by Mitchell and Utkus (2004) that people are found to be more ready to commit themselves to future savings.

5.3 Charging and reserves

This benefit set would allow for cross-subsidies between members. The question arises as to the extent that they should be disclosed or prevented.

An important governance principle to apply would be that the fund should not build up reserves that are not allocated to particular individuals. Such reserves necessarily involve opaque cross-subsidies and usually depend on the discretion of the trustees, which creates uncertainty and the potential for conflicting interests and duties.

Because of the significant differences between people of different ages, it would be desirable for mortality and disability rates to be age dependent. Flat rates would mean that the costs of different funds would differ dramatically because of different age profiles – especially if the equality extended to pensioners, whose higher mortality releases would contribute significantly to reduced active life costs.

It does not seem necessary for any other distinctions to be made. Of the numerous possible categorizations, sex, income, habits (principally smoking and drinking), marital and health status are possibilities. To the extent that they are made, there will be cross-subsidies from those subject to lower mortality for insurance benefits and in the opposite direction when allocating annuitization releases. As illustrated in section 4.3.1, the impact over a lifetime is not that significant and may differ for different groups. Particularly sensitive cross-subsidies may arise if single people are required to pay for spouse's pensions and those without dependent children for orphans' benefits.

In order to avoid any legal uncertainty debate about potential discrimination, legislation should make it clear what types are acceptable and what not.

Simplest administratively would be to permit no discrimination other than for age for the compulsory package of benefits. This is also defensible in that it treats all members of the same age equally, and is likely to provide some subsidies to the less well off.

5.4 Governance

There are a number of possible approaches to ensuring compulsory membership. South Africa is unique in choosing an approach that requires membership of funds to be compulsory to all employees of a single employer if the tax advantages are to be available. This, together with tripartite agreements covering a number of industries, would appear to have proved successful in obtaining a high coverage of the formally employed without legal enforcement. It has the particular advantage of not harassing smaller employers and the informally employed for which membership would appear to be a burden. The Taylor Committee's recommendation was that compulsory membership should be enforced by the same bodies as currently enforce workers' compensation insurance.

Another governance question relates to representation of members and the level of expertise of the governing bodies. There are some, such as Myners (2001) who suggests

that funds should have some professionally trained members of their governing body. This has apparent advantages, but clearly weakens the position of the lay trustees in board meetings. If trustees want independent advice, they would be advised to contract for it rather than to surrender some of their powers. The Taylor Committee recommended that lay trustees continue to control South African retirement funds, as reflected its preference for democratic rather than technocratic governance.

6 CONCLUSION

Retirement funds, in South Africa, are the pivotal institution for the management of the middle class financial lifecycle and its attached risks: death, disability, dismissal and divorce. DC funds do not integrate these benefits in a way that DB funds were able, but there is no going back to the opaque cross-subsidies of DB funds.

This paper recommends that funds be based on a DC accumulation accounts but required to offer a minimum package of pension benefits including enhanced disability benefits. Recent research confirms that a more aggressive approach to rehabilitation and return to work can be justified – particularly in the case of backache and depression. Legislation should require compulsory:

- contributions, but only if withdrawal from the account is permitted for housing and unemployment;
- preservation except for unemployment, when monthly benefit payments should be permitted;
- annuitization of at least a portion of the proceeds at retirement;
- orphans' pensions;
- disability benefits.

Policy reform is also required in a number of areas that create over-insurance, which is both a waste and a moral hazard. There is the need to facilitate the match between life cover offered by retirement funds and that demanded by banks and other money lenders so that people are not required to over-insure themselves. The accident cover offered by workers' compensation arrangements and the Road Accident Fund is often over-insurance and should be abolished; people need cover for death and disability regardless of its cause. Funeral insurance for family members should also be discouraged in order to limit over-insurance that creates perverse incentives.

Reform should also facilitate life and disability cover being paid monthly: tax considerations discourage the former, restrictions on the benefits that can be paid by retirement funds the latter.

REFERENCES

Allen E T (1970) Survivor's Income Benefits *Journal of Risk and Insurance* 37.3: 397-406

Allen S G (2001) More on the Estimation of the Human Capital Depreciation Rate: Technology and the wage structure *Journal of Labor Economics* 19.2: 440-483

Amato P R (2000) The Consequences of Divorce for Adults and Children *Journal of Marriage and Family* 62.4: 1269-1287

Apps P & R Rees (2003) Life Cycle Time Allocation and Saving in an Imperfect Capital Market *Working Paper, Sydney Law School*

Arrazola M & J De Hevia (2004) More on the estimation of the human capital depreciation rate *Applied Economics Letters* 11.3: 145-148

Asher A (2003) Meeting Consumers' Needs *Understanding Actuarial Management: the actuarial control cycle* ed Clare Bellis, John Shepherd & Richard Lyon, Institute of Actuaries of Australia, Sydney: 71-108

Auerbach A J & L Kotlikoff (1991) Life Insurance Inadequacy - Evidence From a Sample of Older Widows *NBER Working Paper No. w3765*

Avellar S & P J Smock (2005) The Economic Consequences of the Dissolution of Cohabiting Unions *Journal of Marriage and Family* 67.2: 315-327

Banks J, R Blundell & S Tanner (1998) Is there a Retirement-Savings Puzzle? *American Economic Review* 88.4 769-788

Becker G S (1993) *Human Capital: A Theoretical & Empirical Analysis, with Special Referen*ce *to Education*. University of Chicago Press

Brantervik A M, I E Jacobsson, A Grimby, T C E Wallén & I G Bosaeus (2005) Older Hospitalised Patients at Risk of Malnutrition: Correlation with quality of life, aid from the social welfare system and length of stay? *Age and Ageing* 34: 444-449

Broe G A (2005) A Longitudinal View of Australian Aged Care Policy: A socio-clinical perspective *Australian Social Policy Conference presentation* www.sprc.unsw.edu.au/ASPC2005/papers/Paper113.ppt

Buetler M, O Huguenin & F Teppa (2004) Why Forcing People to Save for Retirement may Backfire *CESifo Working Paper No. 1458*

Bytheway B (2005) Ageism and Age Categorization *Journal of Social Issues* 61.2: 361-374

Campbell J Y (2006) Household Finance *NBER Working Paper No. 12149*

Conti R M, E Berndt & RG Frank (2006) Early Retirement and Public Disability Insurance Applications: Exploring the Impact of Depression *NBER Working Paper No. W12237*

Comanor W & S Phillips (2002) The impact of income and family structure on delinquency *Journal of Applied Economics* 5.2: 209-233

Cooper D R (2002) Family Fortunes: A Guide to Saving for Retirement *British Actuarial Journal* 8.5: 849-885

Costa D L (1998) The Evolution of Retirement: Summary of a research project *American Economic Review* 88.2: 232-236

Dave D, I Rashad & J Spasojevic (2006) The Effects of Retirement on Physical and Mental Health Outcomes *NBER Working Paper* 12123

Deaton A S (1998) Getting Prices Right: what should be done *Journal of Economic Perspectives* 12.1: 37-46

Deaton A S (2005) *Franco Modigliani and the Life Cycle Theory of Consumption* ssrn.com/abstract=686475

De Jong P R, M Lindeboom & B van der Klaauw (2006) Screening Disability Insurance Applications *IZA Discussion Paper No. 1981* ssrn.com/abstract=884436

Dorrington R E (1989) African Mortality Rates: An initial estimate *Transactions of the Actuarial Society of South Africa* 7.1: 169-180

Duggan M, P Singleton & J Song (2005) Aching to Retire? The rise in the full retirement age and its impact on the disability rolls *NBER Working Paper No. 11811*

Duncan G J (1988) 'The Volatility of Family Income over the Life Course' *Life Span and Development* 9: 317-351

Ehrlich I & J Kim (2005) Social security, demographic trends, and economic growth: Theory and Evidence from the international experience *NBER Working Paper No. 11121*

Fisher H F and J Young (1965) *Actuarial Practice of Life Assurance* Cambridge University Press, London

Fore D C & P B Hammond (2005) Resurrecting the Defined Benefit Pension Plan: A new perspective *PRC Working Paper No. 2005-15* ssrn.com/abstract=830047

Fornero E & E Luciano (ed) (2004) *Developing an Annuity Market in Europe* Edward Elgar

Gomes F & A Michaelides (2005a) Optimal Life-Cycle Asset Allocation: Understanding the empirical evidence *The Journal of Finance* 60.2: 869-904

Gomes F & A Michaelides (2005b) *Asset Pricing with Limited Risk Sharing and Heterogeneous Agents* AFA 2006 Boston Meetings ssrn.com/abstract=687126

Goodman N J & D C Stapleton (2006) Federal Program Expenditures for Working-age People with Disabilities: Research Report *Journal of Disability Policy Studies*, Forthcoming

Gourinchas P-E & J A Parker (2002) Consumption over the Life-Cycle *Econometrica* 70.1: 47-89

Groot W (1998) Empirical estimates of the rate of depreciation of education *Applied Economics Letters* 5: 535-538

Hadjimatheou G (1987) *Consumer Economics after Keynes: Theory and Evidence of the Consumption Function* Wheatsheaf Books, Brighton

Hayward M D, W R Grady, M A Hardy & D Sommers (1989) Occupational Influences on Retirement, Disability, and Death *Demography* 26.3: 393-409

Hellerstein J K, D Neumark & K R Troske (1996) Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations *NBER Working Paper No. W5626*

Hurd M D & S Rohwedder (2006) Some Answers to the Retirement-Consumption Puzzle *NBER Working Paper No. W12057.*

Hurst E (2003) Grasshoppers, Ants, and Pre-Retirement Wealth: A test of permanent income *NBER Working Paper No. W10098*

Jager A, Kops Y & van der Lippe T (2003) The work-family balance for couples in Europe *ch 10 in HWF Research Report 4.2 Ed C Wallace* http://www.hwf.at/project_report04_vol02.htm

Kerrigan G K (1991) The role of COSATU affiliated trade unions in retirement provisions in SA *Transactions of the Actuarial Society of South Africa* IX.I: 177–98

Kristen K & P Malone (2005) Housework and the Wages of Young, Middle-Aged, and Older Workers *Contemporary Economic Policy* 23.2: 224-241

Laibson D I, A Repetto & J Tobacman (1998) Self-control and Saving for Retirement, *Brookings Papers on Economic Analysis* I: 91-196

Larson M & P J Pedersen (2005) Pathways to early retirement in Denmark 1984-2000 *Institute for the Study of Labour Discussion Paper 1575* ssrn.com/abstract=714903

Lilly C (2006) Life Insurance Coverage: How much is the right amount? *TIAA-CREF Institute Trends and Issues 3/06* www.tiaa-crefinstitute.org/research/trends/030106b.html

Madge A (2000) Long-Term Aged Care: Expenditure trends and projections *Productivity Commission Working Paper No.1649* ssrn.com/abstract=270779

Millward J, A Lutte & R Purvis (2005) Depression and the perpetuation of an incapacitated identity as an inhibitor of return to work *Journal of Psychiatric & Mental Health Nursing* 12.5: 565-573

Mitchell O & S Utkus (2004) Lessons from Behavioral Finance for Retirement Plan Design *Pensions Design and Structure: New Lessons from Behavioral Finance* ed O Mitchell S & Utkus, chapter 1, Oxford University Press, Oxford

Modigliani F (1986) Lifecycle, Individual Thrift and the Wealth of Nations *American Economic Review* 76.3: 297-313

Mulligan C B & X Sala-i-Martin (1999) Gerontocracy, Retirement, and Social Security *NBER working paper w7117*

Murphy K & F Welch (1992) The structure of wages *Quarterly Journal of Economics* 107.1: 285–326

Nachemson A (1999) Back Pain: Delimiting the problem in the next millennium *international Journal of Law and Psychiatry* 22.5-6: 473-490

Neal D (1995) Industry-specific Human-capital: Evidence from displaced workers *Journal of Labor Economics* 13.4: 653-677

Nelson J A (1988) Household Economies of Scale in Consumption: Theory and evidence *Econometrica* 56.6: 1301-1314

Noble J H (1998) Policy Reform Dilemmas in Promoting Employment of Persons with Severe Mental Illnesses *Psychiatric Services* 49:775-781

Noonan A E (2005) At this Point now: Older workers' reflections on their current employment experiences *International Journal of Aging & Human Development* 61.3: 211-241

Ogg J (2005) Social Exclusion and Insecurity among Older Europeans: The influence of welfare regimes *Ageing & Society* 25.1: 69-90

Olson M (1965) *The Logic of Collective Action: Public goods and the theory of groups* Harvard University Press, Cambridge, Mass

Polachek S W & W S Siebert (1993) *The Economics of Earnings* Cambridge University Press

Psacharopoulos G (1994) Returns to Investment in Education: A global update *World Development* 22.9: 1325-1343

Rowntree B S (1901) *A Study of Town Life* Longman

Rusconi R (2004) *Cost of Saving for Retirement: Options for South Africa* http://www.assa.org.za/scripts/file_build.asp?id=100000357&pageid=1000000120

Sampson R J & Groves W B (1989) Community Structure and Crime: Testing social disorganization Theory *American Journal of Sociology* 94.44: 774-802

Sanlam Survey (2006) *Annual Survey of Retirement Benefits in South Africa* http://www.sanlam.co.za/eng/aboutus/employeebenefits/surveys/seb+survey.html

Schulz J H (2002) The Evolving Concept of Retirement: Looking forward to the year 2050 *International Social Security Review* 55.1: 85-105

Sener F (2001) Schumpeterian Unemployment, Trade and Wages *Journal of International Economics* 54.1: 119-148

Sigle-Rushton W (2005) Young Fatherhood and Subsequent Disadvantage in the United Kingdom *Journal of Marriage and Family* 67.3: 735-753

Social Security Administration (2006) *Social Security and Supplemental Security Income Disability Programs* http://www.ssa.gov/policy/docs/chartbooks/disability_trends/trends.pdf

Stiglitz J & Yun J (2002) Integration of Unemployment Insurance with Retirement Insurance *NBER Working Paper 9199*

Stock W A & K Beegle (2004) Employment Protections for Older Workers: Do Disability Discrimination Laws Matter? *Contemporary Economic Policy* 22.1: 111-126

Tate D G (1992) Workers' disability and return to work *American Journal Physical and Medical Rehabilitation*. 1992 71.2:92-6

Taylor V *et al* (2002) *Transforming the Present - Protecting the Future* Report of the Committee of Inquiry into a Comprehensive Social Security System for South Africa, Department of Social Development, Pretoria

Thompson L H (1998) *Older and Wiser: The economics of public pensions* Urban Institute Press, Washington DC

Tilse C, D Setterlund, J Wilson & L Rosenman (2005) Minding the money: a growing responsibility for informal carers *Ageing & Society* 25.2: 215-227

Tucker W (1990) *The Excluded Americans: Homelessness and Housing Policies* Regnery Gateway, Washington

Van Weel C & J Michels (1997) Dying, not Old Age, to Blame for Costs of Health Care *Lancet* 350: 1159-1160

Venti S F & D A Wise (1998) The Cause of Wealth Dispersion at Retirement: Choice or chance? *American Economic Review* 88.2: 185-191

Waldfogel J (1998) Understanding the 'Family Gap' in Pay for Women with Children *Journal of Economic Perspectives* 12.1: 137-156

Zaida A & T Burchardt (2005) Comparing Incomes when Needs Differ: Equivalization for the extra costs of disability in the U.K. *Review of Income and Wealth* 51.1: 89-114

APPENDIX

Table for Figure 1a: Children born in twenties

Age	Income after tax		Children	Housing payments	Total Expenses	Accumulated Savings
						0
23	50000	50000	0	0	47500	2500
24	60000	60000	0	0	57000	5500
25	72000	72000	0	0	68400	9100
26	86400	86400	0	0	82080	13420
27	95040	66528	2500	0	61776	18172
28	104544	73181	2500	0	67954	23399
29	114998	80499	5000	0	80499	23399
30	118448	82914	6000	0	76991	29322
31	122002	85401	7000	37500	85401	29322
32	125662	125662	7000	36408	125662	29322
33	129432	129432	7000	35347	129432	29322
34	133315	133315	7000	34318	133315	29322
35	137314	137314	7000	33318	137314	29322
36	138687	138687	7000	32348	138687	29322
37	140074	140074	7000	31406	140074	29322
38	141475	141475	7000	30491	141475	29322
39	142890	142890	7000	29603	142890	29322
40	144319	144319	10000	28741	144319	29322
41	145762	145762	10000	27904	145762	29322
42	147219	147219	10000	27091	139858	36683
43	148691	148691	10000	26302	147171	38203
44	150178	150178	10000	25536	146554	41828
45	151680	151680	10000	24792	145960	47548
46	153197	153197	10000	24070	145390	55356
47	154729	154729	5000	23369	139842	70243
48	156276	156276	5000	22688	139316	87204
49	157839	157839	0	22027	133811	111231
50	159417	159417	0	21386	133327	137321
51	159417	159417	0	0	111942	184797
52	159417	159417	0	0	111942	232273
53	159417	159417	0	0	111942	279748
54	159417	159417	0	0	111942	327224
55	159417	159417	0	0	111942	374700
56	157823	157823	0	0	111782	420741
57	156245	156245	0	0	111625	465361
58	154683	154683	0	0	111468	508575
59	153136	153136	0	0	111314	550398
60	151604	151604	0	0	111160	590842
61	147056	147056	0	0	110706	627192
62	142645	142645	0	0	110264	659572
63	138365	138365	0	0	109837	688101
64	134214	134214	0	0	109421	712894
65	130188	130188	0	0	109019	734063

Table for Figure 1b: Children born in thirties

Age	Income after tax		Children	Housing payments	Total Expenses	Accumulated Savings
						0
23	50000	50000	0	0	47500	2500
24	60000	60000	0	0	57000	5500
25	72000	72000	0	0	68400	9100
26	86400	86400	0	0	82080	13420
27	95040	95040	0	0	90288	18172
28	104544	104544	0	0	99317	23399
29	114998	114998	0	0	107500	30898
30	118448	118448	0	0	107845	41501
31	122002	122002	0	0	108200	55303
32	125662	125662	0	0	108566	72399
33	129432	129432	0	0	108943	92887
34	133315	133315	0	0	109331	116870
35	137314	137314	0	0	109731	144453
36	138687	97081	2500	0	97081	144453
37	140074	98052	2500	0	98052	144453
38	141475	99032	5000	37500	99032	144453
39	142890	100023	6000	36408	100023	144453
40	144319	101023	7000	35347	101023	144453
41	145762	145762	7000	34318	145762	144453
42	147219	147219	7000	33318	147219	144453
43	148691	148691	7000	32348	148691	144453
44	150178	150178	7000	31406	149424	145208
45	151680	151680	7000	30491	148659	148229
46	153197	153197	7000	29603	147923	153504
47	154729	154729	7000	28741	147214	161019
48	156276	156276	7000	27904	146531	170764
49	157839	157839	10000	27091	148875	179728
50	159417	159417	10000	26302	148243	190902
51	159417	159417	10000	25536	147477	202842
52	159417	159417	10000	24792	146734	215526
53	159417	159417	10000	24070	146012	228932
54	159417	159417	10000	23369	145311	243039
55	159417	159417	10000	22688	144630	257826
56	157823	157823	5000	22027	138810	276840
57	156245	156245	5000	21386	138010	295075
58	154683	154683	0	0	111468	338289
59	153136	153136	0	0	111314	380111
60	151604	151604	0	0	111160	420555
61	147056	147056	0	0	110706	456906
62	142645	142645	0	0	110264	489286
63	138365	138365	0	0	109837	517815
64	134214	134214	0	0	109421	542608
65	130188	130188	0	0	109019	563777

CHAPTER 6 - THE RELATIVE INVESTMENT PERFORMANCE OF THE COMMUNITY GROWTH FUND

First published in the South African Actuarial Journal (2005) 5: 1-26

1 INTRODUCTION

The explicit use of ethical or social criteria in the choice of institutional investments grew considerably in the last two decades of the twentieth century. A major question, raised by the growth in number and size of institutions adopting such criteria, has been the effects on investment performance.

In South Africa, the Community Growth Fund (CGF) was the first institution to explicitly adopt social criteria. The balance of this paper is divided into four sections. Section 2 gives an outline of the history and structure of the CGF. The third section looks at the reasons why share prices move both absolutely and relative to each other. The fourth then briefly discusses socially responsible investment. The final part looks at the relationship between the CGF social criteria and investment performance.

The conclusion is that, if anything, and contrary to some prior opinions, the CGF social criteria have had a positive impact on investment performance.

2 HISTORY AND STRUCTURE

2.1 Unions and provident funds

The creation of the CGF was rooted in two uniquely South African conditions.

The first is the South African retirement system. State involvement is limited to the provision of the flat basic old age grant (currently R780 per month), and the regulation of private provision. This is unusual in that there is no mandatory requirement to contribute to state run earnings related retirement provision. South Africa may also have been the first country to provide an explicit legal framework for retirement funds with the passing of the Pension Fund Act, 1956. The thriving private industry that has developed is a

major contributor to national savings[1]. High and volatile rates of inflation, and the limited availability of other inflation protected investments, have meant that equities play an important role in retirement fund portfolios.

The second root is the links between the black trade unions and retirement funds. Black South Africans were prevented from joining trade unions until the early eighties. This changed following the recommendations of the Wiehahn Commission (1979). The pensions arena became the first in which the new movement showed its strength. Government happened, at the same time, to put forward a Bill proposing to restrict pension fund members from withdrawing their interest when leaving the sponsoring employer.[2] The withdrawal benefits were seen as a critical support during periods of unemployment, and the new unions successfully led a campaign that lead to the withdrawal of the Bill.

Retirement funds therefore became a critical issue for the new union leadership. After the campaign against preservation, the major issue of the eighties was the transfer of workers' benefits from defined benefit pension schemes, controlled exclusively by employers, to union designed provident funds. These are typically defined contribution schemes controlled by boards of trustees, half of whom were elected by members or appointed by trade unions. They pay lump sum benefits on withdrawal and at retirement, which were much more favourable to lower paid workers. This is because the benefits on withdrawal are not subject to the penalties previously imposed on early leavers, while the lump sum on retirement allows pensioners to escape the means test that would otherwise have prevented them from drawing the state old age pension.[3]

---

[1] Statistics available from the South African Reserve Bank Quarterly Bulletin suggest that over a third of the total fixed capital stock is owned directly and indirectly by retirement funds. This adds the assets of public and private funds including those on the held by life insurers and allowing for double counting. In the March 2005 bulletin, the numbers can be found in tables S37, 41, 42 and 129.

[2] The Bill emerged from the First report of the Interdepartmental Committee of Inquiry into Certain Specific Pension Matters (The Meiring Committee) 1980.

[3] One large pension fund, advised by the author, saw 1000 people retire in 1989. 980 of these would have gained no net benefit from the pension fund as their state pension would have been reduced by an equal amount (and sometimes more).

Having gained control of the funds, some unions began investigating ways of exercising greater influence over the companies in which they were investing.

