ASSET ALLOCATION DECISION MODELS IN LIFE INSURANCE

by

Alen Sen Kay Ong

A thesis submitted for the degree of Doctor of Philosophy

City University, London
Department of Actuarial Science and Statistics
October 1995

CONTENTS

FIGURES

TABLES

ACKNOWLEDGEMENTS

I am indebted to my supervisors Mr. Philip Booth and Dr. Robert Chadburn for their continued support and encouragement during these past three years, and for commenting on an earlier draft of this thesis.

I am grateful to Mr. Paul Huber for his invaluable help in developing the cash model, and to Mr. Michael Boskov for introducing me to the Jackknife. Thanks are also due to the many others who have contributed to this research in some way, including Professor Steven Haberman and Professor Iain Allan.

Finally, I wish to acknowledge the Association of British Insurers and the Association of Consulting Actuaries for funding this research.

DECLARATION

ABSTRACT

The problem of determining the optimal asset allocation strategies for a non-profit life company is approached from a rational decision-making framework. Initially, a number of methods for analysing investment risk are discussed, from which utility theory is felt to be the most appropriate. Stochastic simulation and numerical optimization methods are employed in order to allow more realistic assumptions to be used in these decision models.

The multiperiod consumption of dividends is dealt with by considering the expected utility of accumulated dividends, or payouts. At first, the case of an open fund is investigated in a static asset allocation framework. In general, the results produced are quite intuitive. At low levels of risk tolerance, the optimal portfolios seem reasonably matched in relation to the liabilities. As the risk tolerance level increases, the preference for matching is seen to reduce. If payouts are measured in real terms, greater proportions tend to be invested in the real asset classes. From a mean-variance perspective, the utility maximizing portfolios generally appear to be efficient. However, imposing insolvency constraints on the objective function can have the effect of shifting some of these portfolios away from the efficient frontier.

In the case of a closed fund, dynamic asset allocation strategies are investigated. Due to the restrictive assumptions it requires, the possibility of applying dynamic programming in this situation is rejected. Instead, it is proposed that the asset proportions be made functions of the duration of the liabilities, so that the expected utility may be maximized in respect of these function parameters. Overall, this appears to produce reasonable results, although the occasional emergence of less intuitive strategies leaves further scope for refining the treatment of multiperiod consumption.

1. INTRODUCTION

1.1 Background and Aims

According to Markowitz (1991), the theory of rational behaviour in respect of investment decisions may be characterized by:

1. the information relating to all the assets concerned
2. the criteria for selecting the most suitable portfolios
3. the procedures for computing these portfolios given the information available.

Although few, if any, may argue against the merits of such an assertion, practical experience however may suggest that investors do not always appear to follow these three steps when making investment decisions. While it is conceivable that a sensible investor will try to gather as much relevant information as possible before making an investment decision, the process of defining the appropriate criteria and its subsequent use in portfolio selection is rarely explicit. Thus, the work of Markowitz (1952, 1991) and other financial economists (see for example, Ziemba and Vickson, 1975) have mainly been concerned with formalizing stages 2 and 3 above. These have led to the general risk-return approaches commonly used in portfolio selection, such as mean-variance analysis.

In principle, the notion of determining investment choices based on well-defined objectives should be just as appealing in the context life insurance funds. But as the circumstances of a life fund are very different from that of pure asset funds, special consideration is needed in the case of the former. For instance, the presence of long term liabilities points to the question of long term investment strategy, rather than short term tactical manoeuvres. Consequently, the allocation by asset class is of greater strategic relevance in the management of a life fund than is the selection of individual securities within each of these asset classes. Moreover, there are a number of interested parties that need to be considered in the case of a life fund. These may include shareholders, policyholders, regulators and possibly others such as management, as well. While shareholders and participating policyholders are mainly interested in the regular distribution of surplus, all parties are concerned with the solvency position of the life office. In general however, these issues tend not to be adequately tackled by portfolio selection models commonly found in the literature.

Therefore, the aim of this research is to investigate a risk and return approach to asset allocation strategies for life offices. However, a more general approach to risk is taken than in the traditional portfolio selection models.

In this thesis, the basic approach used for deriving the optimal strategies is to maximize the expected utility of surpluses distributed, though these strategies are considered in a mean-variance framework for comparison. Where appropriate, the risk of statutory insolvency is also incorporated into the decision-making process. In order to provide a realistic representation of the liability profile, a model of a non-profit proprietary life office is developed and stochastic simulation techniques applied to calculate the shareholders' expected utility. The only stochastic variables used relate to the economic variables of inflation and the asset classes. For this purpose, a stochastic investment model is used. Numerical optimization methods are then employed to obtain the optimal strategies for different liability profiles and risk preferences.

It should be stressed that the aim is 'strategic' rather than 'tactical'. A strategic allocation usually refers to a policy best suited to the decision-maker's circumstances based on some long term view. This contrasts with tactical allocations, which are more to do with profit-making using shorter term predictions. For example, a strategic allocation could be equivalent to an immunized position; but if the investment manager anticipates a rise in yields in the near future, it would be a tactical decision to switch into shorter term gilts. Although both are important aspects in the management of any fund, this study will be restricted to one of strategic allocations alone.

The approach used attempts to generalize some of the more traditional actuarial principles, including those in Pegler (1948) and Redington (1952), by showing how their extreme positions of maximum return and minimum risk may be traded off. As mentioned before, it extends the work of financial economists by giving detailed consideration to long term liability structures. The models developed also allow for the complexities of multiperiod consumption and regulatory insolvency, thus progressing from the single period models of Wise (1984a, 1984b, 1987a, 1987b) and Wilkie (1985). This is facilitated through the application of stochastic simulation techniques which are increasingly being used in asset/liability studies (see for example, Ross, 1989; Hardy, 1993). These life office simulation models, however, are taken a stage further here by actually optimizing the decision variables concerned. In this case, these variables are the proportions to be held in each asset class. Suggestions are then given as to how the multiperiod consumption-investment problems addressed by Samuelson (1969) and Sherris (1992) may be dealt with in a life office context.

1.2 Structure of the Thesis

In the chapter that follows, a survey is given of the literature that has been published in the area of investment decision-making. As many of the approaches previously used may be framed in a utility maximization framework, the concept of utility theory is also analysed critically.

Chapter 3 deals with the investment model used in these investigations. The model is essentially based on that suggested by Wilkie (1986) and the main properties of this model are reviewed. Details are given of how this model may be extended to encompass additional asset classes. Simulation output from the resulting model is then analysed.

Chapter 4 introduces the general methodology used to derive optimal asset allocations for a simple asset fund with no explicit liabilities. Numerical optimization is briefly described and then applied to an asset fund problem. The potential sources of error in the optimization results are also considered.

Chapter 5 discusses how the methodology may apply to a non-profit proprietary life office. The assumptions used in the model office are detailed, including a description of the treatment of cashflows, dividend policy and insolvency. As the issue of multiperiod consumption of dividends will need to be addressed, shareholders' utility is discussed with particular regard to time preference.

Chapter 6 combines the methodology in the previous two chapters to arrive at the optimal asset allocations for a model office issuing only twenty year endowments. This office (Model A) assumes that shareholders will be prepared to provide additional capital to the office in times of crisis, i.e. statutory insolvency. Utility maximizing portfolios are obtained for different levels of initial surplus and risk preferences, with and without constraints on the probability of insolvency. All optimal portfolios are then analysed from a mean-variance perspective.

In Chapter 7, Model B is used. Here, instead of absorbing additional capital from its shareholders, the long term business funds are effectively 'sold' when the office becomes insolvent. As well as a liability profile of twenty year endowments alone, ten year endowments and a mix of index-linked annuities and twenty year endowments are also considered. If the proposed asset allocation methodology is effective, the optimal asset allocation should change with the nature of the liabilities in a broadly intuitive way. The results from using different liability profiles are discussed in this context.

Up until Chapter 7, the optimal asset allocation strategies are obtained assuming that the mix of assets remains constant over the period concerned, in an open fund situation. These optimization models may therefore be referred to as static models. In Chapter 8, the problem of allowing the asset mixes to change over time is discussed. Such models are often described as dynamic optimization models and can be applied to closed funds.

Neither analytical nor simulation based results have been found in respect of optimal asset allocation strategies for complex, static, open fund liability profiles in a utility maximization framework. Furthermore, solutions to dynamic asset allocation problems have previously been limited to certain, simple, special cases.

Finally, Chapter 9 discusses the main conclusions derived from this research.

2. THE ANALYSIS OF INVESTMENT RISK AND UTILITY THEORY

2.1 Introduction

The problem of investment risk has interested economists, financial economists and actuaries for decades and as a result much literature has been published on the subject. This chapter reviews some of the proposed approaches to analysing investment risk.

The chapter begins with the development of actuarial investment principles and immunization theory. Traditional mean-variance analysis is then introduced, followed by a discussion on how this has been extended to include simple liability structures. Assessment of insolvency risk through the use of stochastic simulation is then considered together with more general measures of downside risk.

Following this, utility theory is introduced: a fundamental concept of rational decision-making in the face of uncertainty. Utility theory is not mutually exclusive of the other approaches to investment risk mentioned and underlies the foundations on which most financial economic theories have been built. Due to its importance, both to this thesis and to the subject of risk itself, it was felt appropriate to discuss the theory separately and in some depth.

Based largely on the work in Fishburn (1988: Chapters 1 and 2), a brief historical review of expected utility theory is given. This is followed by a discussion of the main properties of utility functions and an outline of stochastic dominance. Lastly, the risk measures described earlier are analysed in the context of expected utilities, which draws from some of the issues discussed in Booth (1995b).

2.2 Review of Investment Risk

2.2.1 Early Actuarial Investment Principles

The discussion of investment strategy in insurance can be traced back to the nineteenth century, when Bailey (1862) suggested a set of investment principles to the actuarial profession, commonly known as Bailey's canons. He considered adverse fluctuations in asset values to be the primary risk to life funds. As a result, his first canon placed paramount importance on the security of capital. His second canon stated that subject to this, the aim of the life fund should be to earn the maximum practicable rate of interest on its investments. As higher returns may not be accepted at the expense of higher risk, the implication of Bailey's canons is a minimum risk strategy.

Pegler (1948) continued the discussion on " ... *the actuarial principles on which the investment policy of life assurance offices should be founded* ", recognizing some of the shortcomings in the canons proposed by Bailey (1862). Pegler criticized Bailey's emphasis on the security of capital on two grounds. Firstly, he made the valid point that a secure investment which provided a return lower than the yield assumed in the premium basis would be nearly certain to produce losses and was therefore unacceptable. Secondly, he also suggested that an asset which produced a reliable income would also maintain its capital value, hence achieving this objective. With the benefit of hindsight, this latter assertion may seem less appropriate given the relatively poor performance from Consols since the 1950's.

Consequently, Pegler suggested that the components of capital and income should be considered together in terms of total returns. In addition, he felt that taking expectations would automatically allow for the probability of achieving poor returns. Hence, Pegler proposed that the aim of life funds should be to earn the maximum expected return on investments. Nevertheless, Pegler also proposed that investments should be spread over the widest possible range in order to minimize the exposure to adverse outcomes due to common causes.

In an attempt to reconcile the positions of Bailey (1862) and Pegler (1948), Clarke (1954) emphasized the need to balance risk with reward. He essentially agreed with Pegler in that the security of investments should not be of paramount importance but also disputed the proposal that maximizing expected return could adequately allow for this. Clarke believed that neither principle should be held in preference to the other and that the conflict may only be resolved through personal judgement. Furthermore, he suggested that the assets should, as far as is practicable, be matched to the liabilities by duration and currency.

2.2.2 Matching and Immunization

Shortly before Clarke's discussion, Redington (1952) examined the interaction between assets and liabilities, deriving his theory of *immunization*. According to Redington's theory, a fund may be immunized against any loss due to small movements in interest rates if certain conditions hold. The derivation of these conditions is shown below:

Let L_t be the liability outgo (claims and expenses less premiums) for year t and let A_t be the asset proceeds (interest plus maturing investments) for year t. Define V_L and V_A to be the present value of liability outgo and asset proceeds respectively at the force of interest δ, and assume that $V_L = V_A$.

Suppose the force of interest changes from δ to $\delta + \varepsilon$. Let the present value of liability outgo and asset proceeds at this revised force of interest be V_L' and V_A' respectively. By Taylor's theorem, the amount of surplus that may arise from this change in the force of interest would be:

$$V_A' - V_L' = (V_A - V_L) + \varepsilon \frac{d(V_A - V_L)}{d\delta} + \frac{\varepsilon^2}{2!} \frac{d^2(V_A - V_L)}{d\delta^2} + \ldots$$

If the objective is to ensure that no profit or loss is incurred as a result of this movement in the force of interest, then all terms involving ε would have to be zero. This may be achieved when every liability outgo is matched by an asset proceed equal in amount and timing, whereby the fund is then said to be absolutely matched.

However, providing that ε is sufficiently small so that third and higher order terms in ε are negligible, it should be possible to immunize the fund against loss if:

$$\frac{d(V_A - V_L)}{d\delta} = 0$$

and

$$\frac{d^2(V_A - V_L)}{d\delta^2} > 0$$

In other words, the conditions for immunization are satisfied when:

1. the discounted mean term of the asset proceeds and liability outgo are equal
2. the spread of the value of asset proceeds about the mean term is greater than that of the liability outgo.

The theoretical appeal of an immunized strategy is that it guarantees the ability of a fund to meet all its liabilities regardless of the direction in which the interest rate may move. However, there are a number of deficiencies with this approach. The immunized conditions only apply in respect of small movements in interest rates. Traditionally, the theory had been assumed to hold only if interest rate changes could be represented by parallel shifts in the yield curve. Although this may be extended to allow for deterministic non-parallel yield curve shifts (see Reitano, 1994), immunization may still fail in an environment with a stochastic term structure of interest rates.

Another drawback with the application of immunization relates to the availability of assets with adequately long discounted mean terms. For example, the discounted mean term of an irredeemable gilt is approximately $1/d$, where d is the discount rate. As it is unlikely that there would be assets available with much longer discounted mean terms than this, it may not always be possible to immunize long term liabilities, especially when coupon rates are high. The new 'strips' market (see Bank of England, 1995: 228) will, however, enable investors to effectively purchase zero coupon bonds and immunize more closely longer term liabilities. In addition, the theory only applies directly to fixed interest assets and fixed liabilities of the same currency. Hence, it excludes asset classes such as equities and property, and limits the applicability of immunization to with-profit funds.

Even in situations where immunization may be possible, it may not necessarily be the most suitable strategy to pursue. The portfolio would need to be continuously rebalanced in order to maintain the immunized position. This would be impractical and incur substantial transaction costs. As immunization works against profits as well as losses, it may sometimes be desirable to deviate from such a position in anticipation of a particular movement in interest rates, with the aim of making additional profits: immunization provides no framework for trading risk and return.

Despite its limitations, immunization still offers a theoretical basis for asset/liability management and its implications on life and pension funds have been assessed by others including Day (1966). In respect of life funds, his conclusions are broadly intuitive: that a life office should maximize returns subject to being immunized, but may deviate from this position in accordance with the level of free reserves available. This suggestion does not, however, indicate how free reserves may be used to trade risk with return. Day felt that pension funds should largely be invested in equities, both as a means of hedging inflation and because matching by term would be inappropriate in this context.

This issue of real liabilities was later revisited in Fellows (1981), who took a similar view to Day (1966) regarding the unsuitability of bonds as a match to pension fund liabilities. Nevertheless, Fellows argued against the use of equities and property towards this end, due to their volatile nature, and suggested cash as the most appropriate means of hedging real liabilities. With the introduction of index-linked gilts around that time, this would appear to have resolved the question relating to index-linked liabilities. But as pension fund liabilities are not strictly inflation linked, the justification for investing in index-linked gilts is still debatable.

Overall, it would appear that Redington's theory of immunization may at least provide a benchmark by which investment strategies of life and pension funds may be devised. In circumstances where an immunized position is adhered to, it may be viewed as a minimum risk position. But deliberate mismatching may be reasonable if adequate surplus is available, though it remains unclear as to how this additional risk should be quantified and how much risk may be considered acceptable. This thesis provides a framework for answering this question.

2.2.3 Portfolio Selection Models in Finance

The development of modern portfolio theory is best known from the pioneering work of Markowitz (1952). In his analysis of portfolio selection, Markowitz claimed that investors would generally try to maximize the expected return and minimize the variance of return, implying that variance could be used as a suitable proxy for risk. Hence the selection process could be reduced to the consideration of the means and variances of portfolio returns.

Using the means and covariance matrix of returns from the available assets, Markowitz showed how a set of *E-V* (mean-variance) *efficient* portfolios may be constructed. (A portfolio is said to be *E-V* efficient if there are no other portfolios with a higher mean and the same variance or a lower variance and the same mean than that portfolio.) The portfolio that best meets the risk-return preference of the investor may then be chosen from this efficient set or *frontier*.

One attraction of Markowitz's analysis is its simplicity. The method only makes use of the means, variances and covariances of returns rather than the entire distributions. These statistics are reasonably easy to compute from historical data and have obvious interpretations. The approach is also intuitive, supporting the widely held view that diversification is beneficial, and has been found to be useful in portfolio construction. In practice, mean-variance analysis is also not inconsistent with an expected utility approach (see Section 2.3).

Apart from its convenience and intuitive appeal, the mathematics of the efficient frontier has other useful properties, especially that of separability due primarily to Tobin (1958). In his example, Tobin assumed the existence of a risk-free asset and a set of risky assets, and showed that the optimal portfolio for any *E-V* investor will be a linear combination of the risk-free asset and an efficient portfolio of risky assets which is the same for all *E-V* investors. Hence, the problem may be separated into the two stages of deriving this mutually optimal portfolio of risky assets, and choosing the appropriate combination between this portfolio and the risk-free asset to suit the investor's risk preference. Separation theorem also underlies the foundation for the *Capital Asset Pricing Model* or CAPM as derived by Sharpe (1964) and Lintner (1965).

But despite being able to simplify the selection problem of many assets to one of just two assets, the efficient portfolio of risky assets still needs to be computed. As Markowitz's analysis had been aimed at the selection of equity portfolios, a practical difficulty with this approach relates to it requiring the entire covariance matrix of all the available assets. For an equity portfolio with say a thousand securities under consideration, this would mean computations involving over half a million covariance estimates. Hence, the model in that particular form was still impractical for such purposes given the limited computational capabilities available at that time.

Sharpe (1963), however, simplified the portfolio selection model by assuming that the returns of various securities are related only through some common index, I. Hence, if there are n securities available, then the return from security i, may be represented by:

$$R_i = a_i + b_i I + e_i \qquad i = 1, ..., n$$

where a_i and b_i are parameters and e_i is a random variable with mean zero. As indicated by Sharpe, the index may be any factor thought to be the most important single influence on the returns from the securities, such as the stock market index or the Gross National Product. (If the index used is the market portfolio, one may recognize this to be of a form similar to the CAPM.) The index value being a random component itself may be given by:

$$I = a_{n+1} + e_{n+1}$$

where a_{n+1} is a parameter and e_{n+1} is a random variable with mean zero. Such a formulation, usually referred to as Sharpe's diagonal model, effectively reduces the full covariance matrix to a diagonal matrix. This is significantly more efficient than Markowitz's original model as it only requires $3n + 2$ estimates as opposed to $(n^2 + n)/2$ estimates. Using a historical data sample of stock returns, Sharpe also showed that the diagonal model produced very similar results to those derived using the full covariance matrix. The apparent accuracy and efficiency of this model made it a practical alternative to the traditional Markowitz approach.

From an actuarial perspective, modern portfolio theory put Pegler's first two principles on a more solid footing. It encourages higher expected returns and aims to reduce the variance of returns through diversification. But in contrast to Pegler's principles, the mean-variance approach acknowledges the conflict between the two more explicitly and hence, the need to trade off risk with return. However, it only goes as far as identifying the efficient set of portfolios, relying on judgement to select the optimum portfolio.

Although the original mean-variance models were exclusively concerned with asset funds, the approach can be generalized to incorporate certain liability structures. In the following section, it will be shown how these models relate to asset/liability studies and the actuarial concepts of matching.

2.2.4 Actuarial Portfolio Selection Models

The main aim of Wise (1984a) had been to examine the subject of matching from a more mathematically rigorous standpoint. He denoted the expected net cashflows by the row vector, $\mathbf{l}^{\mathrm{T}} = (l_1, \dots, l_m)$, where l_j is the liability outgo at time j, and the set of base assets by n linearly independent vectors, $\mathbf{e}_1, \dots, \mathbf{e}_n$, where $\mathbf{e}_k^{\mathrm{T}} = (e_{k,1}, \dots, e_{k,m})$ and $e_{k,j}$ is the expected asset proceeds from the kth asset in year j. Defining the amounts held in each base asset by the vector, $\mathbf{x}^{\mathrm{T}} = (x_1, \dots, x_n)$, the vector of asset proceeds is $\mathbf{a} = \mathbf{x}^{\mathrm{T}}\mathbf{E}$, where the matrix $\mathbf{E} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$. If the accumulation factor to time m is F, then the amount of ultimate surplus, $S = F(\mathbf{a}) - F(\mathbf{l})$.

From this definition of accumulated surplus, Wise proposed two main forms of matching:

1. minimizing $E(S^2)$ with respect to $\mathbf{x}$.
2. minimizing $E(S^2)$ with respect to $\mathbf{x}$ subject to constraining $E(S) = 0$. (This Wise referred to as the unbiased match)

Hence, $E(S^2) = 0$, is a special case whereby the asset and liabilities are absolutely matched.

Although the cashflows in Wise's model span m time periods, the asset portfolio is only chosen at the start of the investigation period. Since the model does not allow for any further rebalancing of this portfolio during the period, it may be described as a static model. In this respect, it is similar to Markowitz's model and those which are discussed in Chapters 4, 6 and 7 of this thesis.

Extending his model further, Wise (1984b) investigated the matching strategy of a simple pension fund. When equities and gilts of varying terms were made available, Wise found a significant proportion of the unbiased match to be in equities, with the remainder being in long term gilts. The result seemed to contradict the conclusions reached by Fellows (1981), who suggested that equities would not be the most suitable match to pension fund liabilities. However, this may be explained to a large extent by considering the assumptions underlying Wise's equity model.

In his example, Wise assumed that the gap between the increase in dividends and the increase in inflation is constant and that the dividend yield also remains constant. He acknowledged that such an asset resembles index-linked gilts more than equities, which is mainly why 'equities' feature strongly in that matching portfolio. Nevertheless, Wise's analysis does show that matching by term is appropriate for index-linked liabilities. The difficulty in the case of pension funds however is that index-linked gilts with a long enough term to redemption may not be available in practice.

Wilkie (1985) showed how the work by Wise (1984a, 1984b) is actually a special case of the mean-variance approach. The problem was defined in portfolio selection terms with the liabilities taken as a negative asset, requiring the means, variances and covariances of each 'asset' to be calculated. In addition, Wilkie introduced asset prices into the selection process. By plotting the price-mean-variance or *P-E-V* space, Wilkie demonstrated that Wise's two matching criteria were not in general *E-V* efficient, as defined in modern portfolio theory. The price is an attempt to take into account inter-temporal issues and the trade-off between current and future surplus.

Wilkie also addressed the issue of investor preference. He defined $\lambda = \partial E/\partial P$, to be the marginal trade-off between expected surplus and price, and $\mu = \partial E/\partial V$, to be the marginal trade-off between expected surplus and variance. In doing so, each combination of (λ, μ) would then point to a unique point on the efficient frontier. Thus, an investor could select the optimum portfolio by choosing values of λ and μ appropriate to his or her needs, although there is no theoretical basis for doing so.

A further enhancement, referred to as the *k*-solvency region, had been suggested by Wilkie and involved the requirement that $P(\text{surplus} < 0) \leq \alpha$. Assuming that surplus is normally distributed with mean, E and variance, σ^2, this would then be equivalent to requiring $E - k\sigma > 0$, where $P(Z \leq -k) = \alpha$ and Z has a standard normal distribution. Although this does not generally lead to a single portfolio, it imposes a constraint around a feasible region within which the optimum portfolio may be chosen.

The contribution made in Wilkie (1985) was re-examined by Wise (1987a, 1987b). Wise (1987a) suggested that from the investment manager's point of view, the price of the portfolio is pre-defined and the objective would be to find a suitable mean-variance trade-off. For the actuary, however, the objective would be to optimize the balance between price and variance for a given amount of expected surplus. In Wilkie (1985), rather eccentric prices had been used to demonstrate the inefficiency in Wise's unbiased match. Wise (1987a) showed that this inefficiency would not be of practical significance if realistic prices were used. He also proved that the optimum portfolio comprised a linear combination of the unbiased match, the mean surplus and the level of risk, $-\partial V/\partial P$.

Using different assumptions about current market conditions, Wise (1987b) found the high risk portfolios to be quite sensitive to the varying conditions, unlike his unbiased match. This seemed to echo the concern expressed in Moore (1972) that " ... *the results of any portfolio selection method are very sensitive to the quality of the data.*" Wise felt that sensitivity analysis of the assumptions would be necessary before making any judgements on the results obtained using a mean-variance approach. He concluded that because of this, efficient frontiers may be difficult to justify for practical purposes.

Sherris (1992) reviewed the Wise-Wilkie model, describing how it may fit into a broader utility maximization framework. As in Wilkie (1985), liabilities were included as a negative asset, although Sherris removed the extra dimension of price. Therefore, the objective would be to maximize the expected utility of ultimate surplus for a given amount invested. Using an exponential utility function and assuming that surpluses were normally distributed, Sherris proved that the utility maximization portfolio would always be E-V efficient. Hence, the Wise-Wilkie model was shown to be a special case of the more general utility maximization approach. However, Wise's implied utility function has been found to have some counterintuitive properties (see Booth, 1995b).

In summary, the portfolio selection models discussed in this section have provided a useful contribution to the study of asset/liability modelling. They place matching and immunization in a more general setting and extend traditional portfolio selection models to problems involving liabilities. But like all mathematical portfolio selection models in general, these models are subject to some practical limitations. For instance, complex liability structures and distribution policies would not be feasible in models of this form. Moreover, the models are also path independent, meaning that the optimal strategy is not dependent on the intermediary surplus levels occurring at each of the sub-periods. This would clearly be a crucial aspect in devising the optimal investment strategy for any life office. One approach to dealing with solvency and other complex issues is to use stochastic simulation techniques.

2.2.5 Stochastic Simulation Models

In spite of the advances made with regards mathematical models, it remains a fact that analytical methods tend to rely on some fairly simplistic assumptions in order for these approaches to work successfully. Stochastic simulation methods on the other hand can provide far greater scope for realism by allowing less limiting assumptions to be used. However, greater care is necessary in interpreting results from simulation, in the context of the probability models used. More complex features may be incorporated into the investigations concerned, with the only restrictions, if any, being computational ones. Since the late eighties, the use of simulation techniques for actuarial purposes has become more widespread in the United Kingdom (UK). The publication of an actuarial investment model by Wilkie (1986) and the vast improvements in computational power seen over the last few years have contributed to this trend.

One application of simulation methods has been in assessing the effects of different investment or bonus strategies on the solvency position of a with-profit office (see Ross, 1989; Roff, 1992; Hardy, 1993). Apart from being able to deal with life insurance solvency in a fairly realistic manner, these models highlighted the possible variability in performance that could be expected in the future. In Hardy (1993), simulation techniques were also shown to reveal certain insolvency risks in offices that were not apparent from deterministic scenario testing.

A related application of simulation has been in the development of more sophisticated measures of solvency. Rather than computing the asset/liability ratio to assess solvency as is commonly done, MacDonald (1993) considered the concept of adequacy, meaning the ability of an office to meet all its liabilities if it was to be closed to new business. The notion of relative insolvency on the other hand has been investigated in Hardy (1994a, 1994b). This is an alternative approach to determining regulatory insolvency and may be said to occur when the performance of a particular office moves out of line relative to the industry as a whole. Although much of the research in this field relates to the solvency of life offices, simulation techniques have also been applied to the areas of general insurance (Daykin and Hey, 1990), pensions (Booth, 1995a) and investments (Booth and Ong, 1994).

2.2.6. Downside Risk Measures

The concept of insolvency or ruin probability mentioned in Section 2.2.5 above is just one example of a downside risk measure. Broadly speaking, downside risk refers to risk measured in relation to the incidence and often the intensity of unfavourable outcomes; the definition of an unfavourable outcome being based on some criteria. Clearly, ruin probabilities are only concerned with the incidence of outcomes in which the asset/liability ratios fall below one.

Although ruin probabilities as such do not have any immediate interpretation in the case of a pure asset fund (with no explicit liabilities), the probability of underperforming a target return may instead be used in such portfolio selection problems. For example, if x is a random variable representing the return on an investment and x^* is the target return, the investor could choose one of the following criteria:

$$\text{minimize } P(x \leq x^*) \qquad \text{(Roy's criterion)}$$

$$\text{maximize } x^* \text{ subject to } P(x \leq x^*) \leq \alpha \qquad \text{(Katoaka's criterion)}$$

$$\text{maximize } E(x) \text{ subject to } P(x \leq x^*) \leq \alpha \qquad \text{(Telser's criterion)}$$

These, often referred to as *safety-first* approaches, have been discussed in various sources including Pyle and Turnovsky (1970), Elton and Gruber (1991) and Tse *et al*

(1993). Safety-first criteria are often felt to be intuitive and objective concepts with which risk may be assessed. However, it may sometimes be necessary to consider the intensity with which adverse outcomes occur, as is the case with variances of returns.

One problem though, with using variance as a measure of risk is that it penalizes upside and downside variability equally. A better approach therefore could be to measure the variability of returns only when these returns are below some target level. This measure of downside risk is known as *semi-variance*. Hence, if x is the actual return and h is the target return, then the semi-variance of x may be defined as:

$$\int_{-\infty}^{h}(x-h)^2 dF(x)$$

From this, one may then obtain a mean-semi-variance efficient frontier. Although semi-variance has been suggested by many, including Porter (1974) and Markowitz (1991), to be superior to variance as a measure of risk, the latter has in the past been preferred to the former on the grounds of its mathematical tractability. However, with the immense computer power available at present, such an issue should now be of little consequence.

Clarkson (1989) attempted to generalize the features of investment risk by proposing a set of basic axioms from which investment decisions could be made based on downside risk measures. These axioms may be stated as follows:

1. risk is a function of the probability and severity of adverse outcomes
2. risk may be measured by a weighted function of these adverse outcomes
3. for a given expected return, investors should aim to minimize this risk
4. all investors have a threshold level of risk which must not be exceeded
5. investors should maximize expected returns subject to satisfying this threshold level
6. risk preferences may be accounted for using different weighting functions and/or threshold levels.

Combining axioms 1 and 2, Clarkson's measure of risk may be expressed in the form:

$$R = \int_{-\infty}^{h} W(h-x)dF(x)$$

where x and h are defined as before and the weighting function $W(s) > 0 \; \forall \; s$. Clearly, this is a more general downside risk measure than semi-variance. In fact, if $W(s) = s^c$ and $c = 2$, then R is identical to the semi-variance. Along similar lines, McKenna and Kim (1986) pointed out that setting $c = 1$ is equivalent to measuring the expected shortfall below h, while setting $c = 0$ would imply the probability of underperforming h.

In Section 2.3.4, it will be shown how Clarkson's risk measure, R broadly fits into a utility maximization framework. Requiring the threshold level to be satisfied, as stated in axioms 4 and 5, however, does imply a utility maximization approach but subject to constraints.

2.3 Utility Theory

2.3.1 Historical Background and Critique of Expected Utility

The development of utility theory is thought to have begun in the eighteenth century through the independent works of Daniel Bernoulli and Gabriel Cramer. Prior to this, risky decisions had generally been assessed on the basis of expected returns. However, this had been proved to be inadequate in evaluating gambles such as the St. Petersburg game. In such a game, a fair coin is tossed until a head appears. If n tosses are required, then a sum of 2^n is won. Hence, the expected return from the gamble is infinite, though most people would offer this game at a finite price.

Bernoulli and Cramer explained the St. Petersburg paradox by proposing that it is the expectation of the subjective value, or utility, of outcomes that should be considered, where the utility is an increasing but concave function. This is equivalent to the concept of diminishing marginal rates of substitution in economics. Bernoulli suggested a logarithmic utility function whereas Cramer proposed a power utility function.

Over two centuries later, von Neumann and Morgenstern (1944) devised a very different expected utility theory. Instead of being a subjective value on an outcome, utility was defined as a preference relation on a convex set which is assumed to behave according to the stated axioms of order, independence and continuity. It was from these axioms that expected utility had been derived, rather than being merely stated as was the case with the Bernoullian expected utility. With either derivation, the end result implies that the optimal decision is that which maximizes the expected utility of outcomes.

Despite being the backbone of much finance and economic literature, expected utility theory has been criticized. One of these criticisms relates to how a choice may be influenced by the way a question is asked, sometimes called a *framing effect*. Tversky and Kahneman (1981) demonstrated this in a life and death situation. Given a choice of two treatments programs p and q, individuals generally preferred p to q if the scenarios were phrased in terms of lives saved but favoured q to p when they were framed in terms of lives lost. This appears to contradict expected utility theory, which implies that decisions should depend on the probability distributions of alternatives and not how they are described.

Allais (1953) described a paradox where, given the choice between a certain event a and risky alternative b, most respondents picked a in favour of b. However, when faced with the choice between $c = \rho a + (1-\rho)e$ and $d = \rho b + (1-\rho)e$, where ρ is some carefully chosen parameter and e is a third event, the majority of them preferred d to c, which breaches the independence axiom stated by von Neumann and Morgenstern. This is not really a consequence of framing but may instead be attributed to one's affinity for certainty, commonly referred to as the *certainty effect*.

Other issues which utility theory is unable to address include the existence of nontransitive relations and probability preferences. An example of a nontransitive indifference relation, $\sim$ is when $x \sim y$ and $y \sim z$ but where x is preferred to z. Probability preference refers to an individual's (distorted) perception of probability. Experimental work has suggested that individuals tend to overvalue small probabilities and undervalue large probabilities.

In view of these criticisms, attempts have been made to generalize utility theory so as to account for some of its limitations. One example is *prospect theory* suggested by Kahneman and Tversky (1979). An important difference between this and expected utility theory is that the probabilities as well as the outcomes are weighted. However, more complex methods such as this can be difficult to implement objectively and are beyond the scope of this research.

Furthermore, it is important to appreciate that many of the criticisms directed towards expected utility really only pertain to special circumstances and should not have a material effect on asset allocation decisions in general. For instance, framing effects should not apply when the description of alternatives is fully transparent, which is the case in these investigations. Allais' paradox is not only based on having to choose between risky and certain outcomes, but also relies on rather extreme payoffs being made. Individuals may not be able to relate to these payoffs, particularly as any experiment in the area of utility theory involves hypothetical and not real money. In reality, few asset allocation decisions involve scenarios that may be predicted with certainty and outcomes that have such extreme consequences. Hence, maximizing expected utility should provide a reasonable basis for making consistent investment decisions under normal circumstances.

2.3.2 Properties of Utility Functions

Although utilities may take any functional form, there are a few general properties which most utility functions should satisfy in decisions involving wealth. The utility functions, $U(.)$ should be monotonically increasing ($U'(.) > 0$), meaning that individuals prefer more wealth to less. Secondly, they should be concave ($U''(.) < 0$), which corresponds to investors being risk averse. It is also usual for them to be continuous functions of wealth, though there may be situations where discontinuities may be included on intuitive grounds.

The extent to which an individual avoids risk may be reflected in the risk premium that is required by the individual before a particular gamble is considered to be acceptable. In particular, Pratt (1964) showed that any risk premium is approximately proportional to $\alpha(.) = -U''(.)/U'(.)$, where $\alpha(.)$ had been defined as the measure of risk aversion. The larger this value, the more averse the individual is to risk. Moreover, Pratt distinguished between the concepts of absolute risk aversion (which is equivalent to $\alpha(.)$ above) and relative risk aversion which will be defined in due course. Their meanings may be better understood by considering specific utility functions.

The quadratic utility function has some useful properties, and usually takes the form:

$$U(x) = ax^2 + x$$

The first and second derivatives are $U'(x) = 2ax + 1$ and $U''(x) = 2a$ respectively. In order for the function to exhibit risk aversion it is necessary for $a < 0$. However, the resulting function will then only be increasing for $x < -1/2a$, which implies that the quadratic utility function is only meaningful over this range. From $\alpha(x) = -1/(x + 1/2a)$, it is clear that the level of risk aversion is an increasing function of x while $x < -1/2a$.

In other words, the individual becomes more averse to risk and will demand higher risk premiums at higher levels of wealth, which seems counterintuitive.

The main advantage of using a quadratic utility function is its tractability. By reducing the decision criteria to just a function of the first two moments of the distributions concerned, closed form solutions to many problems may be derived. This also implies that maximizing the expected quadratic utility will lead to a mean-variance decision. Unfortunately, the restricted range in which the quadratic utility function is meaningful and the property of increasing risk aversion make it less appealing, especially in comparison with the exponential, logarithmic or power utility functions.

The exponential utility function is most commonly expressed as:

$$U(x) = -\exp(-ax),$$

where $a > 0$. Unlike the quadratic function, it has the desired properties of $U'(x) > 0$ and $U''(x) < 0$ for *all* real values of x. Another useful feature of the exponential function is that $\alpha(x) = a$, a property normally referred to as constant absolute risk aversion. This implies that the decision making process will depend on the amount invested and not initial level of wealth for a given investment, as illustrated below.

If an individual with an exponential utility function of wealth w invests an amount x at a random rate of return R, the utility maximizing decision is to maximize:

$$E[U(w + xR)] = E[-\exp(-a(w + xR))] = k.E[\exp(-axR)],$$

where $k = -\exp(-aw)$. Hence, the optimal decisions based on maximizing $E[U(w + xR)]$ and $E[U(xR)]$ will be identical.

Another commonly used utility function is the logarithmic function, given by:

$$U(x) = \ln(x)$$

Clearly, the function will only be valid over the range $x > 0$. But apart from this disadvantage, it has the required properties of $U'(x) > 0$ and $U''(x) < 0$. However, the most significant difference between this and the exponential utility function is in the measure of risk aversion. For the logarithmic utility function, $\alpha(x) = 1/x$, which implies decreasing absolute risk aversion, though it also has the property of constant relative risk aversion, defined in Pratt (1964) as $\rho(x) = x.\alpha(x)$. Having constant relative risk aversion means that decisions will depend on the proportion of wealth invested, and not on the starting level of wealth.

For example, if the individual with wealth w investing an amount x has a logarithmic utility function, the objective would then be to maximize $E[\ln(w + xR)]$. If the proportion invested is defined to be, $k = x/w$, then the objective function to be maximized may be expressed as:

$$E[\ln(w + kwR)] = E[\ln(w(1 + kR))] = \ln(w) + E[\ln(1 + kR)]$$

Therefore, given the proportion of wealth invested, the decision-making process will be independent of the initial amount of wealth.

The only other utility function that exhibits constant relative risk aversion is the power function:

$$U(x) = x^c,$$

where $0 < c < 1$. It may be shown that $\rho(x) = 1 - c$. As the measure of relative risk aversion is less than one, it means that power function is also less risk averse than the logarithmic function. If $c = 1$, this is simply a linear utility function and $\alpha(x) = \rho(x) = 0$, which implies risk neutrality.

It has already been mentioned that the utility maximizing decision for the quadratic utility function leads exactly to a mean-variance result. However, it would also be true to say that expected utility from the exponential, logarithmic or power utility functions may be approximated by a suitable function of mean and variance. In fact, there is much empirical evidence to suggest that decisions based on mean-variance functions are almost identical to most of their expected utility equivalents (see for example, Levy and Markowitz, 1979; Pulley, 1983; Kroll *et al*, 1984; Reid and Tew, 1986; Booth, 1995a). However, despite being somewhat justified in practice, mean-variance analysis has been criticized on theoretical grounds; one criticism being its inconsistency with stochastic dominance.

2.3.3 Stochastic Dominance

Put simply, a probability distribution function $F(x)$ is said to stochastically dominate another probability distribution function $G(y)$ if $E_F[U(x)] \geq E_G[U(y)]$, where $U(.)$ is any member of a particular class of admissible utility functions (see Hanoch and Levy, 1969; Huang and Litzenberger, 1988). If this admissible set applies to all non-decreasing utility functions, then F is said to have first degree stochastic dominance over G. The necessary and sufficient conditions for this as given in Hanoch and Levy (1969) are:

$$F(x) \leq G(x) \quad \forall\, x$$

and

$$F(x_0) < G(x_0) \quad \text{for some } x_0.$$

However, if F dominates G for all utility functions belonging to the class of non-decreasing and concave (i.e. risk averse) utility functions, then F is said to have second degree stochastic dominance over G. The necessary and sufficient conditions for this are:

$$\int_{-\infty}^{x} F(t)dt \leq \int_{-\infty}^{x} G(t)dt \quad \forall\, x$$

and

$$F(x_0) \neq G(x_0) \text{ for some } x_0$$

Returning to situations where mean-variance analysis may be inadequate, consider the following example which has been adapted from Hanoch and Levy (1969). Let X and Y be uniformly distributed random variables over the intervals $[b,d]$ and $[a,c]$ respectively, where $a < b < c < d$. In addition, assume that the interval for X is wider than that of Y, i.e. $(d-b) > (c-a)$. As X has a greater mean and variance than Y, both these distributions would appear efficient from a mean-variance perspective. But it is also clear that $F(x)$ is never less than $G(y)$, which by first degree stochastic dominance implies that X will always give a higher expected utility to Y as long as the decision maker prefers more to less. Hence, it is not necessary for a stochastically dominant random variable to have a lower variance than another.

In addition, Hanoch and Levy (1969) showed that the mean-variance criterion is also not a sufficient condition for second degree stochastic dominance. For example, it is possible to have two random variables X and Y, such that $E(X) > E(Y)$ and $V(X) < V(Y)$ even though $U(Y) > U(X)$ for some increasing and concave utility function. However, it may also be shown (see Huang and Litzenberger, 1988) that if $F(x)$ stochastically dominates $G(y)$ for all (increasing and decreasing) concave utility functions, then it implies that $E(X) = E(Y)$ and $V(X) \leq V(Y)$. In this respect, there is at least some theoretical basis for mean-variance analysis, i.e. if one asset is preferred to another by all risk averse investors, then the dominant asset must have the lower variance.

2.3.4 Generalizations of Other Risk Measures

In addition to mean-variance methodology, it may also be possible to describe the other approaches to risk analysis stated earlier in an expected utility framework (see Booth, 1995b). There are two main reasons for doing so. Firstly, it is useful to be able to compare these criteria amongst themselves and with other utility functions by effectively standardizing them. Secondly, some of the more descriptive risk measures will need to be formulated mathematically if they are to be used objectively for decision-making. One example is the set of canons postulated by Bailey (1862).

From Section 2.2.1, it may be recalled that Bailey's first canon requires the security of capital to be maintained. But rather than distinguishing between income and capital, it would seem more sensible and convenient to formulate Bailey's canons in terms of total returns. One possible interpretation of the first canon then is to ensure a non-negative return from the investment. The second canon states that subject to this, the highest expected return should be earned. Therefore, Bailey's implied utility function may be represented by:

$$U(x) = \begin{array}{ll} x & x \geq L \\ x - k & x < L \end{array}$$

where k is a large, positive amount and L is the wealth level associated with a zero rate of return. As long as total returns are non-negative, the insurer will seek the portfolio with the highest expected return, hence a linear (riskless) utility function above $x = L$. If the condition is breached and k is sufficiently large, then no expected return will be sufficient to compensate for the disutility of earning negative returns. So in the utility maximization process, the only portfolios that will be considered are those which minimize the probability of yielding negative returns.

Pegler (1948) on the other hand advocated maximizing expected returns as his first investment principle. Taken on its own this implies risk neutrality, or a utility function which is linear over all levels of wealth. Hence, the utility function could be:

$$U(x) = x \qquad \forall x$$

However, Pegler also encourages the spreading of risk, suggesting some hint of risk aversion in his utility function. But since Pegler is only concerned with diversifiable risk, there does not appear to be any utility function which can represent Pegler's principles precisely. In a sense, Pegler is suggesting positive infinitesimal risk aversion.

With the analysis given by Clarke (1954), it had been proposed that investors should trade risk with return, where risk is given the broad definition of the uncertainty with respect to expected returns. Hence, a range of risk measures could be used including variance, mean absolute deviation, etc. Furthermore, no objective criteria had been put forward for choosing the optimal portfolio. In this sense, Booth (1995b) describes Clarke's approach as being similar to an efficient frontier, rather than a utility maximizing one.

Wise (1984a, 1984b) set out a number of matching criteria based on minimizing the second non-central moment of ultimate surplus. Taking ultimate surplus to mean ultimate wealth, X, the aim is essentially to minimize $E(X^2)$. Since this is equivalent to maximizing $-E(X^2)$, an obvious utility function for such an investor would be the quadratic function:

$$U(x) = -x^2$$

In Booth (1995b), it was noted that Wise's utility function has the property, $U'(x) = -2x$, which means that the function is only increasing for negative amounts of wealth.

Hence, Wise's approach has rather limited appeal both in terms of utility theory, and from a mean-variance perspective as demonstrated by Wilkie (1985).

The safety-first approaches described earlier may also be considered in a utility context. Roy's criterion of minimizing $P(x \leq x^*)$ may correspond to that of an investor with a utility function such as:

$$U(x) = \begin{cases} 0 & x \geq x^* \\ -k & x < x^* \end{cases}$$

where $k > 0$. However, this implies that the investor has a zero marginal utility of wealth at all values of x apart from $x = x^*$.

Pyle and Turnovsky (1970) showed that maximizing x^* subject to $F(x^*) = P(x \leq x^*) \leq \alpha$ (Katoaka's criterion), may be rewritten as maximizing $E(X) + F^{-1}(\alpha).\sqrt{V(X)}$, assuming X belongs to the class of two-parameter distribution functions. If this assumption holds, then the decision-making process will be equivalent to a mean-variance one.

Maximizing $E(X)$ subject to $P(x \leq x^*) \leq \alpha$ (Telser's criterion), may be seen as a utility maximization approach subject to a constraint, where the utility function is linear. It seems unlikely, though, that such a constraint could be incorporated directly into a simple utility function.

If the approach in Clarkson (1989) had been limited to one of minimizing downside risk, $W(.)$ for a given level of expected return, then a utility function which is consistent with this could be:

$$\begin{aligned} U(x) &= x & x \geq L \\ &= x - k.W(x-L) & x < L \end{aligned}$$

where $k > 0$. With a utility function of this form, the extent to which expected returns may be traded against downside risk is determined by the value of k. As k increases, greater importance is placed on reducing downside risk. If $W(.)$ is a quadratic function, then one may recognize this to be consistent with a mean-semi-variance approach.