2.2 Formation of the CGF

From the papers available to the author, the idea of the CGF was first mooted in a meeting between staff of the Cape Town based Syfrets Managed Assets (SMA) and Labour Research Service (LRS) in May 1991. The LRS facilitated contacts with the largely Johannesburg based union head offices, leading to the official launch of the CGF in June 1992.

The CGF is operated by a management company, jointly controlled initially by SMA and Unity Incorporated, a non-profit company set up by the seven sponsoring unions. The management company subcontracted its functions to SMA and its affiliates. The author became the independent acting chairman of the board of the management company and served as Chairman until 2002.

The seven trade unions came from both the large trade union groups. Affiliated to the Congress of South African Trade Unions (COSATU) were the National Union of Mineworkers (by far the largest investor in the fund), the Construction and Allied Workers' Union, the Transport and General Workers' Union, the Paper, Printing, Wood and Allied Workers' Union. The others were from the National Council of Trade Unions (NACTU): the National Union of Food, Wine and Spirits and Allied Workers, the Transport and Allied Workers' Union, and the Metal and Electrical Worker's Union of South Africa. The Memorandum of Association of Unity set out Unity's main business as:

> "To advance the concept of socially responsible investment and to promote the influence of the Trade Union Movement in the economy through:
> - establishing and holding shares in Unit Trust Management Companies;
> - establishing financial vehicles, whether on its own or jointly with other organisations, for personal, Trade Union and Retirement Fund investments …"

2.3 Investment process

There were initially 16 social criteria, weighted as follows:

Table 1: Initial social criteria

Criterion	Weighting %
Job creation	14
Industrial relations	14
Employment conditions	13
Training	7
Equal opportunities	7
Health and safety	6
Product safety	5
Rejected participation in privatisation	5
Profit retention – for expansion	5
Affirmative action	5
Operations located in South Africa	4
Environment	4
Worker participation	4
Disclosure	4
Political profile – not abused	2
Social spending	1
TOTAL	100

SMA nominated those companies that they wished to include in the investment universe. The LRS and Unity then evaluated these companies. The process involved an initial questionnaire to be completed by the company, and then interviews with management and union organisers. It proved difficult at times to get co-operation from companies and some unions not affiliated to the CGF. The LRS then prepared a report that was submitted to the Unity board of directors for approval.

Once approved, SMA was able to buy shares for the CGF. The Unity board was keen to publish the results as quickly as possible, but SMA persuaded them to keep the names of approved companies confidential until they appeared as investments in the published reports of the CGF. SMA had been chosen for its exceptional investment performance in the years preceding the launch and its management was concerned that publishing the names of the companies in which they intended to invest, would push up share prices

before SMA had been able to acquire an adequate holding. The relatively low liquidity of the South African share market provides reason for these concerns.

At the meeting of the board of the management company in July 1992, one month after the launch of the CGF, it was reported that 9 companies, out of SMA's list of 51, had been approved, 7 had been rejected, and 11 had been referred back to the LRS for further investigation. The slow rate of approval, arising partly from the limited resources available and partly from the need to obtain consensus from the volunteer directors of Unity, has been a constant, but inevitable, source of frustration to asset managers.

2.4 Initial responses

The first investor in the CGF was Cyril Ramaphosa, then Secretary General of the ANC who made out his personal cheque at the initial press conference. This was followed the next day by a R1 million investment from the pension fund of the business (rather than labour) oriented Times Media Limited. Its managing director, Stephen Mulholland identified its launch as a "significant departure in the affairs of a trade union". A LRS spokesman[4] was "delighted that Times media has acknowledged the superior potential of socially responsible investment", but hastily pointed out the possibility that the company's industrial relations practices would disqualify it from appearing on the unions' list of approved shares.

Greenblo[5] identified the tensions. Existing participation in provident funds meant that the Unions had implicitly accepted the principle of investment in shares and participation in the capitalist system, even though this flew in the face of a COSATU economic policy document released earlier in 1992. The Unity unions' involvement in the CGF made it more explicit. Greenblo expressed the hope that the CGF approach would break down the "adversarial divide between employer and employee".

Other criticisms arose from a fear that the social criteria would reduce investment performance for the reasons set out in section 4 below.

---

[4] Reported in the Weekly Mail May 8 to 14 1992.

[5] Allan Greenblo: Editorial Comment, Finance Week May 7-13, 1992

2.5 Subsequent developments

As it transpired, the CGF was the forerunner of much greater union involvement in the capitalist system. Unions that initially resisted the CGF's apparent compromises[6] subsequently created union investment companies that actively pursued commercial objectives.

Inflows of R100 million were predicted for the CGF in the first year. Given that union controlled provident funds had some influence over perhaps R20 billion of assets, this did not appear particularly ambitious, but the fund had only R70 million by the end of 1993. Steady growth, usually below expectations, has been achieved, with the fund reaching R900 million by 1998 and currently approaching twice that size.

A number of factors for the disappointing growth can be identified. The initial opposition of some unions, inadequate management resources and numerous changes to personnel were perhaps the main factors. More important to the subject of this paper, was the question whether socially responsible investment would reduce investment returns.

2.6 New investment criteria

In the 1996 report to unitholders, "tougher new criteria for companies seeking inclusion in the share portfolio" were introduced. Eight categories, each given an equal weight in the final score, were to be used.

- Create jobs through innovation and expansion.

  The questionnaire used allocates a maximum of 20 points for this category, 5 are given for growth in jobs, another 9 for expansion plans, 2 for retrenchment procedures and 4 for not creating jobs overseas.

- Training of workers to enhance skills

  2 out of the 9 points allocated to this score were related to training expenditure as a percentage of payroll; the rest were qualitative measures, depending to a significant degree on collaboration with the trade union.

---

[6] The 1992/3 Institute of Race Relations Survey reports that a number of COSATU and NACTU trade unions refused to endorse the fund. (page 335)

- Economic and social empowerment

  20 out of 30 points here relate to bargaining with unions, 5 to donations to relevant social projects, and 5 to partnerships with emerging black businesses.

- Equity through affirmative action within the workplace

  Half the points here are allocated on whether the company scores better than the LRS average for its industry. The rest of the points are qualitative assessments.

- Good conditions of employment

  6 of the 15 points for this category relate to wages, the rest to benefits.

- Sound environmental practices must be promoted

  This is more complicated. Companies not involved in an industry which could have a potential environmental incident automatically get 9 out of 13 points. Companies in other industries can score these 9 if they have suitable policies – losing 1 if they experienced accidents and another if the accident was not catered for in their policies. Plus or minus 3 points are allocated for the desirability or otherwise of the company's products – and one for not having dealings with oppressive regimes.

- High health and safety standards must be applied

  Measures of injuries account for 3 out of 11 points; the rest relate to policies.

- Demonstrate open and effective corporate governance

  The points here are largely allocated in line with King Report (1994) on corporate governance. 3 points are given for the company's response to CGF queries and another 4 for disclosure to unions.

3 DETERMINANTS OF SHARE PRICE MOVEMENTS

3.1 Overall level of prices

The next three sections describe the three underlying forces at work in the determination of share prices. There is the demand for and supply of shares that arises independently of their actual value, and then the factors that drive profitability.

3.2 Demand

Demand for shares will depend firstly on the forces that lead to people saving. This is one of the major concerns of economics, but 75 years of intense debate and research has produced very little in the way of mathematical relationships. Smith (1990) reviews research that shows people tend to save more when:

- they cannot borrow,
- their incomes are rising rapidly,
- their income (current and future in the form of a public pension) is insecure,
- they have less in the way of accumulated wealth,
- they need the money for their own businesses and
- there are some tax advantages in doing so.

They may, but this is unclear, also save in order to leave money to their children.

People may also shift their savings between cash deposits, real assets - largely property, long term fixed interest stock and equities. These shifts will depend partly on inflationary expectations and legislation.

People also save, through retirement funds of various types or directly, for retirement. This aspect at least, we can model, as we can project the amounts people will require for their retirement with some accuracy. Some of the impetus for the higher share prices of the last decade appears to come from this source.[7]

3.3 Supply

The supply of equities will come from new issues of equity, which in turn will depend on matters such as technology, economic growth, labour scarcity, tax and politics (especially privatization or nationalization) and the internal politics of family businesses. Changes to these factors are largely unpredictable. Sales may also occur for reasons such as death duties and retirement expenses, which can be modeled.

Even if we could model the factors that underlie the demand and supply curves, they appear to be greatly influenced by changing perceptions, and it is unlikely that we should

---

[7] See Poterba, J (2004) *The Impact of Population Aging on Financial Markets* NBER Working Papers 10851, National Bureau of Economic Research (http://www.nber.org/papers/w10851.pdf)

find any stable relationship to convert them into prices. I like the description of these curves as "a mere smudge, to which it is vain to attribute elegant geometrical properties"[8].

3.4 Underlying profitability

The other major influence on share prices will be the underlying profitability of the companies that have issued the shares. This in turn will also depend, but in different ways on technology, economic growth, international trade, legislation, taxation, the relative scarcity of capital, and the quality of management. All these relationships are unstable.

This is however not to say that there is no measure of determining a reasonable long-term value for a share. The net asset value (NAV) is not commonly used to price shares, but is useful in providing two (albeit approximate) views of a company's worth. Firstly, if a competitor were to start up from scratch, it would seem that it would have to invest in similar equipment and make similar profits and losses to get to where the company now is. If so, it would have the same net asset value. Of course, inflation distorts this figure, and there are intangible items that enhance a company's future earning capacity: particularly intellectual property such as patents and reputation. On the other hand, newcomers to an industry can avoid the mistakes of the incumbent and focus on its more profitable markets. The point is that more newcomers will be encouraged into an industry as the ratio of price to net asset value increase. This ultimately brings share prices down towards NAV.

Secondly, the NAV also provides an estimate of the break up value of the company, or of its value if it were run down. Again, many adjustments would need to be made. This second argument is of importance when share prices are too low and asset strippers can buy up the shares and make a profit by selling off the assets.

In the economics literature, Tobin's Q - the ratio of share price to net asset value - has been used as a measure of the size of the share bubble[9] as well as a measure of the

---

[8] This is vintage Joan Robinson - a prolific Cambridge economist who died in 1983 – but I have lost the precise reference.

company's success in achieving a competitive advantage - or as a measure of lack of competition[10].

An alternative measure of the underlying value of a share is to discount future dividends at a suitable interest rate. The total return to an investor is given by the dividends and the capital growth. If dividend yields remain constant or fluctuate in a narrow band, the total return is entirely dependent on the dividends as the capital growth is equal to the rate of dividend growth. Actuarial models built by Wilkie (1995) for other countries, and by Thomson (1996) for South Africa, use mean reverting dividend yields to model investment returns.

The discounting of dividend yields becomes a less valuable estimator of a shares value if companies do not pay out a consistent proportion of the profits as dividends. It appears that one of the Miller Modigliani (1961) propositions - that suggests that investors would be happy to see management keep the dividends if the resultant increase in share price is greater than the value of the dividend - is relatively widely accepted, and has lead around the world to a fall in dividend yields. This has been exacerbated in South Africa since secondary tax on companies has penalised companies for paying dividends.

The difficulties with these methods lead to the most popular measure of a share's value, which is the ratio of market price to reported earnings per share. It is normally used as a simple rule of thumb to determine value relative to the whole market. A more rigorous approach would require the earnings to be discounted, as with dividends. Allowance again has to be made for inflation and the need to retain profits to fund growth.

3.5 Individual share prices

3.5.1 Fundamentals and perceptions

The movement of the share prices of individual companies, relative to the average, will depend primarily on the underlying profitability of the company relative to others –

---

[9] In an unpublished paper, Harney, M and E Tower, 2002 Rational Pessimism: Predicting Equity Returns using Tobin's Q and Price/Earnings Ratios http://www.econ.duke.edu/Papers/Other/Tower/Pessimism.pdf, finds that Tobin's Q provides a good indicator of future investment returns. See also Smithers and Wright (2000) which they quote.

[10] Lindenberg and Ross (1981)

rather than on overall supply and demand factors. In the short run perceptions of its relative future are important as are temporary pressures of demand and supply. For example, it is common to speak of an "overhang" of shares such as followed the demutualisation of the largest SA life assurers, which depressed their price for some time. The relative underlying profitability will depend on a number of factors.

3.5.2 Sector

Profitability will depend in many ways on the industry in which a company competes.

- It is probable that the demand for the products produced by companies operating in the same industry will be correlated, whether changes arise from demographics, technology or politics.

- In the same way, the supply of goods offered by competitors affects all companies within a sector.

- Similarly, companies within a sector are likely to find their costs of production are related, whether because of changes in technology or the costs of inputs.

Sectors can be measured, as most stock exchanges allocate companies to their appropriate sectors, but classifications change over time. Data on the rest of the factors mentioned below is only obtainable from the company with some difficulty.

3.5.3 Financial structure

While companies in the same sector may tend to have a similar financial structure, factors other than industry may apply. Growing or family controlled businesses, for instance, may use greater debt. Different managements may also take a different view on future interest rates – or differently hedge against price movements in their revenues or expenses. The most important of the implications of different financial structures are different exposures to interest rate, currency and commodity fluctuations.

3.5.4 Markets

Even within the same industry, companies develop different niche markets, where they may have a more appropriate product, stronger distribution or are better known. Thus, companies may be affected differently by demographic and economic changes – especially to exchange rates if they operate in different currency areas.

3.5.5 Production

Companies within the same industry may also use different technologies. This may lead, for instance, to greater exposure to changes in wage levels or to the development of different patents.

3.5.6 Staff and systems

In spite of the importance of the financial and material assets of a company, or its market strength, it is widely recognised that profits may well depend more on the skills, abilities and motivation of the management and staff, and on formal and informal institutional arrangements within the company.

3.6 Perceptions

It is clear from the above, that each of these headings hides a multitude of other factors, and that it is impossible to accurately model the future relative profitability of any company. The share price also therefore depends on the perceptions of those stock market participants who might consider buying or selling it.

Perceptions are influenced by the perceptions of others. One readily observable result is the herding of investors into observable fashions. Two often mentioned are the alternating attractions between large and small companies and between growth and value shares.

Differences in perception are perhaps indistinguishable from differences in the skill of investors or their mangers in understanding likely changes in share prices. The returns earned by different investors will differ randomly; it may be that more skilled investors can expect to earn a higher return.

3.7 Efficient market hypothesis

Readers may be struck by the difference between this discussion and a common assumption that share markets are efficient: market participants price all assets (and liabilities) such that the expected risk adjusted rate of return is equal. There are some differences as to how the adjustment for risk is made. Many, such as Malkiel (2005), will argue – in his case from that the fact that professional investment managers do not outperform the market averages - that there is good evidence of market efficiency in the USA over the past twenty years or more.

While it may be enormously difficult to outperform the market after fees, few would today argue that investment markets are always and everywhere efficient: recent research has found numerous pricing anomalies (inexplicable returns arise from such varying characteristics as market to book ratios and calendar months). Section 4 of Cantor and Sefton (2002) provides an overview of changes in the economic literature over the past two decades. In particular, one might expect markets with fewer participants, such as the Johannesburg Stock Exchange, to be less efficient.

If markets are always efficient, participants will quickly adjust prices to reflect expected risk adjusted returns, and investors using social criteria will obtain the same return as everyone else. The question of relative investment performance only has meaning if markets are not perfectly efficient in adjusting for social criteria. Unsurprisingly therefore, Sauer (1997) finds that socially responsible investors will not have to "sacrifice investment performance" on the grounds that two socially responsible indices have not underperformed indices representing the whole market over a nine year period. This may not however always be true.

4 SOCIALLY RESPONSIBLE INVESTMENT AND SHARE PRICE

4.1 Screening

In the light of the above, this section looks at the affects that social criteria might have on share prices.

Alperson (1991) traces socially responsible investment back to the 1920s with funds with a church connection prohibiting investments in "sin stocks", mainly related to companies producing liquor and tobacco or involved in gambling. This screening approach has been extended to armament manufacturers, serious polluters, companies found offensive for a range of causes and the Sharia (Muslim) avoidance of companies charging interest. This is the basis for the Johannesburg Stock Exchange's Social Responsibility Index introduced in 2004[11].

Another approach to screening is based on companies' corporate governance structure. Although the underlying reason for such screening may be economic, the connection

[11] Details can be found at http://www.jse.co.za/sri/background_criteria/background_criteria.doc

does not detract from the ethical principles underlying good governance. A positive relationship has been found in a number of studies. One of the most extensive can be found in Gill (2001), who examines and ranks 495 companies in 25 emerging markets and shows strong correlations between corporate governance and share price performance. He finds performance in terms of both share price, and of underlying profitability measures, is related to his index of corporate governance and social criteria.

The screening approach will affect the demand for a particular company's share – and may alter the return. Temporary and permanent effects can be noted:

- If the number of investors applying the screening is increasing, demand for an offensive share will decline as may its price – and vice versa. This is clearly a temporary phenomenon.
- If large numbers of investors permanently avoid a particular company, then its price may well be permanently depressed – relative to its earnings. If the company survives, it will necessarily provide higher returns. Investors that ignore the moral stigma attached to such companies may be said to be beneficiaries of moral arbitrage. The reverse affect might conceivably apply to a company regarded as a moral exemplar.

The CGF criteria do act as a screening device for shares. At first blush therefore, it would appear that the only possible long-term impact of the screening *per se* would be to reduce the performance of its investments. For the screening to increase performance, it would have to include factors that positively correlated with improved performance, and for the rest of the market to remain in ignorance of these factors.

4.2 Active shareholding

An alternative approach, referred to by Leeman (2002) as the overlay approach, is for shareholders to actively engage the companies in which they invest in order to influence management practices. The most prominent exponent of this approach is the Californian public pension system (Calpers), which has been prominent as a vocal shareholder. Although its portfolios are indexed, it takes an active interest in companies it feels are poorly managed, and Calpers (1995) reports significant improvements in management, and dramatic increases in share prices, as a result.

The CGF also engages the managements of the companies in which it invests. This arises initially with the collection of data from management and union officials, which is often

followed up by negotiations between Unity and the company. Although it has not been consistent, Unity also has a program of encouraging union members to attend AGM's with CGF proxies, and of asking questions. The interaction with the company is, unlike the Calpers interaction, intended to improve the company's social performance rather than its profits. It may however have other side effects.

These interactions may achieve improved performance of management and staff, as mentioned in 3.5.6 above. The Calpers approach may also provide motivation to management to give better returns to shareholders. They may therefore create better returns by avoiding the mistakes and losses that can be caused by bad management.

The CGF approach would be more indirect, and the results may well not always lead to better returns. Managers may, for instance, be reluctant to create jobs for a variety of reasons not related to profitability. One reason might be that the managers fail to distinguish between maximising the rate of return on capital and maximising shareholder returns.[12] Pressure to create jobs may then lead to profitable new ventures or reductions in their required rates of return. Whether the latter will lead to an increase or reduction in profit depends however on the company's cost of capital.

The impacts of all the CGF's social criteria are ambivalent in this way. They may well promote good management practices that will lead to higher profitability and reduce the risk of the company being subject to labour unrest and liability claims. On the other hand, they could be taken beyond the optimum level and lead to expenditure that does not provide a return to shareholders.

It is highly unlikely that the same affects will be observable over many years. Even if it were possible for good managers to achieve a balance between the social criteria and returns, the proportion of good and bad management in the market is bound to change, so changing relative performance.

4.3 Profits and rent

One can distinguish between legitimate profits earned by shareholders for providing capital at some risk to themselves, and monopoly rents extracted from other stakeholders

[12] Chew (1998) reports a panel discussion, where ten senior business people and academics make these assertions on pages 165-188.

because of a company's stronger market position. In Asher (1998), it is suggested that the latter is morally reprehensible.

The impact of rents on share price movements would be similar to the points raised in section 4.1 above.

- An increase in rents would see a temporary increase in share prices, and vice versa.
- To the extent that the market permanently saw rents as reprehensible, or risky, moral arbitrage could lead to higher profits from monopolist firms.

4.4 Fiduciary duties

Given that the CGF is intended as a vehicle for retirement fund investment, the management would have to conform to the fiduciary duties expected of fund trustees. A conservative view was set out as in the Megarry judgment:

> "To do the best they can for the beneficiaries ... and ... take advantage of the full range of investments authorized, rather than narrowing that range, and ... (trustees are) required to consider the need for diversification of the trust investments
>
> ... Accordingly, trustees of a pension fund could not refuse for social or political reasons to make a particular investment if to make that investment would be more beneficial to the beneficiaries of the fund." [13]

Yaron (2001) summarises the objections to this narrow view, and suggests various defences of socially desirable investment, even if it does reduce investment returns.

- The retirement fund rules may permit the use of social criteria.
- To the extent that the additional profits arise from moral arbitrage, this it would seem within the rights of the trustees to refuse such profits if their view is shared by the majority of beneficiaries.

---

[13] Cowan and others v Scargill and others [1984] 2 All ER 750, where the British Coal Board as employer prevented the trustees of their pension fund from using non-investment criteria in asset selection. In this case, the criteria were relatively eccentric. Apart from not wanting to invest in any company investing outside the UK, they also wanted to avoid energy sources competing with coal. Yaron (2001) also points out that the union appointed chairman of the trustees personally argued the case so depriving the judge of more nuanced legal debate.

- A more common view of trust law is that trustees should aim for a reasonable rate of return; it is not necessary to aim at the maximization of risk adjusted returns.

To the extent that social criteria do not reduce returns, and within these constraints if they do, there is plenty of room to apply social criteria.

4.5 Truth, good governance and creativity

Most of the reasoning above led to the expectation that socially desirable investment is likely to under perform, if anything, over the long term. If permanent sources of outperformance are to be found, one would have to have to look for entrenched behaviour by the managers of those companies that fail the social criteria. One such possibility is that the single minded pursuit of shareholder profit so alienates their other stakeholders as to reduce their ultimate profitability.

The issue can be seen as one of justice, the virtue that governs the use of power: in this case that of shareholders over other stakeholders. As Lucas (1980) puts it:

> "Justice is the bond of society... We have been pursuing the wrong political goals - productivity, efficiency, equality - and have neglected the cardinal political virtue of justice, which, together with liberty, is the condition under which I and every man can identify with society, feel at one with it, and accept its rulings as my own."

It might be put that the human rewards of working for a company that adequately cared for all its stakeholders would unleash creativity that overflowed into greater profits for shareholders. Cynically, this could be phrased that the fulfilling of social criteria enabled shareholders to better motivate staff.

In either case, the boot is on the other foot. It is not those trustees that adopt social criteria that fail in their fiduciary duty, but those who place their funds' returns at risk.

5 MEASUREMENT

5.1 Earlier studies

From the above, it is clear that there are many influences on share prices, and that the affects of social criteria are at best going to be small, and may well be transitory. Measuring relative performance may however offer some insight into investment markets, company profitability and the importance of corporate governance. From a

fiduciary perspective, it was important to have some confidence that performance was not significantly affected adversely; from a marketing perspective useful to know if performance was positively affected. Two relatively informal studies examined relative performance in the mid-nineties.

Abdulla (1993) studied the relative performance of the 21 shares approved and the 10 shares rejected by Unity, against various benchmarks over a 5 year period ending 31 December 1992. He found that the approved share outperformed the rejected shares by an average of 8% annually over the period. It would appear that the outperformance was statistically significant. [14]

Asher (1994) looked at a similar time period a little more closely by distinguishing between shares in different sectors. This was to remove the influence of sector on share price - discussed in 3.5.2 above. In 1993, the difference in the annual percentage increase in share price between accepted and rejected shares by sector is shown in table 2.

Again, the results can be argued to be statistically significant (less than one chance in 16 of outperforming in all four sectors).

Table 2: Outperformance of approved shares

Sector	**%**
Gold	63
Industrial	31
Engineering	26
Hotels	10

5.2 Present study

This study looks at the relative performance of the data of 54 companies for which evaluations on the new (post-1996) social criteria could be found in the Unity

[14] First Bowring Consulting and Actuarial Services, 1991 *Survey of Investment Performance of Investment Managers,* Unpublished reported that the standard deviation of the average returns of 17 asset managers over 5 years to December 1991 was 2,5% p.a.. Although the periods are not identical, an 8% would be statistically significant.

Corporation offices. To these were added 7 additional companies that had been rejected[15] although reports were not available.

The relative performance was calculated as the total return on each share less the average return for its sector. Share and sector prices and dividend information was obtained for the years 1998 to 2001 inclusive. The relative performance in each year was calculated, and then averaged for each share.

Only 38 companies had been listed for the whole 4 year period. The relative performance for the other companies was averaged for the period over which they had been listed. There were 45 data points for calendar 1998, 48 for 1999 and 55 from the last two years. It is not thought that this distorts the results significantly.

Table 3: Relative outperformance – 1998 to 2001

A rated – accepted	6%
B rated – pending further investigation	-1%
C rated – rejected	2%

The pattern of the earlier investigations is shown in table 3, and is confirmed.

The differences are impressive, but not significant statistically. There is a probability of almost 30% that these differences are the result of random or other unobserved affects.

5.3 Further analysis

The correlation matrix for the 54 companies with social criteria data was calculated – and shown in appendix 1. The most significant was the positive relationship that between job creation and conditions of employment. Interestingly this relationship did not appear to be related to sector: financial companies in particular appeared in both extremes. It seemed to have no significance for performance.

[15] These have been obtained from the Unity board minutes of 3 March 1997, updated by reports the quarterly "CGF Report" (until December 1998) and "Community Investment Update" (subsequently), sent to unitholders.

5.4 Empowerment

Empowerment had a positive correlation with all of the other social variables – statistically significant in 4 of the 7 cases. This is hardly surprising given that the questions in this category are heavily weighted to relationships with the unions, as are questions in each of the other categories.

Only 5 of the 30 points in this section relate to the form of black empowerment popular from 1994 to 1998, when individual black people were first given soft finance to take significant shareholdings in relatively small companies. Shares in these companies experienced something of a bubble, which was not entirely past at the beginning of the period of this investigation. The impression of the author, when he attended meetings of the Unity board, was that the evaluation of black empowerment companies was initially less critical than those of other companies.

In order to test this possibility, the five approved black empowerment companies (defined by their shareholders) were separately examined. Their average relative return over the period was very low (-14%) given the ending of the bubble. This was not statistically significant (17% chance of error), but sufficiently large to consider excluding them from the sample. Doing this increased the outperformance of the A rated companies from 6% to 9%. The probability that this outperformance is random reduces to 13%.

5.5 Regression on all data

It was assumed that relationships, if they existed, would be more or less linear. A linear regression using the relative returns as the dependent variable, and the social scores as the independent variables was therefore performed. The procedure used was to see the effect of using all the variables, and then progressively eliminating the least significant variables until the remaining variables are significant at least at the 10% level. The results using all the data proved fruitless: no variable were ever found statistically significant even at this level of significance.

5.6 Regression excluding black empowerment companies

The analysis was repeated excluding the black empowerment companies. The results did achieve a significant result for affirmative action. Appendix 2 shows details of the regression. The points allocated for affirmative action were largely from objective

comparisons of the level of black people and women in higher levels of management. This is information not widely available, so that outperformance from this source should not relate much to perceptions but to real underlying causes. Two possibilities are that high scores on this dimension relate to companies with good personnel practice, and that the promotion of blacks and women led to greater profitability for the company because their skills had previously been unrecognised.

5.7 Regression excluding other outliers

The plot of the residuals in appendix 2 shows that the model fails to predict outperformance of up to 100% (1 on the vertical axis).[16] The reasons are relatively easy to explain when one examines the data, which are set out in appendix 4. There are a few companies that have very large outperformances – of up to 103%. The reasons for this outperformance are not likely to relate to the social criteria. A particular problem with methods of linear regression is that extreme outliers of this nature can completely change the relationships between variables.

Two further progressive regressions were therefore performed. The first restricted the outperformance to a maximum of 50%, and a minimum of -30%. It produced very similar results to that of appendix 2. Corporate governance however came close to statistical significance as positively correlated with increasing performance.

The second approach was to exclude these extreme companies altogether. The results are shown in appendix 3. Surprisingly, but not unknown with regressions, affirmative action is no longer significant, but corporate governance and job creation become significant.

Unfortunately, job creation has a negative correlation with performance. Given the large number of job losses within the SA economy over the nineties, it suggests that companies that have responded most quickly to the technological changes at the root of these changes have been more profitable. If so, given the enormous social costs of the losses, it is to be fervently hoped that this situation come to a speedy end.

[16] Residuals refer to results - outperformance in this case - that the regression model fails to explain.

6 CONCLUSIONS

The social criteria adopted by the CGF have, if anything, enhanced its investment performance over the years. It appears that three of the criteria have been important. Affirmative action, which relates mainly to the number of blacks and women in senior positions, and good corporate governance were seen to have a positive effect. Job creation, under which heading is included general business expansion in SA as opposed to internationally, was unfortunately found to have a negative impact.

The discussion of the process of share price determination in this paper suggests that, with the possible exception of corporate governance which could have an ongoing effect on company profitability, social criteria are likely to have small and transitory affects on investment performance.

Acknowledgements

This paper was originally commissioned by the Centre for Working Capital, Washington. Thanks are also due to Arno Loots for gathering the data, Jacky Galpin and Bhulesh Singh for their help in interpretation, the two scrutineers for their useful comments, and Peter Kohlhagen for his sharp eyed proof reading.

REFERENCES

Abdulla I (1993) *The Analysis of the Equity Component of the Community Growth Fund* Labour Research Service, P O Box 376, Salt River

Alperson M (ed) (1991) *The Better World Investment Guide* The Council on Economic Priorities Prentice Hall, New York

Asher A (1994) *Profitability and Socially Responsible Investments - Recent Progress with Union Initiated Projects.* International Executive Seminars (unpublished)

Asher A (1998) Effective and Ethical Institutional Investment *British Actuarial Journal* 4.V: 969-1027

Calpers (1995) *Why Corporate Governance Today? A policy statement* California Public Employees' Retirement System, Sacramento

Cantor A C M & J A Sefton (2002) Economic Applications to Actuarial Work: Personal Pensions and Future Rates of Return *British Actuarial Journal* 8.1: 27-74

Chew D H (ed) (1998) *The New Corporate Finance* (2nd edition) Irwin/McGraw-Hill

King Report (1994) *The King report on corporate governance* King Committee on Corporate Governance, The Institute of Directors of Southern Africa, Johannesburg

Lindenberg E & S Ross (1981) Tobin's q Ratio and Industrial Organization *Journal of Business* 54: 1-32

Leeman M G (2002) *Corporate Citizenship, Fiduciary Responsibility and Shareholder Activism* (http://www.assa.org.za/downloads/events/papers2002/assa%20paper%202002.doc)

Lucas J R (1980) *On Justice* Clarendon Press, Oxford

Malkiel B G (2005) Reflections on the Efficient Market Hypothesis: 30 years later *Financial Review* 40.1: 1-9

Miller M H & F Modigliani (1961) Dividend Policy, Growth and the Valuation of Shares *The Journal of Business* 34: 411-413

Sauer D A (1971) The Impact of Social-Responsibility Screens on Investment Performance: Evidence from the Domini 400 Social Index and Domini Equity Mutual Fund *Review of Financial Economics* 6.2: 137-49

Smith R S (1990) Factors Affecting Saving, Policy Tools, and Tax Reform: A review *IMF Staff Papers* 37: 1-70

Smithers A & S Wright (2000) *Valuing Wall Street: Protecting Wealth in Turbulent Markets*. New York: McGraw-Hill

Thomson R J (1995) Stochastic Investment Modelling: The case of South Africa *British Actuarial Journal* 2: 765- 802

Wiehahn Commission (1979) *The Challenge of the New Industrial Relations Dispensation in South Africa* Government Printer, Pretoria

Wilkie A D (1995) More on a Stochastic Asset Model for Actuarial Use *British Actuarial Journal* V.5: 777-964

Yaron G (2001) *Answers to Common Questions about Pension Trustee Fiduciary Duties and Socially Responsible Institutional Investment* Canadian Shareholder Association for Research and Education (SHARE) (http://www.share.ca/files/pdfs/Trustee%20_digest.pdf)

APPENDIX 1: CORRELATIONS BETWEEN VARIABLES

	Average Performance	*JOB CREATION*	*TRAINING*	*EMPOWERMENT*	*AFFIRMATIVE ACTION*	*CONDITIONS OF EMPLOYMENT*	*ENVIRONMENTAL PRACTICE*	*HEALTH & SAFETY*	*CORPORATE GOVERNANCE*
Average	1,00								
JOB CREATION	0,03	1,00							
TRAINING	0,13	0,21	1,00						
EMPOWERMENT	0,13	0,22	**0,38**	1,00					
AFFIRMATIVE ACTION	**0,20**	-0,19	-0,07	**0,29**	1,00				
CONDITIONS OF EMPLOYMENT	0,03	**0,71**	**0,44**	**0,37**	0,06	1,00			
ENVIRONMENTAL PRACTICE	-0,12	0,05	**0,30**	0,27	-0,22	0,25	1,00		
HEALTH & SAFETY	-0,08	0,05	-0,17	**0,44**	**0,32**	0,15	**0,27**	1,00	
CORPORATE GOVERNANCE	**0,21**	0,04	-0,03	0,34	0,20	-0,13	0,09	0,21	1,00

Significant values are shown in bold.

APPENDIX 2: REGRESSION EXCLUDING BLACK EMPOWERMENT COMPANIES

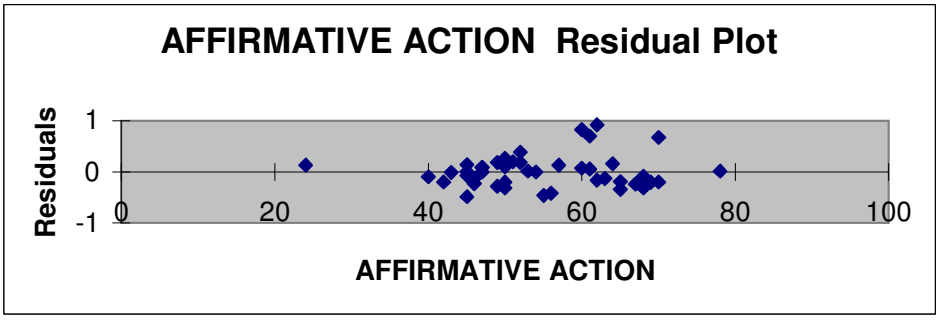
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,286976
R Square	0,082355
Adjusted R Square	0,062407
Standard Error	0,312728
Observations	48

ANOVA

	df	*SS*	*MS*	*F*
Regression	1	0,403747	0,403747	4,128342
Residual	46	4,498743	0,097799	
Total	47	4,90249		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-0,42576	0,243378	-1,74938	0,086896
AFFIRMATIVE ACTION	0,008775	0,004319	2,031832	0,04797



APPENDIX 3: REGRESSION EXCLUDING BLACK EMPOWERMENT COMPANIES AND OUTLIERS

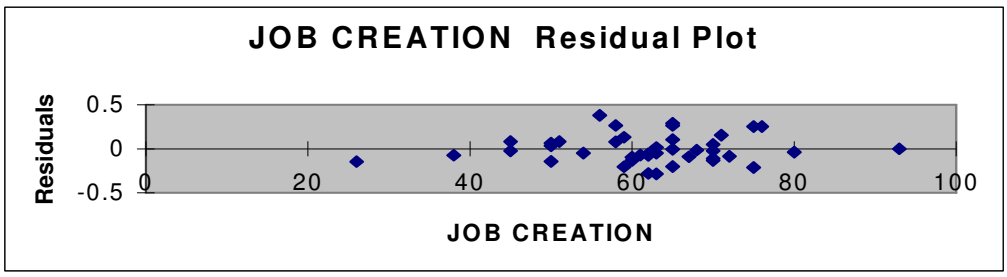
SUMMARY OUTPUT

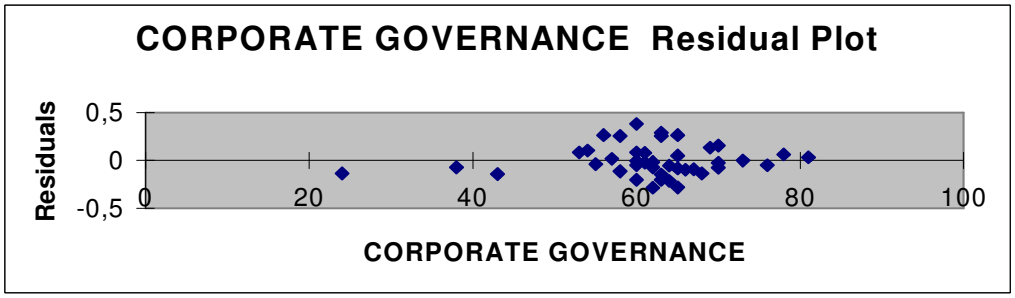
Regression Statistics	
Multiple R	0,406955
R Square	0,165612
Adjusted R Square	0,118577
Standard Error	0,161076
Observations	41

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0,20084	0,10042	3,870425	0,029519
Residual	39	1,011875	0,025946		
Total	41	1,212716			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	0	N/A	N/A	N/A
JOB CREATION	-0,00496	0,001837	-2,69749	0,010267
CORPORATE GOVERNANCE	0,005214	0,00185	2,819276	0,007525





APPENDIX 4: BASIC DATA

COMPANY	Relative return	JOB CREATION	TRAINING	EMPOWERMENT	AFFIRMATIVE ACTION	EMPLOYMENT CONDITIONS	ENVIRONMENT PRACTICE	HEALTH & SAFETY	CORPORATE GOVERN.	AVERAGE	STATUS
ADvTECH GROUP LTD	-3%	68	67	65	45	67	66	67	62	63	A
AFRICAN MEDIA	-32%	10	43	50	76	31	56	65	77	51	B
ALACRITY FIN SERVICES	-21%	65	67	65	47	63	61	52	65	61	A
ANGLOPLAT	0%										C
ANGLOVAAL INDUSTRIES	80%	56	61	63	61	66	67	59	64	62	A
APLITEC: NET 1 APP TECH	103%	60	65	60	62	64	65	50	66	61	A
AVIS SOUTH AFRICA	-25%	62	60	61	46	58	64	53	65	59	B
BARLO WORLD	22%	50	78	86	51	72	93	59	78	71	A+
BARPLATS	19%	75	53	55	49	67	61	52	58	59	B
BIDVEST	7%	65	67	65	47	66	61	53	54	60	B
BoE	-7%	67	64	68	67	66	61	65	67	66	A
CARSON HOLDINGS	-19%	76	65	76	67	74	85	78	65	73	A
CASHBUILD	92%	70	70	76	60	68	68	70	70	69	A
CERAMIC INDUSTRIES	20%	59	63	63	50	67	59	47	69	60	B
CLIENTELE LIFE	-1%	65	67	57	47	65	67	65	60	62	A
DEL MONTE ROYAL	30%	65	67	71	64	61	62	56	63	64	A
DISCOVERY HOLDINGS	-16%	70	68	67	46	68	65	60	58	63	A
ENERGY AFRICA	86%	63	75	69	70	72	43	53	65	64	A
ENVIROSERV HDGS	-35%	50	54	66	56	60	67	64	72	61	A
FEDSURE	-40%	60	60	71	55	67	63	55	63	62	A
FIRSTRAND	-2%	62	67	67	69	69	61	68	70	67	A
FRASER ALEXANDER	-28%										C
GENSEC	-8%	93	72	43	24	70	75	33	73	60	A
GREENWICH GROUP	-46%	67	65	55	50	70	60	58	57	60	A
HIGHVELD	-10%	72	72	55	45	64	65	40	65	59	B
ILLOVO	-5%	60	66	74	65	67	65	60	66	65	A
ISCOR	11%	51	59	57	45	57	60	50	53	54	C

L A RETAIL STORES	-8%	65	60	67	59	59	60	51	59	60	A
LIBERTY GROUP	-13%	70	72	74	68	75	60	60	68	68	A
LIBERTY LIFE	-18%	59	65	67	50	60	60	51	60	59	B
MASSMART	27%	58	56	66	78	73	61	59	56	63	A
MED CLINIC	0%										C
MERCANTILE LISBON BANK	-52%	57	62	58	45	67	68	60	60	60	B
METRO CASH & CARRY	17%	71	72	59	61	69	60	60	70	65	A
M-NET	12%										C
MUTUAL&FEDERAL INS	-3%	62	65	55	69	65	62	64	64	63	A
NANDOS	8%										C
NEW CLICKS HOLDINGS	21%	50	45	45	52	50	54	55	81	54	B
NORTHAM PLATINUM	4%	70	61	64	54	60	60	68	65	63	A
NU-WORLD HOLDINGS	-31%	60	78	55	50	78	62	45	24	57	B
OCEANA FISHING	22%	78	61	83	35	64	75	64	73	67	A
OLD MUTUAL GROUP	5%	26	74	55	53	41	75	45	63	54	B
OMNIA HOLDINGS	-17%	50	45	46	40	55	61	60	43	50	B
POLIFIN	28%	65	65	67	50	60	62	50	65	61	A
PPC	-5%	61	72	70	62	72	96	73	62	71	A+
PRISM HOLDINGS	-5%	63	60	66	68	66	61	52	60	62	A
PRIVEST GROUP	-25%	75	65	67	42	70	65	60	64	64	A
RAINBOW	42%	56	59	60	52	57	60	55	60	57	B
RELYANT	-48%										C
RENTSURE HOLDINGS	-27%	63	60	72	49	65	69	70	62	64	A
REUNERT	49%										C
REX TRUEFORM	-6%	38	60	38	43	57	75	45	38	49	C
SAAMBOU	-19%	65	65	67	65	67	61	53	63	63	A
SAGE LIFE	-15%	80	58	62	68	75	64	55	55	65	A
SANLAM	7%	45	64	62	47	63	67	60	61	59	B
SANTAM	11%	58	59	65	50	58	60	51	61	58	B
SEARDEL INV COR	0%	63	56	69	63	63	61	59	57	61	A
STANDARD BANK GROUP	8%	54	71	74	68	75	70	47	76	67	A
SUN INTERNATIONAL (SISA)	17%	45	72	60	60	47	50	41	60	54	B
THE HOUSE OF BUSBY	-1%	70	55	68	70	60	60	60	70	64	A
WETHERLYS INVESTMENT	21%	76	48	43	57	70	60	52	63	59	B

CHAPTER 7 - MEAN REVERSION IN INVESTMENT MARKETS: THE IMPLICATIONS FOR INVESTORS AND REGULATORS

Forthcoming in the Australian Actuarial Journal

1 INTRODUCTION

This paper attempts to outline what is known about mean reversion in investment markets, and to explore the implications for the regulation of money purchase superannuation funds in Australia. Australian superannuation fund members now have, in the main, choice of both fund and investment mix. As members of accumulation funds, their retirement income depends significantly on how they exercise their investment choices.

While there is a plethora of advice available on investment choice, there is limited agreement - even between experts - on many basic facts. There are, however, a few clear principles that bear repetition and expansion: diversification, expense management, fundamental analysis and - in the light of the mean reversion discussed at length in this paper – avoiding panic. These principles are not unknown, and are captured in some popular books such Bernstein (2002), but they are not universally accepted and still need explanation and defence.

In section two this paper quickly summarises the debate between efficient market theory and behavioural finance. Section three is a more detailed consideration of research into mean reversion, the results of which support a time varying equity risk premium and non-linear reversion to a mean range rather than a point. Together they make the case for fundamental analysis, which is discussed in section four as part of a rehearsal of investment theory. The section also discusses recent theoretical and market developments that take members' age into account. These developments are found to fall short in their consideration of housing, expenses over the life-span and mean reversion. Section five considers regulation and recommends some changes.

In reviewing the literature, an attempt has been made to focus on more influential papers, but the field is too broad to have confidence that everything has been covered. Writing

ten years ago, Campbell *et al* (1997) mention the impossibility of covering all the papers that had then investigated the efficiency of markets, let alone the other subjects covered in this paper.

2 HOW INVESTMENT MARKETS WORK

A full discussion of investment market functioning must include the theories of market efficiency and behavioural finance, with perhaps a mention of econo-physics and evolutionary finance as interesting alternative approaches. Such discussion can, however, be made difficult by dogmatic arguments. Arnott (2004), in a Financial Analysts Journal editorial, makes the point that: "much of our industry works on dogma. And much of that dogma is based on sound economic theory. But theory is only theory: it tells us what should happen, not necessarily what does happen."

2.1 Market efficiency

Belief in market efficiency is one piece of dogma that may be a source of poor advice.

2.1.1 Markets are not perfectly efficient

It is clear that even the large US share market is not perfectly efficient. This is recognised by the wiser proponents of the theory, such as Ball (1995), who merely asserts that it is a reasonable *a priori* assumption that deep and liquid markets with low barriers to entry should be efficient. He mentions a number of anomalies that have not been explained. Of particular interest are:

- Shorter term under-reaction to news that rewards momentum strategies.
- Subsequent price over-reactions and then longer term corrections that reward contrarian investment strategies.
- The weak correlation of observed betas[124] - and the stronger correlation of smaller capitalization, lower market to book ratios, and higher dividend yields - with higher returns on individual shares.

[124] Correlation of individual shares with stock market indices that should in terms of the Capital Asset Pricing Model explain returns.

- Various seasonal patterns of hourly, daily, monthly and quarterly prices. He does not mention the even longer four year cycle, which Nickles (2004) shows to coincide with US presidential elections since the early fifties.

2.1.2 The impossibility of perfect efficiency

Occasional inefficiencies in markets are partly explained in a widely referenced article by Grossman and Stiglitz (1980) who prove the impossibility of completely efficient markets. There must be enough inefficiency in market prices to make it worthwhile for traders and arbitrageurs to operate. Anomalies in relative value occur if there are insufficient numbers of active investors. The profit these investors make keeps them in business, but there is no assurance that they will be uniformly successful.

One implication is that anomalies are not likely to persist for long. They will emerge, become subject to exploitation by traders and arbitrageurs and then may disappear only to reappear some time later. Another implication is that at least some people must believe that it is possible to make trading or arbitrage profits for markets to be efficient. As a dogma, the efficiency of markets is self-defeating. The more it is believed, the more money will be lost.