Nevertheless, the additional axiom of Clarkson's, requiring investors to maximize return subject to a threshold level of risk, cannot be integrated into a consistent utility function shown above. On its own though, the latter axiom implies a linear utility function subject to a constraint, as in Telser's criterion.

2.4 Summary

In this chapter, some of the main criteria and methods for dealing with the problem of investment risk have been reviewed. Early actuarial investment principles tended to be descriptive and lacked mathematical rigour, making them difficult to apply objectively. In these respects, they were bettered by the development of mathematical portfolio selection models and in particular, mean-variance analysis. Although this had originally been aimed at stock selection problems, the approach was later extended to allow for liabilities, bringing it closer to the actuarial principles of matching and immunization.

Unfortunately, analytical methods such as this generally suffer from numerous practical limitations, most of which relate to the restrictive assumptions that these methods require. One solution, though, has been to harness the enormous computing power currently available to facilitate stochastic simulation techniques. Through the use of simulation, complex issues such as life office solvency can now be addressed with far greater realism than had been possible previously.

In this chapter, it was also suggested that these criteria for decision-making may be assessed in terms of utility theory. Many of the approaches to risk mentioned could be reflected by appropriate utility functions, usually piecewise linear utility functions. Those which could not were generally found to be either inconsistent within themselves, or counterintuitive. In particular, mean-variance analysis may sometimes imply utility functions which are counterintuitive, and is therefore capable of leading to inappropriate investment decisions. Empirically however, mean-variance analysis has been found to be an adequate approximation to the expected utility approach in many circumstances.

3. THE INVESTMENT MODEL

3.1 Introduction

As with most investigations involving asset allocation decisions, the results derived will inevitably depend on the assumptions relating to the characteristics of the asset classes involved. It would therefore seem appropriate to devote at least one chapter to the model of the United Kingdom (UK) economy used in this research.

This chapter begins with a brief discussion of the basic requirements for any investment model and how well the chosen model meets these objectives. A description of the model is given and reviews of it which have been conducted by independent sources are considered. Measures taken to allow for some of its limitations and simulation results produced by the final model are then analysed.

3.2 Basic Requirements

3.2.1 The Asset Classes

In order to arrive at the optimal investment strategy for a life insurance fund, it is imperative that the investment model used consists of all the asset categories that would feature in such a strategy. Although it would be impractical to include all possible asset classes in an investment model, it should still comprise those categories which are important to an insurer. Therefore, one criterion for including a particular asset class in the model would be if there is a strong intuitive reason for an insurer investing in that asset class. In addition, the investment instruments used should also be distinct in nature from each other so that the portfolio resulting from the optimization process may be more meaningful. If two or more asset classes have similar characteristics, the relative preferences for these asset classes are likely to be very sensitive to sampling error and model specification.

In practice conventional non-profit life insurance funds have generally been backed by fixed-interest securities with suitable terms to maturity in order to minimize potential losses due to changes in investment conditions. Such a strategy is deeply rooted in the concepts of matching and immunization. According to Redington's theory of immunization (see Section 2.2.2), a set of fixed liabilities may be protected against movements in interest rates if the fund is invested in a portfolio of stocks with the same discounted mean term as the liabilities. So if immunization is at all possible, then this should be able to be achieved by an appropriate combination of the shortest and longest stocks available.

Notwithstanding all the limitations of immunization in practice, it would therefore seem adequate for modelling purposes just to consider two fixed-interest asset classes with extreme terms to maturity, such as cash and Consols. These two asset types also happen to possess much more historical data compared with redeemable gilts, which should lead to more credible stochastic investment models.

As well as needing to hedge fixed liabilities, a typical insurer would also have inflation-linked liabilities to contend with. The most general source of inflationary liabilities comes in the form of maintenance expenses, although these may also arise directly from certain types of business such as index-linked annuities. From a matching perspective, the most suitable investments for these liabilities would be index-linked gilts. Apart from its clear merits as a match to inflation, this asset class has unique characteristics which provide yet another possibility for insurers to gain from the benefits of diversification.

Any investigation into risk-return strategies should include at least one volatile asset. Insurance companies of all types and sizes tend to invest some proportion of their assets in UK equities. The reason for this is twofold. If past experience is anything to go by, equities should provide the best means for maximizing returns, particularly of surplus assets, in the long run. Secondly, they are considered to be a good long term inflation hedge and therefore may be used as a crude match for long term inflationary liabilities. Their characteristic short term price volatility and longer term growth potential distinguish them clearly from the other assets mentioned above.

Hence, the four asset classes of cash, Consols, index-linked gilts and UK equities are arguably the building blocks of a well-diversified fund in the context of an insurance company and are essential components of the stochastic investment model. In addition, the model should also comprise some inflationary measure such as the retail price index, so that inflation-linked liabilities may be incorporated in the investigations.

For the sake of simplicity, it is helpful to limit the number of asset classes to as few as possible. Although overseas equities and property are two asset classes which also tend to be associated with the assets held by life offices, they are not being used here as their inclusion may only really be justified for the purpose of diversification. The difference between overseas equities and UK equities may be attributed to the added currency risk inherent in the former. Nevertheless, it may be sensible to consider this asset class if some of the liabilities themselves are denominated in foreign currencies.

In the case of property, there is difficulty in encapsulating, via an investment model, the unique features of this asset class i.e. its lack of marketability and the fact that a property portfolio is less diversifiable than most other asset classes. If these are not properly accounted for in the model, then property as an asset class may appear more favourable than it would otherwise be in reality. But for these characteristics, its strong growth potential would make property similar in nature to UK equities.

3.2.2 General Criteria

In addition to encompassing a wide enough range of asset classes, the adequacy of the investment model should also be judged on more general criteria. Ideally, it should be based, at least in part, on sound statistical fitting procedures and applied to a sufficiently large and accurate historical data set. The model should be realistic, meaning that the distribution of values produced by the model would need to be consistent with one's own intuition. As the model is to be used in an insurance context, it should be capable of medium or long term projections. Moreover, it should be widely available and independently scrutinized so that the features and implications of the model are well understood.

Having set out these fairly stringent requirements of the investment model, it is worth emphasizing that the process of deriving a long-term econometric model is fraught with difficulties. Data inadequacy is just one of the problems. The underlying mechanism which generates economic variables is unclear and appears to evolve with time. This may imply that the understanding about the way in which the economy behaves cannot necessarily be captured in a statistical model. Consequently, a number of authors including Chatfield (1995: 15-16) and Miller and Newbold (1995) have discussed the issue of uncertainty in time series models and how this level of uncertainty is often underestimated.

It is therefore not surprising that a model has yet to be produced which satisfies all the above requirements. In fact, the only published long term model in the UK at the time of when these investigations were being carried out was that developed by Wilkie (1984, 1986). Although this model has since been refitted using data from 1923 to 1990, the resulting parameter values were merely quoted in Wilkie (1992, 1995a), without any details of the fit being given. These have now been superseded by a range of possible models detailed in Wilkie (1995b) and are based upon data up to June 1994. However, this most recent publication was not available in time to be used here.

Despite the obstacles faced in econometric modelling, Wilkie's 1986 model would on the surface seem reasonable for practical purposes. It has also been reviewed in a number of sources including Kitts (1988, 1990), Geoghegan *et al* (1992) and Huber (1995), and their conclusions will be discussed shortly. As the model probably remains the most widely used and accepted investment model in the actuarial profession, it would seem to be a sensible base case for these investigations.

3.3 Wilkie's Model

3.3.1 Description

As there is more than one version of Wilkie's model currently available, the one actually used in these investigations is outlined here for completeness. The model adopted is identical to the 'Full Standard Basis' as specified in Wilkie (1986) and comprises the retail price index (RPI), UK equities share dividend index and dividend yield and the yield on 2.5% Consols. Hence the asset classes which this model encompasses are UK equities and Consols, together with a model for inflation.

The model for the retail price index, $Q(t)$ is:

$$\nabla \ln Q(t) = QMU + QA(\nabla \ln Q(t-1) - QMU) + QSD.QZ(t)$$

where the backward difference operator, $\nabla X(t) = X(t) - X(t-1)$ and $QZ(t)$ are independent, identically distributed (i.i.d.) unit normal variates. The model implies that the force of inflation follows a first order autoregressive or AR(1) process about a mean of *QMU*, while *QSD* relates to its standard deviation. The autoregressive parameter *QA*, determines the extent to which the current force of inflation depends on the previous years' force of inflation. The parameter values in Wilkie's Standard Basis are:

$$QMU = 0.05, \quad QA = 0.6, \quad QSD = 0.05.$$

The model for the share dividend yield, $Y(t)$ is:

$$\ln Y(t) = YW.\nabla \ln Q(t) + YN(t)$$

with:

$$YN(t) = \ln YMU + YA.(YN(t-1) - \ln YMU) + YSD.YZ(t)$$

where $YZ(t) \sim$ i.i.d. N(0,1). This model implies that dividend yields follow an AR(1) process with lognormally distributed error terms and includes the force of inflation as a transfer function. *YMU* and *YSD* reflect the general level of the mean and standard deviation of dividend yields respectively. *YA* in this model has a similar interpretation to *QA* in the retail prices model, and *YW* indicates the extent to which the force of inflation influences the level of dividend yields. The parameter values are:

$$YMU = 0.04, \quad YA = 0.6, \quad YW = 1.35, \quad YSD = 0.175.$$

The model for the share dividend index, $D(t)$ is:

$$\nabla \ln D(t) = DW.DM(t) + DX.\nabla \ln Q(t) + DMU + DY.YSD.YZ(t-1)$$
$$+ DSD.DZ(t) + DB.DSD.DZ(t-1)$$

with:

$$DM(t) = DD.\nabla \ln Q(t) + (1-DD).DM(t-1)$$

where $DZ(t) \sim$ i.i.d. N(0,1).

Basically, the model suggests that the force of increase in dividends comprises a random component with mean, *DMU*, a moving average of the force of inflation, *DW.DM*(*t*), and an additional weighting of *DX* on the current force of inflation. For example, constraining $DW + DX = 1$ ensures that an increase in retail prices will eventually result in unit gain in dividends. The parameter *DD* determines the relative influence which more recent inflation has in the moving average term. The error terms are lognormally distributed and are linked to the previous year's residual term in the dividend yield model via *DY*. The parameters values are:

$$DMU = 0.0,\ \ DY = -0.2,\ \ DD = 0.2,\ \ DW = 0.8,\ \ DB = 0.375,\ \ DX = 0.2,\ \ DSD = 0.075.$$

The model for the Consols yield, *C*(*t*) is:

$$C(t) = CW.CM(t) + CN(t)$$

with:

$$CM(t) = CD.\nabla \ln Q(t) + (1 - CD).CM(t-1)$$

and

$$\ln CN(t) = \ln CMU + (CA1.\mathrm{B} + CA2.\mathrm{B}^2 + CA3.\mathrm{B}^3)(\ln CN(t) - \ln CMU)$$
$$+ CY.YSD.YZ(t) + CSD.CZ(t)$$

where the backshift operator, B*X*(*t*) = *X*(*t*–1), and *CZ*(*t*) ~ i.i.d. N(0,1). The model consists of two distinct parts: an inflationary moving average component, *CW.CM*(*t*), similar to that in the dividend index model, and a real yield component, *CN*(*t*) which assumes an AR(3) process about a mean, *CMU*. The error terms are lognormally distributed and also include the residual from the dividend yield model. The parameters values are:

$$CMU = 0.035,\ \ CY = 0.06,\ \ CD = 0.045,\ \ CW = 1.0,$$
$$CA1 = 1.2,\ \ CA2 = -0.48,\ \ CA3 = 0.20,\ \ CSD = 0.14.$$

The neutral initial conditions as stated in Wilkie (1986) are used:

$$\nabla \ln Q(0) = QMU,$$

$$Y(0) = YMU.\exp\,(YW.QMU), \;\; YE(0) = 0,$$

$$DM(0) = QMU, \;\; DE(0) = 0,$$

$$CM(0) = QMU, \;\; CN(0) = CN(-1) = CN(-2) = CMU.$$

In addition, it is necessary to specify an arbitrary starting value for the retail price index. It would seem sensible to choose $Q(0) = 1$. The recommended minimum value for $C(t)$, $CMIN = 0.5\%$, is also imposed.

3.3.2 Review of Wilkie's Model

One general requirement on which the original model falls short is the absence of cash and index-linked gilts. The necessary enhancements are dealt with in the next two sections. Nevertheless, Wilkie's model does have a number of useful features. Many of the model parameters such as the long term means can be readily interpreted and may with some degree of caution, be adjusted to suit the needs of the individual user. The model had been fitted with some contribution of personal judgement on the part of its author, as opposed to a purely statistical perspective which is usually the case in Vector Autoregressive (VAR) models. More specific evaluations of Wilkie's model carried out by a number of independent parties are summarized below.

Kitts (1988, 1990) looked at the model for the retail price index alone, being the driving force of the whole model. He concluded that the residuals were in fact correlated, implying that "... *the model undergenerates sustained periods of retail price inflation and deflation*" (Kitts, 1990). He also found the residuals to be not even approximately normally distributed. Kitts remarked that these problems were amplified because the stochastic component in the model is large relative to the systematic component, but also noted that the model would ultimately be most sensitive to long term assumptions like the mean, *QMU*.

A Working Party set up in 1989 to review Wilkie's model published their report in Geoghegan *et al* (1992). This report recommended that a model should be fitted using post 1945 data to reflect the fundamental changes that have occurred since the Second World War. It was also stated that "... *there was little evidence that a better fitting parsimonious model could be estimated using standard Box-Jenkins methodology*", though some concerns were expressed over the fit of the model to the data.

Problems of heteroscedasticity of residuals, random shock effects and non-normality of residuals had also been discussed but the general conclusions were that the use of Autoregressive Conditional Heteroscedastic (ARCH) effects, mixture models and different distributions for residuals would not make a significant difference to predicted values in the long term. Nevertheless, the difficulties in modelling non-stationary time series were acknowledged. In applications, the Working Party suggested that the aim of the model should be to provide the long term means and the covariance matrix of future values, rather than extreme values.

Huber (1995) analysed Wilkie's model from a statistical perspective, uncovering a number of fundamental problems with the model. Apart from problems concerned with non-stationarity and random shocks in the series modelled, he also highlighted some inconsistencies in the data used. These essentially stemmed from including data going further back in time than there was clean data available.

In Huber's analysis, more crucial inadequacies were unveiled relating to the structure of the model itself. Many of the structural parameters including *QA*, *YW*, *DW*, *DX* and *CY* were found to be inappropriate, either because they were unstable over sub-periods of the data in which they were fitted, or because they were shown not to be statistically significant. It was also shown that a more parsimonious model for Consols could be fitted by setting $CW = 0$. The net effect of implementing such changes, however, would result in an investment model which may not make good economic sense. Eliminating *QA* implies that the force of inflation is simply white noise. When *YW*, *DW*, *DX* and *CW* are set equal to zero, the result is that the dividend yield, the dividend index and the Consols yield are no longer linked to inflation. The removal of *CY* also breaks the link between the dividend yield and the Consols yield. This appears to cast doubts as to the appropriateness of the cascade structure in Wilkie's model.

In summary, there appears to be some uncertainty regarding the extreme values produced by the model, the cascade structure and some of the parameter estimates suggested. Although these deficiencies do not detract from the methodology used in this research, the absolute results implied in later chapters should, nevertheless, be treated with a fair degree of caution. If sufficient care is used in interpreting the relative results produced, some benefit may still be derived from applying the model in the determination of appropriate asset allocation strategies.

3.4 Index-linked Gilts

3.4.1 Model Structure

Despite their importance as an asset class, index-linked gilts had not been included in Wilkie's original model. Although no reasons were given in Wilkie (1986), it may be assumed that their exclusion was mainly due to the lack of data available at the time.

Even now, nearly a decade since this model has been published, it is arguable whether sufficient data exists for a long term model of index-linked gilt yields to be fitted with a reasonable degree of credibility.

A possible solution, apart from excluding index-linked gilts from the investigation altogether, could be to make a minor adjustment to Wilkie's model that may allow for this. For example, the share dividend yield, *Y(t)*, less a constant could be used to represent real yields on index-linked gilts, as suggested in Daykin & Hey (1990). Although this form of deterministic adjustment may not critical to the distributions of long term returns, it does impose an artificial relationship between the dividend yield and the real yield. As the latter has repercussions on the determination of the reliable yield calculation in the statutory valuation basis for insurance liabilities, the inclusion of some stochastic variation would be preferable.

It may be recalled that the model for the Consols yield is essentially made up of two components: an inflationary element, *CM(t)* and a real yield element, *CN(t)*. It should therefore be possible to make appropriate adjustments to some of the parameter values in *CN(t)* and so that this version may be used as a crude proxy for the real yield on index-linked gilts, *IL(t)*. Restricting these adjustments to just the parameters concerned with the mean and variance should allow for the main characteristics of index-linked gilts while maintaining the relationship between the real yield and the Consols yield as implied in Wilkie's model. For this reason, it also seems appropriate to use the same error terms, *YE(t)* and *CZ(t)* in both the index-linked gilts yield and the Consols yield models. Thus:

$$\ln IL(t) = \ln ILMU + (CA1.\mathrm{B} + CA2.\mathrm{B}^2 + CA3.\mathrm{B}^3)(\ln IL(t) - \ln ILMU)$$
$$+CY.YSD.YZ(t) + ILSD.CZ(t)$$

where *ILMU* and *ILSD* are the parameters associated with the mean and standard deviation of this model for index-linked gilt yields respectively.

3.4.2 Selection of Parameter Values

In determining the extent to which any changes should be made, historical monthly real yields on index-linked gilts with five or more years to maturity assuming 5% inflation had been assessed (see Figure 3.1). Although this index is only readily available from the beginning of 1986, it may still give an indication as to the general behaviour of index-linked gilt yields.

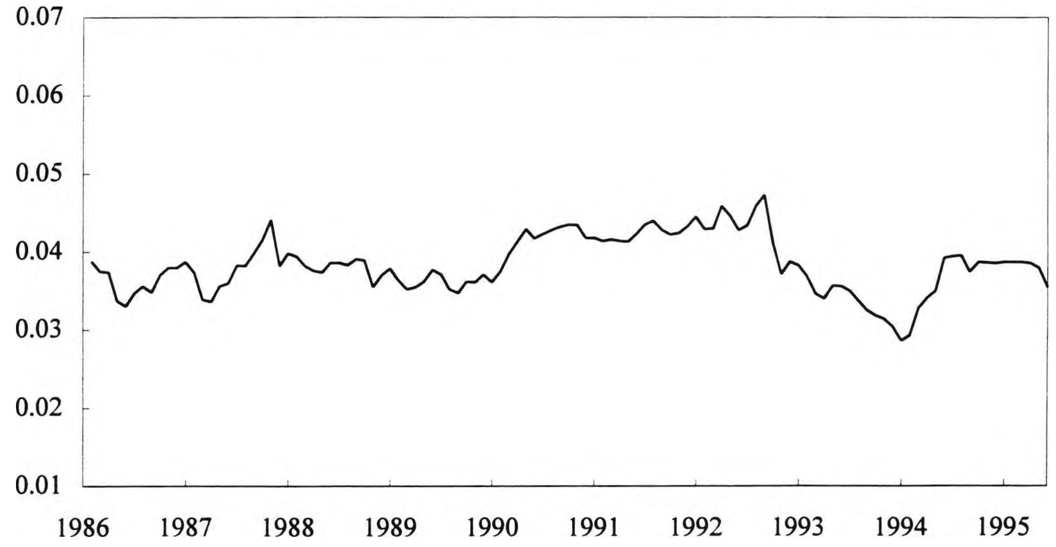


Figure 3.1. Historical monthly yields on index-linked gilts with terms to maturity in excess of 5 years and an inflation assumption of 5%.

In relation to this data, the mean real yield of $CMU = 0.035$, suggested in Wilkie's Consols yield model seems about right (correct to the nearest 0.5%) and its value was retained in the $IL(t)$ model as well, i.e. $ILMU = 0.035$. It is, however, less obvious whether the setting *ILSD* equal to $CSD = 0.14$, attributes variability in the index-linked gilts model appropriately to the historical data. As the data set is too small to give a credible measure of annual variance, prior judgement was used.

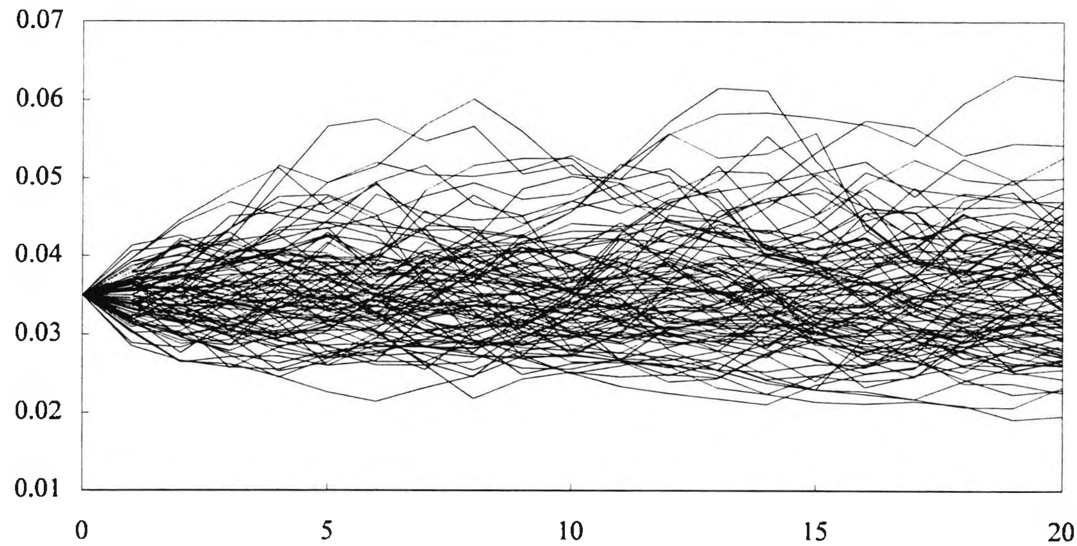


Figure 3.2. 100 simulations of the index-linked gilt model over 20 years: *ILMU* = 0.035 and *ILSD* = 0.07.

By simulating a hundred scenarios of $IL(t)$ over twenty years, a feel for the range of values the model may produce can be obtained. With the recommended parameter values for $CN(t)$ i.e. $ILMU = 0.035$ and $ILSD = 0.14$, the model exhibits fairly high levels of variability, with about a 55% chance that real yields will exceed 5%, a 30% chance of exceeding 6% and a 15% chance of exceeding 7% at some time over the next 20 years. This contrasts with past data which reveals only one month during the period 1986 to 1994 where real yields had risen above 4.5%.

There are arguments for and against retaining these parameter values for $IL(t)$, but on balance a reduction in variability would appear to be more reasonable. Thus, the simulations, repeated for the parameter *ILSD* halved from 0.14 to 0.07, are shown in Figure 3.2. These results seem to be more in tune with the historical yields shown in Figure 3.1. It would also seem sensible to allow real yields to occasionally climb as high as 6%, bearing in mind the limited past experience with this asset class. Hence, the parameter values chosen for $IL(t)$ were: $ILMU = 0.035$, $ILSD = 0.07$.

It is perhaps worthwhile comparing the model for $IL(t)$ above with a basic model for the index-linked gilt yield, $R(t)$, recently published in Wilkie (1995b):

$$\ln R(t) = \ln RMU + RA(\ln R(t-1) - \ln RMU) + RSD.RZ(t)$$

where $RZ(t) \sim$ i.i.d. N(0,1). One notable difference between $R(t)$ and $IL(t)$ is that an AR(1) rather than an AR(3) process was proposed. This is consistent with the Reduced Standard Basis in Wilkie (1986) and subsequent versions of Wilkie's Consols yield model (see Wilkie, 1995a, 1995b). The resulting autoregressive parameter RA of about 0.5 implies much stronger mean reversion properties than the net effect of the three CA parameters used in the model for $IL(t)$, as $CA1 + CA2 + CA3 \approx 0.9$.

Comparing the residual parameters, the value for $RSD = 0.0731$ is remarkably close to the value for $ILSD = 0.07$, while $ILMU = 0.035$ appears to be a reasonable approximation to the fitted value of $RMU = 0.0386$. Wilkie also suggests that the error terms in $R(t)$ are strongly correlated with CZ(t), though not to the extent implied in $IL(t)$. Hence, $IL(t)$ may produce index-linked gilt yields which are more highly correlated with the Consols yield than one might expect in practice (also see Table 3.7 later).

Although this is unlikely to have much impact on the correlation between the returns from these two asset classes, it may limit the ability of either asset class to be diversified against changes in the statutory minimum basis for valuing insurance liabilities. This may in turn reduce the extent to which these two asset classes appear together in any optimal portfolios derived in respect of a life office and should be kept in mind when analysing results produced in later chapters. Nevertheless, the relative stability of index-linked gilt yields and the fact they are not perfectly correlated with the Consols yield may help to constrain this affect.

In conclusion, it must be emphasized that the choice of model for *IL*(t) and in particular the parameter values is somewhat arbitrary and has not undergone any rigorous statistical tests. While the proposed model is far from ideal, data inadequacies limit the usefulness of statistical fitting procedures to a long term model. Even so, it is reassuring that the model adopted is broadly similar to the model fitted by Wilkie, particularly in terms of the implied means and variances. With index-linked gilts, the most important variable which affects the total return is inflation, rather than the real yield. Such crude adjustments may not be so easily justified with other asset classes.

3.5 Cash

The importance of including some form of short term fixed interest assets in these investigations has been briefly discussed, though Wilkie's original model did not incorporate short term interest rates. Daykin & Hey (1990) attempted to alleviate this deficiency by using a deterministic yield curve, in conjunction with the Consols yield produced by Wilkie's model, in their investigations. As with the real yield adjustment, deterministic yield gaps should be avoided even if they appear to behave in a reasonable fashion over a long period of time.

For this reason, an attempt had been made to produce a stochastic model for short term interest rates in Ong (1994), the main results of which are detailed in Sections 3.5.1 and 3.5.2. Using a Box-Jenkins approach, a model was fitted which could be used alongside Wilkie's model, without requiring the latter to be altered in any way. The data used as a proxy for short term interest rates was the discount rate on three-month Treasury Bills.

3.5.1 Preliminary investigations

Considering the purpose of the model, it would seem reasonable, *a priori*, to conduct the fitting procedure using data from 1919 to 1982, similar to the period used in Wilkie (1986). However, a more recent period had been chosen for two reasons. During the Second World War, short term interest rates were fixed by the government at 0.5%. This artificial restriction on cash yields seemed inappropriate for linear modelling. As with other asset categories, the general trend of cash yields has also changed over the past 70 years. Therefore, it should be preferable to use a more recent data set while maintaining a reasonable number of observations with which to achieve a reasonable fit. Although government restrictions on interest rates were eventually removed in 1951, it seemed sensible for modelling purposes to allow some time for yields to revert back to their more usual state. Hence, the data period selected was from 1955 to 1993.

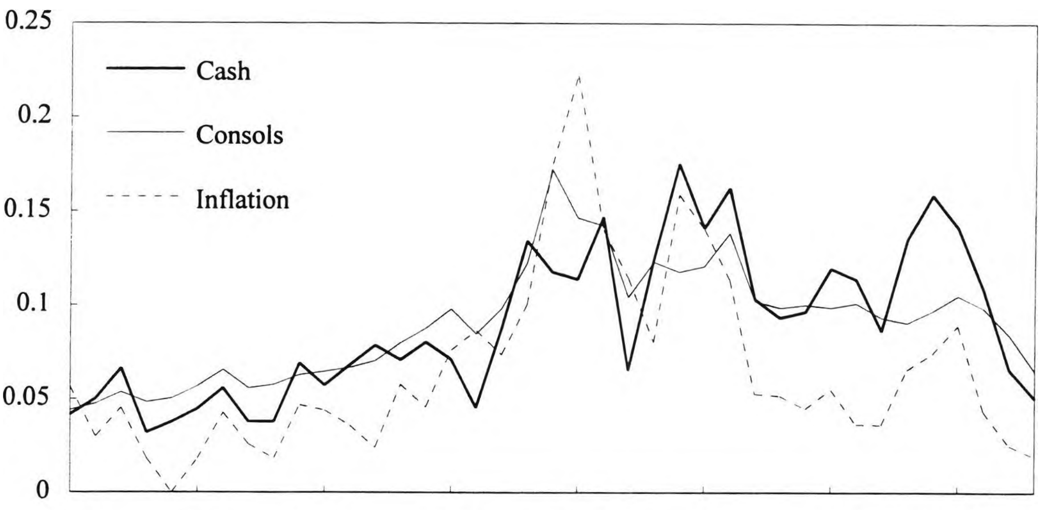


Figure 3.3. Historical cash yield, Consols yield and force of inflation.

Figure 3.3 above shows a graph of cash yields together with the Consols yield and the force of inflation over this period. From the graph, it appears that the series are non-stationary, thus failing a pre-requisite of the Box-Jenkins approach. In Autoregressive Integrated Moving Average or ARIMA modelling, a non-stationary series may often be fitted satisfactorily by differencing the data one or more times. This had been decided against in order to maintain as much consistency with Wilkie's model as was possible.

A more appropriate way of dealing with non-stationarity would be to model the relationship between cash yields and some other series such as Consols or inflation. Looking at Figure 3.3, it would appear that the three series do tend to track each other over the time. If the relationship between cash yields and at least one of the other two series is fairly stable over time, a reasonable fit may be possible by including the Consols yield and the force of inflation as input variables in the model.

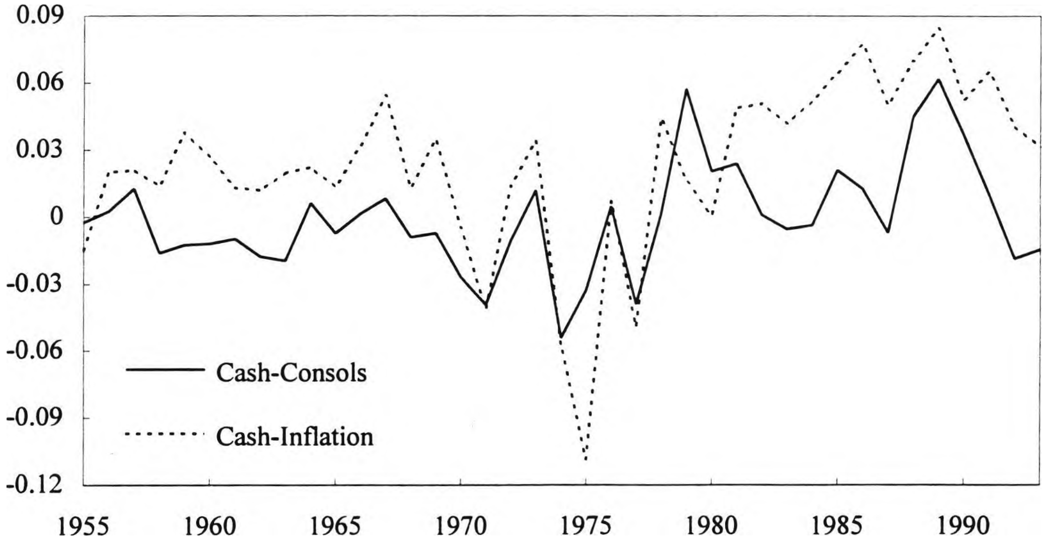


Figure 3.4. Historical [cash yield – Consols yield] and [cash yield – force of inflation].

A time series is said to demonstrate weak form stationarity when its mean and autocorrelation function remain the same over different time segments (see Chatfield, 1989). Figure 3.4 shows how the mean of the series [cash yield] – [Consols yield] does not appear to change much over time which brings it closer to the definition of stationarity. However, the higher variance over the second half of the period may indicate the possibility of heteroscedasticity in this series. With regards the spread between the cash yield and the force of inflation, there appears to be even less evidence of stationarity, with higher means being exhibited over the last decade and exceptionally high variances during the mid-seventies.

Further justification for including transfer functions for the Consols yield and the force of inflation may be observed from the crosscorrelation plots shown in Figures 3.5 and 3.6. In both cases, the highest crosscorrelation occurs at lag zero.

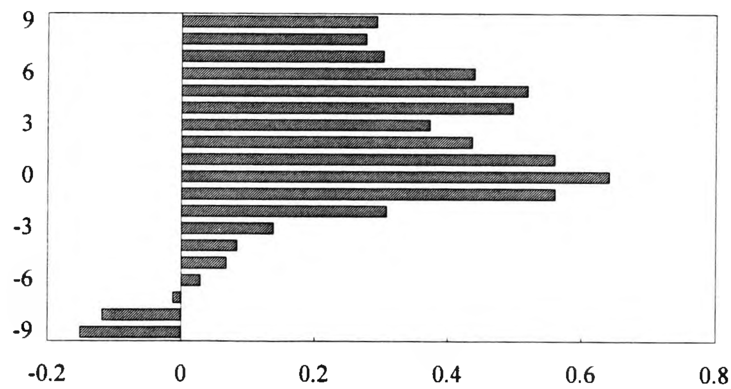


Figure 3.5. Sample crosscorrelation function between the cash yield and the force of inflation.

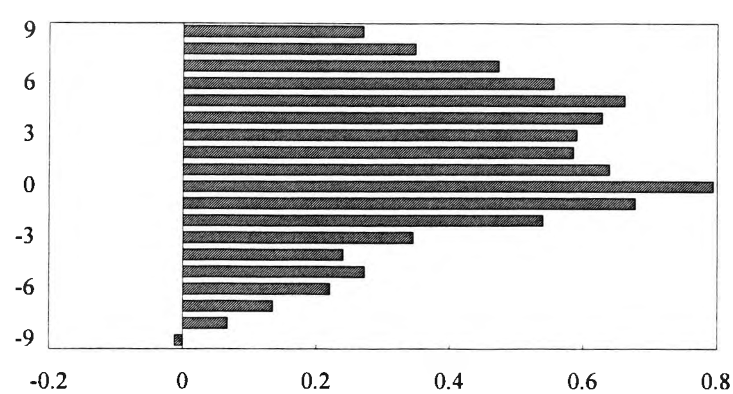


Figure 3.6. Sample crosscorrelation function between the cash yield and the Consols yield.

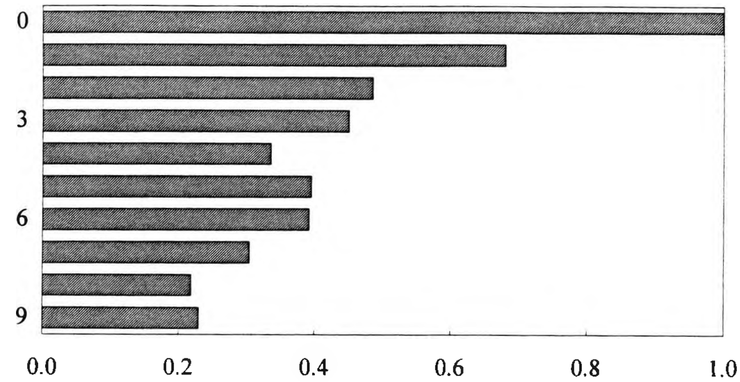


Figure 3.7. Sample autocorrelation function for the cash yield.

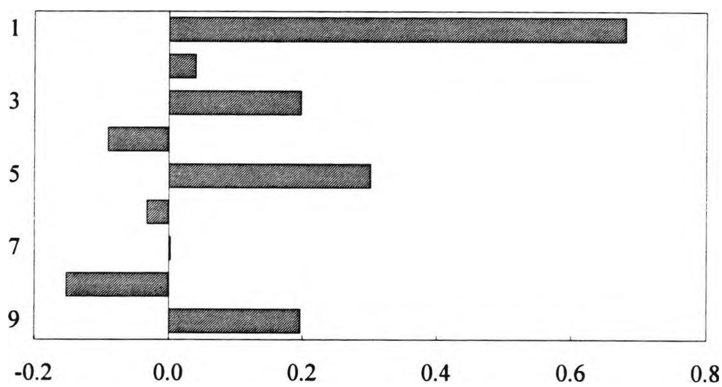


Figure 3.8. Sample partial autocorrelation function for the cash yield.

The autocorrelation function (a.c.f.) and partial autocorrelation function (p.a.c.f.) of the cash yield were then computed over a number of lags and illustrated in Figures 3.7 and 3.8 respectively. From the a.c.f. plot, the lack of stationarity in cash yields is evident whilst the plot of the p.a.c.f. appears to suggest an AR(1) process.

3.5.2 A Model for Cash – Consols

Based on the preliminary investigations, an AR(1) model for the cash yield, $K(t)$, was fitted with the Consols yield and the force of inflation included as transfer functions, the form of the model being:

$$K(t) = \mu + \lambda I(t) + \omega C(t) + \frac{1}{(1-\phi \mathrm{B})} e(t)$$

where $I(t)$ and $C(t)$ are the force of inflation and the Consols yield at time t, with the error term $e(t) \sim$ i.i.d. $N(0, \sigma^2)$. B is the backshift operator, $\mathrm{B}X(t) = X(t-1)$. As with this and subsequent models fitted here, the estimation procedure had been carried out by the method of maximum likelihood using SAS/ETS® (SAS Institute Inc., 1988). The parameter estimates obtained for the above model are given in Table 3.1.

Parameter	*Estimate*	*Approx. S.E.*	*t-Ratio*
ϕ	0.44	0.15	2.84
λ	0.03	0.15	0.20
ω	1.02	0.25	4.12
μ	0.00	0.02	−0.25
σ	0.02	-	-

Table 3.1. Parameter estimates for 5-parameter cash model.

The parameters λ and μ were found to be insignificant and correlated with the parameter ω. Hence, both parameters were discarded and the estimation procedure was repeated for the reduced model:

$$K(t) = \omega C(t) + \frac{1}{(1-\phi \mathrm{B})} e(t)$$

The resulting estimates (see Table 3.2) remained very much unchanged in relation to those shown in Table 3.1. This was also true of the standard deviation of the residuals which had been marginally increased from 0.02189 to 0.02194. The parameters ϕ and ω were not significantly correlated, yielding a correlation coefficient of −0.023.

Parameter	*Estimate*	*Approx. S.E.*	*t-Ratio*
ϕ	0.43	0.15	2.92
ω	1.00	0.06	15.53
σ	0.02	-	-

Table 3.2. Parameter estimates for 3-parameter cash model.

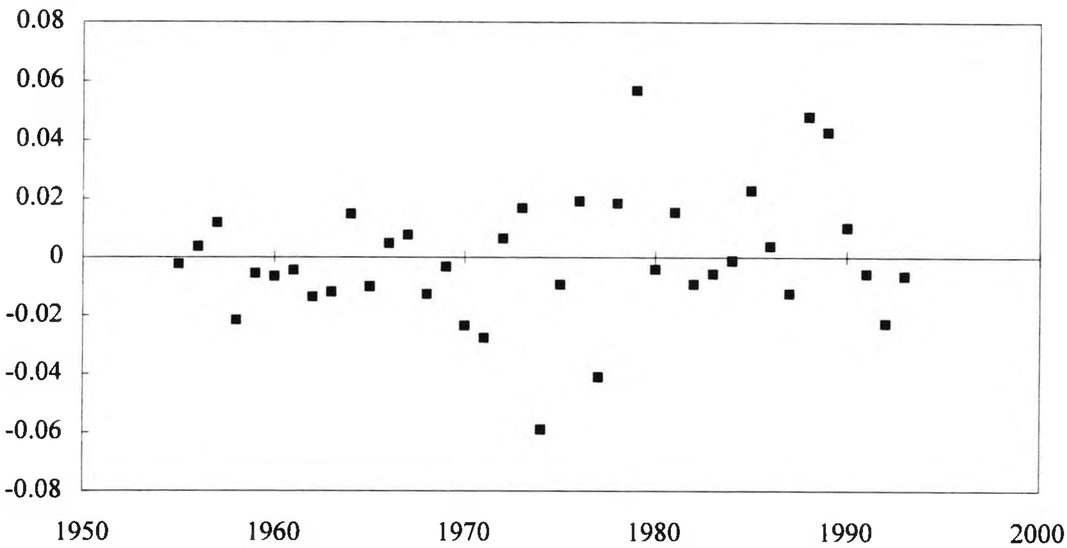


Figure 3.9. Residual plot of [actual – forecast] for $(1 - \phi B).(K(t) - C(t))$.

Looking at Figure 3.9, the residual plot for the reduced model seems satisfactory. It may also be noted that the two outliers of greatest magnitude occurred in 1974 and 1979, two years of economic crisis in the UK. When the residuals were tested for autocorrelation over 6 and 12 lags, the results provided no evidence to suggest that the residuals correlated. In addition, the Box-Ljung Q-test for heteroscedasticity (see Harvey, 1989) was found to be insignificant, giving a Q-value at lag 6 of 2.78.

Summary statistics of the residuals are shown in Table 3.3 below. From the first three moments given, it would appear that the distribution of residuals are roughly symmetrical about a mean of zero. The only indication of non-normality of residuals exists in the relatively high measure of kurtosis obtained, which is mainly due to the presence of the two outliers mentioned earlier. Nevertheless, the residuals still pass the Jarque-Bera test for normality (see Wilkie, 1995b) at the 10% significance level.

Mean	*S.D.*	*Skewness*	*Kurtosis*
0.00	0.02	0.24	1.59

Table 3.3. First four central moments of the residuals for $(1 - \phi B).(K(t) - C(t))$.

Alternative models were also tested by altering the transfer function relating to Consols yield (see Ong, 1994). For each of these variants, the standard deviation of the residuals still remained at about 0.022, indicating that the increased parameterization in the transfer function was not leading to significantly better fits. Hence, the model suggested in Ong (1994) was of the form:

$$K(t) = C(t) + KA.(K(t-1) - C(t-1)) + KSD.KZ(t)$$

where $KZ(t) \sim$ i.i.d. N(0, 1), with parameters values (rounded to one significant figure):

$$KA = 0.4, \quad KSD = 0.02.$$

The neutral initial condition for this model is $K(0) = C(0)$. For practical reasons, a lower bound on the cash yield of 0.5% had also been imposed. When tested by simulation, the model appeared to produce reasonable results.

3.5.3 A Model for ln(Cash/Consols)

Shortly after these results were published, Wilkie (1995a) released for the first time a model for short term interest rates. Although it was not clear exactly which data period had been used, it may be inferred from the text that it probably included the years from 1923 to 1990. The proposed model structure was essentially the same as that of $K(t)$ above, except that the logarithms of the cash yield and the Consols yield were being used, as opposed to the untransformed variables. Using the log transformation means that it is the ratio between the two variables is being modelled rather than the difference. This also implies that the model residuals are assumed to be lognormally distributed.

Regardless of which approach is perceived to be more intuitive, a definite advantage of the log transformation is that the cash yields produced by the model will always be positive. However, as no such transformation had been applied to $C(t)$ in Wilkie's model, using $\ln C(t)$ as an input variable in the cash yield model may be considered inconsistent.

The model for short-term interest rates proposed in Wilkie (1995a) is given by $B(t)$, where:

$$\ln B(t) = \ln C(t) + BA.(\ln B(t-1) - \ln C(t-1)) - (1-BA).BMU + BSD.BZ(t)$$

and $BZ(\mathrm{t}) \sim$ i.i.d. N(0, 1). The parameter values are:

$$BA = 0.75, \;\; BMU = 0.185 \;\text{ and }\; BSD = 0.175.$$

Comparing the two models, it is clear that the parameters values in $B(t)$ are very different from those in $K(t)$. In particular, the autoregressive parameter of $BA = 0.75$ is very high compared with $KA = 0.4$, and the model includes an extra term $BMU = 0.185$, which translates to the cash yield being proportionately smaller on average than the Consols yield. While these discrepancies may be the result of the log transformation, the more likely cause relates to the artificial restrictions placed on cash yields during and shortly after the Second World War.

In order to investigate this further, the following model was fitted over the period from 1955 to 1993 inclusive, where $K(t)$, $C(t)$, $I(t)$ and $e(t)$ are defined as before.

$$\ln K(t) = \mu + \lambda I(t) + \omega . \ln C(t) + \frac{1}{(1-\phi B)} e(t)$$

Parameter	*Estimate*	*Approx. S.E.*	*t-Ratio*
ϕ	0.40	0.16	2.59
λ	−0.44	1.43	−0.31
ω	1.27	0.22	5.82
μ	0.64	0.61	1.05
σ	0.24	-	-

Table 3.4. Parameter estimates 5-parameter log transformed cash model.

Looking at Table 3.4, it may be seen that the parameters λ and μ are not significant. In addition, they were also found to be strongly correlated with each other and the parameter ω. Hence, both λ and μ were felt to be inappropriate and the model was refitted with these parameters set at zero, i.e.:

$$\ln K(t) = \omega . \ln C(t) + \frac{1}{(1-\phi B)} e(t)$$

The resulting parameter estimates for this reduced model are given in Table 3.5. It may be noted that these values of ϕ and ω are very similar to those shown in Table 3.2 for the untransformed series. With a correlation coefficient of 0.057, it would appear that they are also uncorrelated.

Parameter	*Estimate*	*Approx. S.E.*	*t-Ratio*
ϕ	0.40	0.15	2.66
ω	1.03	0.02	41.81
σ	0.23	-	-

Table 3.5. Parameter estimates model for 3-parameter log transformed cash model.

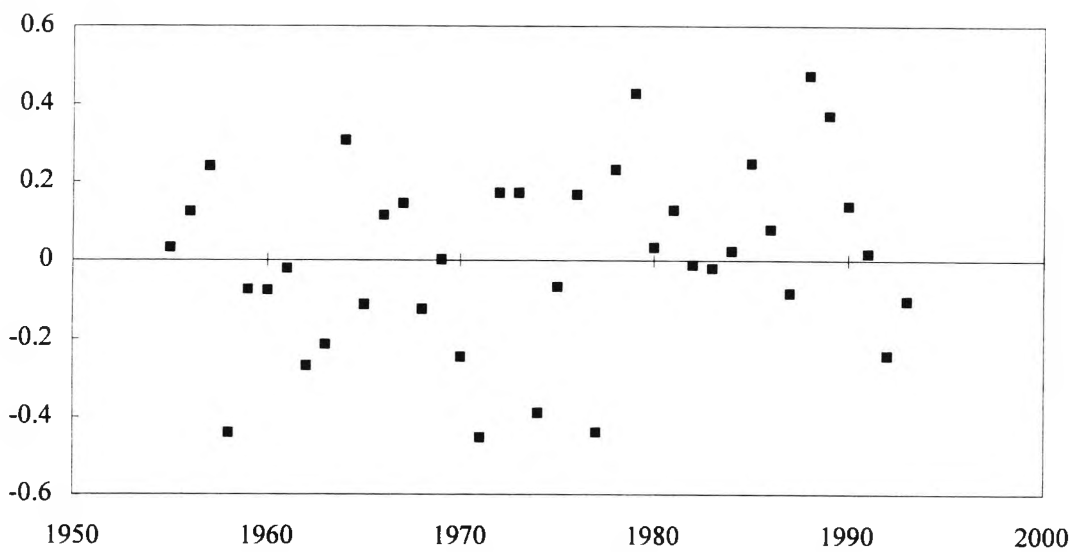


Figure 3.10. Residual plot of [actual – forecast] for $(1 - \phi B).\ln(K(t)/C(t))$.

In relation to the fit of the model, the standard deviation of the residuals has increased slightly from 0.2268 to 0.2315, as a result of removing the parameters λ and μ. The residual plot shown in Figure 3.10 appears to be better than that of Figure 3.9 for the untransformed model, with outliers being less pronounced. As before, the residuals show no significant autocorrelation over 6 and 12 lags. A Q-value at lag 6 of 3.69 also suggests no evidence of heteroscedasticity in the residuals. Summary statistics of the residuals are shown in Table 3.6 below. The log transformation has succeeded in eliminating the sizeable fourth moment from the residuals and appears much more satisfactory compared with the untransformed model. With Jarque-Bera test statistic of 0.37, there is no evidence to suggest the distribution of these residuals is non-normal.

Mean	*S.D.*	*Skewness*	*Kurtosis*
0.00	0.23	-0.21	-0.22

Table 3.6. First four central moments of the residuals for $(1 - \phi B).\ln(K(t)/C(t))$.

Overall, the log transformed model does seem to be more appropriate. The residuals appear to have more favourable properties, particularly in relation to the measure of kurtosis observed. In addition, the transformations do away with the need to impose a minimum on the cash yield. Hence, the model for cash chosen for these investigations is:

$$\ln K(t) = \ln C(t) + KA.(\ln K(t-1) - \ln C(t-1)) + KSD.KZ(t)$$

where $KZ(t) \sim \mathrm{N}(0, 1)$, with parameters values:

$$KA = 0.4 \text{ and } KSD = 0.25.$$

The neutral initial condition for this model is $K(0) = C(0)$.

Overall, this would appear to be the best fitting model in the context of a Box-Jenkins framework. The differences in parameter values when compared against Wilkie's model may be attributed to the different data periods used.

3.6 Simulation Results

The purpose of this section is to analyse the simulation results produced by the investment model. As well as highlighting the main features of the variables concerned, the simulation results should provide a means of checking that the model has been implemented correctly. A thousand simulations are employed in all computations that follow. The same initial random seed is also used throughout, so that the different sets of results may be more comparable.

3.6.1 Means, Standard Deviations and Correlations of Yields

The results from a thousand simulations of the yields after one year are summarized in Tables 3.7. Yields for RPI, cash, Consols, index-linked gilts and equities in this context are defined as $Q(1)-1$, $K(1)$, $C(1)$, $IL(1)$ and $Y(1)$ respectively. The average yields are broadly compatible with the neutral assumptions specified in the model, i.e. 5.17% for the inflation rate, 8.5% for both $K(0)$ and $C(0)$, 3.5% for $IL(0)$ and 4.28% for $Y(0)$.

	RPI	*Cash*	*Consols*	*IL Gilts*	*Equities*
MEAN (%)	5.4	8.9	8.6	3.5	4.4
S.D. (%)	5.1	2.3	0.5	0.3	0.8
CORRELATION					
Cash	0.10				
Consols	0.36	0.27			
IL Gilts	-0.04	0.25	0.91		
Equities	0.32	0.00	0.21	0.15	

Table 3.7. Means, standard deviations and correlation coefficients of yields after 1 year.

Amongst the asset classes, cash yields are the most variable, followed by dividend yields, Consols yields and index-linked gilt yields, which seems reasonable. The only variables which are highly correlated with each other are Consols and index-linked gilts, which is to be expected given that the same error terms are used in both $CN(t)$ and $IL(t)$.

3.6.2 Means, Standard Deviations and Correlations of Nominal Accumulations

In this section, the accumulated or 'rolled-up' amounts of 1 over one and twenty years is examined. The accumulations of Consols and equities are calculated exactly as described in Wilkie (1986). Index-linked gilts are assumed to have a real coupon of 2.5% at all times and a term to maturity of exactly ten years when first purchased. At the end of each year, the stocks are sold as nine year gilts and the proceeds used to purchase ten year gilts with the assumption that the real yield is identical for both nine and ten year gilts. For simplicity, the eight month lag used in practice for indexing coupons and redemption proceeds has not been implemented. Tax rates are assumed to be zero for all asset classes.

	RPI	*Cash*	*Consols*	*IL Gilts*	*Equities*
MEAN (%)	5.4	8.9	8.0	8.8	10.6
S.D. (%)	5.1	2.3	6.3	5.8	21.0
CORRELATION					
Cash	0.10				
Consols	-0.37	-0.28			
IL Gilts	0.93	0.00	-0.01		
Equities	-0.21	-0.01	0.18	-0.14	

Table 3.8. Means, standard deviations and correlation coefficients of nominal rates of return over 1 year.

Results for one year accumulations are summarized in Table 3.8 in terms of annual rates of return. The mean return for each of the variables seems intuitively reasonable, although Consols do not appear to perform as well as one might have expected. This because yields on average have risen from 8.5% to about 8.6% during that year. In contrast to the variability in yields seen in Table 3.7, the asset class with the lowest standard deviation of returns is cash. Overall, the results are very similar to those obtained in Wilkie (1986), except for equity returns which Wilkie found on average to be 12.17%. This discrepancy could be attributed to sampling error as the two estimates of equity returns may still lie within two standard errors of the true mean.

The accumulations over a twenty year period are presented in Table 3.9. However, only the mean accumulations are expressed here in terms of *effective* annual rates of return. It is important to note how this differs from calculating the annual rate of return in each simulation and then taking averages, which was the approach adopted in Wilkie (1986). In the case of inflation say, the estimate calculated here is that of $E[Q(20)]^{1/20}-1$, as opposed to Wilkie's estimate of $E[Q(20)^{1/20}]-1$. The differences between these two approaches may be checked by deriving the analytical distributions for these statistics.† Estimates of standard deviations and correlation coefficients on the other hand are calculated in respect of the actual accumulations.

From this table, the expected return on Consols appears to be more in line with that of cash. This is reasonable given that the model basically assumes a flat yield curve in the long term. The variability in Consols is also very much reduced relative to cash, which seems sensible. Index-linked gilts appear to outperform fixed interest assets though at the expense of having much higher variability than either cash or Consols. Nevertheless, the equity class maintains its position as the asset class with the highest mean and standard deviation of accumulations.

---

† Assuming the neutral initial condition $\nabla \ln Q(0) = QMU$ and if $Q(0) = 1$, then $Q(T) \sim \text{logNormal}(\mu, \sigma^2)$,

where $\mu = T.QMU$ and $\sigma^2 = \left(\dfrac{QSD}{1-QA}\right)^2 \left(T + \dfrac{QA^2(QA^{2T}-1) - 2QA(QA^T-1)(QA+1)}{(QA^2-1)}\right)$.

Therefore, $E[Q(T)]^{1/T} = \exp(\mu/T+\sigma^2/2T)$ and $E[Q(T)^{1/T}] = \exp(\mu/T+\sigma^2/2T^2)$. At $T = 20$, $\mu = 1.0$ and $\sigma^2 = 0.274416$ resulting in $E[Q(20)]^{1/20} - 1 = 5.85\%$ and $E[Q(20)^{1/20}] - 1 = 5.16\%$. These are approximately equal to their corresponding simulated values.

	RPI	*Cash*	*Consols*	*IL Gilts*	*Equities*
MEAN (%)	5.8	9.2	8.8	9.6	11.0
S.D.	1.7	1.9	1.1	3.5	6.4
CORRELATION					
Cash	0.53				
Consols	-0.08	0.28			
IL Gilts	0.99	0.58	0.00		
Equities	0.57	0.29	0.03	0.57	

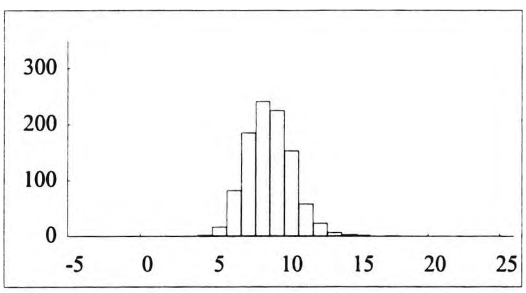
Table 3.9. Means, standard deviations and correlation coefficients of nominal accumulations over 20 years.

Looking at the correlation matrix, accumulations of index-linked gilts appear to be almost perfectly correlated with the retail price index. In addition, cash and equities also exhibit quite a strong positive correlation with both inflation and index-linked gilts over a twenty year period. Consols on the other hand seem to be independent of all the investment variables apart from cash, with which they are weakly correlated.

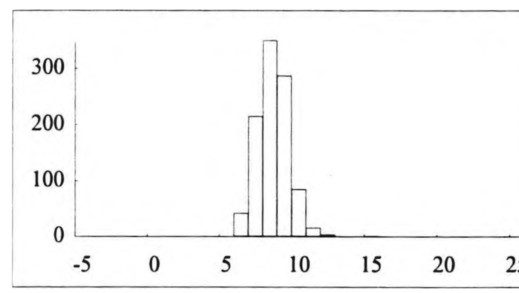
3.6.3 Histograms of Nominal Annualized Returns over Twenty Years

In order to examine the distributions of nominal annualized returns over twenty years, histograms are plotted for each of the variables and shown in Figure 3.11. From these plots, it is possible to get some idea of the range of values which each variable may take. For example, inflation and equities are the only two variables which seem to produce negative 'returns' in the long term. In contrast, the returns on both cash and Consols are rarely less than 5% over a twenty year period.

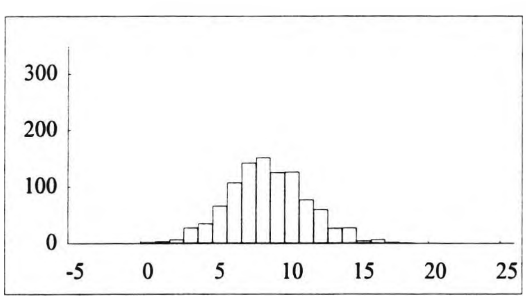
Another notable feature of these plots is that the distributions appear almost symmetric, which is partly to do with the fact that annualized rates of return are being used. If frequencies had been plotted against the accumulated amounts, the distributions would have appeared much more skew. This increased skewness may account for the generally higher effective mean rates of return observed in respect twenty year accumulations compared with one year returns.



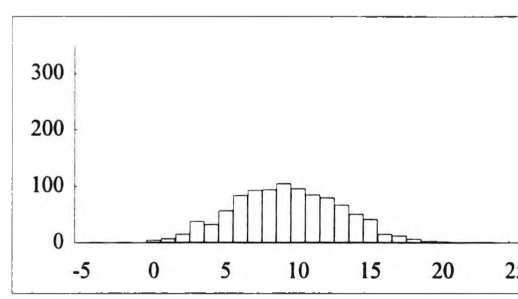
Cash



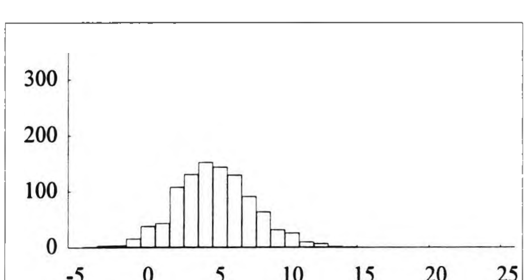
Consols



Index-linked Gilts



Equities



Inflation

Figure 3.11. Histograms of nominal rates of returns (% per annum) from 1000 simulations over 20 years.

3.6.4 Means, Standard Deviations and Correlations of Real Accumulations

The results shown in Tables 3.10 and 3.11 are equivalent to those of Tables 3.8 and 3.9 when calculated in real terms. In each simulation, the accumulated amount is divided by $Q(t)$ before any estimates are calculated. Hence, the retail price index is now redundant.

	Cash	*Consols*	*IL Gilts*	*Equities*
MEAN (%)	3.6	2.9	3.3	5.4
S.D. (%)	5.3	9.3	2.0	21.8
CORRELATION				
Consols	0.66			
IL Gilts	-0.14	0.57		
Equities	0.41	0.41	0.14	

Table 3.10. Means, standard deviations and correlation coefficients of real rates of return over 1 year.

In Table 3.10, the mean real returns are roughly consistent with the mean nominal returns less the mean rate of inflation. Index-linked gilts are now shown to be the most stable asset class in real terms, which was to be expected. Compared with the results in the nominal case, it may be noted that the level of variability has increased quite dramatically in both cash and Consols, reduced considerably in index-linked gilts and remained about the same in the case of equities. The high volatility seen from real cash returns in the short term is an inherent property of the model due to there being no direct relationship specified between cash and inflation.

Looking at Table 3.11, the real accumulations from equities now appear to be slightly less volatile in relation to Consols. This, to a limited extent, may have been anticipated from the estimated correlations between these asset classes and inflation shown in Table 3.9. Over twenty years, the standard deviation of real returns from index-linked gilts is almost negligible when compared with the other asset classes. From the above table, it is interesting note how on average, Consols and cash seem to outperform index-linked gilts in real terms. This is in contrast to the effective average nominal returns over the same period, which are 8.8%, 9.2% and 9.6%, for Consols, cash and index-linked gilts respectively.

	Cash	*Consols*	*IL Gilts*	*Equities*
MEAN (%)	4.1	4.3	3.5	5.2
S.D.	0.99	1.74	0.13	1.67
CORRELATION				
Consols	0.73			
IL Gilts	0.34	0.20		
Equities	0.18	0.16	0.04	

Table 3.11. Means, standard deviations and correlation coefficients of real accumulations over 20 years.

The curious reversal of orderings is possible because the mean real accumulation is not necessarily comparable with the mean nominal accumulation divided by the mean value of the retail price index. In fact, this situation is not different in principle to the earlier discrepancy between mean effective rates of return and effective mean rates of return. From the figures obtained, it may be deduced that the chances of obtaining high nominal returns when inflation is low, must be greatest in Consols and smallest in index-linked gilts. The correlations shown in Table 3.9 also support this explanation.

3.7 Summary

In summary, the investment model described in this chapter comprises the asset classes of cash, Consols, index-linked gilts and equities, as well as the retail price index. The 'Full Standard Basis' specified in Wilkie (1986) had been chosen to represent the retail price index, the Consols yield, the share dividend yield and the share dividend index. With an appropriate reduction of variance, the real yield component in the Consols yield model was felt to be a crude but adequate proxy to the index-linked gilt yield for the purposes of this research.

It was also shown that cash yields could be modelled by a first-order autoregressive process of the log of the ratio between the cash yield and the Consols yield. In arriving at this model, two interesting issues had emerged. Firstly, the results gave little evidence for incorporating inflation directly into the cash model. While it may be perceived in some circles that inflation rates and cash yields are closely linked, the correlation between these two variables has only become prominent since the 1980's. Secondly, there was sufficient evidence from the data observed since 1955 that the yield curve should on average be flat. This contrasts with perhaps the more usual view that the yield curve tends to be upward sloping.

From the simulation results, it would appear that the overall distributions obtained are intuitively reasonable. Over one year, cash returns are predicted to be the most stable in nominal terms whereas Consols seem to be the least volatile over a twenty year period. Index-linked gilts on the other hand are the most stable in real terms over both short and long time horizons. In all cases, equities appear to outperform the other asset classes on average, though they also tend to exhibit the greatest variability as well.

Although the general orderings described above are broadly intuitive, the choice of model structure and parameter values seems more uncertain. For instance, the strong link between equity returns and the retail price index implied by the cascade structure does not appear to be supported by the data. This should be kept in mind when assessing the role of equities in relation to real liabilities. In addition, it would also be sensible to note that the model for index-linked gilt yields had not been fitted with any degree of statistical rigour due to data inadequacies. The slope of the yield curve also appears to be sensitive to the data period chosen, thus introducing greater uncertainty to the relative position of cash in any optimal asset mix which may be obtained.

Despite the uncertainty in the investment model, it is important to remember that this research is aimed at investigating optimal asset allocation strategies based upon rational decision-making, rather than obtaining *the* correct answers. Rational decisions need to take into account the decision maker's future beliefs and hereafter, it will be assumed that this investment model accurately describes these future beliefs. Ultimately, however, the optimal decisions which will be obtained later must be viewed in the light of the assumptions which underlie the investment model and not in absolute terms.

4. STATIC OPTIMIZATION I - AN ASSET FUND

4.1 Introduction

Although the aim, ultimately, is to investigate the optimal asset allocation strategies for life offices, the process involved could potentially become too complex for the results to be interpreted in isolation. Apart from having to contend with the interactions between an investment model, a liability model and a utility function all at once, incorporating a liability structure may severely increase the computation time required, thus putting restrictions on the scope for experimentation.

As the methodology used in this research is applicable whether or not a liability model is included, it would seem instructive to first consider the case of an asset fund in the absence of explicit liabilities. Being fairly transparent, this situation should provide a feel for the appropriateness of the asset model and utility functions. From this position, it should then be easier to understand the more complex case with insurance liabilities involved. Hence, this chapter deals exclusively with an asset fund where no explicit liabilities are present.

In this chapter, consideration is first given to the framework in which the investment decision is being made. The procedure for obtaining analytical solutions is outlined and its limitations are discussed. Numerical optimization is introduced as an alternative to closed form solutions, including a brief description of the methods that will be used throughout this research. These methods are then applied to the framework set out earlier and the results analysed. Problems which may arise in the process are also investigated here. Finally, the accuracy of the results are interpreted with regard to the uncertainty inherent in this form of modelling.

4.2 Utility Maximization

4.2.1 Formulation of the Problem

This section considers the case of a personal investor about to invest a fixed sum of money in a pool of assets, such as a unitized fund. It will be assumed that the investor is required to select the best asset mix for this fund from a given set of asset classes. In order for the problem to be resolved rationally, the optimal decision will need to take account of a number of factors. These generally pertain to the characteristics of the asset classes and the investor's risk preferences.

The joint distributions of returns from the asset classes are obviously crucial to any portfolio selection problem. In arriving at the optimal decision, a rational investor would not only need to ascertain the expected return that could be derived from the portfolio concerned, but also the level of risk that is involved. This issue is related to the period of investment as the behaviour of the asset classes tend to be different over different time periods. Cash, for example, would usually be considered a sound investment over a fairly short time scale. On the other hand, the return from Consols should be more stable relative to cash in the longer term, but not in the shorter term. However, unless one portfolio stochastically dominates all the others, assessing the characteristics of the asset classes is only one part of the portfolio selection problem.

The investor will also have to decide how to trade off risk with return, and this relates to the investor's risk preferences. Although the reason for investigating the asset fund is so that the complexity of a liability structure may be ignored, it is almost inevitable that implicit liabilities will play a role in the decision-making. Unless the investment is being made for no reason at all, the investor's perception of risk should depend on the individual's commitments. For example, if the investment is being made in order to repay a loan of a fixed amount at a fixed date, it would be sensible to look at the returns over the period concerned in nominal terms. The extent to which the investor is affected by returns other than that necessary to repay the loan will determine the shape of the utility function. However, if the investor simply wishes to maintain the purchasing power of the fund, it would then be more appropriate to look at real returns.

Consider the case where the investment is to be made over a one year time horizon. Suppose there are M asset classes available. Let the row vector of accumulation factors be $\mathbf{v}^T = (v_1 \dots v_M)$, where v_i is the random variable representing an amount of 1, accumulated over one year in asset class i. For the investment model described in Chapter 3, the asset classes could be ordered as follows: $\{i = 1,2,3,4\} \equiv$ {cash, Consols, index-linked gilts, UK equities}.

In addition, let the asset mix be represented by the row vector $\mathbf{w}^T = (w_1 \dots w_M)$, where w_i is the proportion invested in asset class i. Clearly, all the w_i's must sum to one. Although it would be common to include the non-negativity constraint of $\mathbf{w} \geq \mathbf{0}$, this will be ignored for the time being. So for a given amount, A to be invested, the accumulated sum, S at the end of the year is defined to be $S = A.\mathbf{w}^T \mathbf{v}$. Hence, the objective is to choose $\mathbf{w}$ which maximizes the expected utility of S, i.e.:

$$\max_{\mathbf{w}} E[U(S)].$$

Throughout this chapter, the investor will be assumed to have an exponential utility function, $U(S) = -\exp(-S/r)$. When written in this manner, r is generally referred to as the measure of risk tolerance of the investor. This is simply the inverse of the measure of absolute risk aversion (see Section 2.3.2). The higher the value of r the greater the tolerance to risk. So an investor with a higher value of r than another, for a given amount invested, would be expected to invest in a more risky portfolio. Use of the exponential function therefore enables us to compare a range of investor risk preferences.

4.2.2 Analytical Solutions

Under certain conditions, it may be possible to obtain closed form solutions for the utility maximizing portfolios. Using an exponential utility function in conjunction with the assumption that all random variables are normally distributed, Sherris (1992) derived a very useful analytical result for just two asset classes and a set of liabilities, L, by defining the ultimate surplus, $S = A.(w_1 v_1 + w_2 v_2) - L$. In fact, solutions for any number of asset classes may be obtained when the problem is defined in such a framework. The derivation below is related to that of Sherris (1992), but pertains to the asset only case with M asset classes.

If the investor has an exponential utility function, $U(S) = -\exp(-S/r)$, then the optimal portfolio is that which maximizes $E(-e^{-S/r})$, or minimizes $E(e^{-S/r})$. Putting $t = -1/r$, the objective function to be minimized, $E(e^{tS})$, may be recognized as the moment generating function (m.g.f.) of S. So whenever the m.g.f. of S exists, it should be possible to obtain analytical solutions to the problem. A situation in which this applies is when all the v_i's may be assumed to be normally distributed. If these assumptions hold, then S would be a linear combination of normal random variables, implying that the distribution of S would also be normal. Hence, the m.g.f. of S may be expressed as $\exp(\mu t+\sigma^2 t^2/2)$, with $\mu = A.\mathbf{w}^{\mathrm{T}}\mathbf{e}$ and $\sigma^2 = A^2.\mathbf{w}^{\mathrm{T}}\mathbf{C}\mathbf{w}$, where:

$$\mathbf{e} = \begin{bmatrix} E_1 \\ \vdots \\ E_M \end{bmatrix}$$

and:

$$\mathbf{C} = \begin{bmatrix} V_1 & \cdots & C_{M1} \\ \vdots & \ddots & \vdots \\ C_{1M} & \cdots & V_M \end{bmatrix}$$

E_i and V_i represent the mean and variance of the accumulation factor for asset i and C_{ij} is the covariance of accumulation factors for assets i and j. So, the problem reduces to one of minimizing $\exp(\mu t+\sigma^2 t^2/2)$, or more simply, minimizing $(\mu t+\sigma^2 t^2/2)$. Hence, the optimal portfolios may be obtained by setting $w_M = 1 - (w_1 + \ldots + w_{M-1})$ and solving the following system of linear equations:

$$\frac{\partial(\mu t + \sigma^2 t^2 / 2)}{\partial w_i} = 0 \qquad\qquad i = 1, \ldots, M-1.$$

For just two asset classes with normally distributed accumulation factors, the optimal portfolio will be (see Appendix B):

$$w_1{}^* = \frac{(r / A)(E_1 - E_2) + V_2 - C_{12}}{V_1 + V_2 - 2C_{12}} \tag{4.1}$$

and $w_2{}^* = 1 - w_1{}^*$. Although this is the simplest case possible, such an expression may be very useful in checking the results obtained using other methods. The approach described may also include various inequality constraints, such as $\mathbf{w} \geq \mathbf{0}$, using Kuhn-Tucker conditions (see Walsh, 1975).

However, the analytical approach does have some drawbacks. The method described above for the exponential function may be of limited use as it would not usually be appropriate to assume that the distributions of the accumulation factors are such that the m.g.f. of S exists. Where most other utility functions are concerned, the more restrictive assumption of normality may be necessary before it is possible to derive simple closed form solutions to the problem. Although quadratic functions do not require these assumptions, they suffer from various criticisms as described in Section 2.3.2. Even when all these limitations do not apply, analytical methods may still break down if complex liability structures are incorporated.

4.3 Numerical Optimization

4.3.1 Introduction

The method used above to obtain optimal solutions analytically is a simple example of classical optimization techniques. Restrictive assumptions were required for the objective function to be written in a form which could be optimized, even though these assumptions may not have been valid. However, an alternative could be to solve for the optimal portfolios numerically. By simulating a large number of scenarios (e.g. 1000) from the investment model, an estimate of the true objective function may be optimized using minimization algorithms.

There are many classifications for numerical optimization routines, depending on whether the objective function is linear or non-linear, whether the problem is constrained or unconstrained, etc. The present problem clearly requires a non-linear optimization routine as the objective function is non-linear in $\mathbf{w}$. If the non-negativity constraints of $\mathbf{w} \geq \mathbf{0}$ are introduced, then a linearly constrained non-linear routine will be needed. Although the intention is to eventually include non-negativity constraints in the problem, it will be shown in due course how this may be achieved by appropriately transforming the decision variables in the unconstrained case. So for the time being, it should suffice just to consider the problem of unconstrained non-linear optimization.

Another issue that arises in choosing a suitable optimization algorithm is whether first (and possibly even second) derivatives for the objective function are available. These may usually be computed using finite difference techniques, though at the cost of additional function evaluations. In addition, this is also conditional upon whether the objective function is sufficiently smooth. As the objective function being considered here is smooth, very efficient classes of non-linear routines involving first derivatives may be used in this particular problem.

4.3.2 Unidimensional Minimization

Before proceeding to discuss multidimensional minimization, it is worth noting that the most of these methods include unidimensional minimization (sometimes referred to as line minimization) as a sub-routine in the overall procedure. Therefore, an efficient unidimensional minimization routine, known as Brent's method (see Press *et al*, 1992), is outlined here.

Brent's method is a hybrid method which combines inverse parabolic interpolation with the golden section search to obtain a minimum point. Brent's method uses the former when the function is sufficiently well-behaved and switches to the less efficient but more robust golden section search when this fails. Given an interval in which a minimum is known to exist, the golden section search is guaranteed to converge to this point (as far as floating point precision will allow). This applies regardless of the behaviour of the function within the interval.

4.3.3 Multidimensional Minimization

The most commonly used multidimensional minimization algorithms generally work by an iterative process of deriving an appropriate search direction and minimizing the function along this direction using a unidimensional sub-algorithm. So for a function, f at the co-ordinate vector, $\mathbf{x}_k$ and given the search direction, $\mathbf{p}_k$, the kth iteration involves minimizing $f(\mathbf{x}_k + \lambda_k\mathbf{p}_k)$ with respect to the scalar, λ_k. Setting $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k\mathbf{p}_k$, the process is then repeated until the specified precision tolerance level is satisfied. The core of multidimensional optimization is described in various sources including Walsh (1979), Scales (1985) and Beale (1988). Details regarding the practical aspects of optimization may also be found in Gill *et al* (1981) and Press *et al* (1992).

Unlike the one-dimensional case, there is no multidimensional minimization routine which is ideal in all situations. The choice of method will usually depend on the nature of the problem and some degree of experimentation. For this reason, three different types of multivariate minimization routines have been considered. The first (Powell's method) is a basic technique which does not require the calculation of derivatives. The other two (conjugate gradient and quasi-Newton methods) involve the use of derivatives in the determination of search directions and are therefore more superior methods. However, they may not be appropriate if first derivatives cannot be computed.

Powell's method

For an N-dimensional problem, the variant of Powell's method described in Press *et al* (1992) begins for a starting vector $\mathbf{x}_0$ by initializing N search vectors $\mathbf{p}_1, ..., \mathbf{p}_N$ to the basis vectors. The function is successively minimized along each of these search directions, resulting in a new point $\mathbf{x}_N$. If $\mathbf{p}_d$ is the search vector which caused the largest decrease in function value, then set $\mathbf{p}_d = \mathbf{p}_{d+1}$, $\mathbf{p}_{d+1} = \mathbf{p}_{d+2}$ and so on until $\mathbf{p}_{N-1} = \mathbf{p}_N$. Next, set $\mathbf{p}_N = \mathbf{x}_N - \mathbf{x}_0$ before $\mathbf{x}_0$ is replaced by the abscissa of the minimum of $f(\mathbf{x}_N)$ along this new search direction $\mathbf{p}_N$. This is then repeated for the updated starting vector and search vectors until the process converges. The reason for displacing the direction of greatest descent is to reduce the chance of linear dependence between the resulting search vectors. In doing so however, the property of quadratic convergence is forfeited.

Conjugate gradient method

The aim of conjugate gradient methods is to find N mutually conjugate search directions so that the minimum of a quadratic function will be found in no more than N line minimizations. For non-quadratic functions, subsequent cycles of N line minimizations should eventually converge quadratically to the minimum. This is a widely used technique with the general form of the kth mutually conjugate search vector being:

$$\mathbf{p}_k = -\mathbf{g}_k + \beta_k \mathbf{p}_{k-1}$$

with the initial condition, $\mathbf{p}_0 = \mathbf{g}_0$ and where $\mathbf{g}_k$ is the gradient vector at $\mathbf{x}_k$. Variations of the method differ in the definition of the scalar, β_k. The routine used actually used here is the Polak-Ribiere variant which defines:

$$\beta_k = \frac{\Delta \mathbf{g}_{k-1}^{\mathrm{T}} \mathbf{g}_k}{\mathbf{g}_k^{\mathrm{T}} \mathbf{g}_k}$$

where Δ is the forward difference: $\Delta \mathbf{g}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$. Derivations of this together with the Fletcher-Reeves and the Hestenes-Steifel variants are set out in Scales (1985).

Quasi-Newton methods

Newton's method for obtaining a function minimum involves searching for a zero gradient. For any quadratic function, the gradient at the point $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{p}_k$ is given by $\mathbf{g}_{k+1} = \mathbf{g}_k + \mathbf{G}\,\mathbf{p}_k$ where $\mathbf{G}$ is the Hessian matrix. If $\mathbf{G}$ is positive definite, $\mathbf{x}_{k+1}$ is the minimum point when $\mathbf{g}_{k+1} = 0$, i.e. $\mathbf{p}_k = -\mathbf{G}^{-1}\,\mathbf{g}_k$. Under these conditions, the minimum can be obtained in a single iteration. With non-quadratic functions, this procedure will need to be carried out iteratively. However, there is no guarantee that the Hessian matrix will be positive definite at all points. The intention of quasi-Newton methods is to build up a good approximation to the inverse Hessian matrix through a series of iterations, without the computational burden of requiring the true Hessian at any stage. In addition, the process begins with a positive definite symmetric approximation, $\mathbf{H}_0$ (usually the unit matrix, $\mathbf{I}$) and updates successive approximations, $\mathbf{H}_k$ in such a way that they remain positive definite and symmetric. This ensures a downhill movement in each iteration as well as quadratic convergence near the minimum. The variant chosen here is the Broyden-Fletcher-Goldfarb-Shanno updating formula (see Scales, 1985), which is given by:

$$\mathbf{H}_{k+1} = \left[\mathbf{I} - \frac{\Delta \mathbf{x}_k \Delta \mathbf{g}_k^{\mathrm{T}}}{\Delta \mathbf{x}_k^{\mathrm{T}} \Delta \mathbf{g}_k}\right] \mathbf{H}_k \left[\mathbf{I} - \frac{\Delta \mathbf{x}_k \Delta \mathbf{g}_k^{\mathrm{T}}}{\Delta \mathbf{x}_k^{\mathrm{T}} \Delta \mathbf{g}_k}\right] + \frac{\Delta \mathbf{x}_k \Delta \mathbf{g}_k^{\mathrm{T}}}{\Delta \mathbf{x}_k^{\mathrm{T}} \Delta \mathbf{g}_k}$$

4.3.4 Bound Constraints and Global Optimization

Most techniques for handling constraints may be used in conjunction with any method of unconstrained optimization. When faced with a constrained optimization problem the general aim, as stated in Walsh (1979), is to reduce it to an unconstrained problem or to a sequence of such problems.

In the case of imposing simple bound constraints on the decision variables, such as the requirement for non-negative asset proportions, this may be achieved by transforming the decision variables as suggested by Nash (1979) and Walsh (1979). One approach could be to set $w_i = a_i^2$, for $i = 1, 2, 3$, and $w_4 = 1 - (w_1 + w_2 + w_3)$. Then the unconstrained optima for a_i will give the positively constrained optima for w_i. The main drawback is of course that w_4 may still be negative. If this occurs, then switching w_4 for the w_i with the largest weight and a single re-run should give the required result.

In addition to the occasional inconvenience of having to repeat the optimization once, transforming the variables is not the most numerically efficient way of addressing constraints as it increases the extent of non-linearity in the problem. Nevertheless, it had been noted by Box (1966) that despite not being 1 : 1, such transformations would still yield correct results as additional local optima would not be introduced as a consequence. Moreover, as the method seemed to work well in this situation, it was adopted in preference to the complexity of programming a more sophisticated linearly constrained algorithm.

The three unconstrained algorithms outlined in Section 4.3.3 were implemented as prescribed in Press *et al* (1992), with the above mentioned adjustments to allow for the non-negativity constraints. Where required, forward differencing was used to obtain first derivatives. In terms of efficiency, Powell's variant performed more poorly compared to the two gradient methods, as had been expected. Between the gradient methods, there did not appear to be a dominant routine although the quasi-Newton method occasionally gave rise to roundoff errors causing $\mathbf{H}_k$ to become nearly singular or non-positive definite. Given that the conjugate gradient method seemed to work well in these circumstances, it was felt unnecessary to attempt the recommended Cholesky factor modification on the existing quasi-Newton algorithm in order to correct this (see Scales, 1985 and Press *et al*, 1992).

Nevertheless, the results which will be shown in Section 4.4 were obtained using all three methods, primarily as a means of checking that the processes were properly converging to the global minimum. In addition, various different starting positions were also used for this purpose. Although global optima cannot generally be guaranteed, a variety of starting positions and routines can usually give a good indication as to whether the global optimum has been reached.

However, bearing in mind that all the asset proportions or weights have to be non-negative and sum to one, it may be feasible to construct a multidimensional grid of all admissible combinations to an acceptable resolution. With just four asset types, a four dimensional grid in steps of 5% would only require 1771 function calculations. For example, a co-ordinate on such a grid could be $\mathbf{w}^T = (0.40, 0.45, 0.00, 0.15)$.

This is remarkably efficient as each accumulation, (which makes up the bulk of the computation), only needs to be performed once regardless of the required number of utility functions to be used. These utility functions may then be applied to each accumulation for a mere fraction of the total effort. Having constructed the grid for a given utility function, the globally optimal weights correct to the nearest 5% may be obtained by simply sorting the grid-points. From here on, this will be referred to as the '*grid approach*', and will be used as a check for global optimization alongside any of the three optimization algorithms. The grid approach is in fact similar in principle to the method employed in Booth (1995a).

4.4 Results from the Optimization Process

4.4.1 One Year Case in Nominal terms

The aim here is to obtain the optimal asset mix for a fund over one year applying the same set of simulations considered in Section 3.6 earlier. The fund is assumed to have an initial amount of 1 invested with no explicit liabilities involved. As the exponential utility function is being used, the objective function may be optimized for different values of risk tolerance r, to see how the optimal mix may change with different risk preferences. Optimization routines are used to compute $\mathbf{w}$ numerically subject to $\mathbf{w} \geq \mathbf{0}$. This process is similar in principle to that found in Booth and Ong (1994).

The optimal asset mixes rounded to the nearest percent for various values of r are given in Table 4.1, including the means and standard deviations of the accumulated amounts. A graphical summary is shown in Figure 4.1.

r	*CASH*	*CON*	*ILG*	*EQ*	*Mean*	*S.D.*
4	0	0	0	100	1.1060	0.2098
2	8	0	12	80	1.1026	0.1670
1	46	0	13	41	1.0959	0.0859
1/2	66	0	13	21	1.0925	0.0461
1/4	75	0	13	12	1.0909	0.0303
1/8	80	0	13	7	1.0901	0.0238
1/16	82	2	12	4	1.0894	0.0209
1/32	80	8	10	2	1.0885	0.0188
1/64	80	10	9	1	1.0882	0.0183

Table 4.1. Optimal portfolios for the 1 year case in nominal terms, with the means and standard deviations of the accumulated funds.

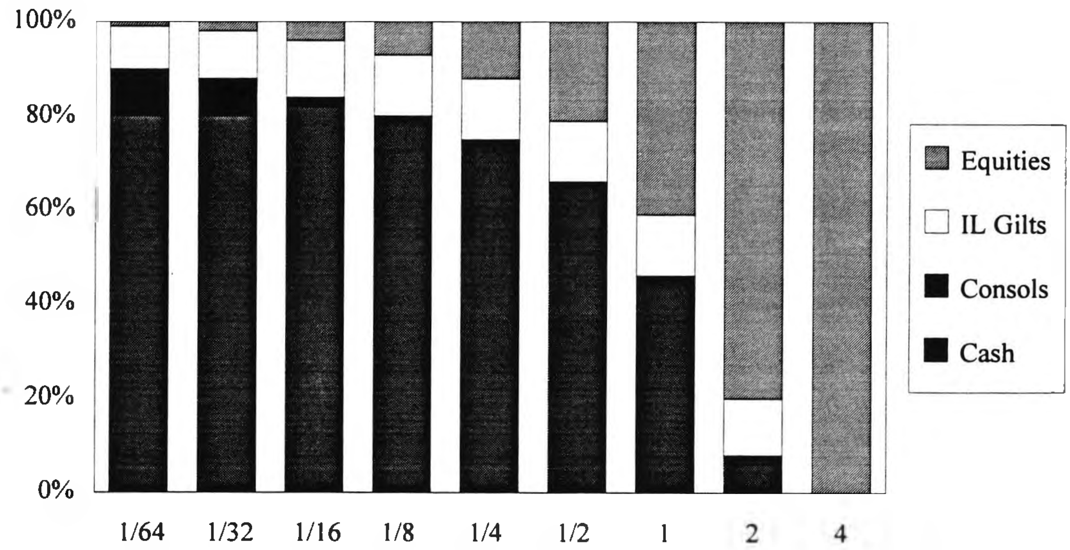


Figure 4.1. Optimal portfolios for the 1 year case in nominal terms at various values of r.

As the reason for using a range of risk tolerance parameters is merely to compare different risk strategies, the method for selecting these risk parameters was arbitrary. In this case, having found that the value of $r = 1$ resulted in a reasonably diverse portfolio, subsequent values of r were simply increased and decreased by multiples of two until the optimal mixes appeared to reach some limit.

Taking into account the main properties of the asset classes produced by the investment model, these results appear to be reasonable. The optimal strategy for the investor with the highest risk tolerance is found to be 100% in equities and this proportion decreases steadily as the level of risk tolerance reduces. Cash on the other hand is the asset class with the lowest variability, and its optimal proportion can be seen to increase as r is decreased. However, the proportion in cash seems to reach a limit of about 80% because the portfolio can be made more stable without reducing the expected return through diversification. This may be also seen by comparing the standard deviation of a fund entirely invested in cash (0.023) with the standard deviations of the portfolios which are optimal at values of r below 1/8, shown in Table 4.1. Another feature worth noting from the results is how insensitive the optimal proportions in index-linked gilts are to the value of r, remaining between 9% and 13% when r takes values of 2 or below.

Figure 4.2 is a graph showing the position of the utility maximizing portfolios in relation to those that are efficient from a mean-variance (*E-V*) perspective. The horizontal and vertical axes represent the means and standard deviations (S.D.) of the accumulated fund respectively. The *E-V* efficient frontier was not computed exactly but had been estimated in the following manner. From the set of 1771 portfolios tested using the grid approach, those portfolios which had a higher standard deviation but with the same or a lower mean than any other portfolio present were discarded, leaving the 'efficient' portfolios. A line drawn through these portfolios therefore is an approximate *E-V* efficient frontier. The fact these mixes were computed in steps of 5% accounts for the lack of smoothness seen in the frontier.

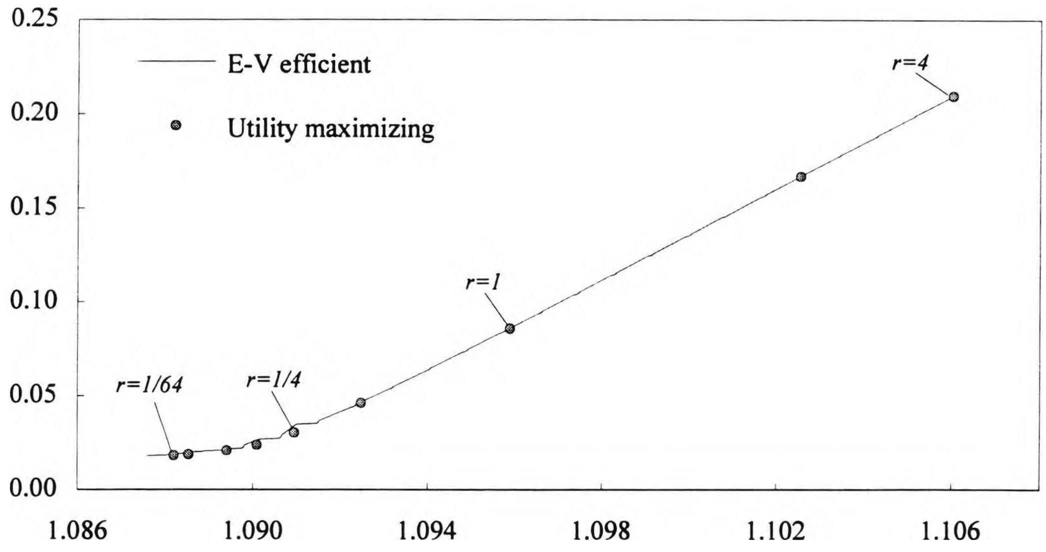


Figure 4.2. Graph of S.D. vs. Mean for *E-V* efficient and utility maximizing portfolios: 1 year nominal accumulations.

Looking at the graph it may be noted that the utility maximizing portfolios all seem to lie on the *E-V* efficient frontier. However this is hardly surprising given how closely the exponential utility function may be approximated by a suitable function of mean and variance. This approximation is made even closer by the fact that the distributions of accumulated amounts from the investment model over a one year period are nearly symmetric.

4.4.2 Twenty Year Case in Nominal terms

Next, consider how the optimal portfolios over twenty years compare with those over one year. In this situation, an issue arises which did not apply in the one year case. Suppose that for a given asset mix, an amount is to be invested without making any withdrawal over a twenty year. At the end of the first year, there are two possible interpretations regarding such a strategy. The total accumulations earned from each asset class could be i) reinvested in the same asset class, or ii) pooled and reinvested in the same proportions as those in which the fund was invested at the start of the first year. Algebraically, if $x_{m,t}$ is the accumulation of an amount 1 from asset class m over the period $t-1$ to t, and w_m is the proportion invested in asset class m, then for an initial investment of A, the accumulated fund of S at the horizon date could be either:

$$S = A.\sum_{m}(w_m.\prod_{t} x_{m,t}) \qquad \text{i)}$$

$$S = A.\prod_{t}(\sum_{m} w_m x_{m,t}) \qquad \text{ii)}$$

The first situation relates to a 'buy and hold' strategy which may be reasonable when no contribution or consumption is made to or from the fund during the period concerned. This is what essentially happens when a person chooses a particular asset mix in a group of unit trusts without issuing any switching instructions thereafter. Hence, the resulting asset mix at the horizon date may be very different from that chosen at the start.

The second strategy however would seem more sensible particularly if money may be added to or withdrawn from the fund. Due to the fact that the asset mix at time t will be the same as at the start of the investment period, the treatment of new money at time t would be straightforward. This in fact is a discrete time approximation to a continually rebalanced fund. As this approach would seem more general, it will be used in the twenty year case here.

r	*CASH*	*CON*	*ILG*	*EQ*	*Mean*	*S.D.*
64	0	0	14	86	8.1011	5.8861
32	0	0	18	82	8.0858	5.7440
16	0	0	24	76	8.0476	5.5341
8	16	0	17	67	7.7938	4.6621
4	25	24	0	51	7.1641	2.9844
2	17	48	0	35	6.6452	2.0050
1	19	58	0	23	6.2770	1.5019
1/2	24	61	0	15	6.0387	1.2777
1/4	29	60	0	11	5.9273	1.2119
1/8	31	59	0	10	5.9023	1.2042
1/16	34	56	0	10	5.9153	1.2276
1/32	37	54	0	9	5.8943	1.2312

Table 4.2. Optimal portfolios for the 20 year case in nominal terms, with the means and standard deviations of the accumulated funds.

Similar to the one year case, the optimal mixes over twenty years were obtained numerically and these results are shown in Table 4.2. From the table, a number of similarities to that of the one year case may be seen. Equities remain the main asset class for high risk strategies and their optimal proportions reduce consistently with decreasing r. Low risk portfolios are predominantly in fixed interest assets as before, although the mix between cash and Consols now favours Consols which makes intuitive sense as their returns are more stable over a twenty year term.

One notable difference is the trend in the index-linked gilt proportions over various levels of risk tolerance. These do not feature in the low/medium risk portfolios and only appear when r takes very high values. Being the asset class with a mean and standard deviation second only to equities may partly explain their presence in high risk portfolios. The reason for failing to appear in lower risk strategies is probably due to their unfavourable balance between risk and reward in relation to the other asset classes, when the control variable is denominated in nominal terms.