In spite of this, market efficiency is widely taught in university courses and used as a useful, if not always challenged, assumption on which to base models for asset prices. Welch (2000) asked a sample of over 100 financial and economics professors whether they believed in the efficiency of markets or not. Over 90% believed that equity markets were efficient and followed random walks in the short run. Fewer than half thought that markets showed negative auto-correlation over a three to five year horizon and that the value of the equity risk premium changed over time. Most appeared unworried by the anomalies mentioned above.

2.1.3 But one cannot beat the market

A lack of efficiency means that it is possible to profit if there is an element of luck (arising from the absence of traders and arbitrageurs) to supplement good judgment. Whether the investment performance of managers persists over time has been intensively researched. Carhart's (1997) findings remain largely unchallenged. There is, at best, weak evidence of persistent superior performance arising from funds that seem able to quickly capture new evidence in the market, or better interpret current information. There is however much stronger evidence of persistent underperformance. Part of the

underperformance can be explained by excessive trading; the rest must come from the purchase or sale of investments at uneconomic prices. One possibility is that managers are trading at an uneconomic price for some personal benefit at the expense of the members. Another is that they have failed to consider the fundamental value of the assets that they are trading.

There is also evidence that recognisable groups of individuals perform significantly worse than average. Barber and Odean (2000 and 2001), for instance, describe a group of individual investors that lose by trading too often: they buy shares that subsequently do worse than average and sell shares that do better, and also waste significant amounts on brokerage. The problem is more acute for males and the unmarried.

2.2 Behavioural finance and bounded rationality

If markets were efficient, they would have to be dominated by rational investors (or sellers of assets). Behavioural finance is based on the view that investors are not only ignorant about the future but not entirely rational in their processing of relevant information, or at least as Simon (1983) puts it, they use bounded rationality.

2.2.1 Credulity

Daniel *et al* (2002) provide a detailed review of behavioural theories of pricing. They suggest that investor credulity lies at the root of the anomalies in efficient markets. They further explain credulity mainly in terms of "limited attention" and over-optimism, both widely attested in the psychological literature and every day experience. (In more recent work on over-optimism, however, Hoelzl and Rustichini (2005) find that it is limited to tasks that are perceived as relatively easy. People can become excessively pessimistic when faced with more difficult tasks.)

The limited attention and processing power of investors is used by Daniel *et al* to explain the under-reaction of share prices to news, the persistence of arbitrage opportunities in markets with fewer participants and the greater returns from smaller capitalization stocks with a limited following of analysts. It can also explain herd behaviour if investors place too much weight on information that comes from other market participants, and not enough on their other sources of information.

Over-optimism could explain some of the excesses of price momentum before it reverts, and may explain some of the relative failure of growth and high market to book

strategies. Companies in these categories may well perform better than others, but relative underperformance arises because they do not do as well as expected, perhaps by management and analysts. Daniel *et al* also report on research that suggests that investors are more credulous about intangible information than accounting data. It may be that these anomalies are also better explained by limited attention than over-optimism.

Cooper *et al* (2004) confirm overreaction theories of short-run momentum and long-run reversal. They also find that momentum profits depend on the state of the market and that up-market momentum reverses in the long-run, which is consistent with over-optimism.

2.2.2 Under-reaction

Under-reaction to announcements and subsequent momentum over a period of months represents perhaps the most potent evidence of inefficiency in markets: prices do not reflect all publicly available information. Ball (1995) argues that the over-reaction is smaller than might be expected: markets are almost efficient. Other attempts have been made to explain the anomaly. Lewellen (2002), for instance, suggests that it can be explained by the correlation structure of markets, but the response by Chen and Hong (2002) shows that his arguments are not valid.

Over-optimism may have a role in this under-reaction. Schwartz and Steil (2002) report that investment managers often appear to think that their competitive advantage lies more in their alternative interpretations of the same information, rather than quickly absorbing new information. Anecdotally, this view arises from the mistaken belief that the market is always efficient in processing information. It may lead managers to neglect the importance of quickly processing new information.

2.2.3 Defences of rationality

In the face of the behavioural critique, there have been a number of attempts to explain the anomalies and defend investor rationality, but none appear - at this point - to be entirely successful. Brav and Heaton (2002) suggest that rather than see market anomalies as irrational, they could be explained in terms of the rational Bayesian adaptation of new information into *a priori* knowledge. Aloysius (2005) uses similar arguments to distinguish stochastic from model risks and to suggest that "ambiguity aversion" explains the equity risk premium puzzle. Ball (1995) suggests that the alternative behavioural theories are too disparate and the patterns too transient to provide

an alternative. In one sense he appears to be right: behavioural explanations do find patterns in the statistical noise in sequences of prices, but they do not seem sufficiently predictable to exploit.

2.2.4 Relevance of irrationality

That brings the discussion to the relevance of behavioural finance to investment practice.

It does not justify communal or government setting of prices. Daniel *et al* (2002) note that the proponents of the efficient market theory make their arguments as a defence against government intervention. Daniel *et al* make the point that this is a brittle and unnecessary defence line. If market participants can be irrational, so can government officials. Government intervention in the mechanism of price formation is not justified by market irrationality.

The irrationality of some market participants represents both a threat and an opportunity to rational investors. Irrational noise interferes with the pricing mechanism, and means that rational investors – if forced to liquidate assets at an unfavourable time – will lose money. This applies even more strongly to rational investment managers who may lose mandates to less rational competitors when prices are out of line with fundamental value. To the extent that rational investors and investment managers are able to buy when assets are irrationally cheap and sell when they are irrationally dear, there are profits to be made.

One of the major and avoidable risks to uninformed superannuation members is that they will credulously accept the recommendation of over-optimistic managers and irrational analysts and advisors, and sell cheap and buy dear.

2.2.5 Other theories

Hens and Schenk-Hoppé (2005) describe the theory of evolutionary finance, which considers the evolution of markets and the adaptation and survival of investment strategies in a manner analogous to the adaptation and survival of biological species. Empirical observation includes the strategies of investors, and not just the pricing patterns that make up the stuff of technical research and underlie the theories of efficient markets and behavioural finance. Econo-physics, as for instance described in Ball (2004), similarly describe the functions of markets rather than prices.

Both do offer useful insights into the working of markets and the formation of prices.

3 MEAN REVERSION

While the insights of the efficient market hypothesis and behavioural finance are frequently incorporated into investment advice, the same does not seem to be true of what is known about mean reversion.

This section therefore looks in some detail at the results of academic research into whether prices are a random walk or can be expected to revert in some way to some mean or fundamental value. Formula 1 states this more rigorously for Δx_t the change in price to time t.

$$\Delta x_t = \mu_0 + \mu_1 t + (\alpha - 1)x_{t-1} + \sum_{j=2}^{p} \beta_j \Delta x_{t-j+1} + \varepsilon_t \qquad (1)$$

A trend stationary series of price changes would have non-zero values for μ_0 and μ_1; while a covariance stationary series would have non-zero values for some of the β_j's. A value of α that differs from one means that the series is not stationary at all, but is mean reverting. Mean reversion thus means that the expected change in the price depends on the current price and can be expected to revert in some way to a longer term mean value.

It is concluded below that there is significant theoretical and empirical evidence of non-linear mean reversion from market extremes. The reversion is not likely to be to a point, but to some level within a range. Discussion begins with currency markets as they provide the necessary background to discuss mean reversion in other markets.

3.1 Currency markets

Identical commodities should sell at identical prices in different markets – after allowing for costs of transport. Rogoff (1996), in a valuable literature review, outlines a puzzle that – in the short run at least - they do not. In that paper, he suggests that while economists all believe that purchasing power parity (PPP) should hold in some form or other, it had taken some 400 years of research to find persuasive data to demonstrate that prices do tend to revert to the expected mean. The pace of reversion, which he set at 15% pa, is however sufficiently slow to represent a puzzle – even if "one confines attention to relatively homogenous classes of highly traded goods".

There are obvious explanations. In the first place there are frictions: not just transport and other trading costs, but tariffs and other barriers to trade. There is also the likelihood that traders and arbitrageurs may not be able, or find it worthwhile, to respond to price differentials until they become relatively large. The idea can be expressed as a "band of inaction" around PPP, with mean reversion applying outside this band.

There are a variety of models that have been tried to fit this pattern. In one of the more elegant Taylor *et al* (2001), confirmed and adapted by Paya *et al* (2003), use an exponential smooth transition auto-regression (ESTAR) model. A one period simplified model of their approach is given by formula 2:

$$y_t = \mu + [1 - e^{-\alpha(y_{t-1}-\mu)^2}][(y_{t-1} - \mu)] + \varepsilon_t \qquad (2)$$

where μ represents a productivity adjusted purchasing power parity and ε_t represents an error term. The exponential weight is limited to the range [0,1] to prevent it from blowing out. Their formula also allowed for a variety of lags over longer periods; the resultant fit for a range of European currencies and the US dollar over 22 years to May 2001 suggests faster mean reversion than reported by Rogoff, especially for the larger shocks away from PPP. For a 10% shock, half lives (representing the time taken to move half way to a long term equilibrium) vary from 13 to 41 months; for a 40% shock the half lives vary from 1 to 17 months. The authors explain the differences in the parameters of the formula for different countries by differences in the ease of arbitrage. This, in turn, depends on geography and institutional structures (which would be taken to include culture).

Amongst this literature, Taylor (2001) provides a useful analysis of two pitfalls in measuring mean reversion. The first is the problem of time averaging. He finds, for instance, that measuring currency rates or returns at monthly intervals will overstate the half life by some 50% if the true half-life is greater than a month, and by more if it is less. The second issue is that of non-linearity, which leads to increasing errors as the band of inaction widens. He uses what he calls the simplest, non-linear model where the reversion is linear outside the band of inaction. This "threshold autoregression" (TAR) requires estimates of the thresholds as well as of the parameters of the reversion.

There is another independent strand of research into mean reversion and PPP sparked off by O'Connell (1998) who argued that the long term linear mean reversion was

"overvalued" because of cross-sectional dependence or collinearity. High inflation had simultaneously affected most of the countries in the samples and thus created spurious relationships. This strand ignores what would appear to be the superiority of the non-linear tests, but the non-linear researchers, in turn, appear to ignore these other influences that should be considered in the modelling process.

After referring to the non-linear developments, Taylor and Taylor (2004) say that "the idea of long-run PPP now enjoys perhaps its strongest support in more than thirty years, a distinct reversion in economic thought." It certainly appears that recent research has resolved the puzzle of PPP. Small deviations from PPP present little in the way of profitable arbitrage opportunities, and can remain for some time. Large deviations from the productivity weighted PPP do, however, produce a speedy and measurable mean reversion.

3.2 Interest rates

Unlike PPP for currencies, there is no natural theoretical level to which interest rates can be expected to revert. There are however good theoretical reasons to expect the real rate of interest to be within a band of some -2% to perhaps 10%. Below rates of -2%, there will be opportunities for arbitrage as is it possible to hold cash if there is no inflation, or non-perishable goods if there is. If real interest rates were 10% in a developed economy (where the balance sheets of the banks exceed GDP and there are numerous other loans and rental on property to pay), total real interest and rent payments could account for 20% of GDP. If half the people in a country are lenders and renters and the other half rentiers (who receive interest and rent), the borrowers will be paying an average of 40% of their gross incomes in interest and rent. While perhaps not impossible, it is difficult to envisage such a situation being sustainable.

In the vast literature on interest rate modelling (most of which covers nominal interest rates), two strands can be examined briefly. One, such as Mankiw and Miron (1986) finds that there is insufficient evidence to reject the hypothesis of a random walk. On the other hand, many interest rate models, such as Vasicek (1977) and Cox, Ingersol and Ross (1985) assumed linear reversion to a mean. These models did not initially fit observed data well, but have been developed to fit more closely.

Some subsequent work appears to have shown that the relationship is not linear. Jones (2003) shows that attempts to measure the parameters precisely are largely dependent on

the hypothesis being tested. He provides an interesting illustration of how results can differ widely depending on alternative Bayesian prior distributions. The problem arises from the paucity of data at extremes. One can believe that each extreme event is likely to be unique in its causes and the way in which markets return to normality. If so, it provides no evidence of future mean reversion.

Kapetanios *et al* (2003) apply an ESTAR model to interest rates in the major OECD economies with success. Nominal interest rates too, therefore, appear to be mean reverting as they approach extreme values. There is less research on the real interest rate as determined by deducting inflationary expectations from nominal rates, but Lai (2004) finds evidence of mean reversion in the US.

3.2.1 Term structure

Seo (2003) finds a non-linear mean reversion in the term structure of interest rates using a TAR approach. He ascribes these plausibly to transaction costs, which prevent investors from realizing arbitrage opportunities. His adjustment coefficients which describe the mean reversion are regime-dependent[125], which makes the model more complex.

Chan and Cheung (2005) produce a TAR model of Australian long and short term interest rates with three regimes based on thresholds of the difference between the two rates. The model may well be over-parameterized, but has mean reverting properties.

3.2.2 Credit risks

Jarrow and Turnbull (2000) bring together default and market risks to model credit risks. They refer to the need, arising from banking regulation, to capture the risks of default, downgrade and spread. The first two are clearly related and are clearly cyclical. Collin-Dufresne and Goldstein (2001) confirm that firms adjust their leverage over time to find an optimal level with the result that credit spreads are mean reverting. Prigent *et al* (2001) find evidence of non-linear mean reversion of the indices, as does Bhanot (2005), who confirms that survival bias does not change this result.

---

[125] The parameters of the model switch between different regimes (or states of the world) in a random fashion. Regime shifting models often fit the data well because of their fatter tails and ability to reflect clumps of increased volatility.

Some idea of the consensus view can be obtained by considering pro-cyclicality. This is the problem that backward looking, risk-based regulatory capital will rise after defaults and spreads increase, and then be too high for the improvement to come. Allen and Saunders (2004) survey the literature and say that they see a consensus on the basic idea of addressing pro-cyclicality, but little agreement on model and policy specifics.

3.3 Inflation

As with real interest rates, there is no obvious long term level for inflation and inflationary expectations, except perhaps the targets set from time to time by central banks. Given that explicit inflation targeting is relatively recent, there is not enough data to confidently test this possibility. There are reasons for believing that the inflation rate cannot fall much below zero – real interest rates would rise too high, and the redistribution between different groups of people become untenable. Inflation rates can however explode, as has been experienced too frequently.

As with the other time series investigated, there are many independent strands of research on inflation. Much of the earlier economic work does not consider mean reversion at all, and where it does, only considers linear characterizations, which from the earlier discussion, are unlikely to be satisfactory. The few reported attempts that have tried non-linear models have been more successful. Baillie *et al* (1996) analyse monthly inflation for 10 different countries, and find strong evidence of long memory with mean reverting behaviour for all countries except Japan. Arghyrou *et al* (2004) try a variety of linear and non-linear models on UK inflation rates over the last third of the previous century. That they have attempted a number of models suggests that there is data snooping (so finding spurious relationships), and the large number of parameters that they use suggests some over-parameterization. In spite of this, their rejection the null hypothesis of a random walk for inflation appears plausible.

3.4 Equity markets

This brings us to the interesting question of whether there is a band of inaction for equity markets. As with PPP, there are a number of ways of modelling the "true" value of a share, with prices deviating from this value to the extent that the assumptions in the particular model do not hold.

One can make the assumption that all markets are competitive and free of trading or arbitrage opportunities. If the markets for every company's products and its assets and liabilities were perfectly efficient, the mean value of a share would be the book value of the net tangible and intangible assets (adjusted appropriately for inflation). If the price were higher, an arbitrageur could raise capital for a new company of the same sort, buy the same assets and then sell the shares for a premium. If the price were lower, then the companies would not invest in replacement assets and would pay their entire cash flows back to their shareholders until the productive capacity in the market was sufficiently reduced.

One can relax the assumption that the underlying markets are efficient, and consider the position if only the investment markets are efficient, and can produce an unbiased estimate of companies' earnings. Under these circumstances, every company would generate earnings at a risk premium to the risk free interest rate of suitable duration. The duration would depend on the period for which the earnings were predictable. If one makes the additional assumptions required for the capital asset pricing model[126], then the margin required by investors over the risk free rate can be determined. If one can also ignore the impact of depreciation and the possible need for investment to maintain the earning power of the assets, a fairly accurate estimate of the value of the company can be obtained by discounting the earnings at a risk-adjusted rate.

If alternatively, we assume that retained earnings are used to increase dividends in real terms at some predictable rate, the fair value of shares can be determined by the dividend yield. One obvious problem is that optimal dividend policy changes with the tax regime, which provides a major obstacle to the assumption that the long term payout ratio has been constant.

3.4.1 Empirical relationships

There is clear evidence of mean reversion in price to earnings ratios and market to book ratios (Tobin's q) as described, for instance, in Harney and Tower (2003). They used a value for the market to book ratio that adjusted for inflation. Cavaglia and Moroz (2002) find that industry share prices tend to converge to long term dividend yield, interest rates

[126] These include a stable correlation structure for all shares and sufficient market participants who can invest or borrow at the risk free rate.

and earnings levels. Cutler *et al* (1991) find significant reversions to the dividend yield in Australia, Canada, the UK and some US periods, but not in the other countries they model. Thomson (1992) finds significant mean reversion of dividend yields in South Africa.

There is also evidence that share prices in different countries and sectors revert to their average relative to the global market. Balvers *et al* (2000) find reversion from 18 countries to a portfolio representing the average value for the 18 countries when testing annual data. Stotz (2004) finds evidence of mean reversion in 50 large European shares using a combination of profits and book values as explanatory variables.

Futures contracts are relatively easily priced using the spot price of the underlying asset and assumptions about interest rates and dividends, or storage costs. Even in this market, however, there are costs to arbitrage, and Monoyios and Sarno (2002) find an ESTAR mean reversion to the theoretical price that produces a superior fit to a linear mean reversion to underlying stock prices.

There is research, such as Narayan and Smyth (2005), that rejects mean reversion, but this might be accused of being naïve in that the authors only look for reversion to an absolute value, and test on daily data which may well hide longer term reversion.

The conclusion is that there appears to be mean reversion from the extreme of equity market values based on earnings yield, dividend yields and market to book ratios.

3.4.2 Bubbles

The corollary of mean reversion in the extremes is that prices occasionally reach extremes. Such extremes would be irrational bubbles if the expected return on risky assets fell below the risk free rate, although they would only ensure the presence of arbitrage profits if there were a way of ensuring that no money is lost.

Bubbles can be explained by herd behaviour as for instance in Hirshleifer (2001) and Hirshleifer and Teoh (2003). Herd behaviour may arise from irrational psychological biases that involve placing excessive reliance on the opinions of others, or from rational incentives not to take a contrarian view. Kim and Nofsinger (2005) provide evidence of the latter applying to institutional investors. The theory of rational bubbles is famously illustrated by Keynes, who described investment as a beauty contest where the prize goes to the person who best guesses the preference of the majority.

Allen and Gale (2000) look for other explanations and suggest that the bubbles in Japan and in Scandinavian countries during the eighties and nineties arose from excessively generous credit that exposed the banks to moral hazard by investment opportunists. If this were true it would be the banks rather than the investors who were irrational, unless the value of government support exceeded their credit losses, in which case government officials supplied the irrationality. On the other hand, the rupture of other bubbles may reflect deeper causes as suggested by Kindleberger (1988) for the crash of 1929.

There are other reasons why share prices may be at a level unjustifiable in terms of their intrinsic value. Some arise because some investors will not sell shares that they know to be over-priced. Owners who are involved in the management, or who have family and other ties to a company, may chose not to sell over-priced shares, while other investors will not sell for tax reasons. Passive investors may also be limited, by prior decision, to buy and hold shares in their index. Woolley and Bird (2003) criticise passive investors for buying shares that were clearly overpriced during the 2000 tech bubble. They cite evidence that share prices rise if they are included in a share index favoured by passive investors as evidence of market inefficiency. They do not raise the problem that this effect is aggravated by share price indices that are based on capitalization rather than tradability, although this problem has been addressed in the last few years.

Restraints on short selling can also explain bubbles. Short selling is particularly risky, as a "bear squeeze" can require a purchase at inflated prices when it comes to delivery. Perhaps the most colourful of these in recent years was the manipulation of the silver price in the late eighties. The Bunker Hunt brothers were the most famous participants in a consortium that drove the silver price up 650% in one year as short sellers attempted to cover themselves. The gold price bubble of the same years may well have been an irrational parallel. Williams (1995) describes aspects of the legal case against the Hunt brothers and their collaborators; Abolafia and Kilduff (1988) discuss the regulatory manipulation that broke the bear squeeze by limiting the number of contracts that an individual could hold.

Gilchrist *et al* (2005) develop a model of bubbles, assuming restraints on short sales and the incentive of companies to issue more shares when prices are high, and to invest the money in expanding their operations. They investigate the 2000 tech bubble, and find that companies behaved as might be expected from their model. They did not find excessive and wasteful investments took place because companies restricted their share

offerings in order to maintain the price. Woolley and Bird (2003), however, suggest that European communication companies did make wasteful investments; a large part was overpaying governments for licences.

While bubbles can be inflated by over-optimism, panics are also not impossible if there are times when rational investors are unable to buy or where portfolio insurance or regulatory capital requirements require sales even though markets are already depressed. The extent of the 1987 share market crash can be explained by programmed selling. Barker (1999) mentions other incidents of the forced sale of equities, and purchases of fixed interest stocks, as a result of capital requirements.

Bubbles and panics can therefore occur in investment markets.

3.4.3 Comparative performance

If markets do mean revert, such relationships should be exploitable. Malkiel (2004) looks at the success of technical and fundamental models to predict returns. He uses dividend yields, price earnings ratios and a model based on interest rates and Tobin's q, and finds evidence of out-performance.

These results are not disproved by findings, such as by Becker *et al* (1999), that mutual fund managers on the whole are unable to benefit from market timing. The average investor must do as well as the market, and these findings merely show that the mutual fund managers in the study are no different from the average investor.

The performance test for mean reversion would be to see whether those investors (and raisers of capital) that attempt to read market levels do better than those that do not. The performance of asset allocation hedge funds could provide such a test. Capocci and Hubner (2004) confirm that earlier analyses appear to have overstated the success of hedge funds, but they do find that one quarter of US hedge funds managers appear to be successful in finding arbitrage opportunities. Amongst these are funds with a market timing strategy. Do *et al* (2005), in contrast, find no evidence of over-performance in Australian hedge fund managers. In summary therefore, the formal evidence that there are opportunities for identifiable groups of investors to exploit market timing is weak.

There is however strong evidence that companies raise more money when the market is high - as found by Gilchrist *et al* (2005). This produces something of a puzzle in that they appear to have more knowledge about the level of the market than institutional investors.

The fact that institutional investors face incentives to herd does not mean that they do not have an incentive to attempt to outperform by small – and measurable - margins.

Outside of the academic research, there is perhaps no puzzle: everyone knows that fundamental investors such as Warren Buffett outperform the market. The problem is that their services are not generally available to all at a fixed price[127] or they are not easy to identify. There has apparently been no academic attempt to measure whether Buffett's success has been due to chance or not, which perhaps creates a puzzle in itself.