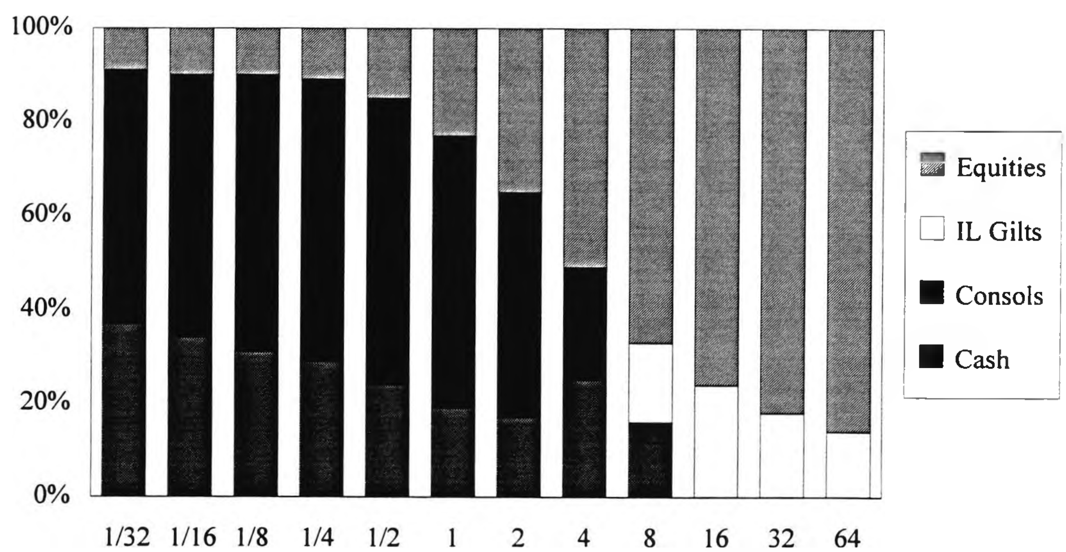


Figure 4.3. Optimal portfolios for the 20 year case in nominal terms at various values of r.

The effect of rebalancing the fund at the end of each year may also be seen to alter the pattern of optimal mixes. Looking at Figure 4.3 it may appear as though the range of r has not been extended far enough to allow the highest risk strategy to emerge, presumably 100% in equities as in the one year case. While $r = 64$ may not be the most extreme level of tolerance possible in this situation, increasing r to infinity would not produce an optimal mix with 100% in equities. This is because the mean accumulation of an all-equity portfolio is 8.090, which is smaller than 8.1011 when equities only make up 86% of the fund with the balance being in index-linked gilts. Rebalancing can therefore be seen to remove the linear relationship between the expected accumulations over the twenty year period.

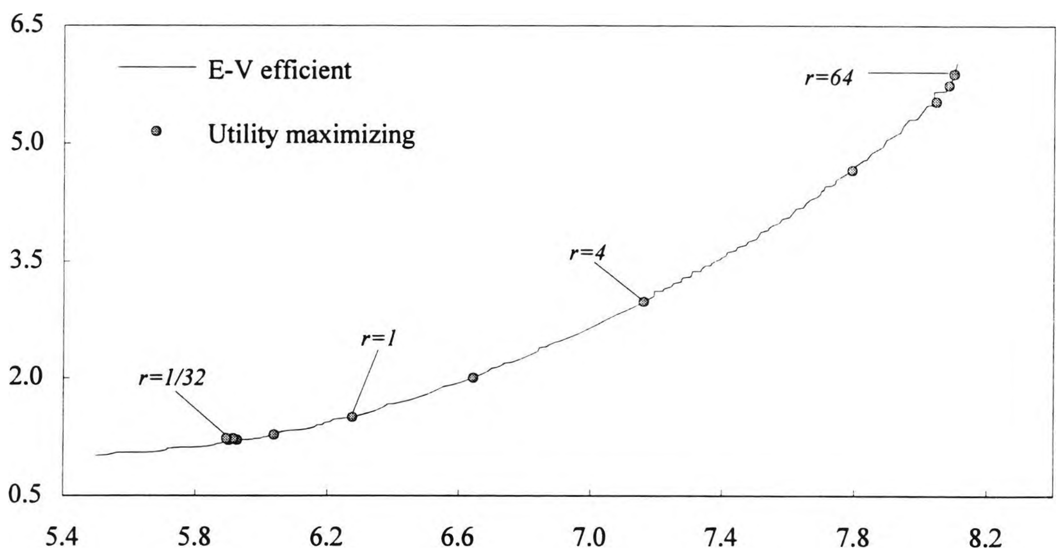


Figure 4.4. Graph of S.D. vs. Mean for *E-V* efficient and utility maximizing portfolios: 20 year nominal accumulations.

As before the *E-V* efficient frontier and the utility maximizing portfolios may be plotted on the same graph which can be seen in Figure 4.4. The bunching of the high risk utility maximizing portfolios at the top right-hand-side of the diagram confirms that they are very close indeed to the ultimate high risk portfolio. This may also be seen from observing the scatter plot of all the feasible portfolios in steps of 10% (see Figure 4.5). The uppermost point represents the 100% equities portfolio while the point just below and to the right of it represents 90% equities and 10% index-linked gilts. The 100% equities portfolio is clearly inefficient in a mean-variance framework.

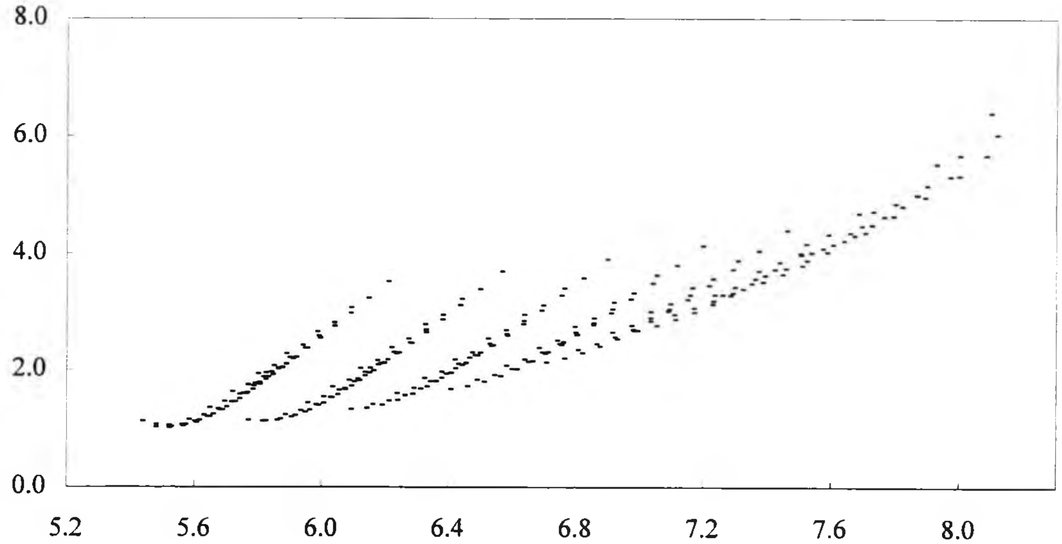


Figure 4.5. Graph of S.D. vs. Mean for all 286 portfolios: 20 year nominal accumulations.

Returning to Figure 4.4, it appears that at least two of the utility maximizing portfolios (r = 1/16, 1/32) lie above the efficient frontier. This may be checked by considering the means and standard deviations given in Table 4.2. For low levels of risk tolerance, the optimal mixes would seem to defy the rules of mean-variance efficiency. When r = 1/8, the optimal fund has a mean and standard deviation of 5.9023 and 1.2042 respectively. When r is halved to 1/16, the mean and variance now rise to 5.9153 and 1.2276 respectively. At r = 1/32, their respective values are 5.8942 and 1.2312.

In order to ensure that the results had been computed correctly, the optimization procedure was checked by appropriately rescaling the objective function and trying additional starting positions with all three optimization routines. The function values in the neighbourhoods of the supposed optima were computed and all were found to be less than these optima. As a result, it was felt that the possibility of computational error had been sufficiently small for it to be ruled out.

In an ideal situation, the results would be compared with analytical solutions for confirmation. However, this was not possible here being a twenty year annually rebalanced portfolio. Furthermore, this divergence from mean-variance efficiency could be due to the effect of third or higher order moments, which would not be taken into account using a closed form solution to the problem. An alternative might be to create a simple situation where analytical solutions are both feasible and valid, and to then observe whether or not a similar phenomenon could occur.

4.4.3 A Test Case with Normally Distributed Returns

Consider the situation where there are only two asset types available and where the objective is to maximize the expected utility of returns over a given period without rebalancing during that period. Let it also be assumed that the returns over the period are normally distributed and that the utility function is of the exponential form. In this situation, the solutions may be derived analytically which, in theory, should be identical to those obtained numerically.

Two sets of a thousand standard normal pseudo random variables were generated as proxies to these asset returns. The first two moments from these samples were as follows: E_1 = 0.007488, E_2 = −0.02190, V_1 = 0.97708, V_2 = 0.95138, C_{12} = −0.00955. Measures of skewness and kurtosis were also computed for these two samples to ensure that they were approximately normal. Inserting the values given above into equation (4.1) would therefore yield the optimal proportion to be invested in asset type 1. This was done for a range of values of r and the solutions were then compared with those obtained through numerical optimization methods (see Table 4.3).

For each risk tolerance level (except zero) in Table 4.3, two sets of figures are given. The first set of numbers relate to the analytical optima and the ones below them in brackets refer to the numerical optima. For example, in the case when r = 25.6, the optimal proportion in asset type 1 obtained from the formula is 87.97% which corresponds to a mean and standard deviation of 0.0040 and 0.8763 respectively. The numerical optimum here is 88.06% in asset type 1 which gives a mean of 0.0040 and a standard deviation of 0.8771.

r	w_1	*Mean*	*S.D.*
25.6	87.97	0.0040	0.8763
	(88.06)	(0.0040)	(0.8771)
12.8	68.66	-0.0017	0.7416
	(68.73)	(-0.0017)	(0.7419)
6.4	59.00	-0.0046	0.7039
	(59.07)	(-0.0045)	(0.7041)
3.2	54.17	-0.0060	0.6941
	(54.27)	(-0.0060)	(0.6942)
1.6	51.75	-0.0067	0.6917
	(51.89)	(-0.0067)	(0.6918)
0.8	50.55	-0.0071	0.6910
	(50.67)	(-0.0070)	(0.6911)
0.4	49.94	-0.0072	0.6909
	(49.64)	(-0.0073)	(0.6909)
0.2	49.64	-0.0073	0.6909
	(47.51)	(-0.0079)	(0.6913)
0.1	49.49	-0.0074	0.6908
	(43.64)	(-0.0091)	(0.6954)
0	49.34	-0.0074	0.6908
	-	-	-

Table 4.3. Optimal portfolios, means and standard deviations for normally distributed returns

The analytical results seem reasonable, with means and standard deviations decreasing monotonically as r decreases. However, when the solutions are computed numerically, the problem outlined in the twenty year case emerges. As the risk tolerance level is reduced, both sets of solutions remain very similar until r starts to fall below 0.4, when the optima begin to diverge. At low values of r, the numerically derived solutions again lead to funds which are not mean-variance efficient.

Given that the distributions of returns are normal in both assets types, a utility maximization approach should yield the same solutions as those derived using a mean-variance approach. The analytical solutions shown in Table 4.3 seem to confirm this. But due to the assumption regarding normality of returns, these solutions should also be the same as their numerical equivalents. Assuming that the numerical optimization has been carried out correctly and to adequate precision, there must be another explanation for this discrepancy.

A utility function may be thought of as a means of attaching weights to individual outcomes. In a risk neutral situation, equal weighting is given to each outcome, thus making the expected utility proportional to the expectation. Now if the decision maker is risk averse, then greater weights will be attached to the more severe outcomes. As the level of risk aversion increases, the expected utility becomes more and more dependent on the lower (i.e. left) tail of the distribution of outcomes. This is when the estimation of expected utilities by simulation can begin to fail. A finite sample will generally provide a poor description of the extreme values of an underlying distribution.

In the case of the exponential utility function, as $r \rightarrow 0$, the expected utility calculated from n simulated outcomes tends towards $1/n$ times the utility of the worst outcome. As the value of the worst outcome from a set of simulations is a particularly unreliable quantity, so too is the expected utility. When r is moderately small, the credibility of the expected utility can be improved upon slightly by increasing the sample size.

However, this is generally not an appropriate remedy for much smaller values of r as the number of simulations required to maintain the same level of credibility tends would not be computationally feasible.

The conclusions which may be drawn from the last two examples is that the expected utilities and hence the optimal asset mixes may be unreliable at very low levels of risk tolerance, in some cases leading to inefficient portfolios. Although this was only tested here for the exponential utility function, in principle similar problems may occur in other utility functions with low tolerance to risk. Computing the means and variances may give an indication of such a problem occurring, as would other measures such as the ratio of the expected utility to the utility of the worst outcome.

4.4.4 Twenty Year Case in Real terms

In the previous cases involving one and twenty year time horizons, the expected utilities were calculated from the nominal accumulations. While some investors may have specific reasons for considering nominal returns, in a world of inflation, the desire to achieve good real returns may be more sensible. It would therefore be logical to repeat the twenty year case in Section 4.4.2 with accumulations measured in purchasing power terms. The results in real terms are given in Table 4.4 and Figure 4.6 below.

The optimal portfolio when the investor has high tolerance to risk is made up of equities and Consols, the two highest yielding assets in real terms. Although Consols actually give a lower mean and higher variance than equities, a combination of the two is still efficient as there is little correlation between these two assets. As r decreases, the proportion in equities diminishes to be replaced by index-linked gilts. Cash also begins to take over from Consols, being less volatile in real terms. Cash is highly correlated with Consols which is possibly why these two asset classes rarely appear together.

r	*CASH*	*CON*	*ILG*	*EQ*	*Mean*	*S.D.*
64	0	12	0	88	2.7572	1.5063
32	0	14	0	86	2.7564	1.4828
16	2	15	0	83	2.7538	1.4440
8	16	8	0	78	2.7438	1.3544
4	29	0	0	71	2.7250	1.2546
2	10	0	29	61	2.6390	0.9717
1	1	0	52	47	2.5232	0.7002
1/2	0	0	68	32	2.3822	0.4608
1/4	0	0	80	20	2.2523	0.2935
1/8	0	0	88	12	2.1581	0.2003
1/16	0	0	92	8	2.1091	0.1639

Table 4.4. Optimal portfolios for the 20 year case in real terms, with the means and standard deviations of the accumulated funds.

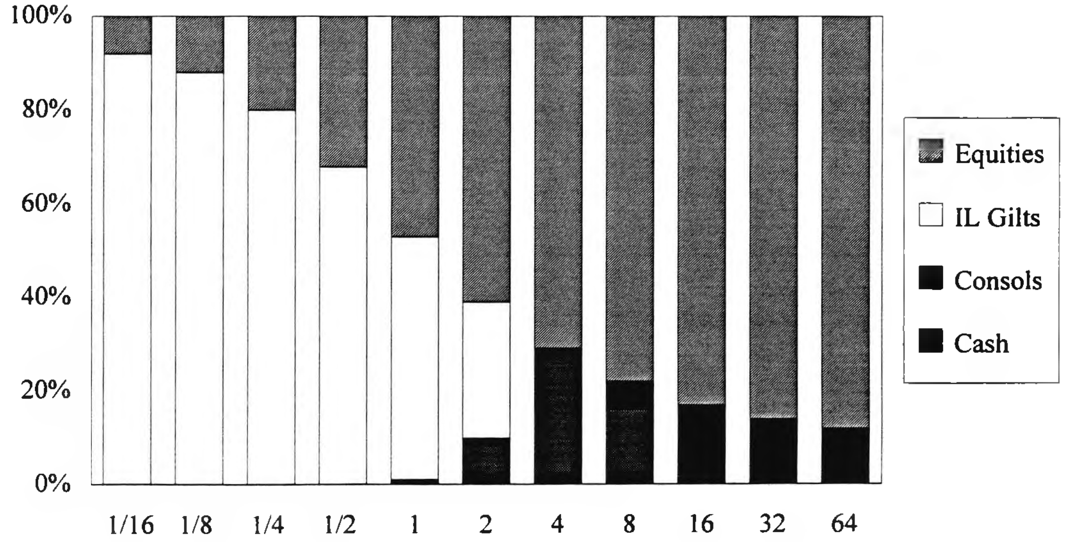


Figure 4.6. Optimal portfolios for 20 year case in real terms at various values of *r*.

At lower risk tolerance levels, cash is displaced by index-linked gilts and the proportions in the latter continue to rise as the value of r reduces. Small holdings in equities are still preferred at low values of r because they are uncorrelated with index-linked gilts, thus helping to reduce variability in the overall portfolio while providing better returns.

In terms of mean-variance analysis, the optimal portfolios appear reasonable. Both the mean and variance reduce monotonically as r reduces. Nevertheless, bearing in mind the problems that occurred in respect of nominal accumulations, the optimal portfolios at lower risk levels should be interpreted with some degree of caution. If r had been reduced further, some deviation from mean-variance efficiency may have resulted.

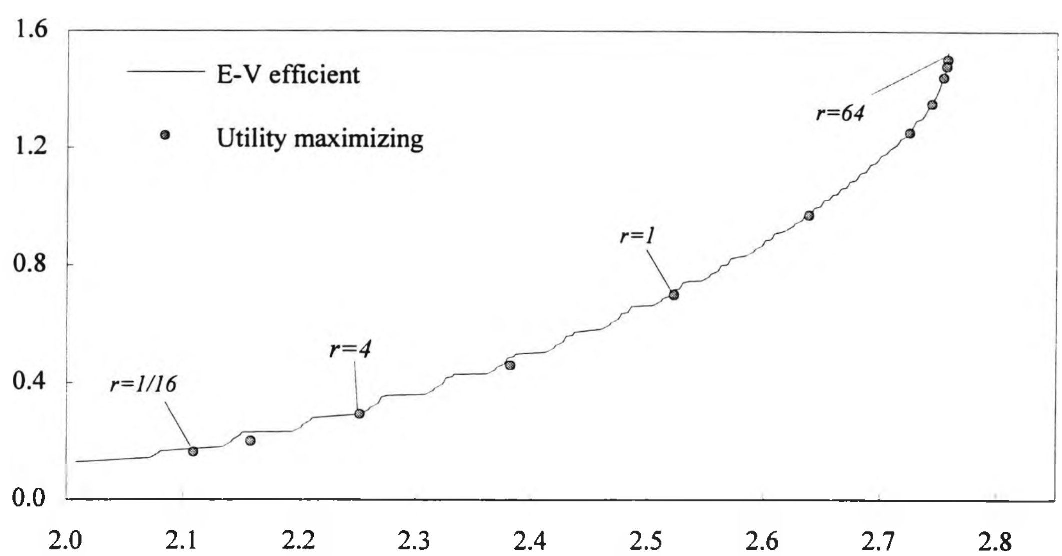


Figure 4.7. Graph of S.D. vs. Mean for *E-V* efficient and utility maximizing portfolios: 20 year real accumulations.

Although the values of r in the nominal and real case are not directly comparable, it may be seen that the two sets of portfolios are distinct and that their differences are intuitively reasonable. But the fact that they are so different does bring into question which set of optimal portfolios should one be looking at. In deciding on this, some consideration must be given to the reason for making the investment. If say the fund is to be used to repay some form of fixed interest loan, then nominal amounts may be more sensible. On the other hand, if the fund is a proxy for a pension fund, then real amounts should be used. As the choice of transformation on the accumulated amount depends on the purpose of the fund, the 'asset only' utility maximization process does after all involve liabilities, albeit implicitly as in the case here.

4.4.5 Twenty Year Case in Real terms: a Truncated Distribution

So far, one of the conclusions that may be drawn from the situations studied earlier is that the utility maximizing portfolios all seem to be approximately mean-variance efficient. While this could be expected in the one year case given the nearly symmetric distributions produced by the investment model, it was interesting to observe how close the twenty year optimal portfolios were to the mean-variance efficient frontier. One of the reasons for applying utility theory in these investigations was to allow scope for involving more than just the first two moments in the decision-making process. Perhaps these distributions were still not adequately skewed to make a significant difference. One way of increasing the level of skewness could be to use a truncated distribution.

Consider the investor with the same choice of asset classes as before, looking to maximize the expected utility of real accumulations over a twenty year period. But suppose that an opportunity has arisen to invest through a special fund where the accumulated real amount is guaranteed to be no less than a fund earning a guaranteed real rate of return of 3% per annum. In other words, if S is the real accumulation for the normal fund, then the accumulated amount for the special fund is: max $(S,\ 1.03^{20})$. Including such an investment guarantee would do a number of things to the distribution of the accumulations. It would increase the mean while lowering the variance for a given asset mix. More importantly the truncated distribution would be less symmetric.

For example, when 100% of the fund is invested in equities the guarantee increases the measure of skewness from 1.91 to 2.41. This should increase the likelihood of producing utility maximizing portfolios that are not mean-variance efficient.

r	*CASH*	*CON*	*ILG*	*EQ*	*Mean*	*S.D.*
8	0	0	0	100	2.9116	1.5284
4	3	8	0	89	2.8870	1.3894
2	17	7	0	76	2.8422	1.2279
1	27	0	6	67	2.7828	1.0809
1/2	0	0	48	52	2.6008	0.7319
1/4	0	0	64	36	2.4396	0.4954
1/8	0	0	81	19	2.2557	0.2739
1/16	0	0	90	10	2.1348	0.1790

Table 4.5. Optimal portfolios for the truncated 20 year case in real terms, with the means and standard deviations of the accumulated funds.

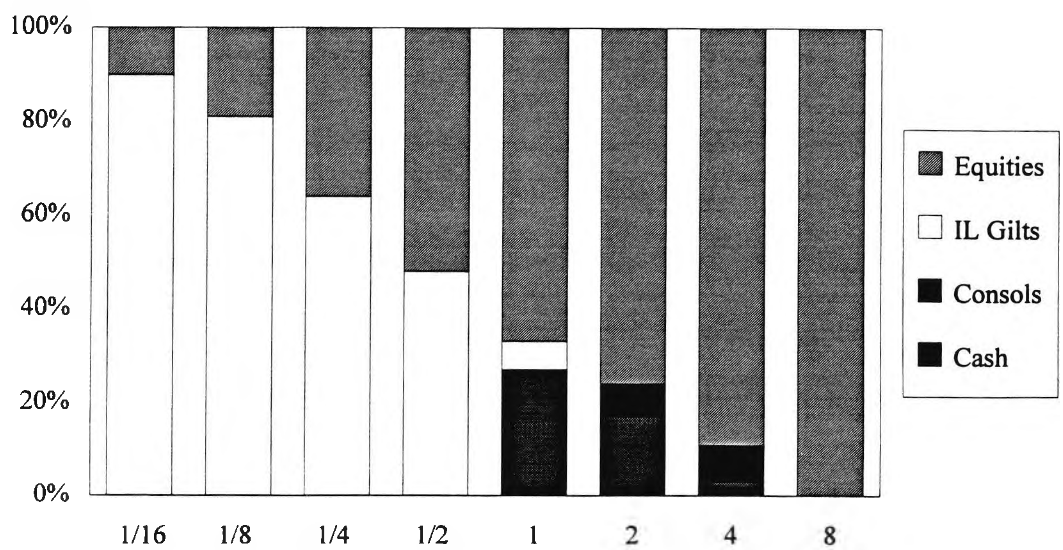


Figure 4.8. Optimal portfolios for the truncated 20 year case in real terms at various values of *r*.

The optimal portfolios for this fund are shown in Table 4.5 and Figure 4.8. The guarantee appears to have made what were previously very risky strategies more favourable. This was to be expected as the guarantee should have a greater effect on the more volatile portfolios. In particular, 100% in equities is optimal at $r = 8$, whereas with the normal fund (see Table 4.4), only 92% were optimal at the much higher risk tolerance level of $r = 64$. Otherwise the general pattern of mixes is broadly similar to those seen in Section 4.4.4.

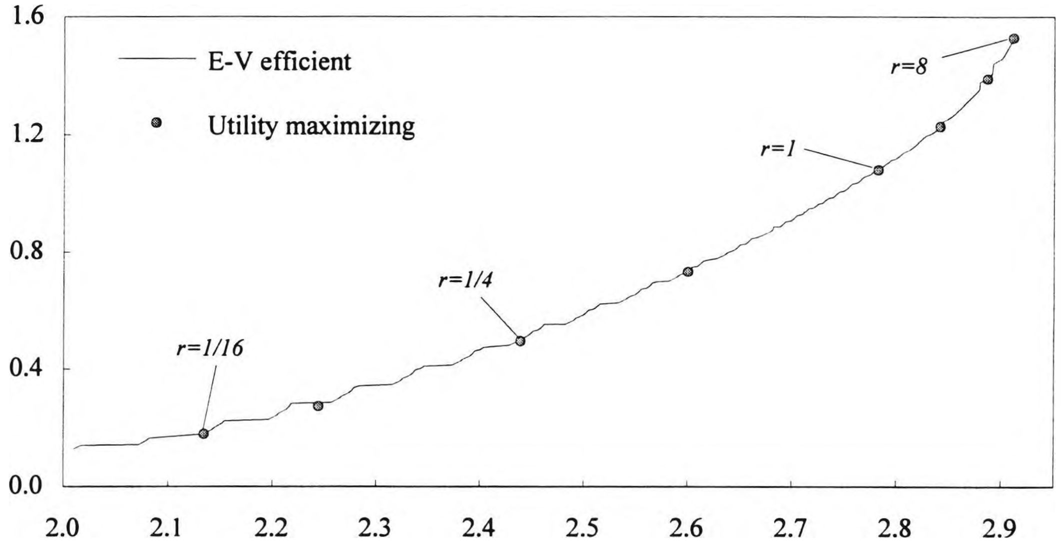


Figure 4.9. Graph of S.D. vs. Mean for the *E-V* efficient and utility maximizing portfolios: truncated 20 year real accumulations.

The efficient frontier and the utility maximizing portfolios may be seen in Figure 4.9. Despite the higher degree of skewness characterized by these distributions, the utility maximizing portfolios still seem to remain very close to the *E-V* efficient portfolios. This would appear to support the proposal made by many, including Levy and Markowitz (1979), that the mean-variance framework is in practice an adequately good approximation to the utility maximization approach. However, the utility approach has the additional feature that it can determine which of the infinite number of *E-V* efficient portfolios is optimal.

4.5 Uncertainty in the Estimates of the Optimal Portfolios

4.5.1 Sources of Uncertainty

The values obtained for **w*** for a given utility function are really only estimates of the optimal portfolios and are subject to a few sources of error. One possible source of error is in the numerical optimization, though all reasonable measures have been taken to ensure that the values quoted above are at least correct to 1%. Another pertains to the accuracy that may be attributed to the number of simulations used and how this relates to the uncertainty in the investment model.

Until now, inaccuracies in computing **w*** have only been shown to occur when the risk tolerance level is small. This had been attributed to the fact that Monte Carlo simulation rarely gives an adequate representation of the tails of a distribution. But regardless of the risk tolerance level, **w*** will still be subject to some degree of sampling error due to there being a finite number of simulations used. It would therefore be very useful to know the extent of error in the production of the point estimates, **w*** due to sampling error. The main aim of this section is to obtain approximate error bounds for the optimal portfolios. Ideally this would entail deriving 95% confidence intervals for **w***. However, as the distribution of **w*** is unknown, the alternative is to compute the approximate standard errors for these estimates.

4.5.2 Jackknife Standard Errors

A robust technique which may be used for estimating the standard error and bias of an estimate is the jackknife as described in Efron and Tibshirani (1993). For a sample of size n, the jackknife involves calculating the estimates for the sample, leaving out one observation at a time. So if $s(\mathbf{x})$ is an estimator from a sample $\mathbf{x} = (x_1 , \dots , x_n)$, the ith jackknife replication of this estimator is $s(\mathbf{x}_{(i)})$, where $\mathbf{x}_{(i)} = (x_1 , \ldots x_{i-1} , x_{i+1} , \ldots x_n)$ is the ith jackknife sample. Defining:

$$\bar{s} = \frac{1}{n}\sum_{i=1}^{n} s(\mathbf{x}_{(i)})$$

the jackknife estimate of the standard error is:

$$\hat{e} = \sqrt{\frac{n-1}{n}\sum_{i=1}^{n}(s(\mathbf{x}_{(i)}) - \bar{s})^2}$$

and the jackknife estimate of bias is:

$$\hat{b} = (n-1)(\bar{s} - s(\mathbf{x}))$$

In the context of the asset fund, $s(\mathbf{x})$ may be interpreted as the estimate for the optimal proportion in a particular asset class, based on a sample of n simulated investment scenarios denoted by $\mathbf{x}$. As $n = 1000$, each $s(\mathbf{x}_{(i)})$ is the optimal asset mix based on 999 scenarios, with the ith scenario left out.

Although the method involves recalculating the sample estimate n times, the resulting jackknife replications, $s(\mathbf{x}_{(i)})$ are likely to be very close to the original estimate, $s(\mathbf{x})$ as only one observation is left out on each occasion. For the purpose of estimating the standard error of an estimated optimal asset proportion, $s(\mathbf{x})$, this may be an advantage in one respect as it should not require many iterations to converge to each $s(\mathbf{x}_{(i)})$, if $s(\mathbf{x})$ is used as the initial starting position. However, this efficiency factor may be offset somewhat by the need to increase the precision level of the optimization algorithm to a sufficient degree, so that the subtle differences between $s(\mathbf{x})$ and $s(\mathbf{x}_{(i)})$ may be accurately reflected in the estimates of bias and standard error.

The jackknife is a reasonable technique in this context as it does not generally make any assumptions about the distribution of $s(\mathbf{x})$. A situation where it may fail, however, is when the distribution of $s(\mathbf{x})$ is not smooth. Although this mainly relates to situations where there are discontinuities in the distribution of $s(\mathbf{x})$, the deficiency may still occur in the case of the asset fund as the optimal asset proportions are constrained to non-negative values. If the optimal proportion in an asset class, $s(\mathbf{x})$ is found to be zero, the removal of one observation, x_i will be unlikely to increase this by enough to produce a positive value for $s(\mathbf{x}_{(i)})$; so the jackknife estimate of the standard error will probably be zero as well, even though the true standard error may be strictly positive. Having said this, the estimated standard error will only be significant if the unconstrained optimal proportion is very close to the constrained optimal, which should be a rare occurrence.

r	*CASH*	*CON*	*ILG*	*EQ*
64	0	0	14.17	85.83
	-	-	(2.98)	(2.98)
16	0	0	24.27	75.73
	-	-	(2.20)	(2.20)
4	24.99	24.22	0	50.80
	(5.93)	(5.93)	-	(1.68)
1	18.80	58.08	0	23.12
	(3.03)	(2.87)	-	(1.20)
1/4	29.01	60.14	0	10.86
	(2.93)	(3.40)	-	(1.43)
1/16	33.83	56.47	0	9.70
	(5.57)	(7.27)	-	(2.89)
1/64	39.42	51.83	0	8.75
	(8.27)	(9.42)	-	(5.03)

Table 4.6. Optimal portfolios and standard errors for the 20 year case in nominal terms

Table 4.6 shows the optimal portfolios and the jackknife estimates of the standard errors for a range of risk tolerance levels for an asset fund over a twenty year period. The estimates of bias were close enough to zero to be ignored. In order to ameliorate the problem of non-smooth estimators, the standard errors shown were estimated on the basis that any asset type which had an optimal proportion of zero was excluded from the procedure. For example, the optimal mix when $r = 64$ is 14.17% in index-linked gilts and 85.83% in equities. The jackknife estimates of their standard errors assuming that only these two assets are available are 2.98%.

A curious pattern may be observed from these results. From the highest risk tolerance level shown, the standard errors in each asset type at first decrease as r decreases. This may be intuitively reasonable as the high risk portfolios are associated with higher variances, which feed through to the standard errors of the optimal portfolios. When r is 1/4, the trend is reversed with the standard errors beginning to increase with decreasing levels of risk tolerance. This is caused by the excessive reliance of the objective function on the extreme tails of the distribution of the accumulated fund, as explained in Section 4.4.3.

This reasoning may also be confirmed by considering the empirical distribution of the jackknife replications. For higher levels of risk, practically all the replications exhibited some variability relative to the original estimate, with the degree of variability being higher for higher levels of risk.

At very low risk levels, most replications remained the same correct to five decimal places with a few replications showing quite considerable variability. In the case when $r = 1/64$ for example, only two replications out of one thousand produced significant variability from the original Consols estimate of 51.8%. The values obtained for these replications were 46.6% and 59.5%, which together account for over 98.5% of the standard error estimate.

4.5.3 Assessment of Uncertainty

Taking an overall view of the results, there appears to be a reasonable amount of uncertainty in the estimates obtained using a thousand simulations. If it may be assumed that the 95% confidence interval for the optimal weights is sufficiently close to $s(\mathbf{x}) \pm 2\hat{e}$, then ignoring the very low risk tolerance levels, these confidence intervals would roughly be between 2.5% and 12% either side of the original estimates, depending on the value of r and the asset class. But even ± 5% may be considered too large to justify higher precision optimization routines when the grid approach in steps of 5% would give results which largely fall within the 95% confidence intervals.

There is however another issue that is worth bearing in mind when considering the level of uncertainty inherent in the optimal portfolio weights. While it is true that the confidence intervals for these weights could be reduced by increasing the number of simulations used, there are doubts regarding the real benefits of such an action. The disadvantage of performing say ten times the number of simulations would need to be outweighed by the ability to obtain more accurate and meaningful results.

It is worth bearing in mind that the simulations are derived from an investment model, which like any other model has its deficiencies. However, this is particularly significant here due to the acknowledged difficulties in modelling economic time series. An infinite number of simulations will only serve to reflect the full features of the model, rather than reality.

A far more sensible approach could be to balance the computational effort involved with the amount of confidence associated with the investment model used. For instance, in these investigations, 1000 simulations had been found to be a manageable sample size to deal with. But in relation to the accuracy of the optimal proportion in Consols say, the standard error obtained at r = 4 is almost 6%. This contrasts with the 95%

confidence interval for the twenty year accumulation in Consols of about $5.4 \pm 0.07 \equiv (1.0873^{20}, 1.0887^{20})$, which is fairly narrow. So while it may be desirable to reduce the standard error on the optimal Consols weight by running more simulations, it may or may not be reasonable to assume the narrower confidence interval for the accumulation factor for that asset class which this implies. However, in assessing the level of uncertainty associated with the optimal proportion in Consols, consideration would also need to be given to the confidence intervals of the accumulation factors for the other asset types as well, as the optimal proportion in Consols is dependent on them too.

Although the level of uncertainty in the investment model does not have a direct relationship with sampling error, performing an extremely large numbers of simulations will tend to result in spurious accuracy. Bearing this in mind, it would be helpful to have some indication of what may be a reasonable number simulations to use. A possible starting point could be to look at the projected mean accumulations and their standard errors, particularly as these are easy to compute and interpret. Only if these standard errors are found to be too large would the simulation number then seem worth increasing. In a mean-variance context, the optimal decisions are very sensitive to the means. Hence, if one is unsure about these statistics, then it would be sensible not to place too much confidence on the results derived from them.

4.6 Summary

In this chapter, it has been shown how investment decisions may be derived in a utility maximization framework. Under fairly restrictive assumptions regarding normality of asset returns and liability cash flows, closed form solutions may be found. However, if more complex asset or liability models are to be used, then numerical methods appear to be the more sensible route to pursue.

Using numerical optimization routines, results for the optimal asset allocation strategies for an asset fund were computed for short durations in nominal terms and long durations in both nominal and real terms. These appeared sensible in relation to the investment model described in the previous chapter. It was also noted that these portfolios were approximately if not on the mean-variance efficient frontier, even when the distributions of the portfolios were clearly skewed.

However, as with all portfolio selection models, the optimal asset mixes are subject to sampling error and model error. Jackknife estimates of standard error in the optimal portfolios for 1000 simulations were found to be between about 1%-5% under normal situations and even greater for very low risk tolerance levels. This indicated that with this number of simulations, accurate optimal mixes would rarely be possible. Unfortunately, increasing the number of simulations may lead to spurious accuracy as the results are ultimately subject to uncertainties which cannot really be reduced without improving the credibility of the investment model itself.

5. INCORPORATING LIFE OFFICE LIABILITIES INTO THE DECISION-MAKING PROCESS

5.1 Introduction

In this chapter, a description is given of how the decision-making process may apply in the context of a UK proprietary life office. It will be assumed throughout that the office only transacts conventional non-profit business, with surplus being distributed to shareholders in the form of dividends. Hence, the asset allocation decisions will largely be based on the expected utility of shareholders' dividends.

It is also proposed that the long-term objectives of the office will be to:

1. distribute smoothed dividends to shareholders
2. maintain the same free asset ratio that existed at the start of the projection period

These aims contradict each other because if the degree of smoothing is high, then there is a greater possibility that the free asset ratio will tend to drift away from its intended level. On the other hand, if excessive priority is given towards maintaining the desired level of free assets, then the resulting dividends may be too volatile. As in practice, an acceptable balance may have to be struck between these two objectives.

In order to incorporate a realistic liability structure into the investigations, the application of a model office will be essential. A description is given regarding the treatment of cashflows, reserves and surplus in the model office. Following this, it would also be appropriate to discuss the issue of shareholders' utility. But before the structure of the model is explained in any detail, it would be helpful to first set out the types of contracts that may be issued by the office, together with the various assumptions made in respect of individual policies and business volumes.

5.2 Liability Profile

5.2.1 Benefit Structure

Two broad classes of business used in these investigations are endowments and index-linked annuities. Unless otherwise stated, endowments are taken to be annual premium non-profit endowment assurances with a guaranteed sum assured, payable either at the end of the year of death or at maturity. Index-linked annuities are assumed to be single premium non-profit temporary annuities with annual payments increasing in line with inflation. The use of index-linked annuity contracts enriches the study somewhat as it increases the number of liability types. In all cases, mortality rates are assumed to be deterministic and correspond to the premium bases stated in Section 5.2.3. Although the assumption of deterministic mortality rates may remove a layer of realism from this study, it does help to minimize the number of stochastic variables involved and should ease the process of interpreting the results from what will be a fairly complex set of investigations.

5.2.2 Expenses and Commission

All contracts incur acquisition costs in the form of initial expenses and commission, which are payable in full at inception. From the second year onwards, maintenance expenses and renewal commission are incurred simultaneously with each premium received or annuity payment made. In all cases, expenses are assumed to increase stochastically each year in line with the retail price index. Commission rates are expressed as fixed percentages of the actual premiums paid. Apart from future inflation rates which are stochastic, these quantities are assumed to be the same as those given in the premium bases below.

5.2.3 Premium Bases

The following assumptions apply to premiums rated at the start of the projection period:

All contracts:-

- interest: 8%
- inflation: 5%
- risk discount rate: 15%

20 year endowments aged 50 at entry:-

- sum assured: 10000
- initial expenses: 100 per policy
- renewal expenses: 30 per policy, increasing with RPI
- initial commission: 50% of first year's premium
- renewal commission: 2.5% of subsequent premiums
- mortality: AM80(2)
- profit criterion: 40% initial commission at risk discount rate
- reserves: net premium valuation at a rate of interest of 6%

10 year endowments aged 50 at entry:- as in 20 year endowments except for

- initial commission: 30% of first year's premium

20 year index-linked annuities aged 65 at entry:-

- annuity: 1000 per annum increasing with RPI
- initial expenses: 100 per policy
- renewal expenses: 30 per policy, increasing with RPI
- initial commission: 2% of single premium
- mortality: IM80(1)
- profit criterion: 50% initial commission at risk discount rate
- reserves: gross premium reserve using 2% real interest rate

For the purpose of relating these policies to those issued in the insurance industry, most of the quantities listed in this section are set in amounts which may reasonably be found in practice. It must be emphasized, however, that no attempt was made to ensure that these assumptions were wholly consistent with industry averages at the time of writing.

5.2.4 Reserves

Realistic reserves are essentially calculated by performing a gross premium valuation using realistic rates of interest. These are taken to be the yields on Consols and index-linked gilts when valuing nominal and real cashflows respectively. However, in order not to capitalize future profits, the office premiums used in the valuations are stripped of profit margins. The adjusted office premiums are calculated on the basis of a zero profit criterion at a risk discount rate equal to the interest rate assumed in the premium basis.

Published statutory reserves are (unzillmerized) net premium reserves calculated using the smaller of a prudent fixed rate of interest and the reliable yield on total assets. The prudent rate of interest is assumed to be 6% when net premiums and benefits are fixed in nominal terms and 1% when linked to the retail price index. Reliable yield is taken to be 92.5% times an average yield (running yield on equities and gross redemption yields for all other assets), weighted by the proportions in the corresponding asset classes. Where single premium annuities are concerned, renewal expenses are treated as index-linked benefits for the purpose of the valuation.

The published statutory rate of interest used here mainly differs from that prescribed in the Insurance Companies Regulations 1994 (see Gallen and Kipling, 1995) in two respects. Firstly, a reduction of 7.5% rather than 2.5% is applied to the average yield. This extra margin of 5% had actually been part of the 1981 Regulations and is maintained here for prudence. Secondly, the 1994 regulations also require the rate of interest used to be no greater than the gross redemption yield on 15 year medium coupon gilts. As this particular asset has not been modelled here, it is not possible to take this part of the regulations into account. The effect of omitting this part of the Regulations from the published basis assumed here is to make this basis less stringent than the Statutory basis on certain occasions. But given the level of conservatism already built into this published basis, such instances should be relatively infrequent.

5.2.5 Business Volumes

Due to the long-term nature of life insurance liabilities, the position of the office at the start of the projection period ($t = 0$) will inevitably depend on past conditions. Therefore, some assumptions will need to be made regarding business volumes over the previous P years, where P is the number of years since the earliest policies currently in force were written. For simplicity, it will be assumed that the number of new policies issued has remained constant over the past P years, but that the per policy amounts have been growing with inflation throughout this period. The rate of inflation in the past is taken here to be constant at 5%. This is in fact consistent with the inflation assumption in premium bases (see Section 5.2.3) and is close to the average inflation rate of 5.13% assumed in the investment model (see Section 3.3.1). As benefits and expenses have been increasing in line with the retail price index, the same may be assumed of the premiums, as long as they have been priced using bases consistent with those stated in Section 5.2.3. Hence, the liability profile of the office is approximately stationary in real terms at the start of the projection period.

With regards to future business volumes, there are three possible situations which may be assumed. The office could remain open to new business throughout the projection period, remain open for a limited number of years and be closed to new business thereafter, or be closed to new business as from the start. If the fund were assumed to be ongoing during the entire period, it would seem a reasonable starting point to assume also that business volumes would continue to grow with the retail price index. As such a fund would have the duration of its liabilities approximately constant over time, it would seem appropriate to analyse this as a static model, i.e. with asset proportions remaining constant over time. This is very appealing as the static models are considerably simpler to deal with than dynamic models.

The main purpose of dynamic models is to accommodate liability profiles which are continually changing with time. This may be due to changing volumes or mix of business, varying solvency margins and so on. Simple dynamic models have been employed in situations similar to a stationary fund, by allowing the asset mix to gradually switch from equities into gilts when the solvency margin falls below some threshold value, as in Ross (1989), Roff (1992) and Hardy (1993). Although such dynamic reallocation strategies may be sensible for with-profits funds, where the average solvency level is likely to be fairly high, they would be of little value if the solvency margin is weak to begin with, which should be the case in most non-profit funds. Further consideration of dynamic models will be given in Chapter 8.

The other extreme to an office open to new business throughout the projection period would be an office closed to new business as at $t = 0$. As the duration of the liabilities will be shortening with time, a dynamic model will be absolutely essential if the asset mix is to reflect the changing liability profile. Unlike the open fund where any projection period would probably be appropriate, it would seem more meaningful in a closed fund to use a sufficiently long projection period to allow every single policy on the company's books to run off. This is feasible here as the contracts being issued have outstanding terms of no more than twenty years.

Another possibility would be to assume that the office will remain open to new business for a short period and then assume it will be closed thereafter. This is also the most complicated situation to consider from an asset allocation perspective. Not only will a dynamic asset allocation strategy be necessary but the projection period will also need extending as it must at least equal the term of the longest contracts plus the number of years in which the fund remains open to new business. While this approach may be a reasonable means for appraising the valuation system (see MacDonald, 1993), it would seem to be an unnecessary complication in this particular decision-making process.

Therefore, only two of the three situations will be considered in the investigations which follow, these being the office open to new business throughout (open fund) and the office which is closed to new business from the start (closed fund). As the open fund is the situation which life offices would be most likely to find themselves, it will be discussed in some detail in Chapters 6 and 7. The more dynamic situation of a closed fund will be considered in Chapter 8.

5.3 General Structure of the Model Office

5.3.1 Global Variables

The following definitions relate to the main global variables used in the projections:

$Real\_Liab(t)$ = realistic value placed on liabilities at time t.

$Stat\_Liab(t)$ = published statutory value of liabilities at time t.

$SMSM(t)$ = statutory minimum solvency margin at time t.

$Acc'(t)$ = total amount of assets before distribution of dividends at time t.

$Acc(t)$ = total amount of assets after distribution of dividends at time t.

$Fund'(t)$ = value of fund before distribution of dividends at time t.

$Fund(t)$ = value of fund after distribution of dividends at time t.

$Dist(t)$ = amount distributed as dividends to shareholders at time t.

Statutory minimum solvency margin is defined to be the 4% of the published statutory liabilities plus 0.3% of the sum at risk. Fund in this context is simply the difference between the total amount of assets and the target amount of free assets. The precise definition of target free assets will be made clear in Section 5.3.2 below.