The question therefore appears to remain open until someone finds a way of identifying successful managers.

3.4.4 Property

Most of the research in this area has been on US listed property trusts, which would be expected to display some equity type characteristics. Stevenson (2002) is the most cited article, and finds that statistics from some markets reject the hypothesis of stationarity. He concludes that mean reversion is slow. He does not however refer to the paper of Okunev and Wilson (1997) that finds significant results for an exponential reversion.

3.5 Possible consequences

In the survey mentioned in section 2.1.2 above, Welch (2000) documented the relatively low level of belief amongst academics in a varying equity risk premium. This contrasts somewhat with the view of Campbell *et al* (1997, 286): "It used to be thought that expected asset returns were approximately constant and that movements in prices could be attributed to news about future cash payments to investors. Today the importance of time-variation in expected returns is widely recognized, and this has broad implications for both academics and professionals…" As their text book is regarded as something of a classic by some, the slowness in the spread of this view about fluctuating risk premiums is disconcerting, but does illustrate the difficulty of developing an informed consensus of views.

---

[127] Berkshire Hathaway, for instance, is a closed fund and the share price partly depends on the demand for Mr Buffett's skills.

It appears that market extremes are relatively short-lived. If investors do not panic by buying at extremely high or selling at extremely low prices, the impact may be relatively small. If there are no fundamental grounds why the extremes should not reverse, then they may be able to benefit from the extremes by buying inexpensively or selling overpriced assets. For the reasons mentioned in 3.4.2, it is normally unwise to participate in short selling.

4 THEORY OF INVESTMENT

This section recapitulates, in the light of these facts about investment markets, the basic principles of investment. Fundamental analysis and diversification are both widely accepted, although not as frequently practised.

4.1 Fundamental analysis

Fundamental analysis attempts to determine the underlying value of an asset by reference to company specific accounting and economic data, and consideration of the future. It can be contrasted with technical analysis that considers only previous price movements, and perhaps turnover, and is purely historical.

Perhaps the most important result of the research on market efficiency, behavioural finance and mean reversion is that fundamental analysis can be shown to identify opportunities for extra profits during those times when the market is inefficient, or subject to behavioural biases.

This appears to be confirmed by Abarbanell and Bushee (1998) whose model includes a variety of firm specific accounting data (such as accounts receivable, margins, labour force productivity etc). Their results do not prove that fundamental analysis will always pay for itself in superior returns; such proof is not possible although the arguments of Grossman and Stiglitz (1980) mentioned in 2.1.2 suggest that it ought to be close. Given its relatively low cost for relatively large fund managers, however, there would be a strong argument for all superannuation funds that can access such managers to adopt it.

Technical analysis may well identify specific shares or bonds or sectors of the investment markets that are anomalously priced, but it is always wise to consider whether there may not be good reasons for the anomalies, when viewed prospectively.

4.2 Diversification and utility theory

Analysis of prospective investment returns should give not only the expected value, but the range of possibilities, which leads to considering how to make the trade off between some measure of risk and return.

The rational approach to this trade-off is to develop some curve of utility of various levels of wealth and weight these by the probabilities of occurrence. This involves a host of brave assumptions on both the utility of levels of wealth and the probabilities of future returns as is pointed out, for instance, by Georgescu-Roegen (1954). The irrationality that underlies much of behavioural finance lies in the fact that people do not have consistent utility functions and are often unreasonable in determining future probabilities. (From a totally different perspective, Allen and McGoun (2000) suggest that investors derive a variety of pleasures from participating in what can be a social as well as an intellectual experience. They refer to the "Motley Fool Investment Guide", but an examination of the daily press would tend to confirm their observations.)

Thomson (2003), however, defends subjective utility theory as a normative method of guiding investment decisions. Whatever other gains there are from investing, there appears to be no better rational approach to choosing between investments with different expected risk and return.

Whatever rational method is used to determine the asset allocation, if it considers risk at all, it is likely that it will result in the diversification of the portfolio. If returns on different assets are not perfectly correlated, a diversified portfolio can produce considerable reductions in risk for relatively little loss of return provided returns on the different assets do not differ markedly.

4.3 Life cycle applications

Recent years have seen a number of applications of utility theory to personal investment decisions. Campbell *et al* (2003), Chen *et al* (2006) and Purcal and Chan (2006) are examples. They take into account not only personal utility and some of the characteristics of the investment markets mentioned above, but also allow for consumption over the financial lifecycle and diversification benefits and risks that arise from wage income. McCarthy (2004) surveys much of the literature and refers to Cocco (2005) and others who incorporate housing in the portfolio choice.

In spite of their complexity, these models are recognised by their authors as relatively crude approximations to reality. They do, however, confirm various intuitions and throw light on some otherwise unanswerable questions. Mean reversion in returns over the medium period means a greater allocation should be made to riskier assets, as does the greater implicit investment in human capital that young people have as a result of their greater future wages. A varying equity risk premium means a varying allocation to risky assets.

Four particular criticisms of these models can be made. They frequently use too high an equity risk premium and under-emphasise its uncertainty; their characterisation of the utility function appears unrealistic and they ignore the insights offered by the growth optimal portfolio (explained below); they do not appear to consider the particular situation of retired superannuation fund members, and they confuse mean reversion and autocorrelation.

4.4 The size of the equity risk premium

There are considerable differences in opinion as to the size of the equity risk premium. Mehra and Prescott (2003) report on two decades of debate on their earlier observation that the premium appeared too high to be explained solely by risk aversion. Not the least is the problems is that of measurement as McCarthy (2004) points out. The historical returns must be estimated between two points in time so that even small changes in choice of period lead to considerable differences in estimates. In Welch (2000) academic economists estimated the arithmetic average long term risk premium as 7% on average. This reduced to 5.5% when the 1998 survey was repeated in 2000 – as reported in Welch (2001). One of the reasons may have been that Welch's original paper reported that the respondents believed that the consensus view within the economics profession was as high as 8%, and that this had "anchored" their views at a higher level. On the more pessimistic side, Wilkie (1995) suggests that, in average conditions, the risk premium may be in the range from 3% to 4% per annum. His calculations make allowance for the re-rating of shares, which has been responsible for much of the out-performance of recent years. This lower estimate is supported by others such as Dimson *et al* (2004).

The difference between these two estimates of the expected risk premium is significant, as can be seen by table 1. While the underlying model is crude its indicates the variability of investment returns and the overlap illustrates the difficulty of distinguishing between

different assumptions. Under plausible assumptions, a difference in return of 2% per annum, over a working career of 40 years, increases the expected outcome by 50%. The chance of a negative return would be reduced by about 75%.

Table 1: Accumulation phase - lump sum

Percentiles	2% return	4% return
1%	22	27
5%	25	34
25%	39	54
50%	57	77
75%	84	127
95%	156	243
99%	207	306
Mean	68	100

Shown is the accumulated value of 1 per annum after 40 years, assuming a lognormal return with standard deviation of 15% - 250 simulations.

Regardless of the shape of the utility function, it is clear that the allocation to equities will depend very much on the investor's view of future equity returns. That the experts differ so dramatically on the mean, but that the variability is not quite as disputed, would tend to support Aloysius (2005) that ambiguity aversion tends to discourage equity investment. For retirement planning purposes, it would seem appropriate to use a lower rather than higher estimate of the equity risk premium.

Over-optimism in these models is aggravated by a failure to take expenses into account. A typical Australian superannuation fund will charge its members administrative expenses of at least 0.5% per annum for equity investment. Tax may reduce the return by another 0.5%. On the other hand, the individual's personal cost of borrowing will be at least 1.5% over the risk free rate. Even if the superannuation fund is able to gear up more cheaply, the net risk premium to be used to estimate the optimum level of gearing within the fund should probably not be more than 1.5% - if the equity risk premium is 3% or lower.

4.5 The growth optimal portfolio

Luenberger (1997) describes the growth or log optimal strategy, which determines the proportion of capital to be invested in each asset by using subjective probabilities of the

returns and a logarithmic utility function. He makes a persuasive argument for the use of the logarithmic rather than another utility function for this purpose. The utility function used in most of the papers that have contributed to the debate on the equity risk premium have used a constant relative risk aversion of the form:

$$U(W) = (W^{(1-\gamma)} - 1) \Big/ (1-\gamma) \qquad (3)$$

where W represents wealth and γ the coefficient of risk aversion, which is taken to be in the range (1,10]. A coefficient of risk aversion close to 1 approaches logarithmic utility. Cocco (2005) uses 5 as being plausible, but this is not intuitive as it implies that every 20% increase in wealth is worth less than half of the previous 20%. It is suggested that this apparent risk aversion is more likely to reflect rational discontinuities caused by the possibility of forced lifestyle changes in the short run as suggested in Asher (1999). Shrikhande (1997) demonstrates that a kink in the utility function of this type can serve to explain the equity risk premium puzzle in the same way as a higher co-efficient of risk aversion.

Younger fund members must necessarily expect their consumption patterns to change in the long run – for lifecycle if not wealth induced reasons. If investment returns do not meet expectations, they are likely to be able to work for slightly longer to make up the shortfall. In such circumstances, any kinks in the utility to be applied to investments ought to disappear. Very risk averse utility is therefore not an appropriate normative basis for investment decision making over the long term.

Luenberger briefly derives the growth optimal portfolio of assets that maximises the logarithmic utility over a single period – given the expected distribution of investment returns from different assets (or - more practically - asset classes). It can be shown that, in the long run, this portfolio dominates all other strategies in the sense that the expected median return is always higher. It is more likely to be optimal the longer the time period for investment. For the youngest members of superannuation funds, with a long term time horizon, it would appear that the growth optimal portfolio is the one that they should choose.

4.6 Older superannuation fund members

Superannuation members face increasingly reduced investment opportunity sets as they age. A reducing life expectancy is often punctuated by two particularly important life decisions.

The first is the purchase of a house. This decision may not have been made for investment reasons (ownership gives benefits not obtainable from renting – some tax related), but the presence of the house and significant debt in the family's portfolio is likely to require a different optimal allocation of assets. On a relatively conservative view of the equity risk premium and investment expenses, the best advice may be to repay the home loan as quickly as possible. To the extent that contributions to superannuation are compulsory, a growth optimal portfolio that takes into account correlations with interest rate movements will be more appropriate. This is consistent with the model of Cocco (2005), who does however not allow for the greater utility that may arise from owning rather than renting.

The second major decision relates to retirement. McCarthy (2004) makes the point that the models do not adequately take retirement fund, social security and tax considerations into account. Of interest in this section are retirement fund members who will want to rely on their retirement benefits from the fund. Such members have two main decision phases: the first no longer to invest in human capital (future career development), the second to finally leave the workforce. Once out of the workforce and drawing down superannuation savings, members' investment choices should be much more conservative.

Table 2, using a volatility one third of that used in table 1, shows the distribution of the residual value of an annuity forty years after it is first drawn. Post retirement benefits at older ages are obviously much more sensitive to volatility and to differences in return. An investment policy even more conservative than that which would produce the standard deviation of 5% used in the chart appears to be indicated. The growth optimal portfolio would appear inappropriate because of the kinks discussed above.

Table 2: Decumulation phase - residue

Percentiles	2% return	4% return
1%	2	23
5%	6	38
25%	15	58
50%	23	76
75%	37	91
95%	55	158
99%	78	204
Mean	27	81

Shown is the residue after an annuity of 1 is drawn from an original amount of 40 units for 40 years, assuming a lognormal return with standard deviation of 5% - 250 simulations.

This significant change to a more conservative investment portfolio in retirement suggests that the current practice of revising asset allocations in retirement (the lifestyle portfolios that reduce exposure to investment risks as people age) may well be appropriate. These are described in Bodie and Treussard (2006), who suggest that they may still recommend too high a proportionate investment in risky assets for those with high coefficients of risk aversion (or kinks in their utility functions).

If there are kinks in people's utility curve that derive from lumpy lifestyle decisions, then it is not correct to use a universal utility curve for all people. It is important to understand each family's circumstances as the level of risk aversion will depend on how close they are to having to make significant changes for financial reasons.

4.7 Exploiting mean reversion and time varying risk premiums

Mean reversion of prices in the extremes is not the same as negative autocorrelation, which looks at changes in the previous period, rather than the level of the market as evident from formula 1. There is also no stable cut-off between short term momentum and long term correction both of which are observed. Ahmed *et al* (2002) for instance, show that a momentum strategy – based on the winning style of the previous year - would have been profitable in the US in the 15 years from 1982.

The models described above take different approaches to mean reversion. Campbell *et al* (2003) allow for auto-correlations and reversion to the dividend yield, Purcal and Chan (2006) allow only for autocorrelation, while Chen *et al* (2006) make no allowance at all. It would seem appropriate to use a more comprehensive model that considers all the factors that have been found to govern mean reversion.

4.8 Expenses and taxes

The need to keep control of expenses and tax does not always seem to receive the emphasis it deserves. The last section of Chapter 4 above calculates the total impact of expenses that are frequently held to be unnecessary but which can be explained by the ordinary member's need for some financial advice.

The costs of trading to individuals are mentioned in 2.1.3 above. The costs to institutions are discussed in Schwartz and Steil (2002). They report on an international survey which shows that soft commissions[128] are widely prevalent and reduce the effectiveness and increase the cost of trading. Glass and Wagner (1998) identify four elements to the cost of trading, which must be considered as a whole as savings of one element is likely to increase costs in another:

- Broker commissions, which are a little higher for soft dollar trades and significantly higher for more difficult trades, and the bid offer spread.
- The market impact, measured by the difference in the price from when the order is placed to when the trade is completed.
- Timing losses that arise from the time the portfolio manager chooses to trade to the time the order is placed.
- Opportunity costs arise when an order is not completed, which can be measured by the product of the change in price and the stocks not traded.

Only brokerage is always a cost; the others may turn out profitable at times. As a measure of the impact of these costs, Glass and Wagner's exhibit 2 is worth repeating (in table 3 below), as it shows that it is possible to significantly reduce the cost of trading. The table also makes their point that the costs must be considered as a whole, as savings of one element is likely to increase costs in another: directed trades (where investment

---

[128] These describe payments in kind and those otherwise not disclosed.

managers are required to use a particular broker) may be cheaper but seem to take longer to implement.

Table 3: Costs of trading

	Directed trades	Non-directed trades
Commissions	-12 basis points	-15 basis points
Market impact	-13 basis points	-24 basis points
Delay costs	-87 basis points	-44 basis points
Total execution costs	-112 basis points	-83 basis points

Exhibit 2 from Glass and Wagner (Plexusgroup)

It is noticeable, and a significant omission, that brokerage costs are not invariably reported in superannuation fund accounts. It also appears that most managers do not measure total trading costs.

Dammon *et al* (2001) discuss how appropriate management of capital gains tax particularly can affect optimal asset allocations and returns. While the rate of tax is relatively low in Australia, capital gains tax on a portfolio that turns over once each year can be as high as 0.75% per annum of the value of the portfolio.

5 REGULATORY RESPONSE

While much has been said about the role of government, this section accepts the limits of Adam Smith: " ... the sovereign has only these duties to attend to; three duties of great importance, indeed, but plain and intelligible to common understanding:

- first, the duty of protecting the society from the violence and invasion of other independent societies;
- secondly the duty of protecting, as far as possible, every member of society from the injustice and oppression of every other member of it, or the duty of establishing an exact administration of justice;
- and thirdly, the duty of erecting and maintaining certain public works and certain public institutions which it can never be in for the interest of any individual, or small number of individuals, to erect and maintain; because the profit could never

> repay the expense to any individual or small number of individuals, though it may frequently do much more than repay it to a great society."(Smith, 1930 edition, 180)

In this context, the role of regulators would seem to be to help protect superannuation members from exploitation and to create the conditions for efficient investment markets that would not arise spontaneously.

This section mentions some recommendations from behavioural finance, outlines the Australian regulatory structure and then makes a few modest recommendations.

5.1 Suggestions from behavioural finance

Three sets of suggestions emerge from the behavioural finance literature; each of which could be taken up by the regulators and required of trustees as suggested later in this section.

5.1.1 Disclosure

Daniel *et al* (2002) are not advocates of government direction of the market, but have some suggestions on:

"minimally coercive and relatively low cost measures to help investors make better choices and make the market more efficient. These involve regulation of disclosure by firms and by information intermediaries, financial reporting regulations, investment education, and perhaps some efforts to standardise … advertising.

"More controversially, a case can be made for regulations to protect foolish investors by restricting their freedom of action or the freedom of those that may prey upon them. Limits on how securities are marketed and laws against market manipulation through rumour spreading may fall into this category.

"There is little cost to requiring companies to provide a standard warning, analogous to cigarette warning labels, to workers of the risks of plunging retirement money in their own company's stock. Regulating the way in which retirement investment options are presented to individuals (i.e. the status quo choice, and how choices are categorized) may have low cost yet may greatly affect lifetime outcomes."

In the light of the observed persistent under-performance of some funds, Blake and Timmerman (2004) conclude that past performance information should be made available to members so that they are able to avoid these under-performing funds.

5.1.2 Option menus

Mitchell and Utkus (2004) take up the question of choices noting that many portfolios tend to be undiversified, that people tend to end with the default investment option particularly if a large number of alternative choices are made, and that they tend to interpret the choices offered as including an element of recommendation. For example, if more equity options are offered, people will put a larger proportion of their assets into equity.

They suggest that the default option should automatically guide investors into optimal investment behaviour. This suggests some type of lifestyle investment pattern as discussed in 4.6 above. Such a pattern would need to be invested in optimal portfolios that progressively sell risky assets and buy indexed-linked bonds as people age, noting the comments that older people are likely to be much more risk averse.

In order to prevent a plethora of options confusing and misleading average investors, they suggest that alternative options should be limited and placed on an alternative investment menu.

5.1.3 Self-monitoring

Behavioural research confirms the human penchant to forget or reinterpret past errors. Biais *et al* (2005) report on experiments that show that self-monitoring of performance provides something of an antidote to overconfidence. One might be more emphatic: if there is any science or management in investment, monitoring the success and failure of decisions is essential.

5.2 Australian regulations

Official responsibility for the regulation of Australian superannuation funds is shared by the Australian Securities and Investment Commission (ASIC), which is responsible for market conduct of the industry, and the Australian Prudential Regulation Authority (APRA), which is responsible for the regulation of the trustees.

ASIC's policy statements govern the content of disclosure; Policy Statement 168 requires disclosure to facilitate the comparison on fees and historical performance.

APRA's Superannuation Circular II.D.1 "Managing Investments and Investment Choice" requires trustees to develop diversified investment strategies that place neither the fund nor the member exercising the choice at excessive risk. It also discusses the implementation of the strategies and, of particular interest to this paper, the reporting of investment performance.

There are also a number of industry, professional and educational bodies with their own principles of investment and standards for the provision of advice. They should also be seen as part of the regulatory system, and the recommendations below apply also to them.

5.3 Recommendations

The following ideas for the development of these regulations may be considered.

5.3.1 Disclosure

The recommendations from behavioural finance recorded in sections 5.1.1 and 5.1.2 and should be incorporated as part of good practice.

In addition, it would not be out of place to require superannuation funds to report on the exact assets held by their various options. Australia is almost unique in not requiring mutual funds to report on their asset holdings[129].

5.3.2 Monitoring performance

In monitoring investment performance, trustees ought to consider the success of each element of the investment decision: asset allocation, stock selection and dealing. Such monitoring should include performance attribution of the form used by Daniel *et al* (1997) that measures the relative contribution of asset allocation and stock selection. It would be even more useful if investment managers were to document the reasons for their particular decisions. This would help to prevent the hindsight bias described in the behavioural literature.

---

[129] Personal communication from Bloomberg management.

The success of stock selection can be further analysed by the approach set out in Day *et al* (1994), which measures the profitability, over recent quarters, of particular decisions to buy and sell. It provides an important short term measure of investment ability and efficiency, especially relevant in view of the tendency reported in 2.2.2 for managers to underestimate the possibility of inefficiencies in information flow. The mis-pricing that may arise is likely to be relatively short, so the wisdom of a particular transaction should soon be apparent. It would, in some respects, be surprising if an investment manager saw opportunities more than a few weeks before the market price responded.

Given that excessive dealing and high dealing costs have been identified as a significant contribution to under-performance, it is also important to monitor the dealing costs identified by Glass and Wagner (1998) and listed in 4.8 above.

Such reporting should also include the impact of tax on net investment returns.

In the interests of transparency, the results of these analyses ought to be made available to members and potential members – perhaps after a suitable delay to protect the intellectual capital of the investment managers. Given the limited interest and ability of members to absorb large quantities of data, such performance reports should be summarized before they are distributed, with fuller reports available on request or on websites.

This more thorough reporting can be contrasted with the performance reports currently given to members. These are normally given for a variety of periods often from 3 months to 5 years which make little or no allowance for different styles. The shorter periods are entirely useless and inappropriately focus attention on a set of random numbers. The longer terms are of greater value but often do not make appropriate allowance for tax.

APRA's investment circular will hopefully encourage trustees to develop a more thorough approach to performance measurement. It should perhaps be made more specific.

5.3.3 Poor performers

The publication of past performance league tables is appropriate, as suggested by Blake and Timmerman (2004), so that investors can avoid under-performing funds. While the worst funds are often included in a table, editorial invariably focuses on winners rather than losers. The reason is presumably that losers are also advertisers. The losers may not

be driven out by competition if they operate through a marketing channel that does not require comparisons to be made.

It would appear that it should fall to the regulator (APRA in this instance) to identify poor performers and ensure that the shortfall is not due to dishonesty or gross incompetence.

It may also be appropriate for the regulator (ASIC in this case) to actively discourage the use of the more risky interest bearing assets that are particularly active in property financing. They are particularly inappropriate for superannuation funds given that they commonly offer limited diversification and the returns are skewed downwards: a maximum interest rate and a possible 100% loss. Advertisements ought, at very least, to disclose a measure of the risk being taken. (Something along the lines of: "The underlying properties have to fall in value by 5% to affect the value of your investment, and by 10% to reduce it to zero".)

5.3.4 Risk appetite

Superannuation fund members are normally given the choice between a range of risk and reward tradeoffs. More recent developments are for products to be related to age.

The existing products also do not appear to take the investment theory set out above into account:

- They do not appear to offer a growth optimal portfolio for younger investors. This might well incorporate some implicit gearing to the extent permitted by law.
- The difference between homeowners and non-homeowners does not appear to be recognised.
- Most of the portfolios have predetermined asset allocations, giving the investment managers limited opportunity to respond to extreme market events.

It is not clear that these issues should be resolved by regulatory intervention. It would appear to be mainly an issue of ongoing research and education, of which this paper is an element.

5.3.5 Benchmark portfolio to discourage panic

One of the functions of prudential regulators is to avoid banking crises. Preventing extremes in investment markets is generally seen as beyond their capabilities. There may

however be some parallels between the role of government in banking crises and in investment market extremes. People at least need to be reassured.

The damage done by a bank run is obvious. Superannuation fund members - who "panicked" and bought foreign technology shares at the peak of the last bubble may have lost half their lifesavings. They may not however make news.

As discussed by Abolafia and Kilduff (1988), regulators do have a direct interest in the effect of price bubbles on lenders that have contributed to the bubble. They will also respond at various stages of a panic. Are earlier interventions possible? To ask the regulator to warn of excesses is unrealistic. The benefits are putative, the risk to reputation far too high.

It is possible, however, that there would be some merit in developing a model growth optimal portfolio that reflected the consensus of economists, asset consultants and investment managers. The portfolio could be published by the regulator or an industry body and be justified as a "public work" in Adam Smith's terms. The consensus could be based on a survey that did not lead to the identification of the individual participants, so preserving any competitive advantage. The published results would be of the anticipated returns and correlation matrices of the major asset classes and of the resultant model growth optimal portfolio[130]. It would include a measure of dispersion in both the expected returns and model portfolio. Such a publication would have a number of advantages:

- It would be resolutely prospective, so providing a constant reminder that historical returns are not estimates of the future. It ought also to encourage more fundamental analysis as it is necessary to estimate future returns.
- It would also provide a constant reminder for the desirability of diversification even where expected returns differ significantly.
- Behavioural research suggests that it would encourage people's views and actions towards conformity. This is the desired response for uninformed investors. Even in the case of the academics reported in Welch (2001), it led to more reasonable

[130] There should be constraints on short sales and borrowing as neither are legally permitted. Luenberger shows that the mathematical result for the growth optimal portfolio can normally be shifted away from the most risky assets with a minimal loss of expected return, but this a technical detail.

estimates of the future. It may aggravate the pressures of informed professional investors, but might also provide an environment where shorter term deviations from the herd would be easier to explain.

- Because it would have to be published with the necessary health warnings, it would remind older investors of their need to reduce risk.

6 CONCLUSION

Extracting a consensus view of the current state of investment theory is made difficult by dogmatically held views and a vast output of often unrelated research.

While markets may be more efficient than we could expect, it is clear that there are momentum and bubble effects that are better explained by the bounded rationality of investors. These anomalies in efficient market theory are well explained by behavioural finance, but exploiting the resulting noise requires fundamental analysis of the merits of each investment. Fundamental analysis can yield profits from both asset allocation and stock selection during those times when markets are inefficient. The benefits of active asset allocation when markets are mispriced do not appear to be emphasised sufficiently.

In protecting uninformed superannuation investors with investment choice, regulators need to ensure that they are given information in a way that encourages sensible behaviours. A particular gap in the current framework would appear to be a thorough monitoring of performance. Poor performers particularly ought to be identified. It is also suggested that the development and publication of a consensus portfolio would encourage all the desirable elements of a sensible investment strategy: fundamental analysis, adequate diversification and a greater sense of comfort with the choices made.

REFERENCES

Abarbanell J S & B J Bushee (1998) Abnormal Returns to a Fundamental Analysis Strategy *Accounting Review* 73.1: 19-46

Abolafia M Y & M Kilduff (1988) Enacting Market Crisis: The Social Construction of a Speculative Bubble *Administrative Science Quarterly* 33.2: 177-193

Ahmed P, LJ Lockwood & S Nanda (2002) Multi-style Rotation Strategies *Journal of Portfolio Management* 28.3: 17-29

Allen F & D Gale (2000) Bubbles and Crises *Economic Journal* 110.1: 236-255

Allen D E & E G McGoun (2000) Hedonic investment *Financial Services Review* 9 : 389-403

Allen L & A Saunders (2004) Incorporating Systemic Influences into Risk Measurements: A survey of the literature *Journal of Financial Services Research* 26.2: 161-191

Aloysius J A (2005) Ambiguity Aversion and the Equity Premium Puzzle: a re-examination of experimental data on repeated gambles *Journal of Socio-Economics* 34: 635-655

Arghyrou M, C Martin & C Milas (2005) Non-linear inflationary dynamics: Evidence from the UK *Oxford Economic Papers* 57.1: 51-69

Arnott R D (2004) Is Our Industry Intellectually Lazy *Financial Analysts Journal* 60.1: 6-8

Asher A (1999) Loading members with Investment Choice *Journal of Pensions Management* 5.1: 76-85

Baillie R T, C F Chung & M A Tieslau (1996) Analysing Inflation by the Fractionally Integrated ARFIMA--GARCH Model *Journal of Applied Econometrics* 11.1: 23-40

Ball P (2004*) Critical Mass: How one thing leads to another* Random House

Ball R (1995) The Theory of Stock Market Efficiency: Accomplishments and limitations *Journal of Applied Corporate Finance* 8.1: 4-18

Balvers R, Y Wu & E Gilliland (2000) Mean reversion across National Stock Markets and Parametric Contrarian Investment Strategies *Journal of Finance* 55.2: 745-772

Barber B M & T Odean (2000) Trading Is Hazardous to Your Wealth: The common stock investment performance of individual investors *Journal of Finance* 55.2: 773-806

Barber B M & T Odean (2001) Boys will be Boys: Gender, overconfidence, and common stock investment *Quarterly Journal of Economics* 116.1: 261-292

Barker M D (1999) *Some Thoughts on Systemic Risk* Biennial Convention, Institute of Actuaries of Australia

Becker C, W Ferson, D H Myers & M J Schill (1999) Conditional market timing with benchmark investors *Journal of Financial Economics* 52.1: 119-148

Bernstein W (2002) *The Four Pillars of Investing* McGraw-Hill

Bhanot K (2005) What Causes Mean Reversion in Corporate Bond Index Spreads? The impact of survival *Journal of Banking & Finance* 29.6: 1385-1403

Biais B, D Hilton, K Mazurier & S Pouget (2005) Judgemental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental Financial Market *Review of Economic Studies* 72.2: 287-312

Blake D & A Timmermann (2004) *Performance Persistence in Mutual Funds*: *An independent assessment of the studies* prepared by Charles River Associated for the Investment Management Association Financial Services Authority, London

Bodie Z & J Treussard (2006) Making Investment Choices as Simple as Possible but Not Simpler *Working Paper* ssrn.com/abstract=900005

Brav A & J B Heaton (2002, Competing theories of financial anomalies *Review of Financial Studies* 15.2: 575-606

Campbell J Y, A W Lo & A C MacKinlay (1997) *The Econometrics of Financial Markets* Princeton University Press

Campbell J Y, Y L Chan & L M Viceira (2003) A Multivariate Model of Strategic Asset allocation *Journal of Financial Economics* 67: 41-80

Capocci D & G Hubner (2004) Analysis of Hedge Funds Performance *Journal of Empirical Finance* 11: 55–89

Cavaglia S & V Moroz (2002) Cross-industry, Cross-country Allocation *Financial Analysts Journal* 58.6: 78-97

Carhart M M (1997) On Persistence in Mutual Fund Performance *Journal of Finance* 52.1: 57-82

Chan W S & S H Cheung (2005) A Bivariate Threshold Time Series Model for Analyzing Australian Interest Rates *Mathematics and Computers in Simulation* 68.5-6: 429-437

Chen J & H Hong (2002) Discussion of "Momentum and Autocorrelation in Stock Returns *Review of Financial Studies* 15.2: 565-573

Chen P, R G Ibbotson, M A Milevsky & K X Zhu (2006) Human Capital, Asset allocation, and Life Insurance *Financial Analysts Journal* 62.1: 97-109

Cocco J F (2005) Portfolio Choice in the Presence of Housing *Review of Financial Studies* 18.2: 535-567

Collin-Dufresne P & R S Goldstein (2001) Do Credit Spreads Reflect Stationary Leverage Ratios? *Journal of Finance* 56.5: 1929-1957

Cooper M J, R C Gutierrez & A Hameed (2004) Market States and Momentum *Journal of Finance* 59.3: 1345-1365

Cox J C, J E Ingersoll & S Ross (1985) A Theory of the Term Structure of Interest Rates *Econometrica* 53: 385-407

Cutler D M, J M Poterba & L H Summers (1991) Speculative Dynamics *Review of Economic Studies* 58.3: 529-537

Dammon R M, C S Spatt & H H Zhang (2001) Optimal Consumption and Investment with Capital Gains Taxes *Review of Financial Studies* 14.3: 583-616

Daniel K, M Grinblatt, S Titman & R Wermers (1997) Measuring Mutual Fund Performance with Characteristic-Based Benchmarks *Journal of Finance* 52.3: 1035-1058

Daniel K, D Hirshleifer & S H Teo (2002) Investor psychology in capital markets: evidence and policy implications *Journal of Monetary Economics* 49: 139–209

Day N, S J Green & J Plymen (1994) Investment - Assessing a Manager's Skill and Monitoring the Risk *Journal of the Institute of Actuaries* 478: 69-118

Dimson E, P Marsh & M Staunton (2004) Irrational optimism *Financial Analysts Journal* 60.1: 15-25

Do V, R Faff & J Wickramanayake (2005) An Empirical Analysis of Hedge Fund Performance: The case of Australian hedge funds industry *Journal of Multinational Financial Management* 15.4-5: 377-393

Georgescu-Roegen N (1954) Choice, Expectations and Measurability *Quarterly Journal of Economics* 68.4: 503-534

Gilchrist S, C P Himmelberg & G Huberman (2005) Do Stock Price Bubbles Influence Corporate Investment? *Journal of Monetary Economics* 52.4: 805-827

Glass S & W Wagner (1998) The Dynamics of Trading and Directed Brokerage *Journal of Pension Plan Investing 2:* 53-72 (The journal appears defunct but the article is in the investment readings of the Institute of Actuaries of Australia, and may be available from the Plexus group.)

Grossman S & J Stiglitz (1980) On the Impossibility of Informational Efficient Markets *American Economic Review* 70: 393-408

Harney M & E Tower (2003) Predicting Equity Returns using Tobin's q and the Price Earnings Ratio *The Journal of Investing* 12.3: 58-69

Hens T & R K Schenk-Hoppé (2005) Evolutionary Finance: Introduction to the special issue *Journal of Mathematical Economics* 41.1-2: 1-5

Hirshleifer D (2001) Investor Psychology and Asset Pricing *Journal of Finance* 56.4: 1533-1597

Hirshleifer D & S H Teoh (2003) Herd Behaviour and Cascading in Capital Markets: A review and synthesis *European Financial Management* 9.1: 25-66

Hoelzl E & A Rustichini (2005) Overconfident: Do you put your money on it? *The Economic Journal* 115: 305-318

Jarrow R A &S M Turnbull (2000) The Intersection of Market and Credit Risk *Journal of Banking and Finance 24.1*: 271-299

Jones C (2003) Nonlinear Mean Reversion in the Short-Term Interest Rate *The Review of Financial Studies* 16.3: 793-843

Kindleberger C P (1988) The Financial Crises of the 1930s and the 1980s: Similarities and Differences *Kyklos* 41.2: 171-186

Kapetanios G, Shin Y & Snell A (2003) Testing for a Unit Root in the Nonlinear STAR Framework *Journal of Econometrics* 112.2: 359-379

Kim K A & Nofsinger J R (2005) Institutional Herding, Business Groups, and Economic Regimes: Evidence from Japan *The Journal of Business* 78.1: 213-242

Lai K S (2004) On Structural Shifts and Stationarity of the *ex ante* Real Interest Rate *International Review of Economics and Finance* 13: 217-228

Lewellen J (2002) Momentum and Autocorrelation in Stock Returns *Review of Financial Studies* 15. 2: 533-564

Luenberger D G (1997) *Investment Science* Oxford University Press

Malkiel B G (2004) Models of Stock Market Predictability *Journal of Financial Research* 27.4: 449-459

Mankiw N G & J A Miron (1986) The changing Behaviour of the Term Structure of Interest Rates *Quarterly Journal of Economics* 101.2: 211-228

Mehra R & E C Prescott (2003) The Equity Premium in Retrospect *Handbook of the Economics of Finance* Edited by G M Constantinides, M Harris and R Stulz, Elsevier

McCarthy D (2004) Household Portfolio Allocation: A review of the literature. Presented at the International Collaborative Forum, ESRI, Tokyo

Mitchell O & S Utkus (2004) Lessons from Behavioral Finance for Retirement Plan Design *Pensions Design and Structure: New Lessons from Behavioral Finance* ed Mitchell & Utkus, Oxford University Press, Oxford

Monoyios M & L Sarno (2002) Mean Reversion in Stock Index Futures Markets: A nonlinear analysis *Journal of Futures Markets* 22.4: 285-314

Narayan P K & R Smyth (2005) Are OECD Stock Prices Characterized by a Random Walk? Evidence from sequential trend break and panel data models *Applied Financial Economics* 15.8: 547-556

Nickles M D (2005) Presidential Elections and Stock Market Cycles: Can you profit from the relationship? *Journal of Contemporary Business Practice* 7.3 http://gbr.pepperdine.edu/043/stocks.html

O'Connell P J G (1998) The Overvaluation of Purchasing Power Parity *Journal of International Economics* 44: 1–19

Okunev J & P J Wilson (1997) Using Nonlinear Tests to Examine Integration between Real Estate and Stock Markets *Real Estate Economics* 25. 3: 487-503

Paya I, I A Venetis & D A Peel (2003) Further Evidence on PPP Adjustment Speeds: The case of effective real exchange rates and the EMS *Oxford Bulletin of Economics and Statistics* 65.4: 421-437

Prigent J L, O Renault & O Scaillet (2001) An Empirical Investigation into Credit Spread Indices *Journal of Risk* 3.3: 27-55

Purcal S & H Chan (2006) Living Life Optimally with a Mean-Reverting Price of Risk *Institute of Actuaries of Australia Financial Services Forum* http://www.actuaries.asn.au/PublicSite/pdf/2006/fsf06_paper_purcal_chan_living%20life.pdf

Rogoff K (1993) The Purchasing Power Parity Puzzle *Journal of Economic Literature* 34: 647-668

Schwartz R A & B Steil (2002) Controlling Institutional Trading Costs *Journal of Portfolio Management* 28.4: 39-49

Smith Adam (1789) *An Inquiry into the Nature and Causes of the Wealth of Nations* (1930 edition) Methuen, London

Seo B (2003) Nonlinear Mean Reversion in The Term Structure of Interest Rates *Journal of Economic Dynamics & Control* 27 11-12: 2243-2265

Shrikhande M M (1997) Non-addictive Habit Formation and the Equity Premium Puzzle *European Financial Management* 3.3: 293-319

Simon H A (1983) *Reason in Human Affairs* Basil Blackwell, Oxford

Stotz O (2004) How to Profit from Mean Reverting Risk Premiums? Implications for stock selection *Journal of Asset Management* 5.3: 192-202

Stevenson S (2002) Momentum Effects and Mean Reversion in Real Estate Securities *Journal of Real Estate Research* 23.1: 7-64

Taylor A M (2001) Potential Pitfalls for the Purchasing-Power-Parity Puzzle? Sampling and specification biases in mean-reversion tests of the law of one price *Econometrica* 69.2: 473-498

Taylor A M & M P Taylor (2004) The Purchasing Power Parity Debate *NBER Working Paper No.10607*

Taylor M P, D A Peel & L Sarno (2001) Nonlinear Mean-Reversion in Real Exchange Rates: Toward a solution to the purchasing power parities problem *International Economic Review* 42.4: 1015-1042

Thomson R J (1992) Stochastic Investment Modelling: The case of South Africa *British Actuarial Journal* 2: 765-901

Thomson R J (2003) The Use of Utility Functions for Investment Channel Choice in Defined Contribution Retirement Funds: I: Defence *British Actuarial Journal* 9.3: 653-709

Vasicek O (1977) An Equilibrium Characterization of the Term Structure *Journal of Financial Economics* 5: 177-188

Welch I (2000) Views of Financial Economists on The Equity Premium and Other Issues *Journal of Business* 73.4: 501-537

Welch I (2001) The Equity Premium Consensus Forecast Revisited *Working Paper, Cowles Foundation for Research in Economics* Yale University, New Haven, CT http://cowles.econ.yale.edu/P/cd/d13a/d1325.pdf

Wilkie A D (1995) The Risk Premium on Ordinary Shares *British Actuarial Journal* 1: 251-293

Williams J (1995) *Manipulation on Trial: Economic analysis and the Hunt silver case* Cambridge University Press

Woolley P & R Bird (2003) Economic Implications of Passive Investing *Journal of Asset Management* 3.4: 303-312

CHAPTER 8 - SMOOTHING INVESTMENT RETURNS

First published as Chapter 8: "Smoothing Investment Returns" in Retirement Funds in Scary Markets (2007), ed H Bateman, Edward Elgar.

1 THE NEED FOR SMOOTHING

Other things being equal, we expect people to want a predictable and smooth income stream in retirement. This would enable them to fund regular consumption and to budget for less regular expenses.

In spite of this both private and public superannuation schemes have been changed, over the past twenty years particularly, from defined benefit (DB) to defined contribution (DC). Investment risk is thereby transferred from the sponsor (employer or state) to the members. The benefits paid are therefore not necessarily predictable or smooth. The issue has been discussed, inter alia, by Knox (1993), Khorasanee (1995), Andrew (1995) and Blake (2000).

This chapter addresses the question of how DC funds can provide smoother investment returns. In the next part, we look at the difference between long and short term investment risks and the problems faced by DB funds in coping with changes to the social and economic environment. This inflexibility, and their opaqueness, can be seen as causes of their current unpopularity. The rest of the chapter then describes the smoothing algorithms.

2 THE MANAGEMENT OF RISK

2.1 Longer term risks

The transfer of risk from sponsors to fund members can be seen to have a long term and a short term component.

This is perhaps best illustrated in the context of mortality risk, which is also transferred from DB scheme sponsors to DC members. It applies to the costs of the death benefits payable to members' dependants, and to the risks of longevity. The relatively short term risks of fluctuations in experience can be absorbed by insurance contracts and life

annuities. These can be purchased from life insurance companies or self insured by the fund. Longer term fluctuations require changes to premium rates. The distinction between the long and short term risks effectively depends on the financial capability of sponsors and insurance companies to provide guarantees. They can do so for short term risks, but not for the long term.

The issue of financial capability is also relevant to long and short term investment risks. The risks are however much larger. Estimates of the long term net real rate of return vary from about one per cent to five per cent annually. The lower return is that available from indexed linked stocks – after inflation, taxes and expenses. The upper end of the range arises from the estimates of Dimson *et al* (2004) of the likely equity premium. The equity premium arises from a number of sources: the inherent risk and the costs of investing probably the most important. Dimson's estimate of 3.5 percent for the equity premium is lower than that reported by Welch (2000) in a survey of economists, but probably more realistic. With typical contributions and expenses, the difference between one and five per cent would translate into a DC pension of between 35 and 130 per cent of average income when spread over a 60 year span. This compares with a variability of perhaps 10 per cent each way in the value of a life annuity as a result of unexpected changes to mortality over the life of a typical pensioner.

If long term investment returns are lower than expected, members of DC funds will have to contribute more during their working lives, accept lower pensions or work for longer. The latter strategy can be very effective: each year of additional contributions will increase a pension by some 10 per cent. The major role for investment smoothing comes in providing members with more time to adjust their plans for retirement in line with variations in long term investment returns.

2.2 Cross subsidies and inequities in DB funds

While DB funds have provided apparently predictable pensions for many of their members, the traditional DB design does not provide a model for smoothing. Even if the sponsoring employer is able to guarantee accrued benefits, future accruals must always be open to renegotiation because it is impossible to guarantee investment returns for ever. The long term variability in investment returns will produce differences to contribution rates of the same order as differences in DC benefit levels. There is however no objective mechanism to make the adjustments to prevent inequities arising.

In employer sponsored scheme there would be no inequity if the salary packages of new employees reflected the *ex ante* fair value of the promised benefits, and the employer absorbed the profits and losses arising from deviations in investment returns from those expected. Neither condition is usually met. The first condition would require salary packages to be adjusted for the value of the pension benefits accruing. It is however practically impossible to distinguish between groups of members, such as different generations of recruits. This creates cross subsidies between the groups, the value of which is seldom if ever determined.

The failure to calculate the value of the cross subsidies between different groups of employees also means that little attempt is usually made to determine whether the employer is bearing a fair share of the costs. This opaqueness and imprecision in the allocation of costs opens the way for strong opportunistic parties to benefit at the expense of others either by manipulating changes to benefit levels or by frustrating equitable changes. The beneficiaries are likely to be shareholders, but could also be senior management or union officials. This uninsurable moral hazard is explored in more detail in Asher (2000).

These issues are writ large in a state sponsored scheme. Even if the relationship between contributions and benefits is determined initially in such a way as to be fair, economic and demographic changes will soon create unintended, and frequently implicit, cross subsidies between different groups of contributors. Making changes is politically fraught as the acrimony of current debate on the issue bears witness.

If the rules of DB funds were never changed, they would give rolling guarantees that extended for the lifetime of every cohort of new members: an infinite guarantee. This is clearly unreasonable. It would appear that the only way of setting the rules to ensure that no interested group will wrest an unfair advantage for itself is to make the cross-subsidies and investment guarantees explicit. This would convert DB funds into complex DC arrangements with employer guarantees of accrued benefits, and regularly reviewed future accrual rates. The future accruals would be a form of guaranteed deferred annuities.

Modern finance and accounting, with their increased focus on market values, have begun to identify the cross subsidies that arise in DB funds, and the unacceptable risk that guarantees bring to the finances of the sponsor. It is perhaps not surprising that growing

numbers are being converted into DC funds, which have four clear advantages over traditional DB designs.

- Members' entitlements are clearly set out and not subject to management discretion or political maneuvering.
- The relationship between contributions and benefits is clearly fair, and free of significant but implicit cross subsidies.
- The risks to the sponsor are not inherent in the benefit design, and any guarantees can be explicitly managed.
- Members are free to participate (if there is investment choice) in the equity premium, so possibly reducing contributions or increasing benefits.

2.3 Managing risks in retirement

In the long run, the equity premium means that investments in equities produce lower pension costs. The price is significant short term volatility: falls of 30% and more within a year have been commonly experienced.

Once retired, it is likely to be difficult for fund members to re-enter the employment market to make up for investment losses. Members therefore need protection against such significant falls in income after retirement. This protection should probably begin before the retirement date, as plans for retirement presumably take some time to implement. Superannuation funds therefore need investment instruments that offer protection against investment fluctuations over this relatively long, but finite, period.

The strategy employed by most DC funds and their members is to use a combination of investments in order to give an optimum mix of security, inflation protection and participation in the equity premium.

To the extent that the equity premium arises from the correlation of share prices with the universe of consumption, it may be possible to develop other assets that share this risk without the short term volatility. One possibility would be to develop investments that provide a cash flow that could match pensions, and was linked to a wage index. This is the approach taken by the new Swedish notional DC national scheme. Asher (1994) suggests a method for utilising housing finance.

In the absence of these instruments, the next part of this chapter describes how volatile investment returns can be smoothed to produce a more acceptable income flow, and provides a family of algorithms for this to be done fairly and efficiently.

3 CURRENT METHODS OF SMOOTHING

3.1 Lifestyle switching to guaranteed assets

For the reasons discussed above, many DC arrangements, even if they are largely invested in equities before retirement, arrange for members' investments to be switched to fixed, or inflation protected, annuities at retirement.