5.3.2 Cashflow Projections

Let the projection period be the time interval $[0, H]$, where H is the horizon date. Initialize the total amount of assets at the start of the projection period to be:

$$Acc(0) = (1 + sm) \times (\text{net premium reserve at the prudent fixed rate of interest})$$

where *sm* (the initial solvency margin) is an input value for specifying the initial amount of assets. The net premium reserve is basically the same as *Stat_Liab*(0), except that the reliable yield is not used in its calculation. Therefore, the valuation is independent of the asset mix held at time 0. Next, initialize the fund value at the start to be:

$$Fund(0) = [Stat\_Liab(0) + SMSM(0)]$$

Define the free asset ratio (as stated in objective 2, Section 5.1) at the start to be ρ, where:

$$\rho = \frac{Acc(0) - Fund(0)}{Fund(0)}$$

Note that since ρ is a function of *Stat_Liab*(0), its value will depend on the asset mix held at time 0. This will be the target free asset ratio over the entire projection period. It then follows that the target amount of free assets at time t is:

$$Target(t) = \rho \, . \, [Stat\_Liab(t) + SMSM(t)]$$

For simplicity, all valuations and cashflows including premiums, expenses, claims and payments of dividends are assumed to occur at the start/end of each projection year. Define the net cash flow at time t, $Ncf(t)$, to be the total of premiums less expenses and commission in respect of all business, new and in force, between time t and $t+1$, less all benefits payable in respect of business in force during the period $t-1$ to t. In addition, define $i(t)$ to be rate of return earned between time $t-1$ and t. Hence, for $0 < t \leq H$, the accumulated amount of assets before distribution will be:

$$Acc'(t) = [Acc(t-1) + Ncf(t-1)].[1 + i(t)]$$

As the fund is the difference between the total assets and the target amount of assets:

$$Fund'(t) = Acc'(t) - Target(t)$$

At this stage, it should be pointed out that the fund is only used as an aid for determining the amount of surplus distributable to shareholders. However, as the process of establishing dividend policy is reasonably complex, it is perhaps worth giving it separate consideration (see Sections 5.3.3 to 5.3.5). Hence, assuming for the moment that $Dist(t)$ may be obtained from the computations performed thus far, the fund value after distribution will be:

$$Fund(t) = Fund'(t) - Dist(t)$$

It also follows from the definitions that the relationships below will hold:

$$Acc(t) = Acc'(t) - Dist(t) = Fund(t) + Target(t)$$

5.3.3 Determination of Surplus

On a simplistic level, the amount of surplus available for distribution could be described as the realistic amount of surplus which has accrued since the previous distribution date. However, there are many other aspects which need to be considered before the appropriate amount of dividends may be declared. Two of these relate to the calculation of realistic surplus. First, it would be prudent to not to capitalize on future profits arising from the margins assumed in the premium bases. This could be avoided by calculating the realistic gross premium reserves using office premiums which have been stripped of profit margins. Second, in order to try and maintain the target free asset ratio, it would be helpful to define the increase in surplus net of that needed to for this purpose. Therefore, a more sensible definition of realistic surplus arising may be:

$$SA(t) = [Fund'(t) - Fund(t-1)] - [Real\_Liab(t) - Real\_Liab(t-1)]$$

Moreover, there are legislative requirements regarding minimum solvency levels which must be accounted for in determining distributable surplus. Again from the perspective of maintaining the target free asset ratio, one definition of distributable surplus could be:

$$DS(t) = Fund'(t) - [Stat\_Liab(t) + SMSM(t)]$$

So in the absence of smoothing, the office may reasonably wish to declare:

$$\max \{\min [SA(t), e.DS(t)], 0\} \tag{5.1}$$

where $0 \leq e \leq 1$. (It may be necessary to use a value for e of less than unity to prevent the statutory free assets from being depleted too frequently.) But in order to satisfy the objective of distributing smoothed dividends, additional complexity will have to be introduced into the procedure.

The method of smoothing adopted here involves a weighted average of the unsmoothed dividends (see expression 5.1) and some smoothing component. As it would be logical to apply more smoothing when the solvency position is strong and *vice versa*, the weights will be defined in terms of the existing level of statutory surplus. Define $R(t)$ to be the ratio of statutory surplus to total assets before distribution:

$$R(t) = \left(\frac{Acc'(t) - [Stat\_Liab(t) + SMSM(t)]}{Acc'(t)} \right)$$

Let $f(R(t))$ and $[1 - f(R(t))]$ be the weights placed on the smoothing component and the unsmoothed dividends respectively. It then follows that the function f would need to be monotonically increasing and bounded between 0 and 1. While the choice of function is fairly arbitrary, it seems intuitive that f should also be concave. The expression used here is:

$$\begin{aligned} f(R(t)) &= 0, & R(t) \leq 0, \\ &= k.[1-(1-c.R(t))^2] & 0 < c.R(t) < 1, \\ &= k, & c.R(t) \geq 1, \end{aligned}$$

where $0 \leq k \leq 1$ and $c > 0$. The parameter k may be seen as the maximum extent of smoothing permissible whereas c may be interpreted as the rate at which this maximum may be reached. The values of k, c and e (see expression 5.1) were selected on the basis of their ability to smooth the dividends by as much as possible, without excessive deviation from the target free asset ratio.

This had been done by carrying out twenty year projections of an open fund with just twenty year endowments in force and comparing the average free asset ratios at $t = 20$ with those at the start, for a range of asset mixes. The parameter values which appeared to result in the smallest variation from the target free asset ratio were: $k = 0.5$, $c = 0.5$ and $e = 0.9$. It should be noted that the method used for choosing these values is rather crude and by no means implies the optimal values. They are just a possible combination of parameter values which seem to broadly satisfy the required objectives. Unless otherwise stated, it should be assumed that these values will be used in all the investigations hereafter.

If the smoothing component is assumed to be the average inflation adjusted dividends in the previous two years, the smoothed dividends may be expressed as:

$$SD(t) = f(R(t)).\tfrac{1}{2}Q(t).(\mathrm{B} + \mathrm{B}^2)[Dist(t)/Q(t)] + [1 - f(R(t))].\max\{\min[SA(t), e.DS(t)], 0\},$$

where $\mathrm{B}X(t) = X(t-1)$ and $Q(t)$ is the value of the retail price index at time t. This would follow the pattern described in Blake (1990). However, unless dividends declared at $t = 0$ and $t = -1$ are specified, it is clear that this expression for $SD(t)$ only applies for $t > 2$. So it will be assumed that for $SD(2)$, the smoothing component is just the previous year's inflation adjusted dividend and that no smoothing is applied in $SD(1)$. On the surface, it would seem as though the actual dividends distributed, $Dist(t)$, may be taken to be: $\max[SD(t), 0]$. But before any dividend is declared, a final check on solvency is required.

5.3.4 Insolvency

The life office is assumed to be technically insolvent when the total amount of assets at any time is less than the published statutory reserves plus the guarantee fund (taken to be one third of the statutory minimum solvency margin). So when the following inequality:

$$\max[SD(t), 0] \leq Acc'(t) - [Stat\_Liab(t) + SMSM(t)/3] \tag{5.2}$$

is satisfied, it would be appropriate to set:

$$Dist(t) = \max\,[SD(t), 0].$$

If condition (5.2) does not hold and the right-hand-side is positive, then:

$$Dist(t) = Acc'(t) - [Stat\_Liab(t) + SMSM(t)/3].$$

The office is only said to be technically insolvent when the right-hand-side of the inequality (5.2) is negative.

When an office becomes technically insolvent, two possible courses of action are to:

A. require a capital injection from the shareholders to bring the solvency margin up to the guarantee fund so the office remains in business;

B. transfer the portfolio of liabilities to another company and distribute the remaining surplus, if any, to the shareholders.

Rather than trying to determine which course of action would be more appropriate under which circumstances, separate treatment will be given to the two liability models, A and B. The implications of the different models in these investigations are important not only in the practical sense but also impact on the choice of utility function (see Section 5.4) and the optimization procedure (see Section 6.1).

Hence, if the office follows course of action A at the point of insolvency, clearly:

$$Dist(t) = Acc'(t) - [Stat\_Liab(t) + SMSM(t)/3].$$

It is being assumed that shareholders have an unlimited source of capital with which to support the business in times of difficulty. This also means that the fund will remain a going concern, at least until the horizon date.

In the case of Model B, no presumptions are made regarding the ability of shareholders to raise additional capital. When the guarantee fund is breached, the entire portfolio of liabilities is assumed to be taken over by another office at a cost equal to the realistic reserves. Any positive surplus remaining is distributed to shareholders as a final dividend and the office ceases to operate thereafter. Hence, if insolvency occurs at time I, then:

$$
\begin{aligned}
Dist(t) &= \max\,[Acc'(t) - Real\_Liab(t), 0], & t &= I, \\
&= 0, & t &> I.
\end{aligned}
$$

Realistic reserves are used to calculate the takeover value because they were found to be good approximations to what might be considered fair values for the liabilities, i.e. statutory reserves less the present value of future profits (PVFP). Empirical evidence for this may be seen by computing [realistic reserves] ÷ [statutory reserves – PVFP] in the case of twenty year endowments, assuming future conditions follow those used in the premium basis. Although the ratio was fairly close to unity for durations in excess of about 5 years, it was shown to be a poor approximation at earlier durations. But given that the fund will always be either stationary in real terms or running-off, this effect at early durations will be hardly noticeable, as the ratio will be weighted by relatively small reserves when the approximation is poor and *vice-versa*. For example, the reserve weighted ratio for the twenty year open fund is 1.018, even though the ratio for contracts at duration 1 are –0.304. The approximation is even better in the closed fund as the average duration of contracts increases with time.

5.3.5 Value of Shareholdings at the Horizon Date

In situations where the office remains in business up to the horizon date, the value of the shareholdings at that point will need to be ascertained. Whether this is interpreted as being some present value of future dividends, assuming shareholders retain their equity beyond the horizon date, or the market value, if the shares are to be sold immediately, is of little relevance as the two amounts should in theory be identical. But as it would seem impractical to model market values, a crude approximation may also be to use:

$$Dist(H) = \max\,[Acc'(H) - Real\_Liab(H), 0],$$

as a proxy for the value of the shareholdings at the horizon date just before dividends for period H are declared. Arguably, this amount may be considered fairly conservative as it does not explicitly incorporate goodwill and other factors that may be thought to affect share prices. However, as any such adjustment would be arbitrary, this issue will not be pursued further. A possible implication of ignoring this from the decision maker's perspective is that the relative advantage of remaining solvent in liability model B may not be quite as great as it might otherwise have been. But relative to all the other factors involved, it seems unlikely that this simplification would have a significant impact on the decision-making process.

5.4 Shareholders' Utility

5.4.1 Time Preference

Due to the nature of the problem defined in Chapter 4, there was little debate regarding the amount to which the utility function should be applied. The asset fund had been accumulated in respect of a fixed horizon date with total consumption taking place at this point. However, in attempting to create a realistic life office model, it has been necessary for dividends to be declared regularly. Otherwise, the life office fund could build up artificially over time, thus increasing the solvency margin as it approaches the horizon date.

Since dividends are declared annually, the utility maximization process no longer involves the straightforward task of calculating expected utilities of ultimate wealth, as in the asset fund. Instead, utilities will need to be considered over a number of time epochs, i.e. $U(Dist(1), \ldots, Dist(H))$. This in fact is an example of a multiattribute utility function, with time representing the different states. Although much research has been carried out in the field of decision theory relating to multiperiod consumption (see Prelec and Loewestein, 1991), the simplest approach is to assume that utilities are time-additive (see Fishburn, 1988):

$$U(x_1, \ldots, x_H) = \sum_{t=1}^{H} U_t(x_t)$$

where $U_t(x_t)$ is the utility function which applies to consumption at time t. In finance literature (see for example Samuelson, 1969), this is often simplified so that for some constant term, b, the relationship:

$$U_t(x_t) = b^t . U_0(x_t), \tag{5.3}$$

holds for all t. In other words, the same utility function is assumed to operate over all time periods, with just a constant discount factor accounting for the difference in timing.

Despite being more tractable mathematically, the discounting of utilities in this manner does not generally lead to intuitive results, especially from an actuarial perspective. The basic principle behind present value models (as used by actuaries) is to assume that an investor would be indifferent to receiving a certain amount, A now, and receiving a certain amount, $A(1+i)^t$ in t years time, where i is a risk free rate of interest per annum. In term of utilities:

$$U_0(A) = U_t(A(1+i)^t), \tag{5.4}$$

However, if this is to be consistent with the concept of discounting utilities as suggested above, then combining (5.3) and (5.4) implies that there must be some constant term, b, for which:

$$U_0(A) = b^t U_0(A(1+i)^t) \;\; \forall\, t. \tag{5.5}$$

A class of utility functions which satisfies equation (5.5) is the power utility function, $U(x) = x^c$, for $x > 0$ and where $c \in (0, 1)$. Here, b would need to take the value, $(1+i)^{-c}$. For the linear utility function ($c = 1$), the factor used to discount utilities is identical to that used to discount amounts of wealth. However, it may be shown that equation (5.5) holds for neither the exponential nor the logarithmic utility function as b cannot be expressed independently of t with either function.

As it may not be sensible to assume that equation (5.3) holds in all circumstances, the results produced when adopting this approach may be meaningless if applied without due care. Therefore, it would be preferable if a more robust method could instead be used to account for investor preference over time.

Ultimately, if the objective is to ensure that some actuarially acceptable time preference relation holds, an obvious approach would be to accumulate the dividends to the horizon date, thus reducing the problem to one of ultimate surplus. As long as the dividends are kept separate from the assets held by the life office, the problem of artificially high solvency margins will not apply. It therefore seems reasonable that dividends be treated in such a manner in the liability model. However, a suitable accumulation rate for this dividend fund will need to be determined.

An interesting issue arises here because the accumulation rate should depend on the reasons for investing in the life company in the first place. In assuming that time preference may be accounted for by accumulating dividends to the horizon date, it is being implied that shareholders are not relying on dividends for the purpose of consumption in the real sense of the word, but rather as a means of achieving good long term growth from their investment. This may not be too unreasonable as most sensible investors would not normally expect company dividends to provide them with a reliable income for managing their cashflow position in the short term. As long term growth would generally seem to be the main aim of shareholders, the accumulation rate may be assumed to be the rate of return earned in the equity market at the time. This is equivalent to saying that the office should only retain surplus available for distribution if it expects to be able to earn a higher return on this surplus than if it were to be invested in equities over the same period. Defining the payout to be the accumulated dividends (including the value of the business at the horizon date):

$$\text{Payout} = \sum_{t=1}^{H-1}\left\{Dist(t).\prod_{k=t}^{H-1}[1+RE(k+1)]\right\}+Dist(H)$$

where *RE*(*t*) is the rate of return earned on equities between time $t-1$ to t. The objective will then be to maximize the expected utility of payouts, using a single-attribute utility function as had been the case in Chapter 4. An extension of this work could involve looking further at intertemporal utility maximization issues.

5.4.2 Choice of Utility Function

Contrary to dealing with one individual's risk preferences, it is in theory necessary for the company to consider the expected utility of all its shareholders. While it would clearly be an impossible task to aggregate utilities of individual shareholders, it is also worth recalling that the purpose of the utility maximization approach is to set out a framework in which consistent decisions may be made in the face of uncertainty. Therefore, it would still be useful just to consider a single utility function for reflecting a particular level of risk tolerance, as this may provide a valuable tool for assessing and making rational decisions.

There is though some philosophical justification for using a single utility function. Intuitively, investors would only purchase shares in the life company if the expected nature of the dividend stream is suitable to their needs. Therefore, any reasonable utility function should appeal to at least one subset of potential shareholders with similar preferences. So having selected the utility function which reflects a particular risk/reward position, the optimal investment strategy for the insurance company should be to consistently maintain this level of tolerance to risk by maximizing the expected utility of payouts. In a sense, this is equivalent to the concept of satisfying policyholders' reasonable expectations in the case of a with-profit office. Thus, maximizing the expected utility of future dividends may be viewed as means of fulfilling shareholders' reasonable expectations.

In practice, it would be a matter for the life office to determine the utility function which best represents the level of risk that would be acceptable to its shareholders as a whole. But for the purpose of these investigations, it would be more useful to consider a range of risk tolerance levels, as had been done in Chapter 4. An obvious choice for this could be the exponential class of utility functions, as it has been shown to be well suited to this task. In addition, the exponential function can accommodate negative quantities which may arise in the case of Model A.

One drawback of the exponential function is that the risk tolerance parameter, r is linked to the initial amount invested, A and is therefore not universal. *Ceteris paribus*, investors will only arrive at the same decision if their values of A/r are the same, rather than if r is the same. A direct consequence is that for a given amount invested, an investor will take a riskier decision if the value of r is increased. But conversely, it also means that a riskier decision will be taken if the amount invested is decreased for a given value of r. The latter is intuitive in so far as risk-averse investors would tend to be more prudent with larger investments than with smaller ones. However, it also means that decisions made for a given value of r under different circumstances may not be directly comparable.

Hence, it follows that due to the relatively large fund sizes associated with the model office described above, the values of r used in the asset fund with an amount of 1 invested will be far too small to be meaningful here. In fact, from preliminary projections, a reasonable range for the level of risk tolerance had been found to be of the order 10^4. So for convenience, the exponential utility function used hereafter is redefined to be: $U(S) = -\exp(-S/10000r)$.

This peculiarity of the exponential function actually relates its property of constant absolute risk aversion, which means that for a given amount invested the optimal decision is independent of the initial level of wealth. If a utility function with constant relative risk aversion is used, such the logarithmic function or power function, then this rescaling would not normally be necessary. The results would then be valid for any investor investing a given proportion of wealth (rather then a given absolute amount) in the shares of the life company. Unfortunately, these two functions may only operate when outcomes are bounded within the set of positive real numbers and would cause difficulties if applied to Model A.

Despite some minor difficulties in using utility theory, it is worth noting that utility maximization is a general approach for decision-making in the face of uncertainty. More common techniques used in portfolio analysis such as the mean-variance approach are also subject to similar shortcomings. These include the treatment of multiperiod consumption and the fact that more than one individual's preferences may need to be considered in the decision-making process. However, as the assumptions which underlie the theoretical justification for the mean-variance approach are not as transparent, these issues are rarely noticed in its applications.

5.5 Summary

Broadly speaking, the aims of the model office described in this chapter are to distribute smoothed dividends to shareholders, whilst maintaining a stable solvency margin over the period of investigation. At the horizon date, the business in force is transferred at a cost equal to the realistic reserve, with any remaining surplus being distributed as a final dividend. In the case of insolvency, two separate courses of action are considered: shareholders could be required contribute more capital (Model A), or the company may be wound up on terms similar to that at the horizon date (Model B). Whichever model is used, all dividends (including the final dividend) are accumulated to the horizon date at the same rate of return earned by equities. The objective function is then taken to be the expected utility of these accumulated dividends, or payouts.

The main limitations of the life office model may be attributed to its divergence from models necessary in practice. Assuming cashflows occur at the valuation dates is one simplification. Apart from the unnecessary complexity that would result if cashflows are assumed to occur midway between valuation dates, the use of an annual asset model would also require approximations to be made when calculating the rate of return earned between such intervals. Items deliberately left out of the cashflows include taxation, surrenders and termination expenses. Although the last item is a minor omission, both surrenders and taxation are significant components in the cashflows of a life office. However, the inclusion of taxation and surrenders would add little to the issues being investigated in this research.

In relation to the utility maximization procedure, there are clearly elements which cannot be fully and realistically taken into account. Multiperiod consumption and shareholders' aggregate utility have only been dealt with pragmatically by accumulating dividends in equities and assuming only one risk tolerance level applies. Although utility functions with constant relative risk aversion have been felt to be superior to the exponential function (due to their self-scaling properties), they may not be used in situations which involve outcomes of negative amounts. Hence, the exponential function, which does not suffer from the latter problem, was felt to be more suitable, although some rescaling of the risk tolerance parameter may be necessary. As long as the utility maximization approach is used in order to arrive at consistent decisions under uncertainty, these techniques should be useful tools in the management of life company's asset portfolio.

6. STATIC OPTIMIZATION II - LIABILITY MODEL A

6.1 Introduction

As the ultimate objective of this research is to investigate the optimal asset allocation strategies for insurance companies, this chapter and the next deal with the investment strategy for an office which is open to new business. The optimal asset mixes for this office are derived using the numerical optimization routines described in Section 4.3. Sensitivity analysis is then performed with regard to some of the assumptions made in the investment model. Finally, the effect of insolvency constraints on these optimal portfolios is investigated and analysed in a mean-variance framework.

All the results derived in this chapter relate to Model A, which has been described in Chapter 5. It may be recalled that in the case of Model A, additional capital is obtained from shareholders whenever the solvency margin of the office falls below the guarantee fund. This capital injection is set at an amount which is just sufficient to meet the guarantee fund. In reality, if shareholders were prepared to provide more capital to the office when in difficulty, it would be sensible if this capital input was sufficient to raise the solvency margin to a more sustainable level, say 5% or 10%, rather than to the exact amount of the guarantee fund. One reason for making this less realistic assumption in Model A is to facilitate the optimization process.

As mentioned in Section 4.3, the most efficient non-linear optimization routines require the calculation of gradient vectors. Providing that the objective function in Model A is a continuous function of the asset proportions, the optimal portfolio may be obtained using these algorithms. However, if the capital injection had been assumed to bring the solvency margin up to a pre-determined level at the point of insolvency, the objective function would no longer be smooth, hence requiring less efficient means of arriving at the optimal portfolios. This in fact is one reason for giving Models A and B separate consideration. Model B is distinct from Model A because of the discontinuity which results when insolvency occurs in the former. In Model B, the liabilities are sold at the point of insolvency. So while Model A can take advantage of gradient methods of optimization, Model B may not. Hence, the case of Model B is discussed in Chapter 7.

6.2 Optimal Portfolios

The optimal asset mixes for Model A are computed by carrying out a thousand simulations of asset and liability cashflows over a period of twenty years. It is assumed that only twenty year endowments are currently in force and that the volume of new business over this period will continue to grow in line with inflation. The amount of assets at the start of the projection period is given by $(1+sm)$ times the prudent value placed on the liabilities, where sm is a measure of the initial solvency margin. In these investigations, just the two situations with sm = 15% and 25% are considered. Utilities of total payouts (accumulated dividends plus final surpluses) are computed for a range of risk tolerance levels, giving separate consideration to nominal and real amounts.

6.2.1 Nominal Payouts

Table 6.1 shows the optimal asset mixes when payouts are measured in nominal terms, assuming sm = 15%. Increasing values of r as one moves down the first column implies riskier investment strategies. The next four columns show the optimal percentages in each of the asset classes: cash, Consols, index-linked gilts and equities in that order. The last four columns give the expected utility, the mean payout, the standard deviation of the payouts and the percentage of scenarios in which the guarantee fund is ever breached, referred to hereafter as the probability of ruin. A graphical representation of these optimal portfolios is also shown in Figure 6.1.

r	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
2	0	100	0	0	-0.00955	127384	47281	62
4	2	97	1	0	-0.07003	128203	48483	57
8	5	79	1	15	-0.22750	144090	77178	36
16	23	32	11	34	-0.43362	174772	146169	48
32	0	0	41	59	-0.60943	215163	245795	99

Table 6.1. Optimal portfolios: *sm* = 15% in nominal terms.

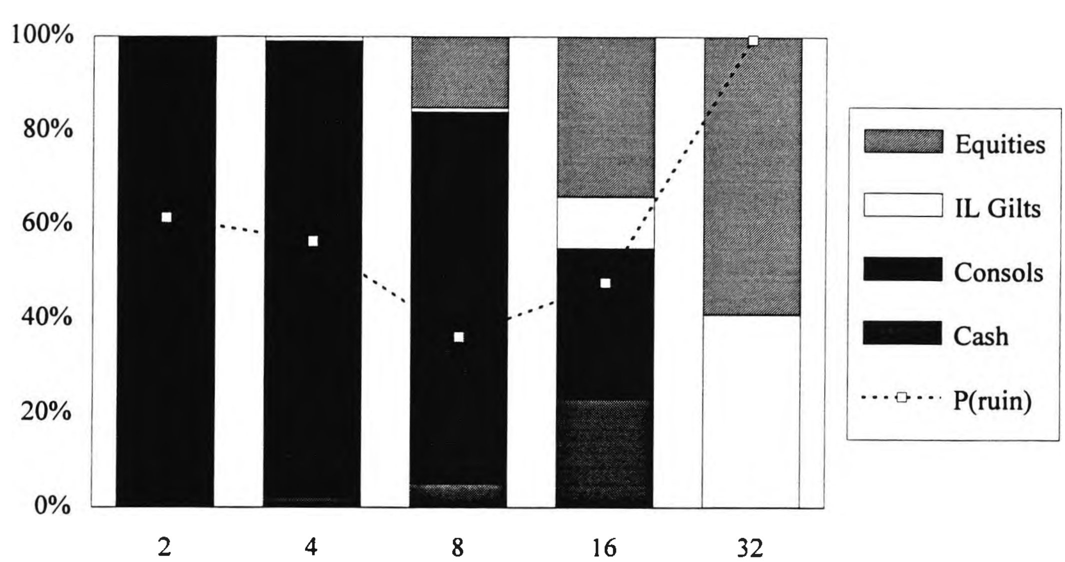


Figure 6.1. Optimal portfolios: *sm* = 15%, in nominal terms.

On the surface, these results seem intuitively reasonable. At the lowest level of risk tolerance shown, the optimal mix is entirely in Consols. These are gradually shifted into cash, index-linked gilts and equities as the risk tolerance parameter is increased.

The mix at $r = 8$ appears to represent a moderate strategy, with about 85% of the fund invested in fixed income assets and the remainder being in real assets. At very high levels of risk tolerance, the optimal strategy is entirely in real assets, with about 40% in index-linked gilts and 60% in equities. This is reasonable as these are the asset classes with the highest expected return over this twenty year period (see Table 3.9). In terms of the means and standard deviations of the payouts, the figures shown in Table 6.1 seem intuitive, with both these quantities increasing monotonically with r.

From Table 6.1, it may be noted that the ruin probabilities observed throughout the spectrum of risk tolerance levels are generally quite high, ranging from 36% at $r = 8$, to 99% at $r = 32$. While it is possible that this may be just a consequence of the investment model, (in which case even greater emphasis should be placed on the relative probabilities rather than the absolute ones computed), the existence of sub-optimal portfolios with much lower probabilities of ruin would appear to refute this view. For example, if 40% of the fund were to be invested in cash with the remaining 60% invested in Consols, the office would only expect to endure a 2% ruin probability. Hence, there must be another reason why such high ruin probabilities are permissible.

A more plausible explanation for this lies in the link between the objective function and the probability of ruin. In the case of Model A, the incidence of ruin has hardly any impact on the amount of the payout. Apart from incurring negative dividends in order to make up the guarantee fund, no real penalty is imposed on the fund by becoming technically insolvent. As long as the amounts by which the fund breaches the guarantee fund are small, these should not have serious consequences on the expected utility, even if they do occur in as many as 62% of the simulations generated, which is the case for the all-Consols portfolio. Ruin is therefore more of an indicator that negative dividends are required, rather than the cause of considerable disutility which would usually be associated with the event of an insolvency in practice. This seems reasonable given that the objective function is defined as a continuous function of the asset proportions.

Table 6.2 is similar to Table 6.1 but with *sm* set at 25% instead of 15%. Initially, one might expect this increase in the initial solvency margin to result in lower probabilities of ruin and possibly more adventurous asset mixes. Comparing these two tables, it is clear that the probabilities of ruin have dropped considerably, now ranging from 12% to 87%, rather than from 36% to 99% as shown in Table 6.1. However, the optimal asset mixes appear to have been rather indifferent to the increase in the value of *sm*.

r	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
2	0	100	0	0	-0.00415	152810	54672	22
4	0	100	0	0	-0.04302	152810	54672	22
8	0	85	0	15	-0.17498	169146	83723	12
16	9	53	3	35	-0.37739	196123	142112	20
32	0	0	42	58	-0.56753	247310	269243	87

Table 6.2. Optimal portfolios: *sm* = 25%, in nominal terms.

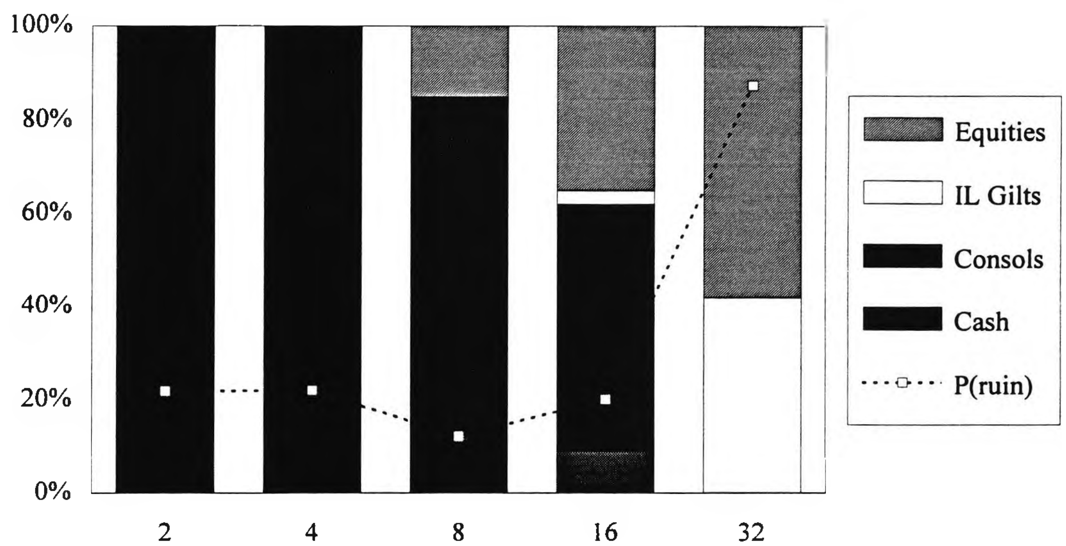


Figure 6.2. Optimal portfolios: *sm* = 25%, in nominal terms.

When analysing these results, it should be noted that increasing *sm* also is similar in principle to increasing the initial amount invested by the shareholders. Since the risk tolerance parameter is only comparable for a *given* amount invested, the values of r in Tables 6.1 and 6.2 may not be directly comparable. However, this will depend on how this additional capital has be raised.

Hypothetically, if it is assumed that the increase in *sm* has come from the same group of shareholders with the same amount of initial wealth, then any differences between Tables 6.1 and 6.2 would be due to the net effect of increasing both the free asset ratio of the fund and the proportion of shareholders' wealth invested. While the former could be expected to yield more adventurous asset mixes, the latter should have an opposite effect. Coincidentally or not, these effects seem to have cancelled each other out to a large extent.

Conversely, if this additional capital is assumed to have been raised by a new group of shareholders, then the initial amount invested by the original shareholders would remain the same, although their proportionate shareholdings would also be diluted from 100% to only 60%. Hence, the original shareholders should only be entitled to 60% of the payouts. In terms of the exponential utility function, multiplying the payouts by 0.6 is equivalent to dividing r by the same factor, as $-\exp[kx/r] = -\exp[x/(r/k)]$. Therefore, the value of $r = 8$ shown in Table 6.1 would be comparable with a value of $r = 8 \div 0.6 \approx 13$ in Table 6.2. On this basis, the optimal portfolios do seem more adventurous when the initial solvency margin in increased.

This latter situation, however, assumes that the value of *sm* accurately reflects the amount of initial surplus which has been contributed by the shareholders. But as the amount of surplus held by an ongoing fund is a function of the valuation basis, interpreting the results as such may be less valid.

6.2.2 Utility Maximization, Immunization and Ruin Probability.

In Tables 6.1 and 6.2, the optimal portfolios at $r = 2$ represent very low risk strategies, and these have been found to be entirely invested in Consols. This seems reasonable given the long term nature of the liabilities and the fact that Consols is the asset class showing the lowest variability of accumulations over twenty years. Nevertheless, it would also be worthwhile examining how this portfolio compares with more traditional approaches to minimizing risk, such as an immunized strategy. From Section 2.2.2, one may recall that immunization had been described as a means of minimizing the effect of small changes in interest rates on the level of surplus in a fund. Hence, at very low levels of risk tolerance, the utility maximizing portfolio should be broadly similar to an immunized portfolio.

Generally speaking, a fund is said to be immunized if the discounted mean term of its asset proceeds is the same as that of its liability outgo. Using a realistic valuation basis with an interest rate assumption of 8.5% for nominal cashflows and 3.5% for real cashflows, the discounted mean term of the liability outgo for an open fund of twenty year endowments is 9.92 years. At the interest rate of 8.5%, the discounted mean term of Consols is the reciprocal of the discount rate or 12.76 years. Assuming cash to have a discounted mean term of zero, this implies an immunized ratio of about 2 : 7 in cash and Consols respectively. On this basis, the discounted mean term of a fund entirely invested in Consols would seem to be too long in relation to its liabilities. This discrepancy between utility maximization and immunization may be explained in terms of the different objectives which these two methods seek to achieve.

Although utility maximization is a general approach to decision analysis, the expected utility in this particular decision model has been defined in relation to the accumulated dividends or payouts. At very low levels of risk tolerance, the aim of this approach is to minimize the variability of payouts. On the other hand, immunization is a theoretical concept which aims to minimize the variability of surplus in respect of a single change in interest rates, after which the portfolio has to be rebalanced if an immunized position is to be maintained. In the context of the type of decision model employed in these investigations, the situation in which utility maximization and immunization would be most comparable is if payouts were to be calculated after one year. i.e. no multiperiod consumption involved and ultimate surplus is only defined in terms of realistic reserves. This may be verified empirically by minimizing the variance of payouts over one year, for which the optimal portfolio (with sm = 15%) is 26% in cash, 62% in Consols and 12% in index-linked gilts. Given that the initial surplus on a realistic basis is about 25%, this portfolio clearly lies within the immunized ratio of 2 : 7 indicated above.

It would, however, seem reasonable to suggest that a portfolio which can minimize the variability of realistic surplus over one year should also be capable of producing very stable dividends in a multiperiod consumption framework. Returning to the twenty year open fund, it may be shown that the variability of dividends in any year will tend to be slightly lower for a portfolio with say 10% in cash and 90% in Consols, than for one entirely invested in Consols. Hence, it seems as though the compounding of dividends is leading to more stable payouts for the portfolio with 100% in Consols.

Initially, it may appear counterintuitive that a strategy which produces more variable dividends than another is nevertheless capable of yielding accumulated dividends which are more stable. In fact, this is analogous to the case of a pure asset fund where the cash had the more stable accumulations over one year although Consols had the more stable accumulations over twenty years. This should be possible if the two stochastic variables are serially correlated to different extents and may be connected with the fact that the investment model assumes Consols yields to be mean reverting. Where the model office is concerned, a sudden rise in the Consols yield may result in a reduction in the amount of surplus available for distribution at a particular valuation date. However, as the yield tends to revert to its mean value of 8.5%, this apparent loss of surplus may be made up at a later date, given sufficient time.

As a result, the optimal portfolios obtained for the twenty year case should be treated with some caution. If the investment model truly reflects reality and if the objective is to minimize the variability of payouts in twenty years time, then the best investment strategy may well be to invest 100% of the fund in Consols. However, if the difference between the immunized position and this minimum variance strategy is a consequence of an invalid assumption about the stationarity of Consols yields, then the 100% Consols strategy may be misleading.

Another approach which is sometimes used to minimize risk is to minimize the probability of ruin. In this decision model, the ruin probability is defined as the proportion of simulations in which the guarantee fund is ever breached during the twenty year projection period. So if the surplus level of a fund is resilient to changes in economic factors, then the ruin probability should be fairly minimal. In other words, an immunized strategy should also be equivalent to a minimum ruin strategy.

A limitation of duration analysis is that it depends on the valuation basis being used, not only in calculating the discounted mean term but also in determining the amount of surplus being held. With this model office, distributable surplus is a function of both the realistic reserves and the published reserves. As these reserves are generally quite different, it may be impossible for the fund to be immunized on both these bases.

Furthermore, ruin is defined in terms of the published reserves which are subject to an upper limit of 6% on the rate of interest used for valuing nominal cashflows. So when Consols yields rise beyond 6%, the fall in Consols prices will not be compensated for by an equivalent decrease in the published reserves. This actually explains why the ruin probability for an all-Consols portfolio is so high relative to most of the other 'higher risk' optimal portfolios shown in Tables 6.1 and 6.2. Conversely, if the fund is entirely invested in cash, it will be resilient to rises in interest rates but vulnerable when interest rates fall. Therefore, it is clear that immunization may not be a feasible proposition in respect of published reserves.

As immunization theory does not necessarily lead to low ruin probabilities in situations which may reasonably occur in practice, this approach would appear to be an inadequate means of assessing whether a fund is secure or not. In the circumstances which have been discussed so far, the stability of payouts was shown to be even less useful as a criterion for determining the level of risk attributable to the office. Overall, ruin probabilities seem to provide a straightforward and informative method for quantifying solvency risk. One of its limitations, however, is that it does not distinguish between the different extents of insolvency. If the ruin probability could be incorporated into the utility maximization framework in a suitable manner, the net result may be a more sensible balance between the interests of shareholders and that of the fund itself. This issue is discussed in greater detail in Section 6.3.

6.2.3 Real Payouts

This section deals with the optimal portfolios when the payouts are expressed in real terms, as shown in Tables 6.3 and 6.4, and Figures 6.3 and 6.4. Comparing these two tables, there appears to be little variation between the optimal asset mixes at sm = 15% and sm = 25%, as was the case when the payouts were measured in nominal terms (see Tables 6.1 and 6.2). As expected, the ruin probabilities are noticeably larger in Table 6.3 than in Table 6.4, due to there being more initial surplus available in the latter case.

From Tables 6.1 and 6.2 shown earlier, it had been pointed out that the ruin probabilities in nominal case did not appear to be strongly linked to the risk parameter, r. Out of the five portfolios shown in each of these tables, the optimal mixes with the lowest ruin probabilities were associated with the value of $r = 8$. This seemed to indicate that a strategy which on average produced higher and more variable nominal payouts to shareholders could also result in an office which is less prone to becoming technically insolvent. From Tables 6.3 and 6.4 though, the correlation between ruin probability and risk tolerance appears more convincing when payouts are assessed in real terms.

r	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
2	1	61	31	7	-0.11634	49733	19514	10
4	6	59	13	22	-0.31283	53340	27076	27
8	16	47	0	37	-0.53442	56385	35154	65
16	0	31	0	69	-0.71379	61024	51707	99
32	0	0	0	100	-0.83289	65317	69677	100

Table 6.3. Optimal portfolios: *sm* = 15%, in real terms.

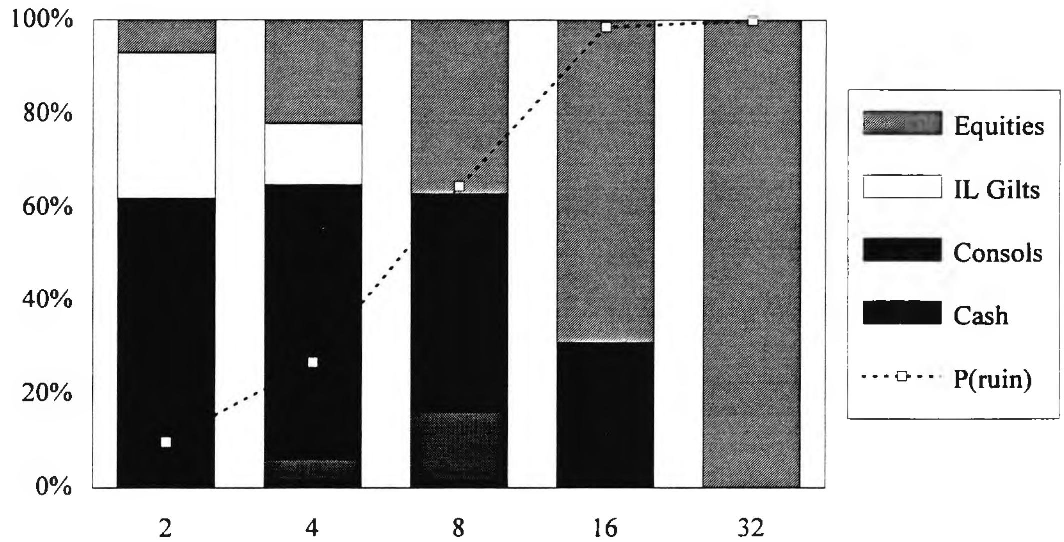


Figure 6.3. Optimal portfolios: *sm* = 15%, in real terms.

r	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
2	3	57	36	4	-0.07743	58641	20746	0
4	2	56	20	22	-0.25306	63123	29845	3
8	8	52	0	40	-0.47498	67709	40842	31
16	3	38	0	59	-0.67133	70723	51331	74
32	0	0	0	100	-0.80767	76318	75880	99

Table 6.4. Optimal portfolios: *sm* = 25%, in real terms.

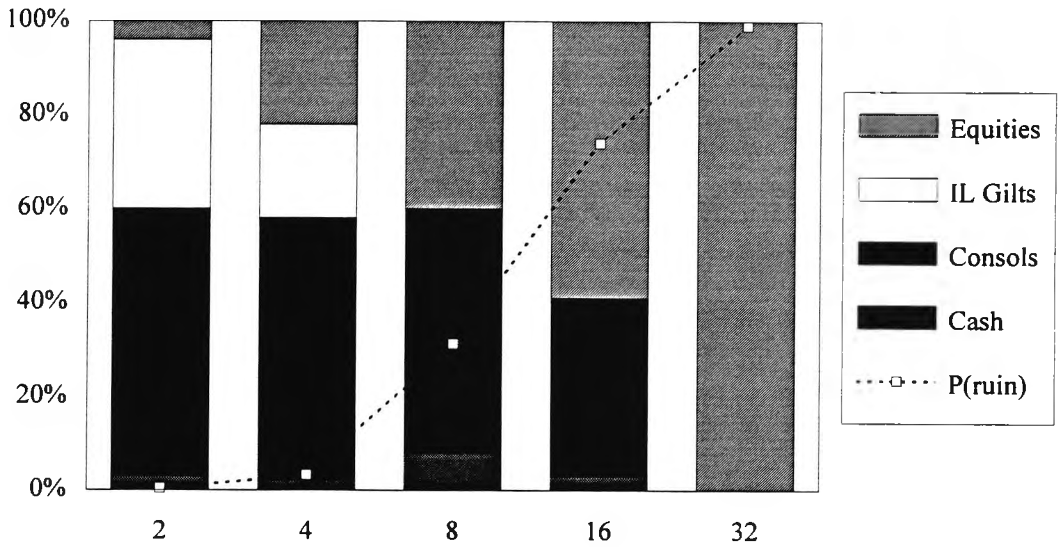


Figure 6.4. Optimal portfolios: *sm* = 25%, in real terms.

Looking at Tables 6.3 and 6.4, the ruin probability increases monotonically with *r*, which is a more intuitive result. As in the nominal case, Consols are the main asset class for the low and medium risk strategies with cash only playing a minor role in these portfolios. The most notable difference between the nominal and real case is in the proportion of index-linked gilts held at the lower risk tolerance levels. The higher proportions held in this asset class when payouts are expressed in real terms is due to the need to earn stable real returns from the solvency margin. But as index-linked gilts also produce the lowest expected accumulations in real terms, their proportions eventually decline towards the higher risk strategies. Another difference between the nominal and real case is in the generally larger proportion of equities held in the case of the latter. This may be a consequence of the strong correlation between equity returns and inflation assumed in the investment model.

6.3 Precision and Efficiency

All the optimal portfolios shown in Section 6.2 were computed using the conjugate gradient method. In order to keep computation time to a minimum, the routine had been run initially with a moderately high precision tolerance level. (The higher the precision tolerance level, the less accurate are the results). As a check that the portfolios obtained were in fact global optima, the results were also compared against all the feasible combinations of asset mixes in steps of 10%. On a few occasions, one or more of these combinations produced expected utilities which were higher than those obtained using the optimization algorithm. In these instances, the precision tolerance level was decreased and the routine re-run, leading to more accurate optima. As a result, every single portfolio shown in the tables above has an expected utility at least as great as any portfolio obtained by the grid approach steps in of 10%.

For the initial level of precision, the number of function evaluations n and iterations I for each of the twenty variants are given in Table 6.5. For example, N15-32 indicates the variant involving nominal payouts, sm = 15% and r = 32. This required 6 iterations and 128 function evaluations. The variations in computational effort generally depend on the proximity of the starting position to the true optimum and whether the function is well scaled. If a function is poorly scaled, it means that the function is relatively 'flat' and the process may converge prematurely. Rescaling the function or decreasing the precision tolerance level usually helps to ameliorate this problem.

Variant	*n*	*I*	*Variant*	*n*	*I*
N15-02	144	8	R15-02	147	6
N15-04	146	7	R15-04	85	4
N15-08	212	10	R15-08	147	6
N15-16	204	8	R15-16	204	4
N15-32	128	6	R15-32	76	4
N25-02	100	6	R25-02	78	4
N25-04	84	5	R25-04	118	5
N25-08	104	5	R25-08	67	3
N25-16	68	3	R25-16	78	3
N25-32	57	3	R25-32	62	3

Table 6.5. Number of function calculations and iterations required for the various situations.

Looking at Table 6.5, an average of about 115 function calculations were required for the optimization process to converge. This compares with the grid approach (in steps of 10%) which needs 286 function calculations. It has been explained in Chapter 4 how the grid approach requires little additional computational effort when applied to more than one utility function, as most of the computation time goes towards calculating the amounts to which the utilities are applied. This economy also applies when computing the utilities of real and nominal amounts. Hence, it should only take about a quarter of the time to derive the optimal mixes for the five risk parameters in both nominal and real terms using the grid approach than it would using the optimization routine.

It then follows that with just four asset classes involved, the only potential advantage of using numerical optimization routines would be in achieving greater precision. If much more accurate results were to be required, the grid approach would not be feasible. The only sensible means of achieving this would be to use optimization routines set to very low precision tolerance levels. However, bearing in mind the potential uncertainty that could be associated with these results, the apparent benefit from the additional accuracy obtained using such optimization routines may be spurious. One approach that may be used to assess the extent of this uncertainty would be to conduct sensitivity analysis on the optimal portfolios.

6.4 Sensitivity Analysis

In previous chapters, it had been noted how the investment model may influence the conclusions one may draw from the results if these are not interpreted carefully. It is therefore important to gain some idea about the sensitivity of the results obtained so far to the investment model used. While it would be ideal to make comparisons between different types of models, Wilkie's model still remains the only widely available investment model in the actuarial field. In view of this, the alternative may be to assess the sensitivity of these results to changes in the model's parameter values. Nevertheless, it is worth noting that this will not allow for uncertainty in the model structure itself.

6.4.1 Parameter Alterations

The sets of parameter alterations which have been chosen for this purpose are first detailed here before their resulting optimal portfolios are shown in Section 6.4.2. An obvious place to begin would be to consider parameter estimates obtained using a more recent data period, as published by Wilkie (1995a). These estimates were based on the period 1923-1990 and any differences between this version and the Standard Basis are shown in Table 6.6. The last three parameters, *BMU*, *BA* and *BSD* refer to Wilkie's short term interest rates model. As the Standard Basis did not include such a model, the comparison is made against the parameters, *KA* and *KSD* described in the cash model from Section 3.5. $KMU = 0.0$ is implicit in this cash model.

Parameter	*Standard Basis*	*Wilkie (1995a)*
QSD	0.050	0.040
YW	1.350	1.950
YMU	0.040	0.038
YA	0.600	0.500
YSD	0.175	0.160
DMU	0.000	1.35%
DY	-0.200	-0.175
DB	0.375	0.550
DSD	0.075	0.060
CMU	0.035	0.031
CA1	1.200	0.900
CA2	-0.480	0.000
CA3	0.200	0.000
CY	0.060	0.150
CSD	0.140	0.175
BMU	0.000	0.185
BA	0.400	0.750
BSD	0.250	0.175

Table 6.6. Parameters alterations specified in Wilkie (1995a).

The main effects of these changes may be summarized as follows:

1. Reducing *QSD* lowers the variance of the force of inflation.
2. The net effect of increasing *YW* and reducing *YMU* is to reduce the long-term average dividend yield by about 0.1%.
3. Increasing *DMU* raises the long-term average annual rate of increase in dividends by about 1.5%, while lowering *DSD* reduces its variance.