A common variation on this approach, called "lifestyle investing", is described by Booth and Yakoubov (2000). The switch to more stable investments is phased in over a few years, rather than made at a single date. As they find however, the success of such a strategy depends on the availability of suitable long term indexed linked instruments, and foregoes some of the equity premium. It does however reduce the volatility of the final benefits.

Cairns *et al* (2003) describe a further theoretical development that they call "stochastic lifestyling", which dynamically matches retirement income to the member's final salary. Given that final salary can be distorted by recent promotions, or a reduction in working hours related to ill health or impending retirement, it is not clear however that this is an appropriate benchmark.

3.2 Delaying action

One approach to smoothing investment returns is to delay the decision to distribute investment returns until the end of some time period. This is not as alarming as it first appears. Unit trusts will normally delay transacting until the end of a day. Many unit linked DC funds will allocate units at the end of a month, while non-linked funds will apply the same crediting rate for a year at a time. In almost every case, however, they will reserve the right to calculate unit prices, or change the crediting rate, at some intermediate time if market conditions change and some participants may otherwise be disadvantaged. This could happen for instance if market values dropped dramatically. In this instance, failing to recalculate unit prices will mean that departing members are given more than a fair share of the assets – to the detriment of remaining members. The

inequity may be aggravated, and the solvency of the fund perhaps threatened, if some members are able to elect to leave in order to take advantage of the artificially high unit price. This risk of some members making elections that harm others is called anti-selection in the actuarial literature.

Delaying action should however be seen more as an administrative convenience than a method to address market fluctuations.

3.3 Smoothed bonuses

Many DB and DC funds historically operated through policies with life insurance companies. The companies declared smooth bonuses on these policies to distribute investment and other profits. DC funds often declare their own smoothed bonus or crediting rates.

One actuarial method of determining the bonus is to begin with a calculation of a smoothed value of assets. Head *et al* (2000) provides a description of their development and justification in the context of DB funds. One common approach is for the value of shares to be determined by discounting future dividends at the actuarial valuation rate. An alternative is to assume that share prices will (immediately) return to a level that reflects some long term average dividend yield. Other approaches to smoothing involve some averaging of the market value of assets.

To the extent that these actuarial values for the assets differ from their market values they are, arguably, unrealistic. Actuaries who use the method respond that their smooth actuarial value represents a more realistic estimate of the long term value of assets than the market price. Market prices may well be overly influenced by the needs or views of those who happen to be transacting at that time.

Even if this were true, three major problems arise. The first is the possibility of anti-selection. Incoming policyholders or members benefit if the difference between market and actuarial value is positive; maturing policyholders and exiting members gain when it is negative. They thus have the incentive to make decisions that disadvantage other members.

The second is that when market values are below actuarial values, the fund is exposed to the risk that a combination of further falls in market values and an outflow of funds may lead to an unacceptable drop in the benefits available to the remaining members, or even

to insolvency if the fund is entirely exhausted. This problem may be further aggravated by anti-selection. Fund trustees and company managers are therefore understandably reluctant to allow a shortfall of assets to become significant. This reduces their ability to cushion falls in benefit payments – which is the original reason for smoothing.

The third problem is more subtle. As a result of the reluctance to pay benefits in excess of the market value of the underlying assets, most smoothing involves limiting the increase in the crediting rate when market values rise rapidly. This creates a tendency for smoothing reserves to build up with no obvious methodology of release. In the absence of protection for policyholders or members, the resulting surplus is open to expropriation by stronger and opportunistic parties. Asher (1991) discusses one such incident, where an "orphan estate", built up from the contributions of previous generations of policyholders, was expropriated by a new shareholder.

Clay *et al* (2001) discuss UK with-profit policies and make various recommendations as to disclosure and the protection of vulnerable parties. They believe however that smoothing necessarily requires discretion on the part of the fund's governing body, in order to ensure equity and solvency. The discussion of their paper reflected widespread doubt as to whether the three problems discussed above can be resolved without using an objective algorithm for smoothing.

3.4 Alternative algorithms

Thomson (1997) suggests an algorithm that optimises measures of smoothness and solvency in order to produce both an investment policy and a crediting rate for a DC fund. The results appear to avoid negative bonuses (which are not easily understood by members), but produce more volatile bonus rates than would be common practice.

Khorasanee and Ng (2000) propose an arrangement where the contributions made by or on behalf of active members are explicitly reallocated to retiring members in order to allow for smoother returns. As pointed out in the discussion by Sze, this approach does not appear to be fair, particularly if investment returns are poor and the fund is declining.

Blake (2003) investigates the benefits of a smoothing algorithm that truncates higher and lower returns so that bonuses lie within a limited range. He does not however provide a justification for the method.

All three methods are vulnerable to anti-selection.

4 A NEW ALGORITHM

This chapter suggests algorithms that smooth using a set of forward contracts of different durations.

4.1 A series of forward contracts

We can take an investment policy issued by a life insurance company as an example. The benefits of a member of a DC fund are identical from an investment perspective.

The maturity payout of a unit linked policy depends on the market price of the units attached. The payout would be smoothed if the maturing policyholder were to enter into series of forward or future[131] contracts, each for some of the units attached to the policy, in each of the months prior to maturity. For purposes of illustration, assume the smoothing will take place over 60 months.

The remaining policyholders in the life fund would be the obvious counterparties to forward contracts. When the payouts on policies maturing, or likely to claim, in the next five years is small relative to the size of the fund, then the fund as a whole could be the counterparty. Alternatively, policyholders wishing to increase their exposure to volatile investment markets could be specifically allocated as counterparties.

Younger policyholders may be liquidity constrained - unable to borrow - and may want greater exposure to higher yielding equities. Such greater exposure seems appropriate as found, for instance, by Cairns *et al* (2003). An important caveat is that such findings depend critically on assumptions as to borrowing rates and the size of the long term equity premium. Long forward positions gear up exposure to equity markets, and a significant fall in market values would reduce surrender values considerably. If one assumes that equity values will recover however – as is normally the case – young policyholders in the accumulation phase would be compensated by being able to accumulate new units at a low price. It would be necessary however to make this patently clear to them in advance in order to avoid unhappiness at unexpectedly low shorter term returns.

---

[131] A future contract is listed, a forward is not. They are otherwise substantially the same.

The allocation of the long position would depend on the surrender values of the policies being sufficiently large to serve as margins. Given the illiquid nature of the forwards, the margin required would be significantly higher than that required on futures contracts regularly marked to market. Determining the size of the margins required is a problem not addressed in this paper. The objective nature of the algorithm nevertheless allows for it to be addressed explicitly, unlike the implicit algorithms of discretionary smoothing.

The payouts on policies maturing (or likely to claim) in the next five years may become large relative to the size of the fund for a particularly aged workforce. In such a case, it may be necessary to enter into future contracts in a suitable future's market, reduce the fund's equity exposure or reduce the smoothing.

Contract design may be dominated by tax considerations, which may also lead to participation in the forward contracts by the insurance company's shareholders.

4.2 Pricing forwards and futures

Determining forward prices is relatively simple. Entering into a contract to buy an asset at a particular price at a future date is equivalent to borrowing money to buy the asset now, and repaying the loan at that date. Similarly, the counterparty could sell the asset now and put the money on deposit. The forward price of the asset should therefore be the current market price plus interest for the period.

The rate of interest to use will depend on market conditions, the bargaining power of the participants, tax, expenses and credit risks. The terms on similar listed future contracts will provide a first estimate for the interest rates to use. Because the market in long-term future contracts (over a year or two) is relatively thin, short term rates will have to be used to give an indication of the important considerations. The starting point in determining a fair rate would normally be the rate of return on riskless government stock – the standard assumption of financial economics. Those who take the short position in future contracts are able to participate in the market at this rate, so should be given a higher rate for participating in these illiquid forward contracts. They are also taking some counterparty risk, suggesting that a further some margin should be added.

Policyholders taking the long positions would probably be subject to borrowing rates significantly higher than the risk free rate – even though their position would be secured by the surrender values of their policies.

This would suggest that the pre tax interest rates should perhaps be somewhat higher than those for good quality corporate debt. The rate on home loans might be considered as a retail analogy for a secured loan, and thus function as a ceiling.

4.3 Smoothed maturity value

The smoothed payout at maturity of those units committed to forward contracts would be determined by the following formula –

$$Maturity \quad value \quad = \sum_{t=1}^{n} U_{-t} P_{-t} (1 + i_{-t})^{t} \qquad (1)$$

Where:

maturity takes place at $t = 0$

n is the smoothing term

U_{-t} is the number of units committed to forward contracts at time $-t$

P_{-t} is the market price of the units at time $-t$, which includes reinvested dividends.

i_{-t} is the spot rate of interest of term t at time $-t$

Insurers, or the policyholders, could choose an appropriate value for the smoothing term n, and formula for U_{-t}.

An obvious formula would be:

$$U_{-t} = 1/(t+1) * [TU_{-k} - \sum_{t=k}^{n} U_{-t}] \qquad (2)$$

Where:

TU_k is the total number of units allocated to the policy at time k.

For policies with no premiums during the smoothing term, this would provide for an equal number of units to be committed to future contracts in each of the n periods. Where premiums are still being paid during the smoothing term, this formula would produce an increasing the number of units as the policies near maturity. If a more equal weighting over the period was wanted in such cases, an appropriate adjustment to formula (2)

would need to be made for the units that were expected to be bought by the unpaid premiums.

4.4 Anti-selection

The approach described here allows for the determination of fair, market consistent, surrender values without the anti-selection risk that some policyholders will be able to trade their units on terms disadvantageous to others. This is because the forward prices can be worked backwards to determine their current market value for purposes of surrender. The surrender value at time k before maturity will be:

$$SV_k = \sum_{t=k}^{n} U_{-t} \; P_{-t} \; \frac{(1+i_{-t})^t}{(1+i_{-k})^k} + [TU_{-k} - \sum_{t=k}^{n} U_{-t}] * P_k \qquad (3)$$

Where:

SV_k is the surrender value at time k.

Time 0 would still be the planned maturity date, 0 < k < n

Another type of anti-selection problem arises if new policyholders wish to take out a policy with a term of less than n, the smoothing term. A similar problem exists if an existing policyholder wishes to pay an additional premium in the last n years of the policy's term. These are analogous to a reverse surrender, and equation (3) would then form the basis for the number of units to allocate to the policy. The additional premium would be equivalent to the surrender value. There would be any number of ways of determining U_{-t}, the number of units that would be committed to forward contracts at the time the additional premium was paid. One approach would be for no immediate allocation, that is $U_{-t} = 0$ for all values of t less than k. An alternative would be for a proportionate allocation; that is $U_{-t} = TU_{-k}/n$ for all values of t less than k. TU_{-k} has to be determined as it is the number of new units to allocate in respect of the additional premium.

4.5 Effects on investment returns

It can be seen that the smoothed return is equivalent to that obtained by "lifestyle" disinvesting from equities, and buying zero coupon fixed interest assets as maturity approaches. As discussed above, this forgoes the equity premium as retirement

approaches. The difference is that this algorithm provides for gearing in the initial years of the policy that allows for the recapture of the equity premium that will be foregone later. It is unlikely that the two will exactly counterbalance each other, but they should be of broadly the same value *ex ante*.

The proposed algorithm has other advantages over the lifestyle approach because most of the transactions should be internal (between generations of policyholders). This reduces the administration and its costs, saves the costs of brokerage on dealing, and saves the difference between the interest rates on borrowing and investments.

5 CREATING ALTERNATIVE INSTRUMENTS

This model can also be used to create new investment instruments.

5.1 Inflation linked assets

Inflation linked instruments of appropriate term can be created by the formula:

$$Maturity\ \ value = \sum_{t=1}^{n} U_{-t}\ P_{-t}\ (1 + r_{-t})^{t}\ (CPI_0 / CPI_{-t}) \qquad (4)$$

Where:

CPI_t is the consumer price index at time t

r_{-t} is the spot real rate for term t at time $-t$

This means that the internal lifestyle investments obtain an inflation linked return, and that the policyholders acting as counterparties borrow at an inflation linked rate. The former is not a new instrument; the latter is. This removes the risk to both parties of unexpected changes to the rate of inflation.

5.2 Other linkages

There is no reason to stop there. Inflation could be replaced by a wage index, GDP per capita, or any other suitable index that allowed pensioners to participate in a country's overall prosperity. In each case, the rate of interest would need to be suitably adjusted to take into account perceptions of the relative risk and expected growth of the index.

The expected rate of interest used would be higher because of the systemic risk involved. Younger policyholders might be prepared to pay the higher rates because they would be hedging their exposure to the volatility of the growth in wages or GDP per capita.

5.3 Dividend links

Another method would be to use an index of dividend growth and the dividend yield. This yields the formula:

$$Maturity\ value = \sum_{t=1}^{n} U_{-t}\ P_{-t}\ (1+dy_{-t})^{t}\ (D\ index_{\ 0} / D\ index_{-t}) \qquad (5)$$

Where*:*

dy_{-t} is the dividend yield on a share index, and

$D\ index_t$ is measures the rate of growth in dividends on those shares.

This method gives a rate of return of somewhat more than the dividend yield plus the rate of growth of the dividend index. This method is suggested, not so much for its smoothing benefits, but because it can be reconciled with the traditional actuarial approach to dividend smoothing discussed earlier. The units used in this chapter first need to be redefined so that that the dividends are no longer reinvested in the units. This allows us to formulate a new price related to the dividend index:

$$P_t^1 * dy_t = D\ index_t$$

Equation (5) can then be reformulated as follows:

$$Maturity\ value = \sum_{t=1}^{n} U_{-t} P_{-t}^{1}\ (D\ index_0 / D\ index_{-t}) + accumulated\ dividends$$

The middle term can then be developed:

$$\sum_{t=1}^{n} U_{-t} P_{-t}^{1} \ (D\ index_{\ 0} \ / \ D\ index_{\ -t})$$

$$= \sum_{t=1}^{n} U_{-t} \ (D\ index_{\ 0} \ / \ dy_{-t})$$

$$= P_0^1 * dy_{\ 0} * \sum_{t=1}^{n} U_{-t} \ / \ dy_{-t}$$

$$\leq P_0^1 * dy_{\ 0} \ / \ (Average \quad dy) \qquad (6)$$

$$\approx P_0^1 * dy_{\ 0} \ / \ (Long \ \ term \ \ dy) = Actuarial \quad Value$$

The conclusion to be drawn from this development is that the actuarial value is an approximation to an average of forward contracts. The inequality (6) is greater if the dividend yields vary significantly. The traditional approach effectively removes the benefits of dollar cost averaging, but is otherwise not necessarily unfair.

6 AN ILLUSTRATION

The graphs below examine the effects of applying this algorithm in South African over the past 3 decades.

6.1 Smoothing achieved

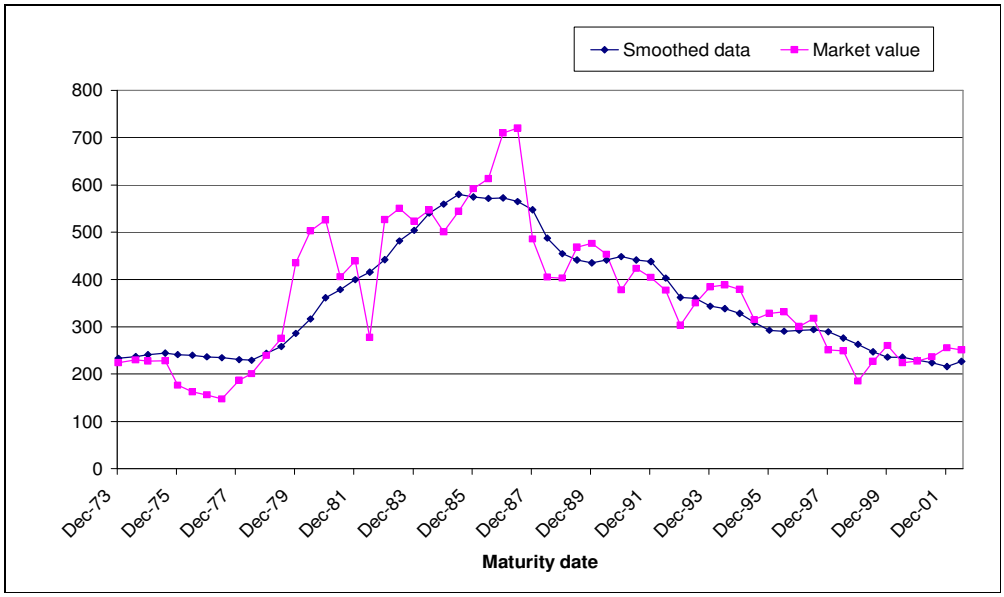
Given relatively high and volatile rates of inflation, nominal spot rates were inappropriate so equation (4) in section 5.1 was used. Real rates of interest were not available over the period, so dividend yields were used instead. A value of n = 60 months was used.

The results shown in figure 1 overleaf are encouraging. It shows payouts at six monthly intervals of an investment of 1 per month for ten years. The large payouts in the eighties especially reflect two decades of inflation in excess of 15% p.a. Maturing smoothed policyholders would have received much smoother returns than linked policyholders. They would have been protected from much of the share market slump of the seventies, and the sharper stock market cracks of 1982, 1987, 1992 and 1998. They would not however have benefited from all the higher prices of the late seventies and mid eighties.

The results do not appear to be particularly sensitive to adjustments in the smoothing term, real yields or formula.

This smoothing was to be expected. If we were using a single premium policy and the forward contracts were spread equally over 60 months, the monthly variance in payout (ignoring the effect of interest rate changes) would be reduced to 1/√60 or about 13% of that of an unsmoothed return.

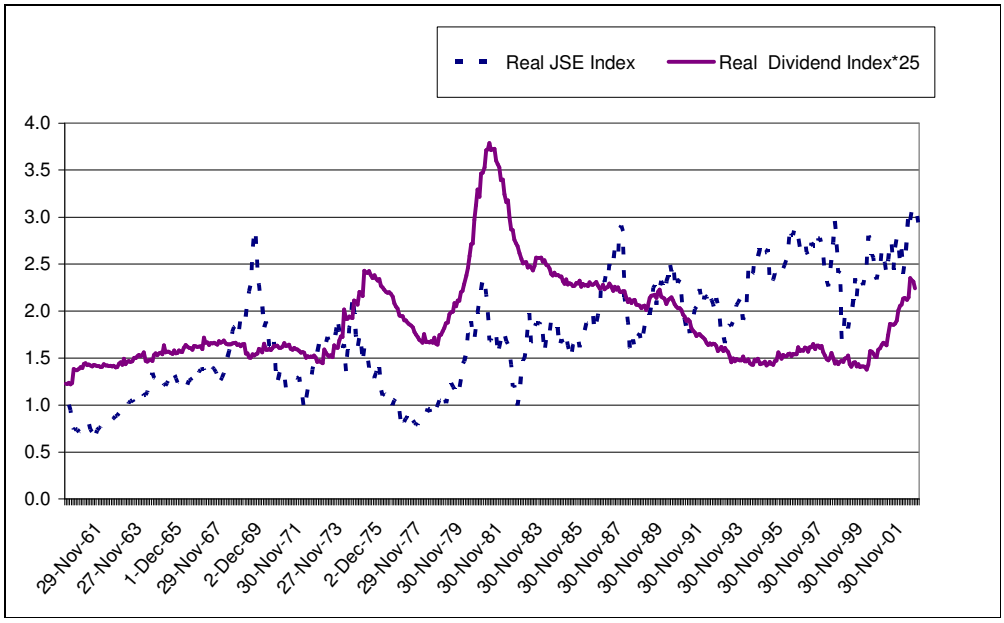
Figure 1: Smoothing (d/y + inflation) - 10 year monthly JSE investment



6.2 A problem with actuarial values

The actuarial values based on formula (6) seemed a reasonable approximation to this smoothing algorithm. Figure 2 shows the real values of the JSE All Share index, and the actuarial value of the assets if the assets are valued by treating the real dividend index as perpetuity at 4%. The actuarial value may be made more or less consistent with forward pricing. Smooth it is not because the dividend index is far from smooth. It may be smoother in more diversified markets, and there may be ways of ad hoc adjustment, but it does not appear suitable for smoothing in South Africa.

Figure 2: Smoothing using d/y + growth



7 OTHER ISSUES

7.1 Guarantees

Traditional with profit policies offer an underlying and increasing guarantee as well as smoothing investment returns. The discussion of Clay *et al* again raised long standing debates as to the value to policyholders of these guarantees, and the common failure to explicitly cost them. This algorithm will significantly reduce the cost of investment guarantees. The exact reduction in costs must be left to further research, but the reduction in variance makes it clear that it will be substantial. Of course, the smoothing is relatively short term – five years in the illustration. Guaranteeing longer term rates is not really possible.

7.2 Application to retirement funds

This algorithm can obviously be applied to DC retirement funds, life office annuities and pensions. The method of determining a surrender value would apply to withdrawals and early retirements. The method can be applied to the smoothing of lump sum payments and (with greater administrative complexity) to the smoothing of pension payments. Monthly smoothing of pensions would perhaps find little favour, but an annual

smoothing (where amounts were disinvested annually and spread over the year) could well provide an acceptably smooth income.

There seems no point in applying it to DB funds, however, as the employer would effectively be on both sides of the forward contracts.

7.3 Investment choice

The method can be applied to policies or funds that offer investment choice. The choice of fund will be the prerogative of the policyholders holding the long position in the forwards, as the maturing policyholders have locked in their returns for the period. The relative risk of the investments chosen might then effect their margin requirements and perhaps limit their ability to change their investment elections. Keeping track of such elections would add another, but not insuperable, layer of complexity to the administration of the funds.

8 CONCLUSION

Smoothing is recognised as a useful element in managing the risks of investing for retirement. The main mechanisms used in the past have been DB pension funds and smoothed bonus policies. They both suffer from being unclear as to the effects of their operation, and open to abuse.

Smoothing with the forward algorithms suggested here provides an objective and fair manner of smoothing that efficiently mimics lifestyle investment strategies. The contracts can be marked to market so they are not vulnerable to anti-selection, and do not lock members into long term contracts. It also allows for the creation of a number of alternative linkages that can be used to smooth income in retirement.

When developing alternative smoothing algorithms, both theoreticians and practitioners could use the formula developed here to check on fairness and resistance to anti-selection.