4. Reducing *CMU* lowers the long-term mean of the Consols yield by 0.4%, while raising *CSD* increases its variance.
5. Introducing *BMU* and reducing *CMU* lowers the long-term average cash yields by about 1.75%, while lowering *BSD* reduces its variance.

	RPI	*Cash*	*Consols*	*IL Gilts*	*Equities*
MEAN (%)	5.6	7.3	8.4	9.3	12.1
	(5.8)	(9.2)	(8.8)	(9.6)	(11.0)
S.D.	1.3	1.2	0.9	2.6	6.4
	(1.7)	(1.9)	(1.1)	(3.5)	(6.4)
CORRELATION					
Cash	0.41				
	(0.53)				
Consols	-0.11	0.19			
	(-0.08)	(0.28)			
IL Gilts	0.99	0.46	-0.01		
	(0.99)	(0.58)	(0.00)		
Equities	0.54	0.23	0.04	0.54	
	(0.57)	(0.29)	(0.03)	(0.57)	

Table 6.7. Means, standard deviations and correlation coefficients of 20 year nominal accumulations using Wilkie (1995a).

Using the same random seed as before, the simulation results over twenty years for this revised version may be summarized in Table 6.7. The corresponding figures for the Standard Basis obtained from Table 3.9 are shown in brackets. When compared with the Standard Basis, noticeable changes have occurred in the means and variances of the accumulations. The mean accumulation rate for cash has fallen from 9.2% to 7.3% per annum. The mean accumulation rates for inflation, Consols and index-linked gilts have fallen to a lesser extent and the mean accumulation rate on equities has increased from

11.0% to 12.1%. The standard deviations of the twenty year accumulations factors have decreased for all variables, most notably for inflation, cash and index-linked gilts.

Another basis for sensitivity testing could involve the use of parameter estimates which are a stated number of standard errors above or below their best estimates. In Wilkie (1984), it had been suggested that 1.5 standard errors (s.e.) would be appropriate degree of uncertainty to adopt for this purpose. However, rather than testing the numerous parameter changes this would involve, only the three alterations shown in Table 6.8 will be used here. These parameter alterations will be tested individually.

Parameter	*Best Estimate*	*Estimate+1.5 s.e.*
QA	0.600	0.750
YMU	0.040	0.045
CMU	0.035	0.045

Table 6.8. Three alternative parameter values suggested in Wilkie (1984).

The effect of increasing in *QA* is to lower the tendency for the force of inflation to return to its mean value of *QMU*, making it more unstable. Increasing *CMU* and *YMU* on the other hand only serves to raise the mean real yield component in Consols and the mean share dividend yield respectively. It is also worth noting that the *QA* parameter is a structural parameter, rather than a residual parameter as in *CMU* and *YMU.*

The final set of parameter changes used here relates to a recent review of Wilkie's 1986 model by Huber (1995). On the basis that Wilkie's structural parameter values were appropriate, Huber recalculated the residuals over the out-of-sample period from 1983 to 1993 and the set of residual parameter alterations implied by this analysis is shown in Table 6.9. The changes will be applied simultaneously and have should have the following effects: a decrease in the variance of the force of inflation, an increase in the mean growth rate of the dividend index, a decrease in the mean yield on Consols, and increase in the variance of the Consols yield.

Parameter	*Standard Basis*	*1983-93*
QSD	0.050	0.0250
DMU	0.000	0.0188
CMU	0.035	0.0318
CSD	0.140	0.3000

Table 6.9. Parameter alterations based on 1983-93 errors

Therefore, the five sets of parameter alterations discussed so far could be referred to as:

1. Wilkie (1995a)
2. $QA = 0.75$
3. $YMU = 0.045$
4. $CMU = 0.045$
5. 1983-93 errors

6.4.2 Optimization Results

For each of these five variants, the optimal mixes at $r = 8$ assuming $sm = 15\%$ were then recalculated with payouts measured in nominal and real terms, as shown in Tables 6.10 and 6.11 respectively. This particular risk parameter was chosen because it appeared to represent a sensible risk position based on the results shown in Table 6.1.

In Table 6.10 the range of optimal proportions in Consols and equities is rather wide. 'Wilkie (1995a)' requires 45% and 55% in Consols and equities respectively, compared with 79% and 15% respectively for the Standard Basis. Similarly, increasing *CMU* to 0.045 produces an optimal mix with 92% in Consols and no equities. It is also interesting to note how similar the optimal mix from '1983-93 errors' is to that of 'Wilkie (1995a)', perhaps indicating the influence which the past ten years' data may have had on the latter's parameter values.

Model	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Ruin %*
Standard	5	79	1	15	-0.22750	36
Wilkie (1995a)	0	45	0	55	-0.18420	91
QA = 0.75	0	90	0	10	-0.22961	51
YMU = 0.045	1	76	0	23	-0.20708	49
CMU = 0.045	8	92	0	0	-0.14326	39
1983-93 errors	0	51	0	49	-0.11805	82

Table 6.10. Optimal portfolios for six versions of Wilkie's model: *sm* = 15%, in nominal terms, at *r* = 8.

Model	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Ruin %*
Standard	16	47	0	37	-0.53442	65
Wilkie (1995a)	0	0	0	100	-0.41852	100
QA = 0.75	0	73	0	27	-0.53481	60
YMU = 0.045	0	42	0	58	-0.49770	94
CMU = 0.045	58	41	0	1	-0.45117	0
1983-93 errors	0	0	0	100	-0.34298	100

Table 6.11. Optimal portfolios for six versions of Wilkie's model: *sm* = 15%, in real terms, at *r* = 8.

In real terms, the results seem even more erratic. This is a little surprising given the cascade structure of the model. One would expect real amounts to be fairly insensitive to parameter changes involving the retail price index. But this has clearly been offset by changes in other parameters, given that 100% in equities is optimal for 'Wilkie (1995a)' compared with only 1% in equities when $CMU = 0.045$. As with Table 6.10 though, '1983-93 errors' produces optimal mixes which are similar to those of 'Wilkie (1995a)'.

The high degree of sensitivity in these optimal asset allocations has a number of implications in decision-making. For a given utility function and liability profile, the optimal asset mix will always be conditional on the asset model. Regardless of how well a model may be perceived to represent economic variables, there will always be some degree of uncertainty relating to the model structure and its parameter values.

In the case of Wilkie's model, updating the model parameters using more recent data led to moderate changes in the distributions of the asset classes. However, judging from the results above, it would seem that even a 1% change in the mean accumulation rate per annum may significantly affect the optimal mixes obtained. Although this points to the need for an investment model whose parameters are more stable over time, it seems doubtful whether an alternative model could be constructed which would be able to cope with the inherent instability in economic variables.

Another implication relates back to the issue of precision discussed in Section 6.3. It was mentioned that the only apparent benefit of using optimization routines with just four asset classes was in achieving greater precision. However, given the uncertainty in the optimal portfolios obtained, the grid approach now seems to be the most sensible method for analysing asset allocation strategies under these circumstances. As well as being able to guarantee global optima within the precision level specified, the grid approach is also capable of incorporating ruin criteria into the analysis, which otherwise would have been extremely difficult to implement satisfactorily.

6.5 Incorporating Ruin Criteria

In all the optimization problems considered so far, the sole objective has been to maximize the expected utility of shareholders. From a practical perspective, though, the portfolio selected should also aim to satisfy the other interested parties involved, such as policyholders and regulators. Broadly speaking, this means ensuring that an adequately low ruin probability is maintained. One way of achieving this could be to select the portfolio with the highest expected utility subject to the constraint that the probability of ruin does not exceed a specified limit. This will be considered shortly in Section 6.5.2. Including constraints of this form, however, may take the analysis away from the traditional utility maximizing framework.

A feature of the grid approach is that it enables all the portfolios included to be ranked by expected utility or any other measure. This could be particularly useful if another factor such as the probability of ruin needs to be taken into account, but not treated as a strict constraint. For example, the optimal or highest ranking portfolio may have an unacceptably high ruin probability, whereas say the fifth ranking portfolio might give a much more acceptable probability of ruin. In this respect, the fifth ranking portfolio may be preferred to the optimal one.

Clearly, this form of analysis is also not wholly consistent with a utility maximizing approach either, allowing additional scope for subjectivity which is both an advantage and a disadvantage. The ranking approach is flexible in that it encompasses portfolios which may be more acceptable in practice but could also have been overlooked through a constrained or unconstrained utility maximizing approach. However, this also means that individual consideration needs to be given to all these grid portfolios, which may be an overwhelming number in total. Therefore, in order to make this approach more practicable, one could limit the portfolios under consideration to say the ten with the highest expected utilities, as described below.

6.5.1 Top Ten Ranking Portfolios

The series of tables which follow in this section relate to the top ten portfolios ranked by expected utility, at values of r = 2, 8 and 32. This is performed for sm = 15% and 25%, and with payouts expressed in nominal and real terms. In each table, the expected utilities, means and standard deviations of payouts, and ruin probabilities are also given next to the corresponding optimal portfolios.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	0	100	0	0	-0.00955	127384	47281	62
2	10	90	0	0	-0.00973	129079	50103	42
3	20	80	0	0	-0.01024	130739	54146	25
4	0	90	10	0	-0.01039	132205	56726	40
5	10	80	10	0	-0.01091	133994	61002	21
6	30	70	0	0	-0.01112	132537	59380	9
7	0	90	0	10	-0.01182	137370	64042	47
8	20	70	10	0	-0.01183	135976	66339	9
9	10	80	0	10	-0.01218	139212	67669	31
10	0	80	20	0	-0.01225	137569	70549	24

Table 6.12. Top ten portfolios: sm = 15%, in nominal terms, at r = 2.

Table 6.12 above refers to the top ten portfolio at r = 2, with sm = 15% and payouts measured in nominal terms. Here, the point made earlier about neighbouring portfolios possibly having better overall characteristics is quite apparent. The asset mix giving the highest expected utility is that of 100% Consols, even though 62% of the simulations have become technically insolvent within the twenty year projection period. On the other hand, shifting 30% of these assets into cash would leave the resulting portfolio ranked sixth out of 286, but with a more favourable ruin probability of 9%. Given the choice between these two portfolios, an investment manager may well prefer the latter. Hence, considering ruin probabilities may allow for the fact that the smooth objective function does not deal adequately with insolvency risk.

Looking at the means and standard deviations of the payouts, it is fairly clear why the optimal mix is entirely invested in Consols. As a relatively low risk tolerance parameter is being used, there is a strong affinity for portfolios with low standard deviations of payouts. From Table 6.1, it would seem that the asset mixes which give lower ruin probabilities are sub-optimal because they also tend to produce more variable payouts. This emphasizes the difference between the objective of stabilizing payouts and other approaches for managing risk such as immunization and minimizing ruin probabilities.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	10	70	0	20	-0.22803	150018	89136	37
2	0	80	0	20	-0.22810	147727	84799	50
3	10	80	0	10	-0.22857	139212	67669	31
4	0	90	0	10	-0.22898	137370	64042	47
5	0	80	10	10	-0.22901	142467	76018	29
6	20	60	0	20	-0.22908	152327	94169	23
7	20	70	0	10	-0.22940	141046	72212	14
8	0	70	10	20	-0.22972	153389	98407	35
9	10	70	10	10	-0.22979	144450	80674	14
10	10	60	10	20	-0.23093	155732	103422	21

Table 6.13. Top ten portfolios: *sm* = 15%, in nominal terms, at *r* = 8.

When r takes the value of 8 represented by Table 6.13, the ruin probabilities from the top ten portfolios appear more uniform when compared with Table 6.12, although a reasonable reduction in ruin probability may be secured by choosing say the seventh ranking portfolio rather than the optimal one. The means and standard deviations of the payouts imply that the trade-off between these two statistics is key in determining the ranks, as opposed to an affinity for either high means or low standard deviations. When $r = 32$ (see Table 6.14), there is clearly no justification for choosing any portfolio other than the optimal one on grounds of ruin probabilities. Although there seems to be some relationship between the rank and the mean payout, the optimal portfolio is not the portfolio with the highest mean payout. Therefore, this risk position still represents a compromise between the variance and the expected return of the payouts.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	0	0	40	60	-0.60943	215385	246260	100
2	0	0	50	50	-0.60998	213034	241962	99
3	0	0	30	70	-0.61007	217443	251278	100
4	10	0	30	60	-0.61036	211017	234218	99
5	10	0	40	50	-0.61070	208699	229624	98
6	10	0	20	70	-0.61125	213023	239614	100
7	0	10	30	60	-0.61168	207949	227985	99
8	0	0	60	40	-0.61174	210405	238523	98
9	0	0	20	80	-0.61182	219238	257030	100
10	20	0	20	60	-0.61184	206546	222449	98

Table 6.14. Top ten portfolios: sm = 15%, in nominal terms, at $r = 32$.

When the office has more surplus to begin with, the ruin probabilities should be lower and this is evident in the case where sm = 25% (see Tables 6.15 to 6.17). However, the same feature noted in Table 6.12 also appears in Table 6.15, with the ruin probability of the optimal portfolio being much larger than that of the nine others shown. A more reasonable alternative might be the third ranking portfolio, with a ruin probability of only 3%. If achieving a low ruin probability is paramount, then the seventh portfolio may be ideal, with approximately none of the one thousand scenarios resulting in ruin. On closer inspection though, an unusual ranking seems to have emerged. The seventh portfolio is clearly mean-variance inefficient compared with the eighth portfolio even the former has a higher expected utility. This is probably due to the problem of using a very low risk tolerance parameter discussed in Section 4.4.3.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	0	100	0	0	-0.00415	152810	54672	22
2	10	90	0	0	-0.00424	154109	57225	10
3	20	80	0	0	-0.00446	155542	61170	3
4	0	90	10	0	-0.00460	157107	64219	9
5	10	80	10	0	-0.00483	158849	68491	2
6	30	70	0	0	-0.00485	157237	66431	1
7	20	70	10	0	-0.00523	160862	73950	0
8	0	90	0	10	-0.00530	163496	72838	13
9	40	60	0	0	-0.00549	159326	73059	0
10	10	80	0	10	-0.00551	164984	76229	5

Table 6.15. Top ten portfolios: $sm = 25\%$, in nominal terms, $r = 2$.

At $r = 8$ (see Table 6.16), there is a slight tendency to hold asset mixes with low standard deviations of payouts, in preference to those with higher expected payouts. Nevertheless, some benefit may still be derived from picking say the sixth or seventh ranking portfolio, as these have much lower ruin probabilities. At the highest risk level, all portfolios shown in Table 6.17 perform just as poorly in terms of ruin probabilities.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	0	90	0	10	-0.17543	163496	72838	13
2	0	80	0	20	-0.17577	174959	95543	13
3	10	80	0	10	-0.17593	164984	76229	5
4	10	70	0	20	-0.17661	176898	99750	6
5	0	80	10	10	-0.17704	168266	84997	6
6	20	70	0	10	-0.17723	166610	80682	1
7	10	70	10	10	-0.17818	170189	89785	1
8	20	60	0	20	-0.17822	178965	104679	3
9	0	70	10	20	-0.17875	180249	109291	7
10	30	60	0	10	-0.17915	168570	86350	0

Table 6.16. Top ten portfolios: *sm* = 25%, in nominal terms, *r* = 8.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	0	0	40	60	-0.56756	247809	270244	89
2	0	0	50	50	-0.56796	245130	265504	81
3	10	0	30	60	-0.56801	243236	257421	84
4	10	0	40	50	-0.56822	240521	252170	75
5	0	0	30	70	-0.56854	250017	275641	94
6	20	0	30	50	-0.56890	235864	239137	69
7	20	0	20	60	-0.56893	238580	244906	80
8	0	10	30	60	-0.56901	240014	250794	85
9	0	10	40	50	-0.56916	237270	245259	75
10	10	0	20	70	-0.56923	245414	263228	91

Table 6.17. Top ten portfolios: *sm* = 25%, in nominal terms, *r* = 32.

When payouts are measured in real terms, the ruin probabilities seem to increase with r, as seen by comparing Tables 6.18 to 6.20. However, at low levels of risk, the ruin probabilities are relatively more uniform within the top ten portfolios. In the case where $r = 2$, the ruin probabilities lie between 2% and 17% when payouts are in denominated real terms compared with 9% and 62% when payouts are denominated in nominal terms. While this may be a consequence of the more diverse portfolios resulting from the need to hold real assets, it also means that there is comparatively less benefit in selecting from any of the nine lower ranking portfolios.

At higher risk levels (see Table 6.19 and 6.20), the scope for choosing other asset mixes is reduced further as all the ruin probabilities are very high. The need to hold real assets remains but with the emphasis now shifting away from index-linked gilts into equities. At $r = 32$, the all-equities portfolio is optimal (see Table 6.20). Looking at the means and standard deviations of payouts, it is clear that this represents the ultimate high risk strategy. At this extreme level of risk tolerance, the chance of ruin is near certain.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	10	60	30	0	-0.11677	48417	17177	7
2	0	60	30	10	-0.11683	50359	20754	11
3	10	60	20	10	-0.11727	50563	21290	8
4	0	70	30	0	-0.11739	48361	17496	14
5	0	70	20	10	-0.11791	50603	21822	17
6	20	50	30	0	-0.11793	48503	17555	2
7	0	60	40	0	-0.11799	48126	17039	11
8	20	50	20	10	-0.11805	50576	21319	3
9	10	50	30	10	-0.11827	50394	20970	5
10	20	60	20	0	-0.11905	48602	18378	3

Table 6.18. Top ten portfolios: sm = 15%, in real terms, $r = 2$.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	10	50	0	40	-0.53448	56822	36602	75
2	20	40	0	40	-0.53461	56886	36692	70
3	0	60	0	40	-0.53479	56770	36750	79
4	0	50	10	40	-0.53504	56634	36208	75
5	20	50	0	30	-0.53506	55265	31852	45
6	10	60	0	30	-0.53533	55255	32081	55
7	30	30	0	40	-0.53533	56953	37031	65
8	10	40	10	40	-0.53541	56704	36404	71
9	30	40	0	30	-0.53549	55266	31941	36
10	10	50	10	30	-0.53557	54973	31044	43

Table 6.19. Top ten portfolios: *sm* = 15%, in real terms, *r* = 8.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	0	0	0	100	-0.83289	65317	69677	100
2	10	0	0	90	-0.83350	64133	64319	100
3	0	10	0	90	-0.83359	63963	63697	100
4	0	0	10	90	-0.83381	64152	65031	100
5	20	0	0	80	-0.83452	62879	59108	100
6	10	10	0	80	-0.83457	62728	58428	100
7	0	20	0	80	-0.83467	62583	57889	100
8	10	0	10	80	-0.83472	62912	59713	100
9	0	10	10	80	-0.83481	62735	58978	100
10	0	0	20	80	-0.83511	62919	60573	100

Table 6.20. Top ten portfolios: *sm* = 15%, in real terms, *r* = 32.

If additional capital were available initially i.e. *sm* = 25%, the optimal portfolios in real terms would be quite satisfactory at low values of *r*. The utility maximizing portfolio at *r* = 2 shown in Table 6.21 has a ruin probability of 1%, which if accurate in absolute terms should be an acceptable level of risk for most offices. However, given the lack of credibility associated figures of this nature, it would be difficult to justify choosing a lower ranking portfolio on the basis that it has a zero probability of ruin.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	0	60	40	0	-0.07798	57602	19212	1
2	0	60	30	10	-0.07822	60068	23538	1
3	20	50	30	0	-0.07831	58039	19922	0
4	10	60	30	0	-0.07833	57937	19872	0
5	10	50	30	10	-0.07841	60096	23478	0
6	10	50	40	0	-0.07862	57705	19564	0
7	0	70	30	0	-0.07955	57846	20483	1
8	20	50	20	10	-0.07957	60283	24248	0
9	10	60	20	10	-0.07974	60279	24489	0
10	0	50	40	10	-0.07990	60081	24112	1

Table 6.21. Top ten portfolios: *sm* = 25%, in real terms, *r* = 2.

Compared with Tables 6.16 and 6.17, the general level of ruin probabilities seen in Tables 6.22. and 6.23 is also reduced as a result of the higher initial solvency margin. However, the significant equity investment and the high ruin probabilities associated with these levels of risk tolerance would probably render most of these portfolios as being inappropriate in practice.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	10	50	0	40	-0.47498	67704	40799	30
2	0	60	0	40	-0.47518	67737	41136	35
3	20	40	0	40	-0.47535	67675	40748	26
4	0	50	10	40	-0.47569	67404	40218	31
5	10	40	10	40	-0.47621	67398	40281	27
6	30	30	0	40	-0.47623	67672	40982	24
7	0	50	0	50	-0.47685	69356	46209	57
8	20	50	0	30	-0.47695	65679	35736	9
9	10	60	0	30	-0.47700	65771	36164	14
10	10	40	0	50	-0.47710	69384	46187	52

Table 6.22. Top ten portfolios: *sm* = 25%, in real terms, *r* = 8.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	0	0	0	100	-0.80767	76318	75880	99
2	10	0	0	90	-0.80788	75165	70239	97
3	0	10	0	90	-0.80796	74995	69655	97
4	0	0	10	90	-0.80826	75150	70873	98
5	20	0	0	80	-0.80849	73943	64743	94
6	10	10	0	80	-0.80853	73792	64093	94
7	0	20	0	80	-0.80860	73657	63605	94
8	10	0	10	80	-0.80881	73922	65258	95
9	0	10	10	80	-0.80888	73747	64561	95
10	0	0	20	80	-0.80926	73896	66036	97

Table 6.23. Top ten portfolios: *sm* = 25%, in real terms, *r* = 32.

In conclusion, there seem to be definite benefits that may be derived from this form of analysis. Depending on the liability model used, there may be instances when a utility maximizing portfolio may not be suitable in practice. With Model A, the ruin probabilities do not appear to have a consistent relationship with the means and standard deviations of payouts, which largely determine the expected utilities. This is because no real penalty is imposed on the objective function when ruin occurs. By considering a limited selection of high ranking asset mixes, a reasonable compromise between ruin probability and expected utility may sometimes be obtained. In effect, one is proxying a discontinuity in the objective function by doing so. From a general viewpoint, even when the optimum portfolio is acceptable, looking at the top slice of the full range of asset mixes may help to highlight the features which make this portfolio optimal.

6.5.2 Optimal Portfolios with Ruin Constraints

Notwithstanding the usefulness of having fuzzy boundaries when allowing for ruin probabilities (see Section 6.5.1), it may often be necessary to employ more objective methods for making decisions in view of such criteria. As mentioned earlier, one approach could be to maximize expected utilities subject to satisfying a ruin constraint. In this section, only portfolios which have a ruin probabilities of no more than 5% will be considered. A typical feature seen in these results is that some of the optimal portfolios may be identical to those at values of r adjacent to them. This is generally due to the fairly large steps of 10% used in the grid approach.

Figure 6.5 below shows the constrained optimal mixes at various values of r. Imposing the constraint appears to result in a more sensible range of asset mixes, given the relative size and nature of the liabilities. When r takes the value of either 2 or 4, the optimal portfolios are found to be entirely invested fixed-interest assets, as had been seen for the unconstrained case. However, in contrast to those shown in Figure 6.1, these portfolios are consistent with the immunized ratio of 2 : 7 in cash and Consols respectively. As the realistic liabilities form on average about three quarters of the total assets, a minimum of $0.75 \times 2/9 = 17\%$ and $0.75 \times 7/9 = 58\%$ of the total fund must be invested in cash and Consols respectively if it is to be immunized on this basis. The constrained optima at these levels of risk tolerance lie within this range, with 40% and 60% invested in cash and Consols respectively. The fact that much more had been invested in cash than was necessary probably enforces the assertion that with an upper limit on the published valuation rate of interest, lower ruin probabilities may be achieved by investing shorter than implied by the 2 : 7 ratio.

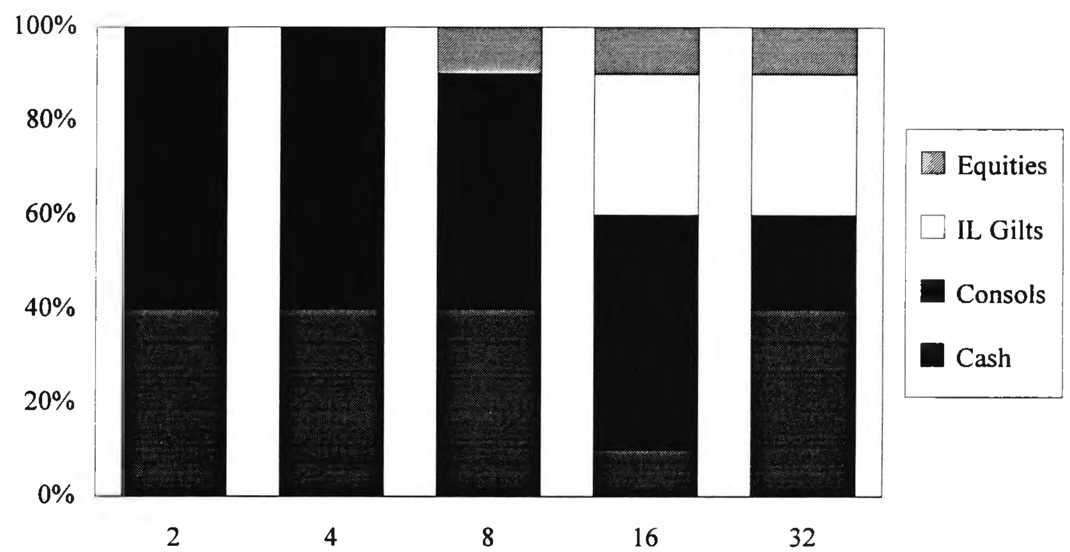


Figure 6.5. Optimal portfolios with 5% ruin constraint: *sm* = 15%, in nominal terms.

As the level of risk tolerance increases, the constrained portfolios tend to become more diverse in order that higher expected payouts may be achieved. However, they are still predominantly invested in fixed income assets as this is necessary to maintain the low ruin probabilities required. These portfolios differ strongly from their unconstrained counterparts in Figure 6.1, which were largely invested in the real asset classes.

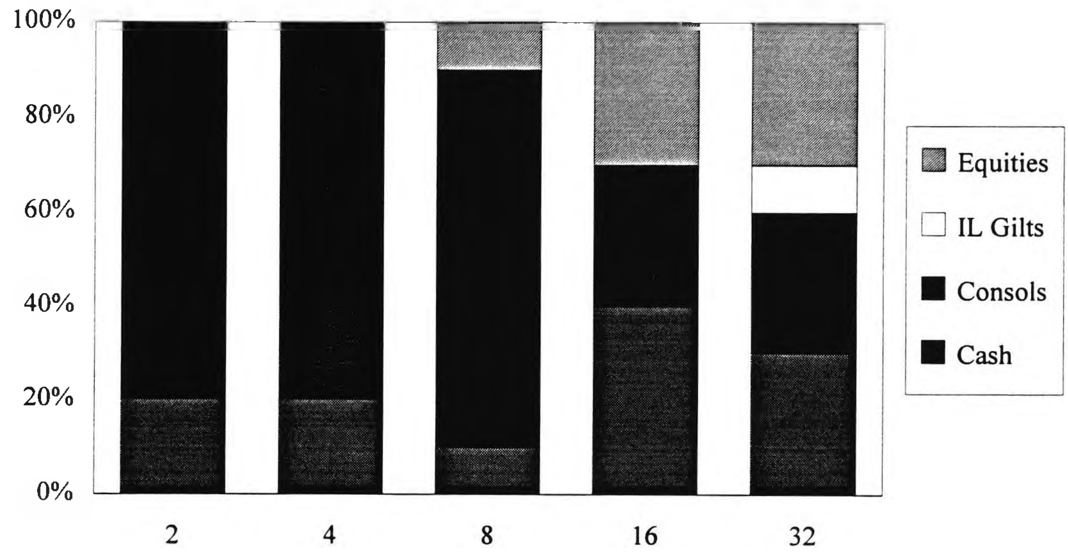


Figure 6.6. Optimal portfolios with 5% ruin constraint: sm = 25%, in nominal terms.

From Figure 6.6, it is apparent that the higher initial solvency margin of 25% may permit more volatile portfolios compared with those in the 15% case. At lower risk tolerance levels, as much as 80% of the assets are allowed to be invested in Consols. While this portfolio may lie within the immunized range on a realistic basis, such a fund is in fact less well matched on a published basis given the ceiling which applies to the valuation rate of interest. This portfolio would not have been able to satisfy the 5% ruin probability if sm had only been 15%. When r is increased to values of 16 or 32, the resulting portfolios are still heavily invested in fixed interest assets, as had been the case with sm = 15%. However, the higher proportions that may now be invested in equities are a result of the additional solvency margin.

Figures 6.7 and 6.8 show the constrained portfolios when payouts are calculated in real terms. When sm = 15% and r = 2, the optimal portfolio comprises 20% in cash, 50% in Consols and 30% in index-linked gilts. Hence, it would appear that the fund is not fully immunized on a realistic basis as this would require at least 17% and 58% of the total assets to be invested in cash and Consols respectively. However, given that the optimal proportions derived are only correct to the nearest 10%, and the fact that the published valuation basis has a tendency to cause portfolios with shorter durations to incur lower ruin probabilities, this mix seems quite reasonable. As index-linked gilts have the most stable real accumulations over this period, investing the free assets in this asset class would seem to be a natural means of reducing the variability of real payouts.

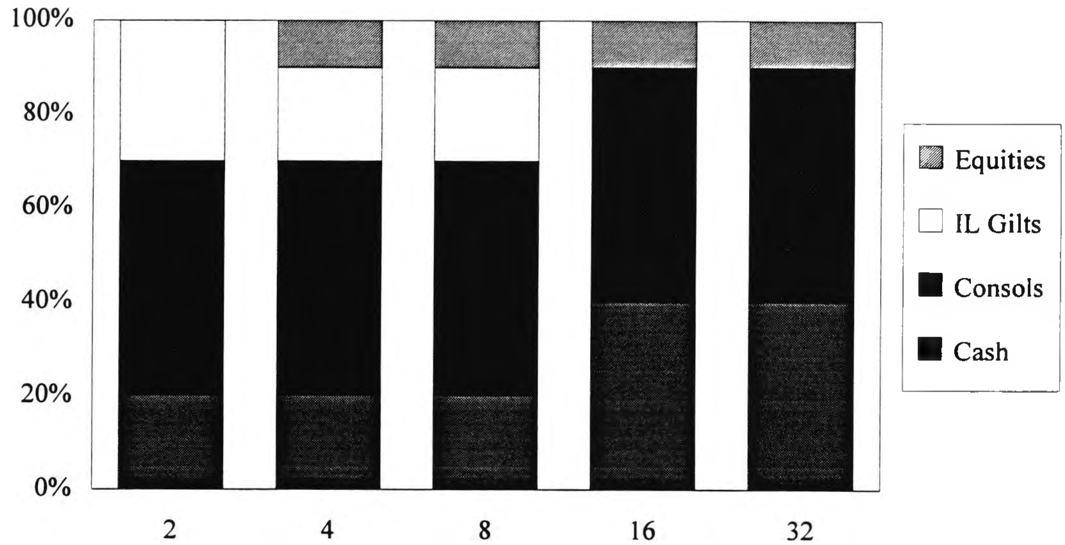


Figure 6.7. Optimal portfolios with 5% ruin constraint: *sm* = 15% in real terms.

As r is increased to 4 and 8, the need to achieve higher expected real payouts causes 10% of the assets to be switched from index-linked gilts to equities. However, as it would not be possible to maintain the 5% ruin criterion whilst investing 20% of the fund in equities, the proportions in equities remains at 10%, even when r is increased beyond the value of 8. At $r = 16$ and 32, higher expected real payouts are obtained by switching from index-linked gilts to cash. Index-linked gilts are less appealing at such high levels of risk tolerance because they yield the lowest expected real accumulations over twenty years. Nevertheless, they are switched into cash rather than Consols because investing more in the latter would lead to unacceptable ruin probabilities.

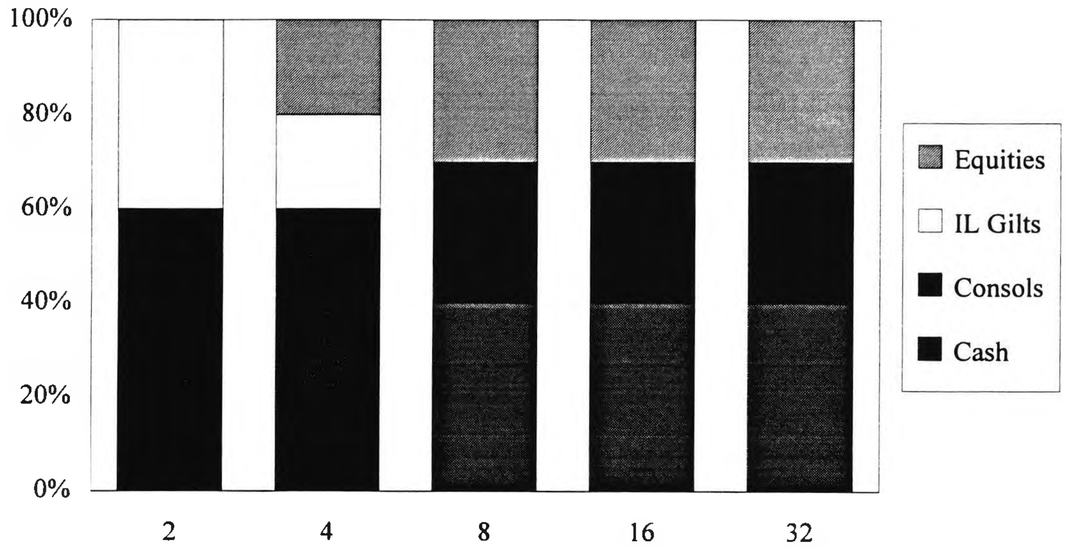


Figure 6.8. Optimal portfolios with 5% ruin constraint: *sm* = 25%, in real terms.

With *sm* = 25% (see Figure 6.8), more scope is available for greater mismatching. For example, the optimal mix at $r = 2$ has 60% in Consols and the remainder in index-linked gilts. i.e. no cash is necessary to satisfy the ruin constraint. As r is increased, the proportions in index-linked gilts are seen to decrease, allowing more of the fund to be invested in equities. When r takes the values of 8 or greater, 30% of the assets are invested in equities. However, in order to maintain a ruin probability of 5% or below with such a high proportion of the fund invested in equities, a large proportion of Consols had to be switched into cash.

Therefore, it has been shown that constrained utility maximization can be an appealing method for practical decision-making. The method is objective and leads to intuitively reasonable solutions. However, a possible drawback of this approach is that it accepts or rejects portfolios based on absolute values of the ruin probabilities. Due to the lack of credibility generally associated with investment models, additional caution may need to be exercised when basing decisions on the tails of the distributions produced by these models. As noted in Geoghegan *et al* (1992), " ... [Wilkie's] *model may not be* *considered appropriate for the estimation of extreme values, such as probabilities of ruin, etc.*" This contrasts with the 'top ten' approach in Section 6.5.1, in which greater emphasis had been placed on the relative ruin probabilities.

6.6 Mean-Variance Efficiency

6.6.1 Efficient Frontiers

Figures 6.9 to 6.12 show for each of the four situations (i.e. sm = 15% or 25%, with nominal or real payouts), the E-V efficient frontier, the unconstrained utility maximizing portfolios and the utility maximizing portfolios subject to a 5% ruin constraint. The main purpose of these graphs is to highlight the extent to which ruin constraints may alter the optimal mixes in relation to the E-V efficient frontier. It also provides a means of checking for any other peculiar trends occurring in the results. As stated in Chapter 4, the line representing the E-V efficient frontier should in theory be a smooth curve. However, as the grid has only been computed in steps of 10%, the curves shown may appear quite jagged. In all cases, the standard deviations of payouts are measured along the vertical axes, while the horizontal axes represent the mean payouts.

Looking first at Figure 6.9, the results seem consistent with those of the asset fund (see for example Figure 4.4), with all the unconstrained utility maximizing portfolios lying on the E-V efficient frontier. However, imposing the 5% ruin constraint on the utility maximizing portfolios appears to alter these portfolios in a number of respects. As a result of the constraint, the portfolios are concentrated in a smaller region near the more risk averse end of the efficient frontier, but away from the extreme points of the frontier. The optimal portfolios at $r = 2$ and 4 have shifted up the curve, whereas the higher risk portfolios seem to have shifted downwards. More significantly, two of these portfolios

(r = 16 and 32) now lie above the efficient frontier, which means that they are no longer efficient from a mean-variance perspective.

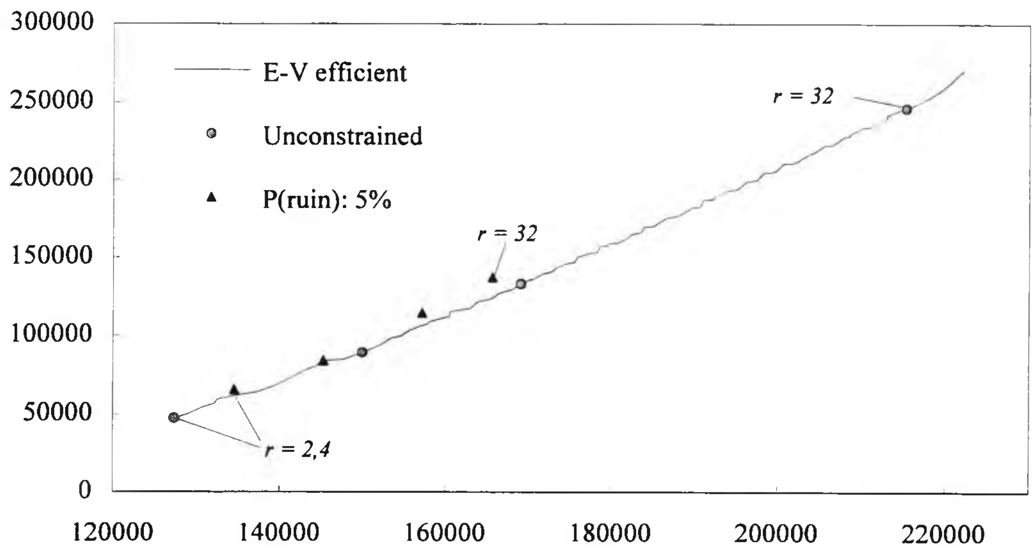


Figure 6.9. Graph of S.D. vs. Mean for *E-V* efficient and utility maximizing portfolios with and without ruin constraints: sm = 15%, in nominal terms.

Increasing the initial level of surplus from 15% to 25% does not really alter the shape of the frontier, as shown in Figure 6.10. However, the additional surplus has reduced the influence of the constraints on the utility maximizing portfolios, resulting in them being shifted in the same manner as before, but to a lesser degree. As a consequence, some of the constrained optimal portfolios may appear to be *E-V* efficient when in fact they are not. For instance in the Figure 6.10, the uppermost triangle (r = 32) is lying virtually on one of the unconstrained optima (r = 16). But by looking at the exact means and standard deviations of payouts resulting from these two portfolios, it is clear that the constrained portfolio must actually be lying above the unconstrained portfolio, and hence above the efficient frontier. Nevertheless, they both seem close enough to being *E-V* efficient for most practical purposes.

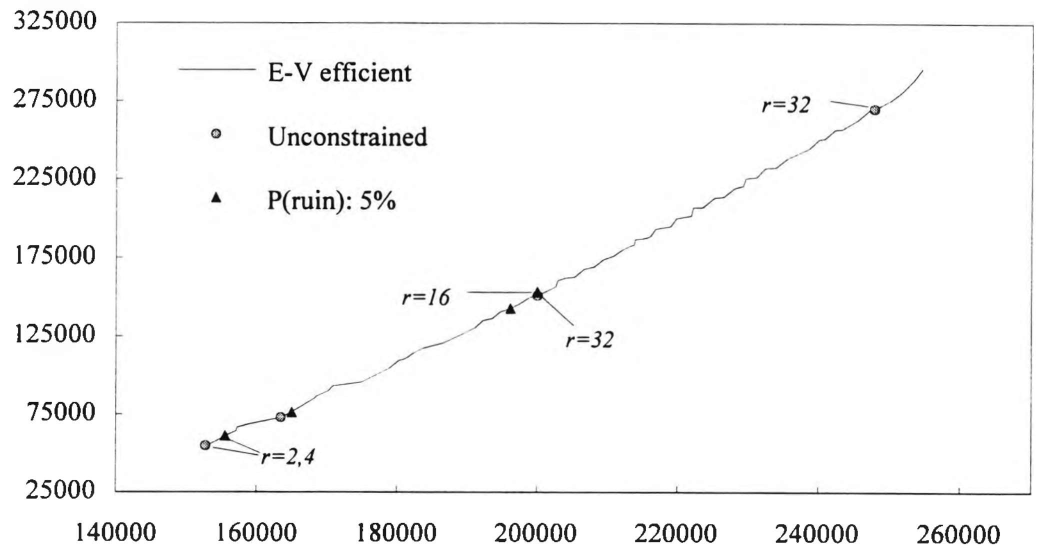


Figure 6.10. Graph of S.D. vs. Mean for *E-V* efficient and utility maximizing portfolios with and without ruin constraints: *sm* = 25%, in nominal terms.

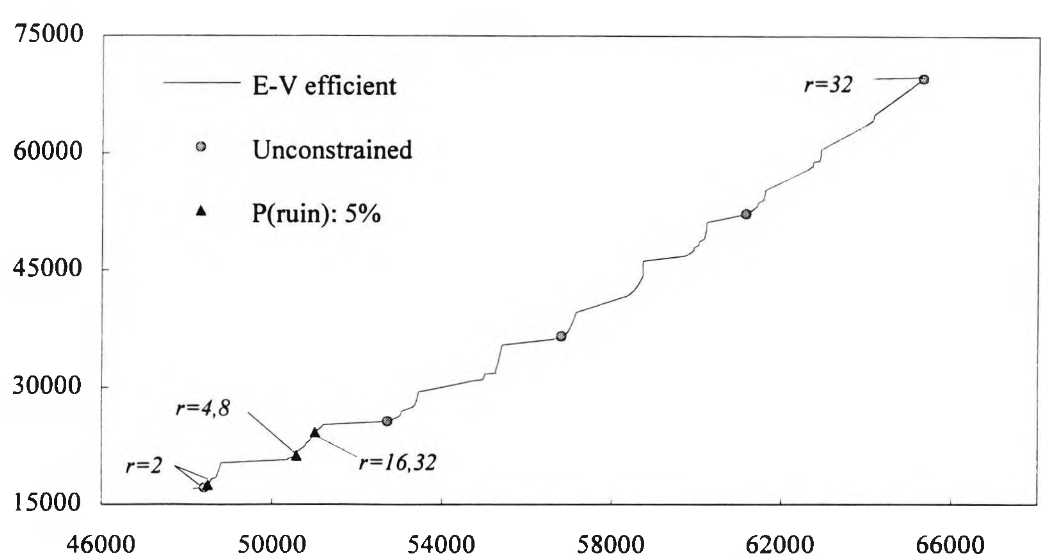


Figure 6.11. Graph of S.D. vs. Mean for *E-V* efficient and utility maximizing portfolios with and without ruin constraints: *sm* = 15%, in real terms.

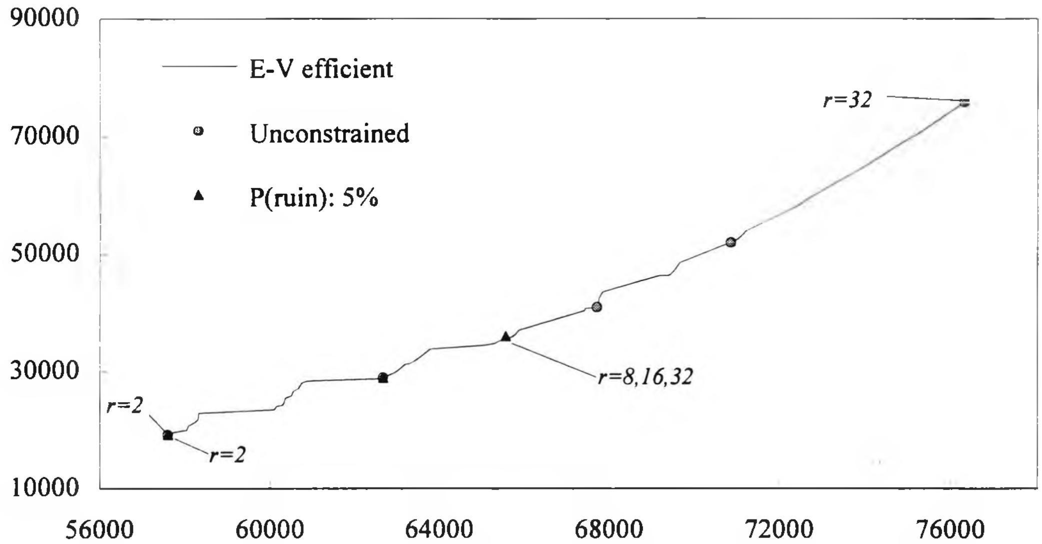


Figure 6.12. Graph of S.D. vs. Mean for *E-V* efficient and utility maximizing portfolios with and without ruin constraints: *sm* = 25%, in real terms.

Figures 6.11 and 6.12 above are equivalent to Figures 6.9 and 6.10 when payouts are expressed in real terms. Comparing these two pairs of graphs, the general shapes of the frontiers in Figures 6.11 and 6.12 are quite similar to their nominal counterparts, despite being somewhat more jagged. Although constraining the utility maximizing portfolios does not seem to have the effect of drawing these portfolios away from the efficient frontier for real payouts, it does seem to be more severe in terms concentrating these portfolios into the bottom left-hand-side of the frontier. This perhaps indicates the stronger relationship between the variability of payouts and the probability of ruin when payouts are measured in real terms, rather than nominal terms.

6.6.2 Scatter Plots of Standard Deviation vs. Mean

Most of the features noted from Figure 6.9 to 6.12 above may be examined further by looking at the scatter plots of standard deviation versus (vs.) mean of payouts for all the 286 portfolios sampled. In Figures 6.13 to 6.16 which follow, each dot represents a single portfolio. The small black dots represent the portfolios which have more than a 5% probability of ruin, while the white triangles represent those portfolios which do satisfy the 5% ruin constraint.

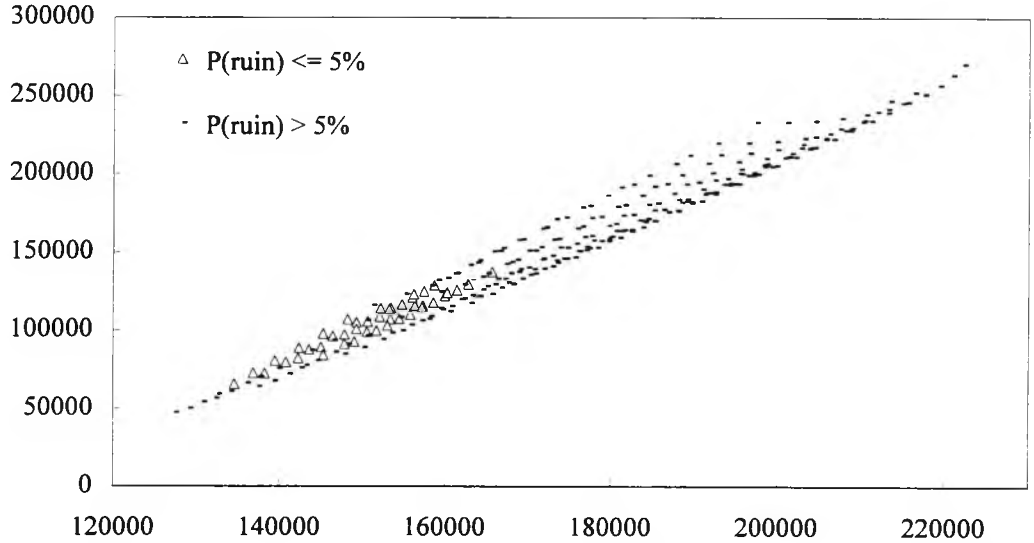


Figure 6.13. Graph of S.D. vs. Mean for all 286 portfolios: *sm* = 15%, in nominal terms.

Comparing Figures 6.13 and 6.14, the office with *sm* = 25% can be seen to possess a larger range of portfolios which satisfy the 5% ruin constraint. This demonstrates why the constrained portfolios from the office with more initial surplus are likely to be closer to the efficient frontier. In addition, the scatter plots appear quite narrow relative to the vertical. This is significant as it means that all possible portfolios are either on or are reasonably close to the mean-variance efficient frontier. In these circumstances, *E-V* efficient frontier analysis may not be of much assistance to the decision-maker. This is because the greater emphasis should then be placed on selecting the most appropriate portfolio from the mean-variance efficient frontier. Such a choice may be made with the aid of utility theory.

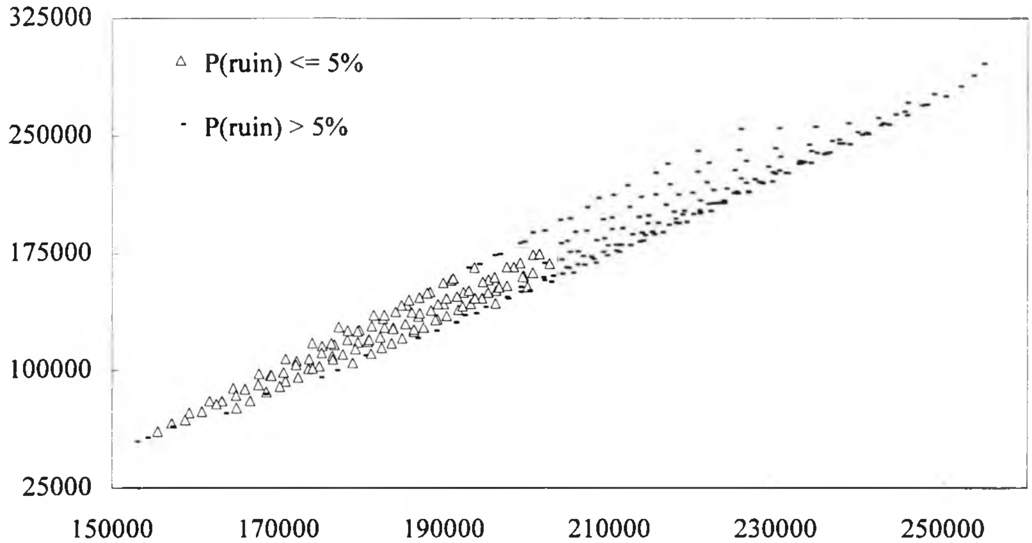


Figure 6.14. Graph of S.D. vs. Mean for all 286 portfolios: *sm* = 25%, in nominal terms.

In real terms, the patterns seen in Figures 6.15 and 6.16 are quite different to those in nominal terms (see Figures 6.13 and 6.14). The variation in the standard deviations of real payouts gradually becomes larger as the mean real payout reduces. Compared with the nominal case, it seems as though greater mileage could be gained from *E-V* analysis at low risk tolerance levels, with or without constraints. Portfolios which satisfy the 5% ruin constraint tend to be grouped near to the efficient frontier rather than away from it.

Another feature of these two graphs is that all the portfolios appear to be grouped into clumps, each being separated by a 10% difference in the proportion of equities held. For example, the portfolio with highest expected real payouts comprises 100% in equities and is symbolized by the solitary point on the extreme right-hand-side of the plots. The clump of three dots just below and to the left of this represents the three combinations of asset mixes with 90% of the assets invested in equities, and so on down to the largest clump which represents the combinations of asset mixes without any investment in equities. Hence, it appears that mean-variance efficiency can be achieved in such cases only by considering the proportions invested in cash, Consols and index-linked gilts. In other words, for a given proportion invested in equities, it should be possible to find a combination of the other three asset classes such that the resulting portfolio is *E-V* efficient.

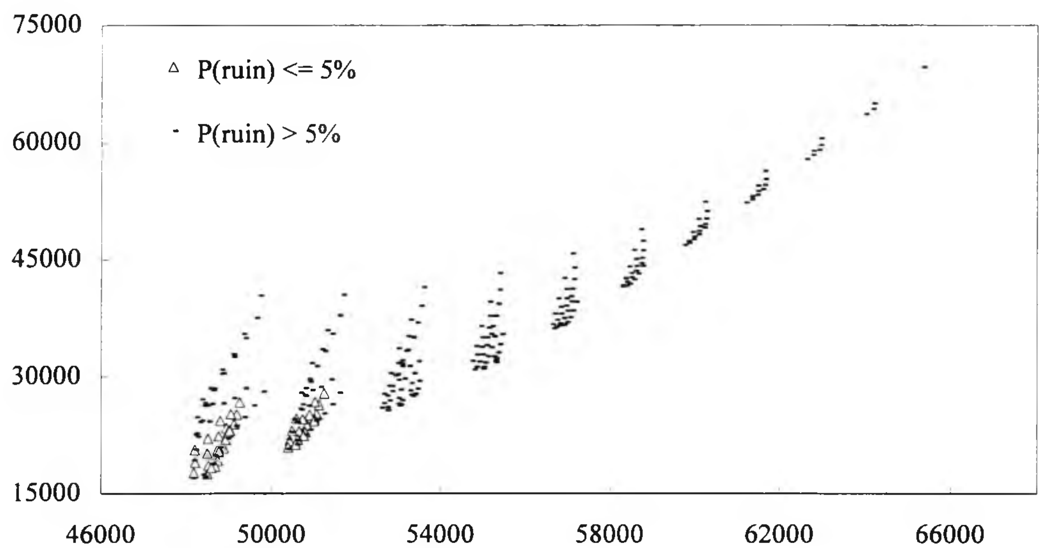


Figure 6.15. Graph of S.D. vs. Mean for all 286 portfolios: *sm* = 15%, in real terms.

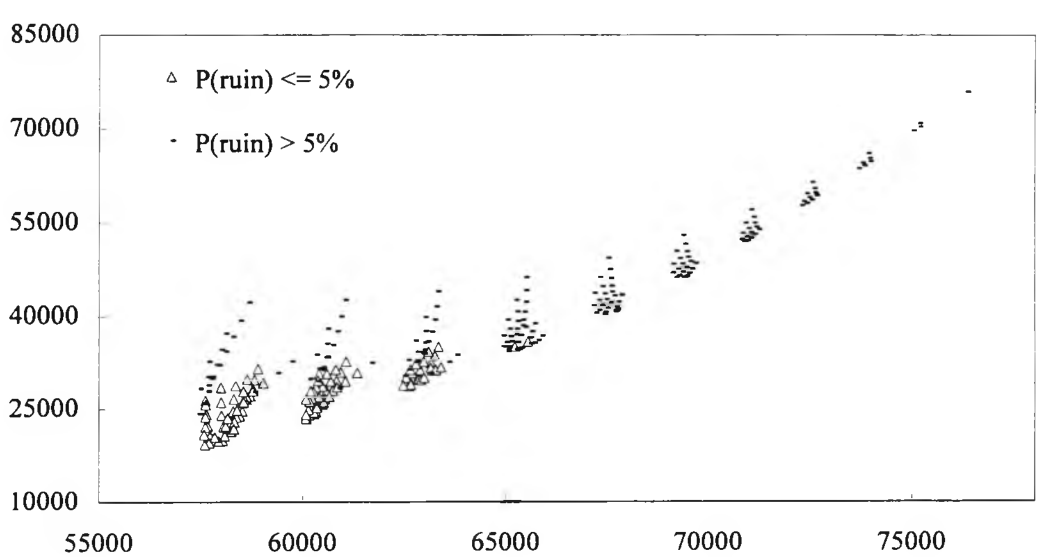


Figure 6.16. Graph of S.D. vs. Mean for all 286 portfolios: *sm* = 25%, in real terms

6.7 Summary

In this chapter, the optimal asset allocation decisions were derived for an ongoing office issuing twenty year endowments, producing some interesting insights into problems of this nature. At high levels of risk tolerance, the results were quite similar to those obtained for the pure asset fund (see Sections 4.4.2 and 4.4.4), with portfolios being predominantly invested in equities. But at lower risk tolerance levels, the presence of liabilities seemed to encourage greater proportions to be invested in Consols than had been the case in the asset fund. This difference was most notable when payouts were computed in real terms.

Compared with an immunized strategy using a realistic valuation basis, the duration of these optimal portfolios at lower values of r were generally found to be too long in relation to the liabilities, despite being shown to produce the most stable payouts. This highlighted the difference between the objective of minimizing the variance of accumulated dividends and immunization, which aims to stabilize the value of one-period surplus. Although there is a close relationship between an immunization and minimum ruin probabilities, the former is less general as it only applies in respect of one valuation basis, which must also be free from limits on the valuation rate of interest.

When the efficiency of numerical routines had been compared with more crude methods such as the grid approach, there did not appear to be a huge advantage in using the former, except for reasons of precision. Sensitivity analysis performed in respect of the investment model parameters showed that the results may be highly dependent on the assumptions made in the asset model, thus supporting the case for the grid approach even further. This more robust method also enabled the probabilities of ruin to be taken account of in the decision-making process. When ruin constraints were introduced into the utility maximization process, the optimal portfolios were more diverse and better matched as a result. However, the merits of ruin constraints were challenged by the more subjective process of selecting the most suitable portfolio from a representative sample of portfolios ranked by their expected utilities. Ruin constraints could sometimes be inflexible and too reliant on the accuracy of the investment model. Subjectivity was felt to be a limitation as well as a benefit of the 'top ten' approach.

From a mean-variance perspective, the utility maximizing portfolios were generally shown to lie on the efficient frontier, though this was not the case with a couple of portfolios which had been subject to ruin constraints. In some of the situations that were considered, *E-V* scatter plots revealed that all possible portfolios in steps of 10% were quite close to the mean-variance efficient portfolios. This strengthens the case for using utility theory to determine optimal investment portfolios because choosing between *E-V* efficient portfolios (which utility theory enables one to do) may well be more important than ensuring that a portfolio is *E-V* efficient, even if one assumes that a mean-variance approach is appropriate.

7. STATIC OPTIMIZATION III - LIABILITY MODEL B

7.1 Introduction

In Chapter 6, a number of investigations were carried out on Model A, for which negative dividends were permitted whenever an insolvency occurred. As the objective function was smooth, it enabled efficient gradient methods to be used in the numerical optimization process. The resulting optimal portfolios were clearly affected by the structure of the liabilities, which distinguished them from the portfolios obtained for the pure asset fund in Chapter 4. However, the impact of insolvency was barely noticeable in the derivation of these optimal asset mixes, requiring ruin constraints to yield more sensible portfolios. This was a direct consequence of assuming a smooth objective function in the decision model.

When the sensitivity of the optimal portfolios to the investment model had been analysed, it became apparent that the benefits of high precision optimization routines were not really merited under the circumstances. With only four asset classes involved, a simpler method, referred to as the grid approach, was found to work just well. In addition, the grid approach could readily handle ruin constraints and would allow the assumption of continuity in the objective function to be relaxed. Hence, it should also be ideal in the case of Model B.

Model B assumes that in the event of a technical insolvency, the office's liabilities are transferred at a cost equal to the realistic reserves at that time (see Section 5.3.4). Any surplus remaining is then distributed as a lump sum to shareholders which is assumed to be invested in equities together with all dividends previously distributed for the remainder of the twenty year projection period. This effectively means that there is a discontinuity on the objective function at the point of insolvency in Model B.

Throughout this chapter, the grid approach is used to investigate the optimal portfolios for Model B. As the payouts are guaranteed to be positive (due to limited liability of shareholders being strictly enforced), the logarithmic utility function will also be used. The office is assumed to be open to new business, as had been the case in Chapter 6. In the first instance, the liability profile considered for Model B will be identical to that investigated for Model A, i.e. involving only twenty year endowment assurances. After comparing the results from the two models, the liability profile will then be extended to include index-linked annuities and ten year endowments, in order to examine the effects of index-linked liabilities and shorter term contracts on the optimal asset allocations.

7.2 Twenty Year Endowments

Apart from the consequences of insolvency, Model B is projected using the same set of simulations and assumptions as in Model A. All possible asset mixes in steps of 10% are computed to obtain the optimal portfolios for initial surplus levels of 15% and 25%, with payouts being determined in both nominal and real terms. Sensitivity analysis is also performed on some of the unconstrained optimal portfolios. Optimal portfolios subject to a ruin constraint and the top ten portfolios are then investigated. Finally, scatter plots of standard deviation versus mean of payouts are shown.

7.2.1 Unconstrained Optima

Figure 7.1 represents the optimal asset mixes to the nearest 10% for five different values of r using the exponential utility function, and the log function. The horizontal axis is ordered from left to right by increasing expected payouts (which is equivalent to increasing tolerance to risk). This ordering serves to highlight the position of the log function in relation to other risk tolerance parameters. The vertical axis on the right hand side of the graph relates to the ruin probability of the optimal portfolios. The vertical axis on the left hand side relates to the proportions in each asset class as in Chapters 4 and 6. For example, at r = 16, the optimal portfolio comprises 30% in cash, 40% in Consols, 10% in index-linked gilts and 20% in equities. As it has a higher expected payout to the optimal mix under the log function, it is shown to the right of this mix. Just under 8% of the scenarios tested for this portfolio lead to insolvency within the twenty year projection period.

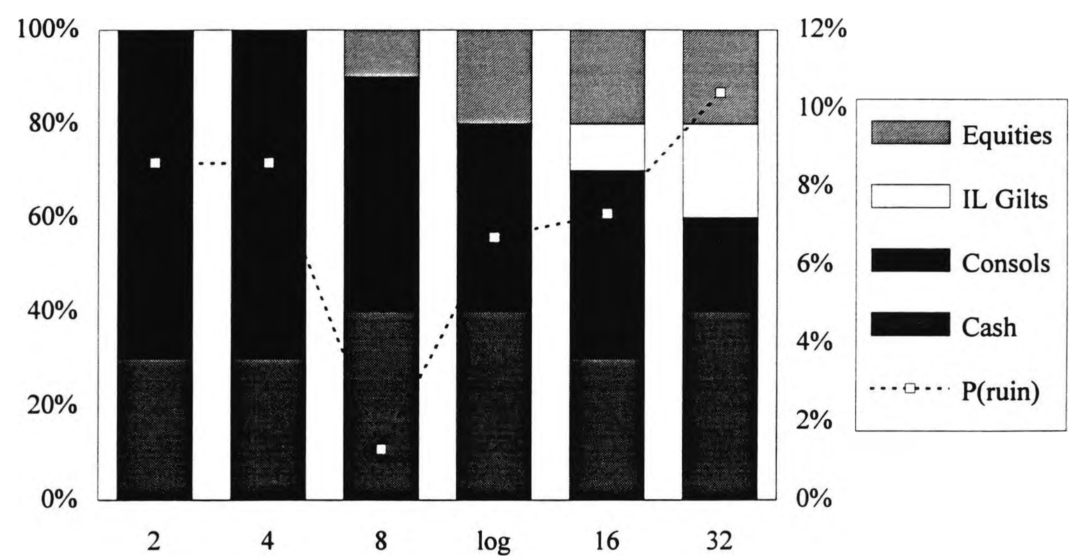


Figure 7.1. Optimal portfolios: sm = 15%, in nominal terms.

In isolation, the asset mixes in Figure 7.1 seem entirely sensible. At the lowest levels of risk tolerance, the liabilities are very closely matched with all the assets being held in fixed interest investments. The 30-70 split between cash and Consols is also well within the range of immunized portfolios described in Section 6.2.2. (On a realistic valuation basis, the immunized ratio is 2 : 7 in cash and Consols, with free assets making up 25% of the total assets.) As the level of risk tolerance increases, the proportion in Consols reduces, only to be taken up by more index-linked gilts and equities. This is to be expected given the ordering of the asset classes by mean accumulation over the twenty year period (see Table 3.9).

In comparison with Model A, the portfolios for Model B shown above appear to be much more restrained from a solvency perspective. With Model A (see Figure 6.1), the optimal mix at $r = 2$ consisted entirely of Consols, giving a ruin probability of 62%. This contrasts with the 30-70 split between cash and Consols seen here which only has a 9% ruin probability. The trend in lower ruin probabilities for Model B persists throughout the entire range of risk tolerance and is even more striking at higher values of r. When $r = 32$, the diverse mix for Model B has an 11% ruin probability, which is much lower than the 100% ruin probability seen in the optimal mix for Model A.

It is important to keep in mind that the probability of ruin for a given asset mix will be identical whether Model A or B is being used. The two models only differ in the treatment of the situation *after* an insolvency occurs. Hence, the only way in which the ruin probabilities may be altered between these two models is by actually changing the optimal asset mixes themselves, which has happened above.

The preference for more solvent strategies in Model B is a direct result of introducing a discontinuity on the payout at the point of insolvency. However, as the discontinuity is implicit, it may be difficult to predict the direction and extent to which each payout is shifted. This will depend on the circumstances prevailing at each time epoch, in each simulation. But as the optimal solutions have tended to move towards portfolios with fewer insolvencies, this implies that the discontinuity in the objective function must generally lead payouts being reduced in the event of an insolvency, i.e. a penalty is being imposed on insolvency. Hence, it may be inferred that it is to the benefit of shareholders if the office remains a going concern, which is intuitively reasonable.

Similarly, the results for an initial surplus of 25% are summarized in Figure 7.2. Due to the lower ruin probabilities experienced by this office, the impact of the discontinuities on the payouts should be less than they were with *sm* = 15%. This is reflected in the optimal portfolios obtained. More Consols are permissible at the lower risk tolerance levels, despite their tendency to increase the probability of ruin. At $r = 32$, the optimal proportion in real assets is now 60%, compared with 40% when the initial solvency margin was 15%. This occurs even though the ruin probability for the former portfolio with *sm* = 25% is greater than that of the latter portfolio with *sm* = 15%.

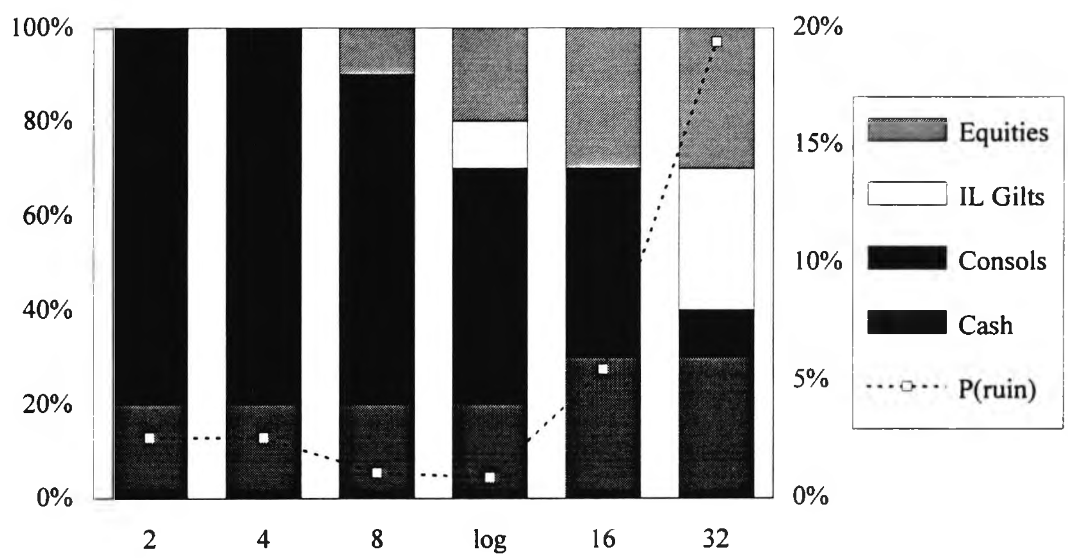


Figure 7.2. Optimal portfolios: *sm* = 25%, in nominal terms.

Figure 7.3 shows the optimal asset mixes when payouts are measured in real terms and *sm* = 15%. The portfolios at lower values of r are quite similar to those shown for Model A (see Figure 6.3), with 30% of total assets being invested between the real asset classes of index-linked gilts and equities. This is consistent with the notion of investing the solvency margin in real assets to stabilize real surpluses, whilst broadly matching the remaining assets to the fixed liabilities, which is the case here. The more diverse portfolios that result from such a strategy also contribute towards lowering the ruin probabilities obtained. This in turn means that less penalties are imposed on the payouts due to insolvencies. However, as *r* increases, the optimal portfolios in the Models A and B diverge quite radically since maintaining reasonably low ruin probabilities is of great importance in Model B. At $r = 32$, the 100% equities portfolio for Model A leads to almost certain ruin. This contrasts with the more cautious portfolio of 80% in fixed interest assets and 20% in equities seen in Figure 7.3.

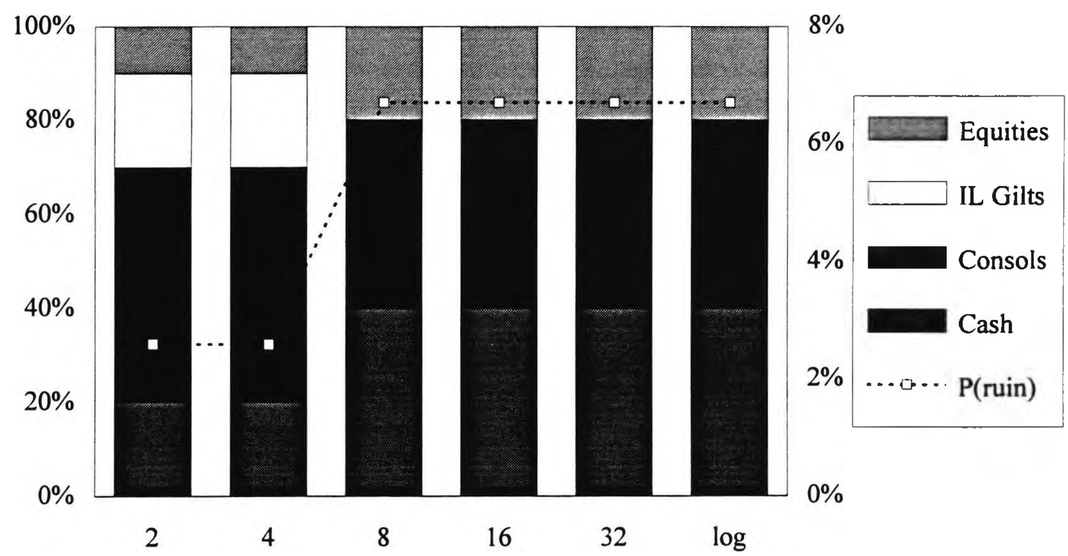


Figure 7.3. Optimal portfolios: sm = 15%, in real terms.

When the office with an initial surplus of 25% is considered in real terms, the pattern of results shown in Figure 7.4 does not differ much from that seen when sm = 15%. Index-linked gilts are only significant at lower levels of risk tolerance and gradually drop out from the optimal portfolios, mainly being replaced by equities, when more risk can be tolerated. Compared with Figure 7.3, equities are generally found in higher proportions which is intuitively reasonable as the higher level of surplus also means that the office is less likely to become technically insolvent for a given asset mix. The importance of avoiding excessive insolvencies in Model B is evident by the low ruin probabilities seen in Figure 7.4.

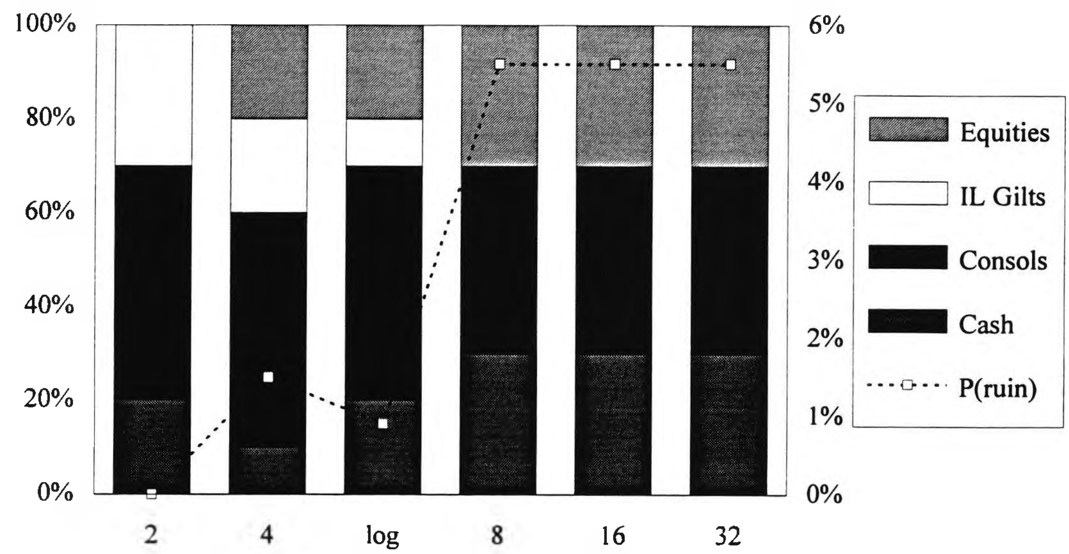


Figure 7.4. Optimal portfolios: $sm = 25\%$, in real terms.

7.2.2 Sensitivity Analysis

Before proceeding any further, it may be worth analysing the sensitivity of the optimal portfolios for Model B to different investment assumptions. In order to be consistent with the results obtained in Section 6.4 earlier, only the portfolio at $r = 8$ with $sm = 15\%$ will be considered here. It may also be recalled that of the five sets of parameter alterations made in Wilkie's model, the version which produced the greatest change in results had been 'Wilkie (1995a)'. Therefore, it should suffice just to consider the sensitivity of the Standard Basis with Wilkie's later version.

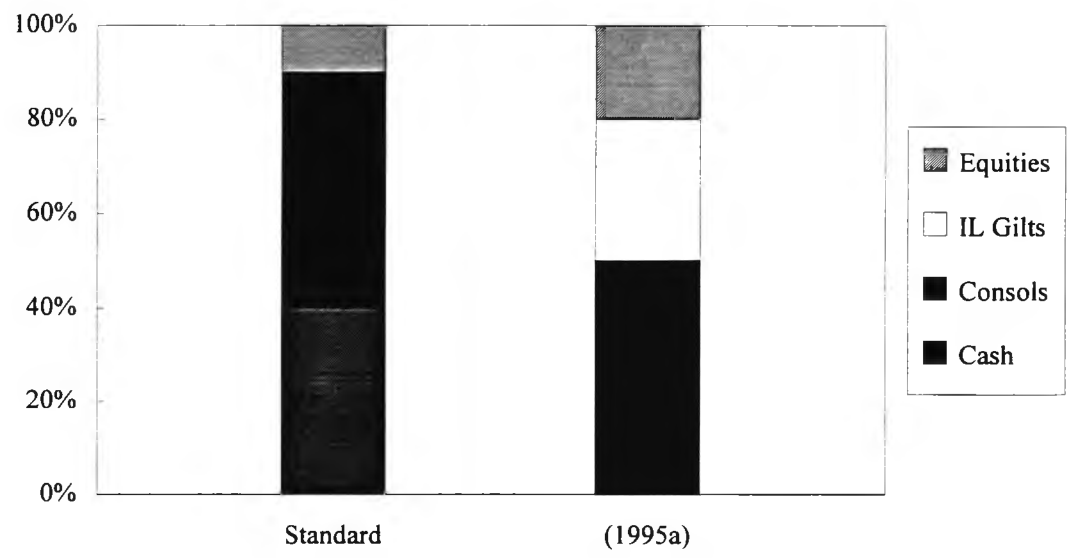


Figure 7.5. Optimal portfolios for two versions of Wilkie's model: sm = 15%, in nominal terms, at r = 8.

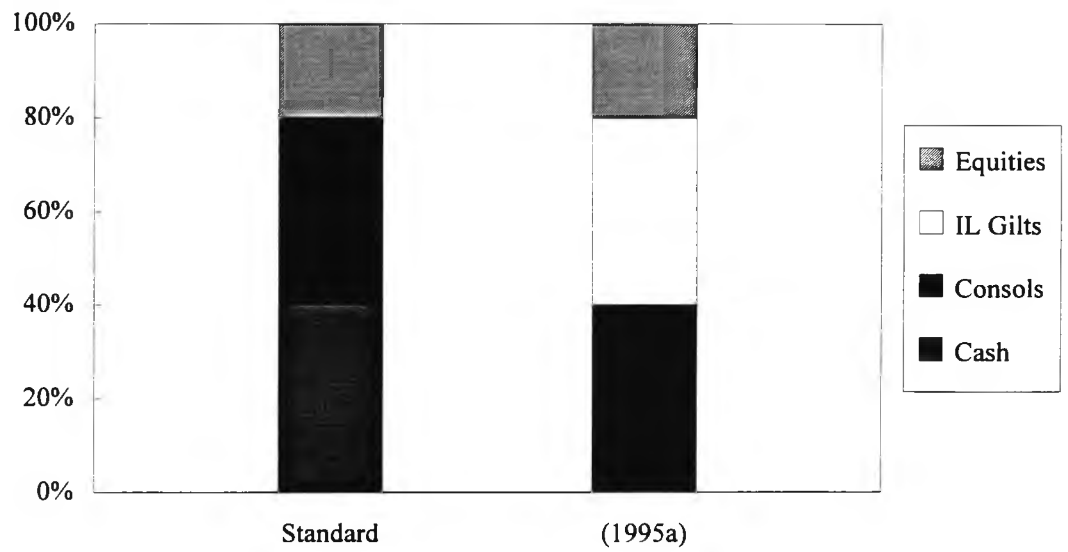


Figure 7.6. Optimal portfolios for two versions of Wilkie's model: sm = 15%, in real terms, at r = 8.

Figures 7.5 and 7.6 show the optimal asset mixes for the two investment models in the case of Model B, with payouts expressed in nominal and real terms respectively. The main difference seen when changing from the standard model to the updated model is the shift from cash into the other three asset classes. This is understandable given that mean cash accumulations over the twenty year period are nearly 2% per annum lower in the latter version. From Section 3.5, it may be recalled that the differences between the two cash models may be attributed to the different data periods used when fitting the models. However, it still means that the investor needs to decide which data set would be more appropriate for predicting cash yields in the future.

Comparing the two liability models, it would appear that Model B is less sensitive to the parameter values in Wilkie's model than Model A. For example, in the case of Model A, the proportions in equities increased from 15% to 54% when payouts were measured in nominal terms, and from 37% to 100% when payouts were expressed in real terms. This occurred despite the fact that the proportions held in cash had only been 5% and 16% respectively. Therefore, these changes must have been due to the increase of 1% in the mean accumulation rate for equities from 'Wilkie (1995a)'. But in Model B, equity proportions only rose by 10% for the nominal case and remained unchanged for the real case, even though a large vacuum had been created by the departure of cash from the optimal portfolios.

This is probably related to the negative effect which insolvency has on the objective function in Model B. Given this feature, it is reasonable that Model B will try to avoid asset mixes which result in very high ruin probabilities. However, it would appear that the ruin probability has not been significantly reduced as a result of the 1% increase in the mean return on equities. Hence, the optimal portfolios in Model B will also be less attracted to these higher yielding equities.

There is a possibility that the use of different model structures or distributions of residual components may bring about increased sensitivity in results produced from Model B. However, due to the very limited availability of investment models in the public domain, a comprehensive sensitivity analysis may not feasible at present.

7.2.3 Ruin Constrained Optima

Depending on the risk preferences of the insurer, it may be desirable that all asset mixes considered meet a strict ruin constraint. As the ruin probabilities in the unconstrained optimal mixes obtained so far are already fairly close to the 5% level, a 1% constraint has been adopted to heighten its impact.

The optimal asset mixes subject to a 1% ruin constraint with *sm* = 15% and payouts expressed in nominal terms are given in Figure 7.7. Despite this stringent restriction, the effect of the ruin constraint in relation to Figure 7.1 has only been to reduce the maximum proportions in Consols from 70% to 50% and equities from 20% to 10%. With *sm* = 25%, as shown by Figure 7.8, the maximum proportions in equities and index-linked gilts were lowered from 30% to 20% and from 80% to 70% in the case of Consols. These relatively minor changes in the asset mixes are a consequence of the unconstrained optimal portfolios already having implicit penalties imposed in the event of ruin.

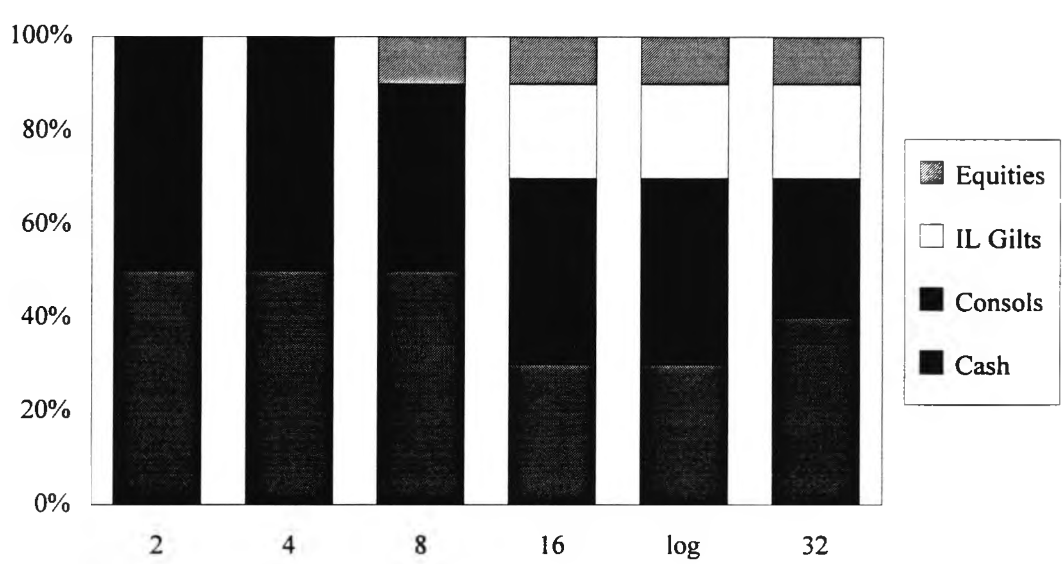


Figure 7.7. Optimal portfolios with a 1% ruin constraint: *sm* = 15%, in nominal terms.

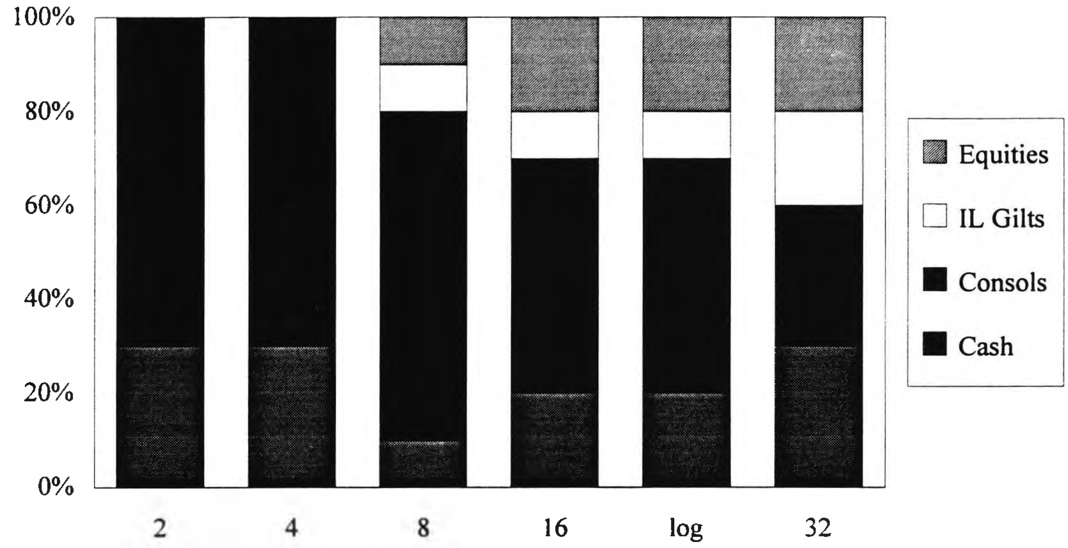


Figure 7.8. Optimal portfolios with a 1% ruin constraint: *sm* = 25%, in nominal terms.

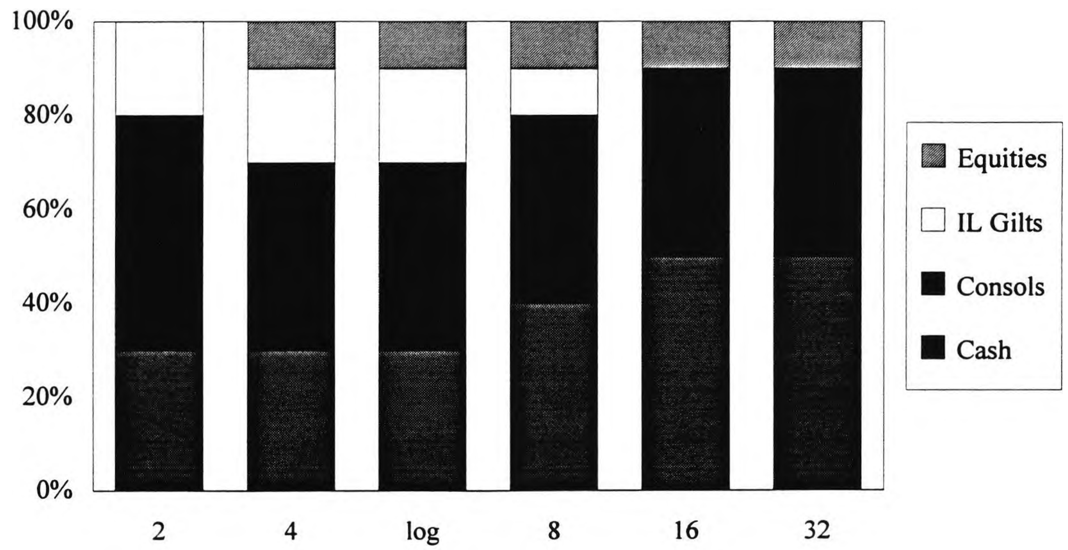


Figure 7.9. Optimal portfolios with a 1% ruin constraint: *sm* = 15%, in real terms.

In Figure 7.9, where *sm* = 15% and payouts are measured in real terms, constraining the ruin probability to be no greater than 1% generally appears to have reduced the proportions of equities permitted, whilst increasing the proportions held in index-linked gilts for the medium risk portfolios, i.e. r = 8 and the log function. The ruin probability is also improved by investing about 10% more in cash at most levels of risk tolerance.

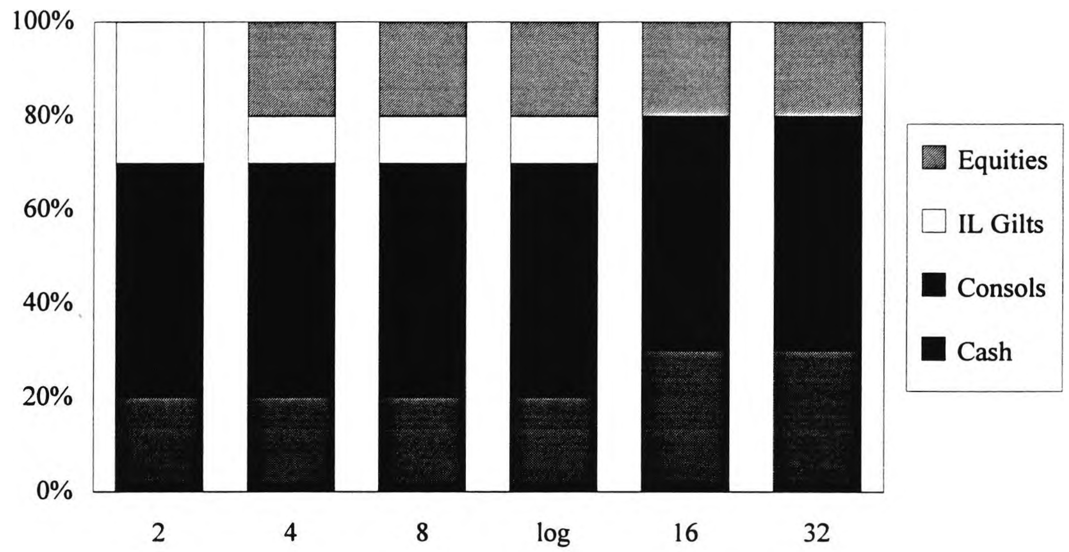


Figure 7.10. Optimal portfolios with a 1% ruin constraint: sm = 25%, in real terms.

When the initial solvency margin is increased to 25%, the resulting constrained optimal portfolios are shown in Figure 7.10 above. As the unconstrained optimal mixes for the log utility function and the exponential function at r = 2 already satisfy the 1% ruin criterion, they are naturally unaffected by this constraint. For all other portfolios, the ruin constraint is satisfied by limiting the proportion invested in index-linked gilts and equities to 10% and 20% respectively. This leaves between 70-80% of the remaining assets to be invested in fixed interest assets, which is a suitable match to the office's fixed liabilities.

7.2.4 Top Ten Portfolios

For reasons given in Chapter 6, it may not always be sensible to disregard a portfolio simply because it fails a particular ruin criterion. From a selection of portfolios, one may be able to choose a portfolio which represents a better compromise between shareholders' utility and ruin probability. Hence, some consideration will be given here to the top ten utility maximizing portfolios for Model B.

As shown in Table 7.1, the optimal mix at $r = 2$ produces a ruin probability of 8.6%. The second highest ranking portfolio on the other hand, with 10% more cash and 10% less Consols, gives a mere 1.9% ruin probability. From a traditional mean-variance perspective, the latter portfolio would actually appear more risky to shareholders. However, it is arguable whether the lower variability of payouts in the former justifies the lower expected payouts and the higher ruin probability that results from it.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	30	70	0	0	-0.01239	131130	60976	8.6
2	40	60	0	0	-0.01266	134091	65920	1.9
3	20	70	10	0	-0.01332	133591	67090	8.5
4	30	60	10	0	-0.01344	137547	72590	1.9
5	20	80	0	0	-0.01353	125773	58935	25.0
6	50	50	0	0	-0.01450	136826	72813	0.3
7	20	70	0	10	-0.01474	138210	74327	14.2
8	30	60	0	10	-0.01480	141955	77662	5.6
9	10	80	10	0	-0.01496	127973	62094	20.8
10	40	50	10	0	-0.01531	140747	79661	0.2

Table 7.1. Top ten portfolios: *sm* = 15%, in nominal terms, at $r = 2$.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	40	50	0	10	-0.23345	145468	84206	1.3
2	30	60	0	10	-0.23407	141955	77662	5.6
3	30	50	10	10	-0.23439	149052	92802	1.2
4	20	60	10	10	-0.23474	145506	86272	5.3
5	50	40	0	10	-0.23671	148040	91398	0.3
6	40	40	10	10	-0.23721	151765	99748	0.2
7	20	50	20	10	-0.23779	152427	103045	2.6
8	30	60	10	0	-0.23804	137547	72590	1.9
9	20	70	0	10	-0.23826	138210	74327	14.2
10	40	40	0	20	-0.23863	156042	106788	6.7

Table 7.2. Top ten portfolios: *sm* = 15%, in nominal terms, at $r = 8$.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	40	20	20	20	-0.63668	167872	141158	10.4
2	30	20	30	20	-0.63678	171338	154603	19.2
3	50	10	20	20	-0.63718	169981	149183	15.8
4	30	30	20	20	-0.63757	165296	133681	9.4
5	20	30	30	20	-0.63759	168817	147386	17.0
6	50	20	10	20	-0.63823	164373	130303	7.6
7	40	10	30	20	-0.63846	172914	162529	25.1
8	40	30	10	20	-0.63871	162064	122778	5.7
9	60	10	10	20	-0.63940	166035	138322	12.5
10	20	40	20	20	-0.63944	162434	126727	10.2

Table 7.3. Top ten portfolios: *sm* = 15%, in nominal terms, at $r = 32$.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	40	40	0	20	11.74338	156042	106788	6.7
2	30	50	10	10	11.74239	149052	92802	1.2
3	30	40	10	20	11.74137	159466	115949	7.3
4	20	50	20	10	11.74074	152427	103045	2.6
5	30	40	20	10	11.74044	155758	110063	0.8
6	40	40	10	10	11.73960	151765	99748	0.2
7	40	50	0	10	11.73689	145468	84206	1.3
8	20	50	10	20	11.73558	155606	109681	12.2
9	20	60	10	10	11.73388	145506	86272	5.3
10	30	50	0	20	11.73359	151899	100368	13.1

Table 7.4. Top ten portfolios: sm = 15%, in nominal terms, with the log function.

Table 7.2 is an example of how the strict application of constraints may not yield the desired effect when the grid approach is being used. For practical purposes, it should be insignificant whether a portfolio is insolvent 10 or 13 times out of a thousand scenarios. But the 1% constraint discards the utility maximizing portfolio in favour of the fifth ranking portfolio. In the case of Table 7.3, the ruin probabilities for the ten portfolios shown are of a similar order and so little benefit is to be gained by choosing a sub-optimal portfolio. However, it would appear from Table 7.4 that there is ample scope for reducing the ruin probability when the log utility function is used.