REFERENCES

Asher A (1991) Actuarial Implications of the Liberty/Prudential Merger *Transactions of the Actuarial Society of South Africa* IX I: 241-267

Asher A (1994) Managing the Financial Costs of Housing *Journal of Actuarial Practice* 2.1: 125-144

Asher A (2000) Fiduciary Duties and Good Faith in the Management of Retirement Fund Surpluses *The South African Actuarial Journal* 1: 1-33

Blake D (2000) Does it Matter What Type of Pension Scheme You Have? *Economic Journal* 110F: 46-81

Blake D: AJG Cairns & K Dowd (2003) Pensionmetrics 2: Stochastic pension plan design during the distribution phase *Insurance: Mathematics & Economics* 33.1: 29-48

Booth P & Y Yakoubov (2000) Investment Policy for Defined-Contribution Pension Members close to Retirement: An analysis of the "lifestyle" concept *North American Actuarial Journal* 4: 1-19

Cairns A J G, D Blake & K Dowd (2003) Stochastic Lifestyling: Optimal dynamic asset allocation for defined contribution pension plans *Discussion Paper PI-0003* The Pensions Institute, Birkbeck College, University of London

Clay G D, R Frankland, A D Horn, J F Hylands, C M Johnson, R A Kerry, J R Lister & L R Loseby (2001) Transparent with-profits: freedom with publicity *British Actuarial Journal* 7: 365-465

Dimson E, P Marsh & M Staunton (2004) Irrational Optimism *Financial Analysts Journal* 60.1: 15-25

Head S J, D R Adkins, A J G Cairns, A J Corvesor, D O Cule, C J Exley, I S Johnson, J G Spain & A J Wise (2000) Pension Fund Valuations and Market Values *British Actuarial Journal* 6: 55-141

Khorasanee M Z (1995) Simulation of the Investment Returns for a Money Purchase Fund *Journal of Actuarial Practice* 3.1: 93-116

Khorasanee M Z & H K Ng (2000) A Retirement Plan Based on Fixed Accumulation and Variable Accrual *North American Actuarial Journal* 4: 63-79

Knox D M (1993) A Critique of Defined Contribution Plans using a Simulation Approach *Journal of Actuarial Practice* 1.2: 49-70

Thomson R J (1997) Strategies for the Management of a Defined Contribution Fund *Transactions of the Actuarial Society of South Africa* XI: 309-343

Welch I (2000) Views of Financial Economists on the Equity Premium and on Professional Controversies *Journal of Business* 73.4: 501-37

CHAPTER 9 - CONCLUSION: METAPHYSICS AND MOTIVATION

This conclusion attempts to outline how the vision given in the introduction, and spelt out in the body of the thesis, can be extended and applied by further research and action.

The earlier chapters have shown how the institutional structure of the retirement fund industry (the legislation, the organizations, the benefits offered) can be reformed to remove injustice and better serve the beneficiaries. This final chapter considers how those who might contribute to reforms could develop the motivation and the wisdom to do so.

This chapter suggests first that that motivation can come from the discovery of vocation. The idea of vocation, or a calling, can be seen as archaic, and based on a different metaphysic to that now dominant in actuarial and economic research and practice, but it provides a way in which energy and courage can be developed and focussed on social reform and other desirable outcomes. While the idea of vocation can be discussed without reference to religion, I find such discourse unsatisfying and inadequate. I therefore include in this section of the chapter some of my own Christian experiences and views. For those who do not share the views, I hope the discussion will help to clarify differences, even if it fails to persuade.

The next section of the chapter turns to the question of wisdom, another term that might be described as archaic, but which includes both the knowledge and personal skills to effect change and the ability to commit to action in the face of uncertainty.

The final section describes, in more detail, possibilities and imperatives for research, practice and regulation.

1 VOCATION

Hall and Chandler (2005) examine vocation as a means of ensuring "psychological success" in a career. They report on other research that correlates objective career

success with people's subjective interpretation of their careers, and suggest that it is simplistic to assume that the subjective follows the objective by some type of *ex post* rationalization. People who set personal goals do achieve some of them and feel better as a consequence. Those who respond to a calling do, as a consequence, create the objective outcome of their own careers. Hall and Chandler find that the basis for vocation can be religious or secular but the idea clearly has religious roots. Chamberlain (2004) describes how sociology has traced the "Protestant work ethic" to a religious understanding of vocation, and how Catholic thought has arrived at a similar view of vocation as having an application to secular as well as religious careers during the past century.

Hall and Chandler also provide some discussion of how people develop a sense of vocation. They mention evidence of unique and applicable gifts; a strong sense of involvement or passion; how the vocation is difficult to find and articulate; and how it provides a sense of purpose. So much can be agreed by secular and religious positions, but animating the concept seems to me to require that people be invited to participate in a particular position, or at least be given enough detail to understand the emotional as well as the intellectual attractions. Hall and Chandler certainly give an example in their paper. It may, therefore, be helpful to share my sense of vocation in the hope that it may provide assistance to others in defining their own.

My involvement in the retirement industry was sparked by a sense of outrage at the unjust treatment of withdrawing members from defined benefit funds, and later flamed by the injustice of the means test and its consequences for lower income people. These two structures – or more accurately the indifference of those who derive a healthy living from these same arrangements and have the power to change things - still have the power to get my adrenalin flowing.

My tendency to anger was however rightly admonished by a friend quoting James[132]:

“So then my beloved brethren, let every man be swift to hear, slow to speak, slow to wrath; for the wrath of man does not produce the righteousness of God.”

While I have often been slow to hear, the bibliographies of the foregoing chapters

---

[132] 1:19 in the New King James version.

indicate that I have been diligent in seeking out the voices - from different disciplines and viewpoints - of those who have written on the subject. I have also heard from a wide variety of participants in the retirement industry, not necessarily from diligence, but from responses to two compulsions. For the first my personal propensities, family, school and career choice spoke as one: professional duty. My father's example was reinforced by the school's motto “Orando et laborando” (by prayer and by work), and the Francis Bacon quote that at one stage appeared on Institute of Actuaries publications:

“I hold every man a debtor to his profession; from the which as men of course do seek to receive countenance and profit, so ought they of duty to endeavour themselves by way of amends to be a help and ornament thereunto."

This should not suggest that I have not gained sustenance and pleasure from family, school or profession. The latter is particularly relevant here, as participation has been repaid by many warm friendships and congenial interactions. What contribution I have made in this area would have been inconceivable had I not felt I belonged inside the profession.

The other compulsion came from the Old Testament prophets and the South African Churches, who both insist that we should listen to the poor. Interactions here were as often intense as congenial, but often more rewarding for it. In particular, I was exposed to two emotionally and intellectually seminal influences in the middle of 1975. The first was the ten day conference of the Anglican Students Federation, the only student organization of the period that had not split into white and black components. I was often left in tears at the intense pain that was expressed; and more were shed as I worked through the second major influence, the SPROCAS (Study Project on Christianity in an Apartheid Society) reports[133]. These were a careful analysis of racial oppression in South Africa and the possibilities for reform. They were written by some 150 largely white and Christian intellectuals. Perhaps because they were also largely English speakers with no influence on government, the main conclusion seemed to be that blacks would have to win their own liberation. The thinking of the black Anglican students, and the SPROCAS participants, was informed by the dialectic of black

[133] Published by the Christian Institute and the South African Council of Churches

consciousness enunciated by Steve Biko:

"The thesis is in fact a strong white racism and therefore, the antithesis to this must, ipso facto, be a strong solidarity amongst the blacks on whom this white racism seeks to prey. Out of these two situations we can therefore hope to reach some kind of balance - a true humanity where power politics will have no place."[134]

James, SPROCAS and black consciousness converged: one should not presume to speak for God or the poor unless specifically called. My calling seemed: become an actuary, "mind your own business"[135], and wait to be called.

In the process, I also made myself available to serve and to undertake some research. For the former, I became an occasional advisor on retirement issues to the Black trade unions, one consequence of which is described in chapter 6 of this thesis. The research began with Asher (1987), which was more an opinion piece on withdrawal benefits than a research paper, but it bears a similar title to this thesis and began my involvement with retirement fund policy: in particular six years as member of the Pension Fund Committee of the Actuarial Society of South Africa followed by as many years on the Ministerial Advisory Committee on Pension Funds. Amidst the more pressing issues of the times, my membership of these committees allowed me to suggest a link of the withdrawal benefit issue to the more pressing question of surplus refunds. This may have provided the seed that led to the 2001 reforms which, *inter alia,* imposed minimum withdrawal benefits on all retirement funds. I do not know how important my role was in this sequence, but it does provide satisfaction to know that I participated actively. I mention it here because it illustrates the long gestation period required for policy reform, and provides a reason to set ambitious long term goals.

In addition to cultivating patience, there is other personal work that may well be necessary in finding and responding to a calling. Luke's beatitudes (6:20-22) all speak to me.

---

[134] Steve Biko (1971) *Black consciousness and the quest for a true humanity* http://www.sahistory.org.za/pages/specialprojects/black-consciousness/biko/writings-humanity.htm

[135] As Paul has it in 1 Thessalonians 4:11

"Blessed are you who are poor, for yours is the kingdom of God.
Blessed are you who hunger now, for you will be satisfied.
Blessed are you who weep now, for you will laugh.
Blessed are you when men hate you, when they exclude you and insult you and reject your name as evil, because of the Son of Man."

While they have a literal meaning as words of comfort, I have always understood them to set standards of virtue. In order: humility, which includes separating from personal and class interests to listen to others; a hunger for justice; a compassion that weeps at evil; courage in the face of opposition.

2 PRACTICAL WISDOM

My sense of calling led to 14 years as Director of Actuarial Studies at the University of the Witwatersrand, Johannesburg. One of the reasons was to give myself time to learn from the academic literature in the fields related to actuarial work. I found it on the one hand surprisingly rich, but on the other disappointingly narrow. The narrowness arises, to my mind, partly from a reluctance to risk controversy even on a technical level, and more especially when linked to practical and moral issues.

The introduction to this thesis quoted Michael Polanyi on the necessity of a "committed and passionate" rather than disinterested approach to finding truth. It seems to me that such commitment is more likely to search widely for assistance and be concerned about applications. It should however be identified and disclosed. My personal commitment is to the Truth that provides meaning and beauty to our private and working lives: intellect and emotion; social, economic and political. My belief is that because they arise from God, the different aspects of our lives co-inhere in each other. This co-inherence[136] means that separating our knowledge into artificial compartments can produce a distorted reductionistic understanding.

I find much economic research and discussion to be reductionistic. As is illustrated in the introduction and chapter 8, the research is frequently narrow and dogmatic. One school of economists has difficulty with explanations of behaviour that are not based

[136] A term that derives from early Christian writers, but I first found it in Williams (1939).

on material self-interest, let alone social and ethical considerations. It seems to me that this materialist economics intersects almost entirely with assertions that true science is value free. Rothbard (1976) deals with this question in his discussion of the views of Ludwig von Mises, one of the more influential proponents of value free economics. He acknowledges the validity of Von Mises's point that one must accept the distinction between how things are and how they ought to be, and that experts in the one are not necessarily experts in the other, but then shows that von Mises's attempt to support both *laissez faire* and value free economics simultaneously is not consistent. Free markets ought to be supported because they embody the value of liberty and appear more likely to be more efficient and reward industry. These are clearly value judgements, as are other views that inequality and failure to provide for everyone's needs, which frequently arise from the operation of a free market, are problems.

Similarly Gouldner (1961), in discussing sociologist Max Weber's influential espousal of the value free, suggests that "the myth of a value free sociology was Weber's way of attempting to adjudicate the tensions ... between science and religion..." Gouldner echoes a suggestion in Polanyi (1962) when he says of the value free notion: "it invites men to ignore the vulnerability of reason to bias, leaves it at the mercy of irrationality." The irrationality arises because neither self-interest nor moral indignation, which often provide emotional drivers for debate, are recognised for what they are.

Conflicts between science, ethics and religion can represent a problem, but seem to me to derive from the relationships of the people concerned rather than the intellectual inconsistencies that emerge from time to time. The solution lies in the rules of engagement rather than insisting on a cessation of debate. "No politics, sex or religion" may be an alternative if participants in a discussion cannot stick to rules of courtesy, but if we want the truth, we have to be prepared to try to understand and interact with reasonable people with whom we disagree.

In applying the distinction between the objective and subjective to actuarial work, Thomson (2004) distinguishes between science, which includes the pursuit of knowledge, and wisdom. He suggests that wise criteria are "essentially subjective", although he is making the case for utility theory to be seen as normative and limits the subjective to the shape of the utility function. It appears to me that, in the absence of a coherent alternative, utility theory more than provides enough criteria to decide

whether a particular investment choice is wise or not. Similarly, it is suggested that justice is a universal standard that requires those exercising decisions of power to consider a definable range of criteria. People may trade the criteria off against each other, but they must all be considered. If they fail to consider all the criteria, the decision is procedurally unjust. And academic research is foolish if it identifies knowledge and techniques but fails to situate them in a framework that can lead to wise decisions.

What then is wisdom? Sternberg (2002) has edited what seems to be the most cited of modern books on wisdom, where a number of the authors lament the decline of the concept. Chandler (1990), for instance, referring to well-known philosophers for support, "argues that our modern conception of the knowing process is a much fragmented and highly restricted version of a once more elaborate conceptual scheme ... of the notion of wisdom and the existence of wise persons." There is however a clear recognition within the book that wisdom integrates the cognitive, the affective and the moral, and makes skilled practical judgements in the face of uncertainty. Wisdom is the object of learning.

I think Chandler is overly pessimistic. In my experience, wisdom is widely valued outside the social sciences. Many of my teachers clearly sought and valued wisdom, and "sound actuarial judgement" is frequently used by members of the profession. With both teachers and actuarial colleagues, there has been a significant intersection with my experience of other Christians, and of sharing together in a search for wisdom. While there is an apparent inconsistency between the biblical injunction that "the fear of the Lord is the beginning of wisdom" (repeated throughout the "wisdom books": Job, Psalms, Proverbs and Ecclesiastes)[137] and the secular method of exposure to alternate views and debate, both include instruction and interaction with a teacher. While not a Socratic dialogue, the raging debates of the book of Job and the volatile intensity of the Psalms do force active consideration of different points of view. Thus it seems to me that controversy is necessary to create the ability to consider alternative positions and to wrestle with them until there is some reconciliation or integration. As such it is necessary for the development of wisdom.

---

[137] See especially Proverbs chapters 8 and 9.

This would also mean that different intellectual and professional disciplines do, or at least should, hold together. Pension fund design must surely be informed by sociology and gerontology as well as the health and management sciences. Modern disciplines however seem to define themselves quite narrowly. Thomson (2004), for instance, also seems to find it necessary to explicitly justify the inclusion of normal actuarial work in a scientific thesis. This is perhaps partly a consequence of the relatively recent introduction of actuarial studies into the university curriculum. I also interpret it as a defence against reductionistic arguments that research is only legitimate if it conforms to a narrow interpretation of the scientific method.

As with justice, I would argue that the value of research is similarly to be considered in terms of a range of criteria. These include the scientific criteria of logical consistency, objectivity and independence and the need to test theory against other theories and against data. The criteria should also include respect for the interests and opinions of others who may be affected; relevance to real world problems and decision making; aesthetics: in which I would include elegance, enthusiasm and clarity; as well as the ability to communicate and persuade.

To those who would argue that these are not the traditional academic objectives, my response would be that their tradition is too short and too narrow. Universities by their name proclaim their founders' objective of the pursuit of universal truth and, by their contemporary nature, are primarily concerned with the professional development of their students, and not the narrow perspectives within which some researchers appear to confine themselves. Specialization does not mean tunnel-vision.

Most importantly theory and practice ought to co-inhere: “good theory is good practice aware of itself” as I was once told. In my experience, Church and duty demanded more than listening. It could be of no value unless there was praxis - practical consequences. The next section from James has it: “be doers of the word and not hearers only, deceiving yourselves.”

Wisdom can of course be found in some of the published economic and actuarial research. My problem could alternatively be framed as a disappointment that so many of the researchers appear so poorly read. Wisdom requires a balance of the materialist economists and the World Bank literature with Herbert Simon and the institutional economists, the International Labour Organization and wise writers such as Thompson

(1998) and Barr (2001) in the modern pensions literature. Some would also benefit from reading Adam Smith - both "Wealth of Nations" (1776) and his earlier book on ethics (1759).

Finally, in considering practical consequences, I would suggest that we need a theory of sin and forgiveness. This is necessary to connect what is with what ought to be and what can be. One of the words translated sin is an archery term that means falling short. It refers to failure in relationships, thinking and working: intellect and emotion; social, economic and political. Some failure is inevitable, but no progress is possible unless we are agreed that there was indeed failure, and that we have the capacity to change ourselves and our world. The Christian concept of grace is that we can be forgiven if we repent of our sins. And while our prime responsibility is to look to our own actions, a concept of sin also permits of a prophetic vocation to hold the powerful accountable for the justice of their actions. For those not persuaded of the Christian position, Braithwaite (2002), writes of the importance of both shame and restoration in a range of non-Christian cultures.

3 FUTURE DIRECTIONS

A few suggestions for researchers, practitioners and regulators are sketched out at the end of each of the chapters. With the whole picture now sketched, I can perhaps be bolder.

For researchers, the main thrust is that there are enormous possibilities for connecting economic research into how people are behaving with actuarial knowledge of what the financial services industry does and how it might be changed. Research should however break with the ideas that people are perfectly rational, or that we will find predictable irrational patterns in their behaviour. Behaviour is always likely to be more complex. While it is worthwhile finding out what people are doing to protect their financial security, it is as important to determine what they should be doing. As discussed in the introduction, contributions to a pension fund are usually made in ignorance of the real needs of orphans, the disabled and the elderly. We need to know what makes for the best protection from disruption and poverty. This is an ongoing quest as the conditions will change over time. Research has not only to look at the present but attempt to tease out future possibilities and develop strategies. As spelt out at the end of Chapter 2 above, such research needs to be based on panel studies.

Practitioners in the pension and life insurance industry are given a range of suggestions in chapters 5 and 8 on how benefits could be better designed to meet the real needs of members. The most common response I have to suggestions of this sort is that there is no demand for the ideas. Most often I believe that this response arises because people are overloaded, and have little time to respond to yet another new demand.

Sometimes the response is informed by the assumption that the purpose of business is to make a profit, and that this is most easily achieved by meeting the obvious demands of the market in the most efficient way. This is reductionistic. Drucker (1977) is I believe closer to the truth and more creative: "the main purpose of business is innovation". The function of management is to know: what is our business? The business is to meet particular needs of customers and potential customers. The needs are met by creating new products and services, and the profit is made by showing customers that their needs are better met. This is an aspect of business that can be particularly exciting, creative and remunerative. I hope particularly that some practitioners will be inspired by this thesis to develop products more appropriate to real needs.

As I understand the duty of regulators, theirs particularly is the duty to prevent injustice: the abuse of power. By regulators, I mean those in compliance roles within companies, especially statutorily appointed actuaries and auditors, and including outside directors, as well as everyone employed by industry and consumer bodies, and government. As to injustice, I have perhaps been too critical in this thesis but have at least spread the criticism widely. Injustice is to be found in civil servants who manipulate pension and tax rules for their own benefit, and not just in Africa. It is to be found in the profession and industry groups that lobby for concessions and compulsion that distorts markets and social structures. Executive management manipulate their income to many times their market value. The most apparently proper of trustees and their professional advisors can milk those they are supposed to serve in a fiduciary capacity. Financial advisors charge too much or give poor advice. Academics would prefer to play with mathematical models rather than address egregious wrongs in the industries they research. This much however is consistent with my understanding of the doctrine of sin, which does not rule out either forgiveness or hope.

Conscious that I may be too negative, I am comforted that established wisdom allows for some critical and negative feelings. As Adam Smith (1759) has it on the topic of indignation: “Resentment seems to have been given us by nature for defence, and for defence only. It is the safeguard of justice and the security of innocence”.

Regulators within companies particularly need courage to take what can be an unpopular line, but I should record warnings that it is important not to embark on an adversarial course without adequate personal support. Jos *et al* (1989), for instance, describe the significant costs paid by those who whistle blow. I do believe that Luke's beatitudes provide the model: listen, weep and hunger before risking oneself, and then private rebuke before open opposition: “the wrath of man does not produce the righteousness of God.”

Also consistent with the nature of sin is the propensity to become more concerned with injustices perpetrated by someone else. Against this, Eliot Spitzer provides a role model. He[138] does not claim to be doing anything more than his job, yet has made a significant impact on US financial services, clearing out conflicts of interest that had become entrenched in what appeared to be the most reputable of banks and insurers. While I do not have research to confirm this, it appears likely that the integrity of any system depends as much on thousands of smaller decisions to eschew or confront injustice - as on the occasional attorney-general with the insight, the courage and the political motivation, to take on an industry. And I am grateful to have worked with many people who have quietly, and not so quietly, made some of these courageous decisions and so protected us all.

Researchers, practitioners and regulators also have an educational role. Three particular areas covered in the thesis stand out for me. The first is to explain the need for insurance; the second to reinforce good investment theory - the value of fundamental analysis; and the third for the centrality in matters of governance of the principle that conflicts of interest must be avoided.

Finally, I ask myself whether these financial issues really justify this rather heavy conclusion, and whether the financial sector is that important in a world of devastating

---

[138] In Eliot Spitzer (2005) Business ethics, regulation and the "ownership society" http://www.oag.state.ny.us/press/statements/Business_Ethics.pdf

poverty, vicious war and global warming. Perhaps not, but it is the area to which I have felt called to offer my gifts. Those of us who labour in this sector can perhaps take some encouragement from the end of the first chapter of James:

"Pure and undefiled religion before God and the Father is this: to visit orphans and widows in their trouble, and to keep oneself unspotted from the world."

REFERENCES

Asher A (1987) Pension Fund Design Transactions of the Actuarial Society of South Africa VII.I: 85-102

Barr N (2001) *The Welfare State as Piggy Bank* Oxford University Press

Braithwaite J (2002) *Restorative Justice and Responsive Regulation* Oxford University Press

Chamberlain G L (2004) The Evolution of Business as a Christian Calling *Review of Business* 25.1: 27-36

Chandler M J (1990) Wisdom in a Post-apocalyptic Age *Wisdom: its nature, origins, and development* ed R Sternberg, Cambridge University Press: 121-141

Drucker P F (1977) *Management* Pan

Gouldner A W (1962) Anti-Minotaur: The Myth of a value-free sociology *Social Problems* 9.3: 199-213

Hall D T & D E Chandler (2005) Psychological Success: When the career is a calling *Journal of Organizational Behavior* 26: 155–176

Jos P H, M E Tompkins & S W Hays (1989) In Praise of Difficult People: A portrait of the committed whistleblower *Public Administration Review* 49.6: 552-561

Polanyi M (1962) *Personal Knowledge: Towards a post-critical philosophy* University of Chicago Press

Rothbard M N (1976) Praxeology, Value Judgments, and Public Policy *The Foundations of Modern Austrian Economics ed* Edwin Dolan, Sheed and War: 89-111

http://www.mises.org/rothbard/value.pdf

Smith Adam (1759) *The Theory of the Moral Sentiments* adamsmith.org/smith/tms/tms-index.htm

Smith Adam (1776) *An Inquiry into the Nature and Causes of the Wealth of Nations* adamsmith.org/smith/won/won-index.html

Sternberg R J (ed) (1990) *Wisdom: its nature, origins, and development* Cambridge University Press

Thompson L H (1998) *Older and Wiser: The economics of public pensions* Urban Institute Press, Washington DC

Thomson R J (2004) *The Development and Application of Actuarial Models* PhD Thesis, University of the Witwatersrand

Williams C (1939) *Descent of the Dove: A history of the Holy Spirit in the church* Oxford University Press