Tables 7.5 to 7.8 show the ten asset mixes with the highest rankings when payouts are measured in real terms. In Table 7.5, where r = 2, a 2.6% ruin probability for the optimal mix may seem adequate. Depending on the affinity for low ruin probabilities, the third ranking portfolio which has a 0.5% chance of insolvency may be obtained by shifting 10% from equities into cash. Switching another 10% from Consols into cash reduces the observed probability of ruin to zero.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	20	50	20	10	-0.11945	50298	21206	2.6
2	20	50	30	0	-0.11948	48187	17458	1.9
3	30	50	20	0	-0.11983	48648	18542	0.5
4	30	50	10	10	-0.12068	50736	22472	1.2
5	30	40	20	10	-0.12099	50601	21953	0.8
6	10	50	30	10	-0.12163	49910	21018	4.9
7	20	60	20	0	-0.12204	48107	18257	3.4
8	10	60	20	10	-0.12207	49812	21302	7.7
9	40	40	20	0	-0.12223	48742	19261	0.0
10	40	40	10	10	-0.12231	50813	22880	0.2

Table 7.5. Top ten portfolios: *sm* = 15%, in real terms, at $r = 2$.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	40	40	0	20	-0.54366	52900	27874	6.7
2	30	40	10	20	-0.54443	52528	26805	7.3
3	50	30	0	20	-0.54487	52918	28504	5.6
4	40	30	10	20	-0.54583	52556	27561	5.7
5	30	50	0	20	-0.54699	52398	27895	13.1
6	20	50	10	20	-0.54709	52039	26510	12.2
7	60	20	0	20	-0.54749	52853	29524	7.0
8	20	40	20	20	-0.54809	51910	26404	10.2
9	50	20	10	20	-0.54888	52469	28712	7.6
10	30	30	20	20	-0.54913	52070	27415	9.4

Table 7.6. Top ten portfolios: *sm* = 15%, in real terms, at $r = 8$.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	40	40	0	20	-0.85071	52900	27874	6.7
2	50	30	0	20	-0.85081	52918	28504	5.6
3	60	20	0	20	-0.85122	52853	29524	7.0
4	30	40	10	20	-0.85148	52528	26805	7.3
5	40	30	10	20	-0.85158	52556	27561	5.7
6	70	10	0	20	-0.85199	52689	30929	10.3
7	30	50	0	20	-0.85205	52398	27895	13.1
8	50	20	10	20	-0.85207	52469	28712	7.6
9	20	50	10	20	-0.85272	52039	26510	12.2
10	30	30	20	20	-0.85284	52070	27415	9.4

Table 7.7. Top ten portfolios: sm = 15%, in real terms, at r = 32.

Rank	*CASH*	*CON*	*ILG*	*EQ*	*E(utility)*	*Mean*	*S.D.*	*Ruin %*
1	40	40	0	20	10.74781	52900	27874	6.7
2	30	50	10	10	10.74682	50736	22472	1.2
3	30	40	10	20	10.74580	52528	26805	7.3
4	20	50	20	10	10.74517	50298	21206	2.6
5	30	40	20	10	10.74487	50601	21953	0.8
6	40	40	10	10	10.74403	50813	22880	0.2
7	40	50	0	10	10.74132	51036	24287	1.3
8	20	50	10	20	10.74001	52039	26510	12.2
9	20	60	10	10	10.73831	50315	22403	5.3
10	30	50	0	20	10.73802	52398	27895	13.1

Table 7.8. Top ten portfolios: sm = 15%, in real terms, with the log function.

The higher risk strategies in Tables 7.6, 7.7 and 7.8, relating to $r = 8$, $r = 32$ and the log function respectively, have the same optimal portfolio despite there being differences in the lower ranking asset mixes. From the overall level of standard deviations of payouts amongst the three tables shown, it may be inferred that the log function represents a lower risk tolerance level than the exponential function when $r = 8$. Unlike the cases where r is 8 or 32, Table 7.8 offers some choice between the low ruin portfolios. In contrast with Model A (see Table 6.20), it is apparent that the optimal portfolio for Model B at $r = 32$ does not represent the ultimate high risk tolerance strategy. There are portfolios which yield higher expected real payouts than the utility maximizing one.

7.2.5 Scatter Plots of Standard Deviation vs. Mean

Figure 7.11 below shows a plot of the standard deviation versus the mean of nominal payouts for all the 286 portfolios tested, with sm = 15%. Portfolios satisfying a 5% ruin constraint are highlighted by the white triangles as in Section 6.6.2.

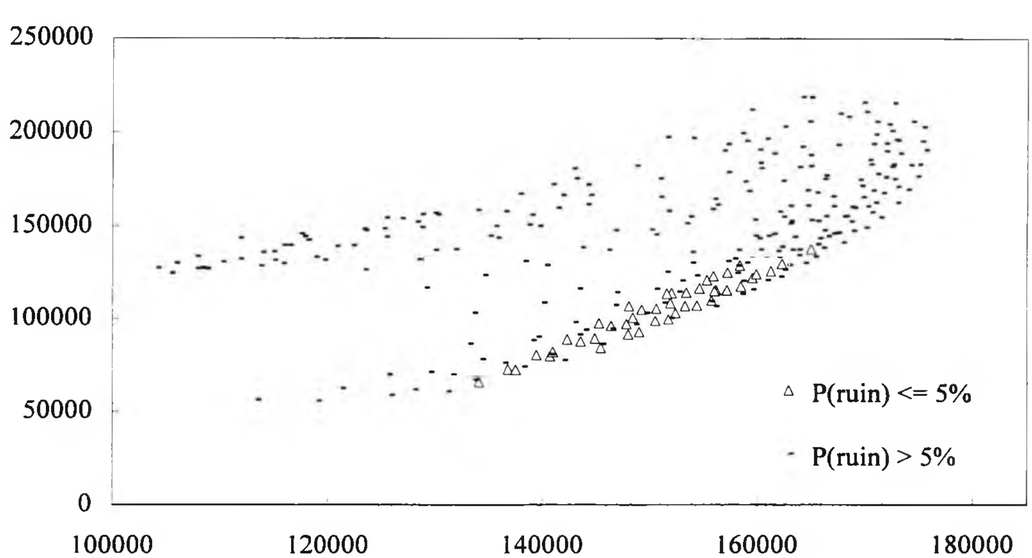


Figure 7.11. Graph of S.D. vs. Mean for all 286 portfolios: sm = 15%, in nominal terms.

One may recognize Figure 7.11 as being essentially similar to the almost linear plot of Model A (see Figure 6.13), but with the upper half of the plot folded back upon the lower half. The extreme point on the lower edge has a mean of about 113,000 and represents 100% in Consols. The extreme point on the upper edge represents 100% in equities. This has a mean of about 104,000. The points in this upper section mainly comprise portfolios with substantial proportions in real assets and are clearly inefficient compared with those below them. From this plot, it is also clear that portfolios which satisfy the 5% ruin criterion are either on or very close to the efficient frontier.

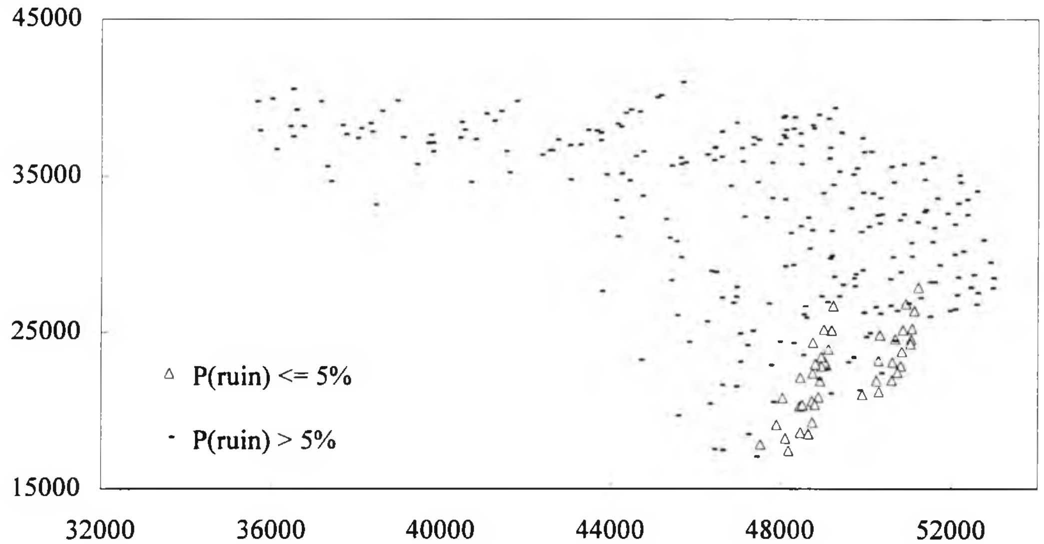


Figure 7.12. Graph of S.D. vs. Mean for all 286 portfolios: *sm* = 15%, in real terms.

Figure 7.12 shows all the asset mix vectors when payouts are expressed in real terms. The vast majority of points appear grossly inefficient. Most of the efficient portfolios also satisfy a 5% ruin constraint: it would seem that the disutility of going insolvent is relatively greater when payouts are measured in real terms as opposed to nominal terms. The portfolio furthest to the left in Figure 7.12 comprises 100% in equities. Those marked by the white triangles are the same subset of portfolios as their nominal counterparts shown in Figure 7.11.

7.2.6 The Dynamics of Insolvency

The only difference between Models A and B is that in the case of the latter, the office is wound up at the point of insolvency. As the office is fairly profitable, one penalty for becoming insolvent is the loss of profits that would have been earned had the office remained in business. Hence, it is not really surprising when the asset mixes which give the best payouts are those which lead to the office staying solvent most of the time. Nevertheless, the plots in Section 7.2.5 perhaps deserves a more detailed explanation.

For simplicity consider the effects of these four asset mixes: 100% investment in each asset class. (It may be assumed that the proportion of inflationary liabilities present is not sufficiently large to be the sole cause for any insolvencies which may occur.) With all assets in cash, the ruin probability is fairly small because asset values are very stable and insolvency really only occurs due to a dramatic or prolonged reduction in cash yields. The low frequency of insolvency means that this asset mix produces very similar payouts to that of Model A.

In the case where all assets are invested in Consols, the probability of ruin is high because the office is very vulnerable to a rise in yields. Reasons for this have been outlined in Section 6.2.2. Although the high frequency of insolvencies reduces the average payouts, the adverse effect is dampened by the fact that the circumstances under which the company will be wound up are generally not unfavourable. The cost at which the liabilities will be transferred is calculated using the current yield on Consols. Hence, despite the fall in asset values due to a rise in yields, the cost of transferring the liabilities will also fall. The only penalty is the loss of future profits.

The asset classes which are most affected by the winding up of the company are index-linked gilts and equities, due to the extremely high ruin probabilities associated with them. One reason for the almost certain ruin in both cases is the very low yields these asset classes possess. In fact, this is so pronounced in the case of index-linked gilts that the statutory surplus at the start of the projection period with *sm* = 15% barely covers the guarantee fund. Despite this disadvantageous start, its position is relatively strong in comparison with equities because of the stable real yields assumed in the asset model. Furthermore, since the office declares no dividends while the statutory minimum solvency margin is breached, the fund should be able to strengthen its solvency margin, providing it actually manages to remain solvent during the first few years.

In the event of insolvency, the amount of final surplus released at that point may be quite substantial. This is because of the difference between real yields and Consols yields. However, as insolvencies tend to occur quite early on in the projection period, the resulting loss of profits is significant. Hence, the net effect is still a reduction in the mean and standard deviation of payouts compared with that of Model A.

With equity portfolios, the extremely high probabilities of ruin are caused by both low and volatile dividend yields. As with the index-linked gilts portfolio, insolvencies tend to occur early on in the projection period, leading to a considerable loss of profits. But due to the high volatility of equity prices, insolvencies sometimes take place when asset prices are significantly depressed. Often in these situations, there is no final surplus available for distribution after the liabilities have been sold. Hence, this further reduces the average payout that emerges from an equity portfolio.

Therefore, the preference in Model B would be for portfolios which comprise a substantial proportion of fixed-income assets in order to maintain solvency. For a low risk strategy, the proportion in cash may be higher than that justified by the usual duration analysis because of the vulnerability of liabilities backed by Consols to high yields, induced by the ceiling on the valuation rate of interest. At high levels of risk tolerance, the concept of matching is less important, although fixed income assets are still necessary to ensure a sufficiently high valuation rate of interest.

7.3 Mix of 20 Year Endowments and 20 Year Index-Linked Annuities

This section considers Model B when a mix of twenty year index-linked annuities and twenty year endowments are being issued. It is assumed that even proportions (by numbers) of index-linked annuity contracts and endowment contracts are in force at all times. However, as the former are larger contracts in terms of amounts, the net effect is an office with predominantly index-linked liabilities on its books. In terms of the prudent valuation at the start of the projection period, the ratio of the reserves held between annuities and endowments is about 5 : 2. Analysing this liability profile will enable the robustness of the methodology to be checked as the optimal portfolios should include a greater proportion of real assets at low levels of risk tolerance.

7.3.1 Unconstrained and Ruin Constrained Optima

Figures 7.13 and 7.14 represent the unconstrained and constrained optimal asset mixes respectively for such a fund with payouts calculated in nominal terms. In Figure 7.13, the unconstrained portfolios at $r = 2$ and 4 consist of 10% in cash, 30% in Consols and 60% in index-linked gilts, which seems to be a reasonably well-matched position. Switching 10% of the assets from Consols into equities appears to reduce the probability of ruin, as seen by the optimal portfolio at $r = 8$. A further reduction in ruin probability is also observed at $r = 16$, when the 10% of assets held in cash is switched for index-linked gilts. However, as greater proportions are invested in equities, the probability of ruin may be seen to rise quite sharply. Introducing a ruin constraint of 1%, though, does not alter the optimal asset mixes significantly (see Figure 7.14), as the unconstrained portfolios are already close to satisfying the constraint. At $r = 2$ and 4, reallocating 10% of the assets from Consols into index-linked gilts appears to produce sufficiently low ruin probabilities, despite having the effect of increasing the variability of payouts. At higher risk tolerance levels, a reduction in the proportion of equities is generally needed to maintain the 1% ruin constraint. Nevertheless, it is still possible for portfolios with 10% invested in equities to satisfy this constraint.

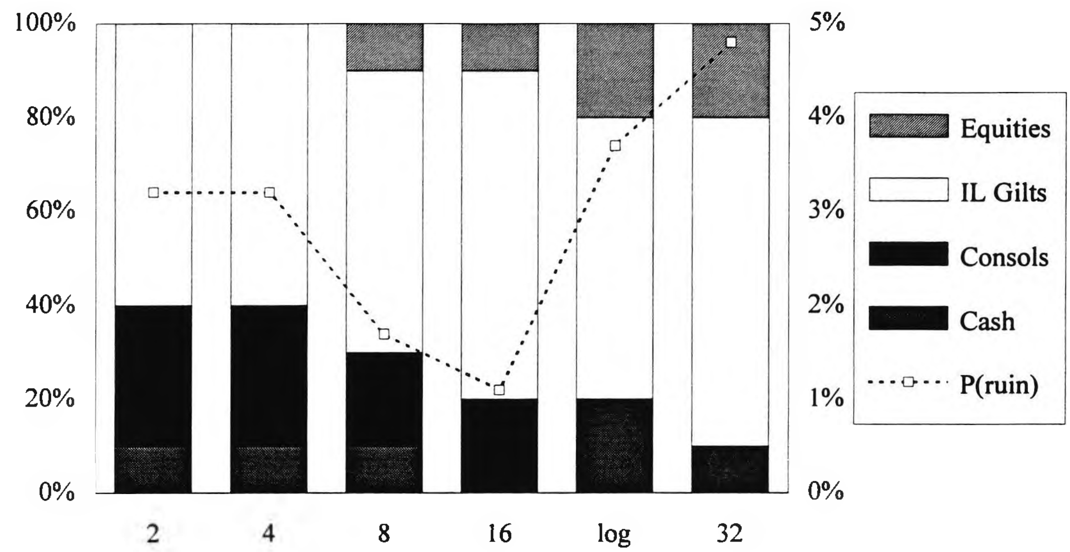


Figure 7.13. Optimal portfolios: a mix of 20 year endowments and 20 year index-linked annuities; *sm* = 15%, in nominal terms.

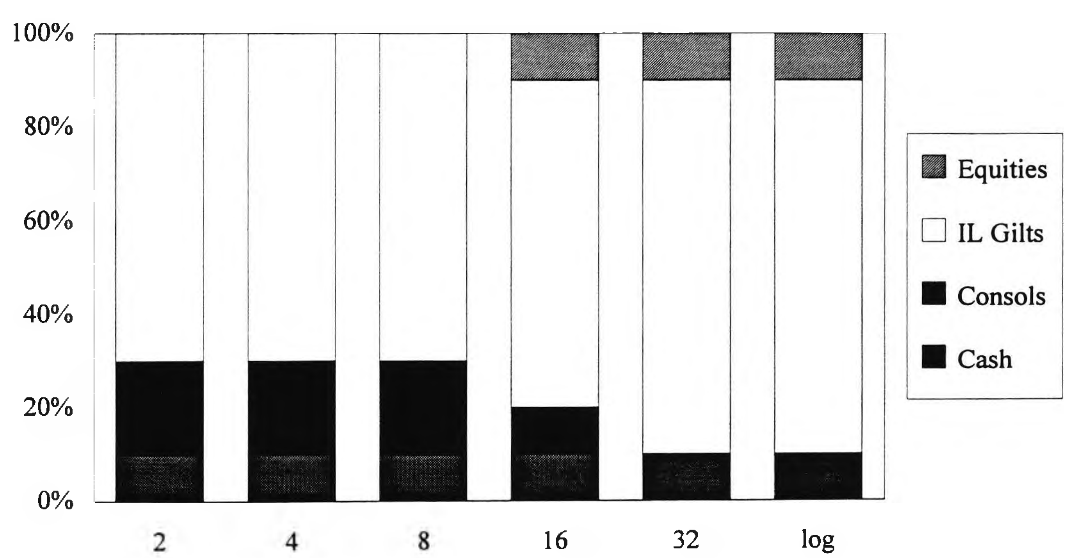


Figure 7.14. Optimal portfolios with a 1% ruin constraint: a mix of 20 year endowments and 20 year index-linked annuities; *sm* = 15%, in nominal terms.

Figures 7.15 and 7.16 summarize the results obtained when payouts are calculated in real terms. For the lowest risk strategy, 90% of the fund are in index-linked gilts with the remainder being in Consols. From the nominal case (see Figure 7.14), it became apparent that a portfolio largely invested in index-linked gilts could produce very low ruin probabilities. But where the unconstrained portfolios were concerned, the need to maintain stable nominal payouts encouraged sizeable proportions of the lower risk portfolios to be invested in fixed income assets. When dealing with real payouts, however, it may be reasonable to expect higher proportions of index-linked gilts to be found in the low risk strategies.

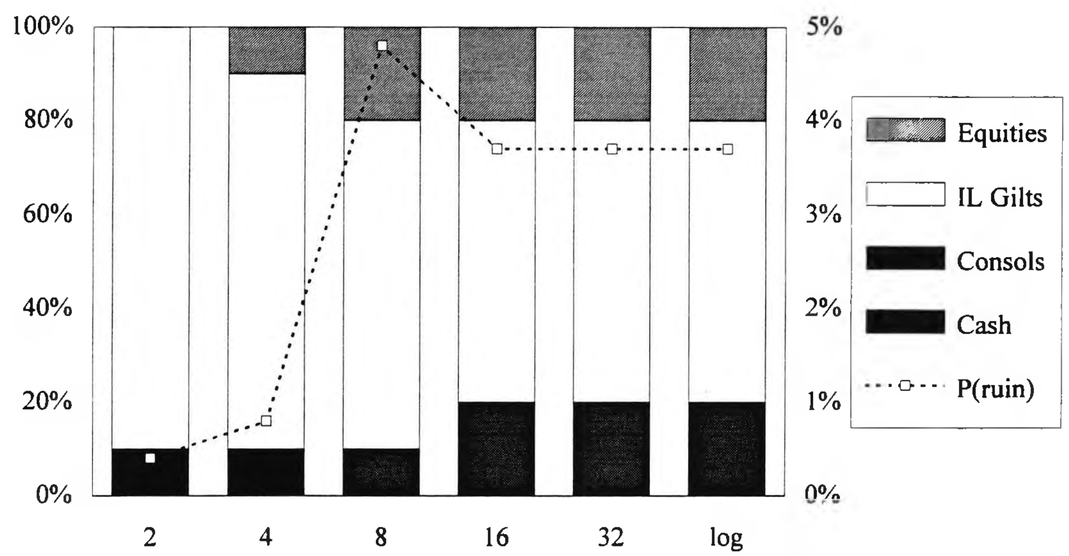


Figure 7.15. Optimal portfolios: a mix of 20 year endowments and 20 year index-linked annuities; sm = 15%, in real terms.

Looking at Figure 7.15, the progression of ruin probabilities with risk tolerance may seem less intuitive, initially. The increasing ruin probabilities between r = 2 and r = 8 are entirely reasonable as more is being invested in equities. But beyond this point, the ruin probability uncharacteristically falls. This is caused by the 10% shift from index-linked gilts into cash, which may be justified by the fact that index-linked gilts yield lower real returns over this period. However, the ruin probability is also reduced as the stable asset value of cash stabilizes the value of a fund with as much as 20% in equities. Imposing a ruin constraint of 1% produces the set of asset mixes shown in Figure 7.16. The first two mixes have obviously remained unchanged as they had previously satisfied the 1% constraint. At higher risk levels, the proportions in equities is again restricted to 10%.

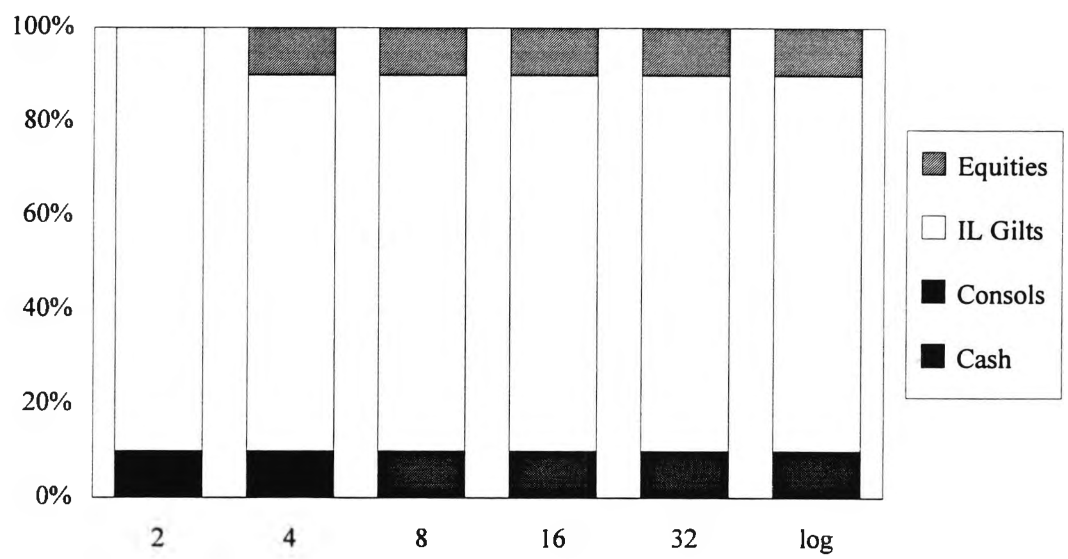


Figure 7.16. Optimal portfolios with a 1% ruin constraint: a mix of 20 year endowments and 20 year index-linked annuities; *sm* = 15%, in real terms.

7.3.2 Scatter Plots of Standard Deviation vs. Mean

Scatter plots of the standard deviation against the mean of nominal payouts for this mix of business are shown in Figure 7.17. From this graph, one may note that the pattern formed by these portfolios is very different from portfolios where the only liabilities were endowments (see Figure 7.11). The portfolios in Figure 7.17 seem to be grouped into a series of strips, parallel to the group satisfying the 5% ruin criterion. Consecutive strips broadly signify a 10% difference in equities, although the groupings become indistinguishable lower down the plot. It is also apparent that the low ruin portfolios are roughly mean-variance efficient.

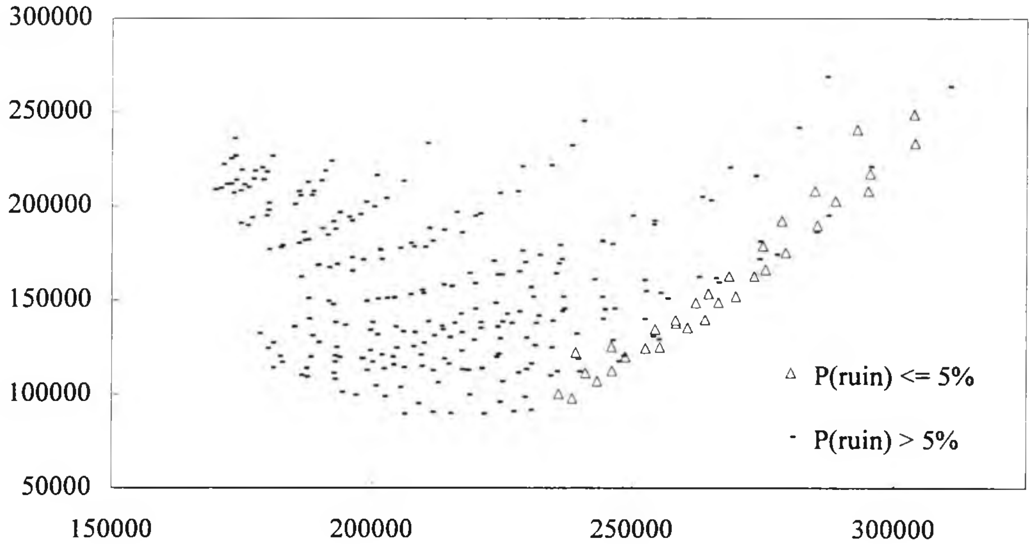


Figure 7.17. Graph of S.D. vs. Mean for all 286 portfolios: a mix of 20 year endowments and 20 year index-linked annuities; *sm* = 15%, in nominal terms.

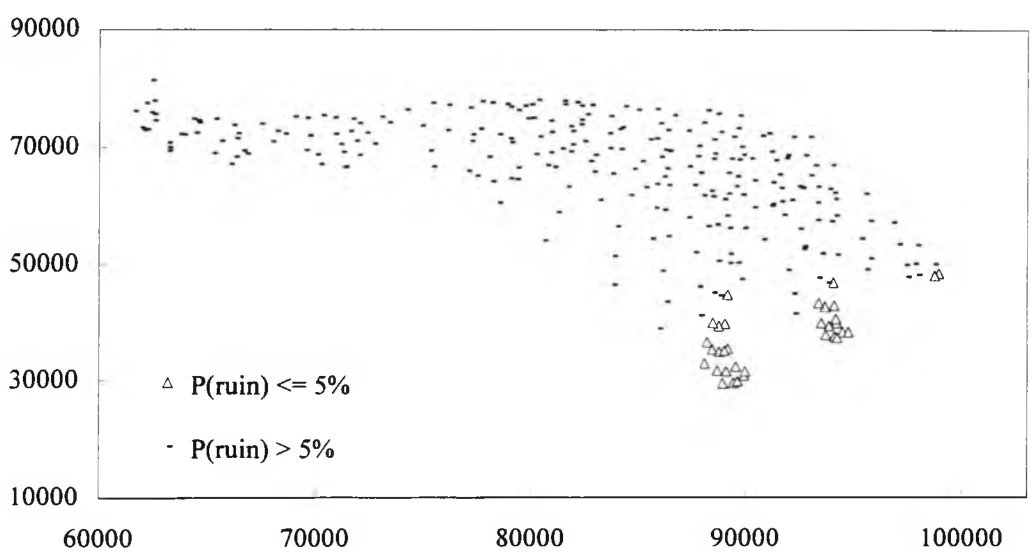


Figure 7.18. Graph of S.D. vs. Mean for all 286 portfolios; a mix of 20 year endowments and 20 year index-linked annuities; *sm* = 15%, in real terms.

When payouts are measured in real terms, the plot shown by Figure 7.18 is vaguely similar to that of Figure 7.12 for endowments alone. The three protrusions on the right hand side of the plot represent the low ruin portfolios with 0%, 10% and 20% invested in equities. Hence, the low ruin asset mixes also seem to be mean-variance efficient in real terms. But unlike Figure 7.17, there is no apparent scope for choosing an efficient portfolio which has a probability of ruin higher than 5%.

7.4 Ten Year Endowments

Another factor in the liability profile which should affect the optimal asset mixes is the term of the contracts in force. If an office were to transact business with terms to maturity shorter than that of twenty year endowments, then it should tend to favour asset mixes which have shorter durations. Hence, this section investigates the optimal asset mixes for Model B when only ten year endowment contracts are being issued.

7.4.1 Unconstrained and Ruin Constrained Optima

When payouts are measured in nominal terms, the unconstrained optimal portfolios with sm = 15% may be represented by Figure 7.19 below. Compared with twenty year endowments, the differences are subtle but evident: the optimal proportions in cash consistently being between 40% and 50% for ten year endowments, in contrast to being between 30% and 40% as shown for twenty year endowments in Figure 7.1 earlier.

From Figure 7.19, it would seem that moving a moderately small proportion of Consols (see r = 4) into equities and index-linked gilts (see the log function) may help to reduce the ruin probability. This probably relates to the maximum rates of interest permitted in the published valuation basis. Although both the real yield and the dividend yield are fairly low on average, the introduction of a small proportion of real assets into a fixed income portfolio will tend to have little effect on the valuation rate of interest used. The resulting fund will have a smaller proportion of total assets invested in Consols and will therefore be more resilient to sharp rises in Consols yields. This diversifying effect also appears to counteract the higher short-term price volatility inherent in real assets. However, if considerably more is invested in equities and index-linked gilts, then this eventually leads to higher ruin probabilities, as seen for the riskier strategies.

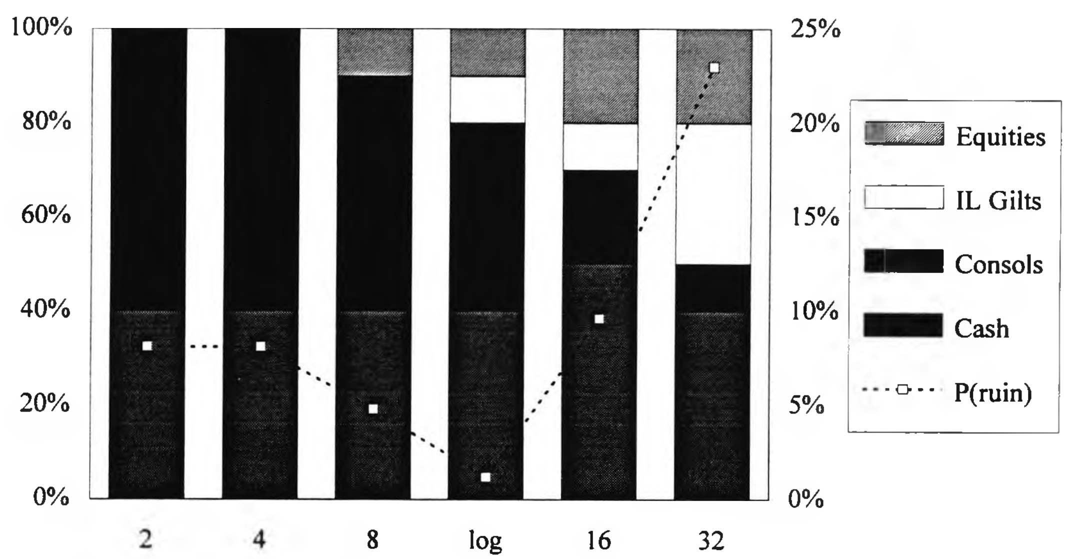


Figure 7.19. Optimal portfolios: 10 year endowments; *sm* = 15%, in nominal terms.

Figure 7.20 below shows the optimal portfolios at each level of risk which satisfies the 1% ruin constraint. When *r* is either at 2 or 4, the 1% ruin criterion may be achieved by shifting 10% of assets from Consols into index-linked gilts. As *r* is increased to 8, the unconstrained portfolio can be made to meet this criterion by switching 20% of the fund from Consols to cash. For the log function, the 20% switch into cash is made possible by disinvesting in both Consols and index-linked gilts. When *r* is 16 and 32, equity proportions are reduced to 10%. The resulting mix of 40% cash, 30% Consols, 20%

index-linked gilts and 10% equities would appear to be a well-matched and diverse portfolio, ensuring a low probability of insolvency and reasonably high average payouts.

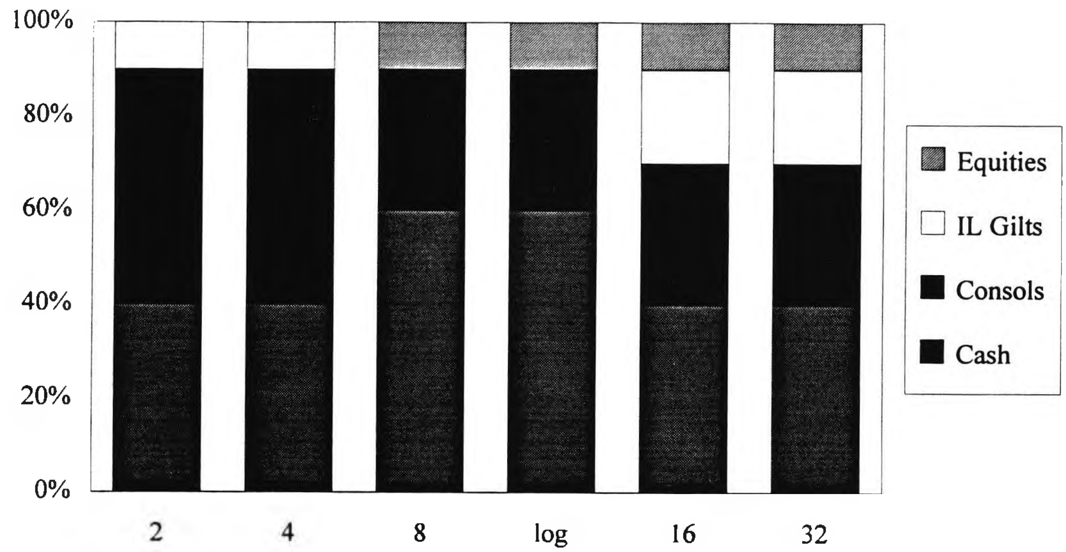


Figure 7.20. Optimal portfolios with a 1% ruin constraint: 10 year endowments; *sm* = 15%, in nominal terms.

Finally, consider the same situation, but with total payouts measured in real terms. The unconstrained optimal mixes are shown in Figure 7.21. This graph shows quite clearly that higher expected real payouts may be achieved by increasing the proportions in cash and equities, while keeping the ratio of nominal to real assets roughly constant. In switching from index-linked gilts into equities, the higher expected real return from the latter more than compensates for the disutility which arises from the increased number of insolvencies, as long as the overall proportion in equities remains low. However, the slight advantage which Consols have over cash in relation to expected real returns does not appear to make up for the higher incidence of insolvencies when greater proportions are being invested in Consols.

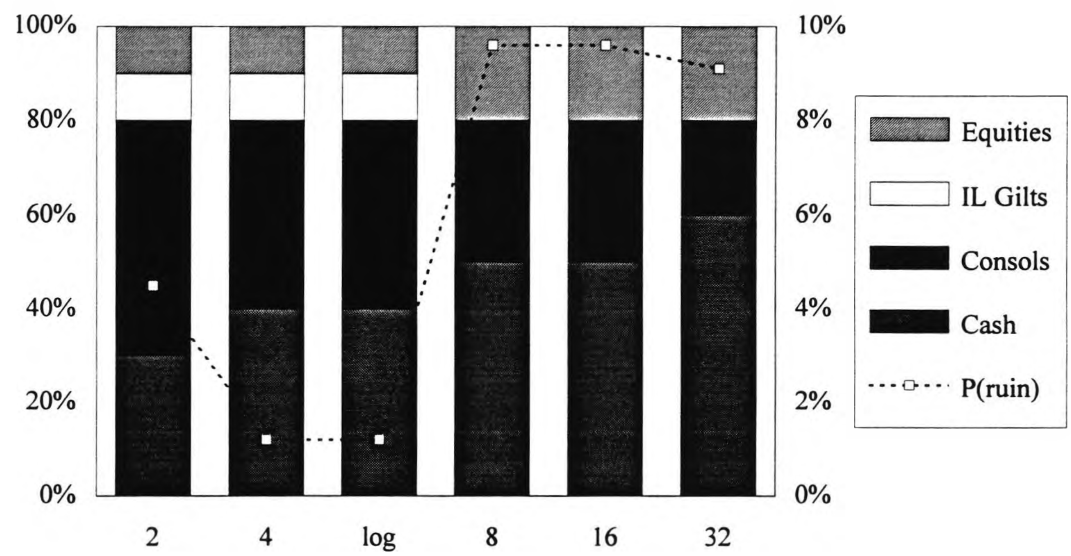


Figure 7.21. Optimal portfolios: 10 year endowments; *sm* = 15%, in real terms.

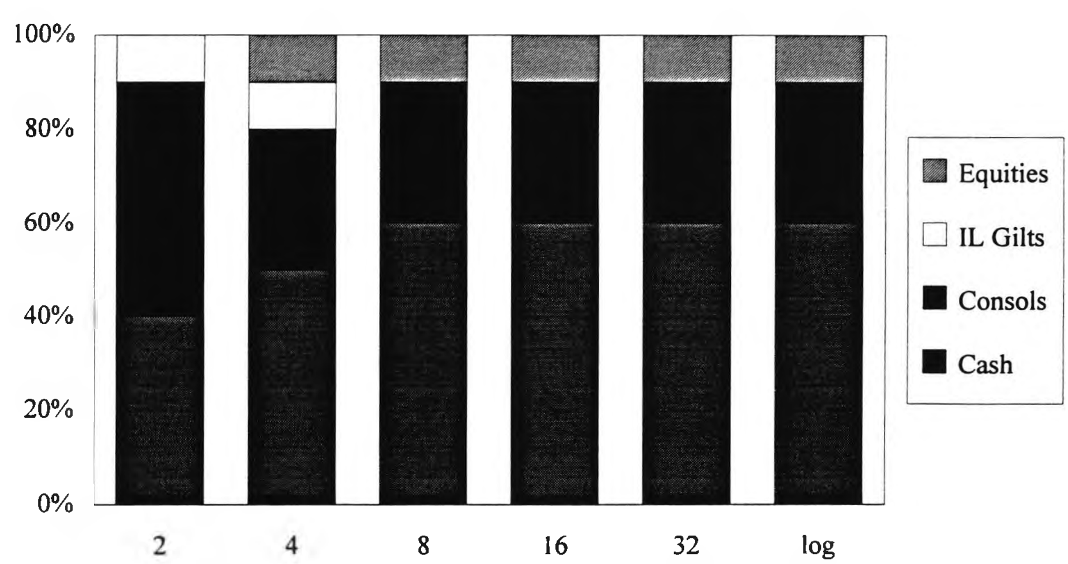


Figure 7.22. Optimal portfolios with a 1% ruin constraint: 10 year endowments; *sm* = 15%, in real terms.

The optimal mixes obtained as a result of imposing a 1% ruin constraint is shown in Figure 7.22. At r = 2, the constraint can be satisfied by switching 10% of the assets from equities into cash. However, this could also been achieved by switching 20% of the fund from Consols into cash, which would result in the optimal constrained portfolio at $r = 4$. This may indicate the stabilizing effect which holding substantial proportions of Consols has on the real payouts. For higher levels of risk tolerance, the same utility maximizing portfolio which meets the 1% ruin criterion is that of 60% cash, 30% Consols and 10% equities.

7.4.2 Scatter Plots of Standard Deviation vs. Mean

The scatter plot for ten year endowments (see Figure 7.23) closely resembles Figure 7.11, involving twenty year endowments. In the plot shown below, it would also appear that all the portfolios with no greater than a 5% probability of ruin lie either on or alongside the efficient frontier. However, the range of portfolios satisfying this 5% ruin criterion also seem to be further up the efficient frontier in comparison with the twenty year case. This is because the portfolios lower down in the plot tend to have larger proportions in Consols.

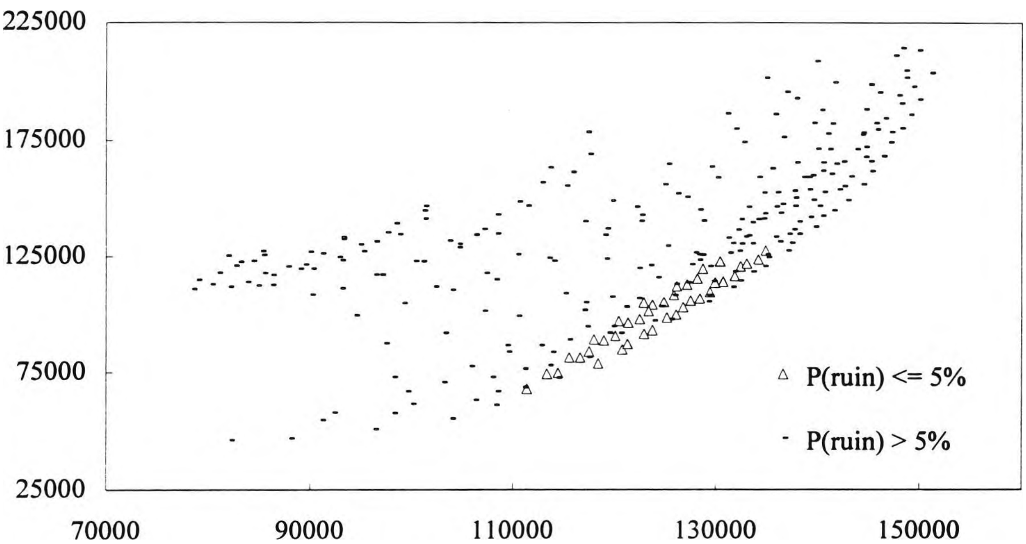


Figure 7.23. Graph of S.D. vs. Mean for all 286 portfolios: 10 year endowments; *sm* = 15%, in nominal terms.

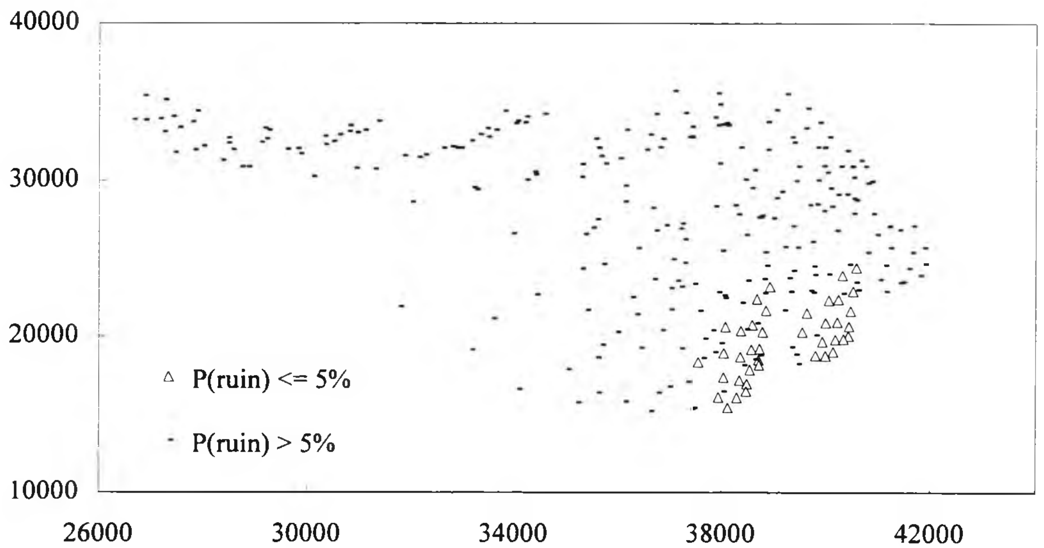


Figure 7.24. Graph of S.D. vs. Mean for all 286 portfolios: 10 year endowments; *sm* = 15%, in real terms.

The scatter plot in real terms, as shown in Figure 7.24, looks quite similar to those relating to other liability profiles (see Figures 7.12 and 7.18). The portfolios satisfying the 5% constraint appear to cluster around the *E-V* efficient frontier, but do not dominate it as had been the case where index-linked annuities contracts were involved.

7.5 Summary

In this chapter, the approximate optimal asset mixes for Model B were obtained using the grid approach and compared where possible with the equivalent results for Model A. The impact of discontinuing the business at insolvency was significant, as the loss of profits resulting from the office being wound up prematurely had a negative effect on the payouts. This led to much lower probabilities of ruin observed in the optimal portfolios for Model B. Where twenty year endowments were concerned, this meant that substantial proportions of nominal assets were preferred at all risk tolerance levels.

When the office was assumed to issue mainly index-linked annuities, the optimal mixes changed dramatically. Index-linked gilts now dominated the optimal portfolios both in nominal and real terms. It was also interesting to see that equities were only held in small proportions, despite their supposed long-term relationship with inflation. The short term volatility of equities and the penalty associated with the higher incidence of insolvencies thus produced seemed to be the reason for this result.

From an immunization perspective, an office transacting only ten year endowments had been investigated in order to compare the durations of the optimal asset mixes with those of an office issuing only twenty year endowments. The results obtained were very much as one may have expected, with higher proportions being held in cash and lower proportions being held in Consols for the office issuing the shorter term contracts.

Overall, the optimal proportions for the different liability profiles seemed intuitively reasonable. The discontinuities in the payouts caused by insolvencies significantly favoured the more solvent asset mixes, therefore leading to more sensible portfolios. This also appeared to reduce the sensitivity of the portfolios to the investment model, increasing the relative importance of liabilities on the decision variables. In these respects, Model B would appear to be superior to Model A. However, the effect of these discontinuities also makes the objective function more difficult to optimize precisely.

8. DYNAMIC OPTIMIZATION MODELS

8.1 Introduction

In preceding chapters, the asset allocation strategy had been optimized on the basis that the proportions in each asset class would remain constant from year to year until the horizon date. While it does not mean that the investor will be prevented from altering the asset mix at a later date, failing to allow for this possibility in determining the initial allocation may lead to sub-optimal strategies. An asset allocation strategy which could be optimized allowing for changes to be made over time should give at least as high an expected utility as that of a static strategy. Nevertheless, it would be useful to isolate the main factors that may give rise to inter-temporal reallocations.

The most important aspect of this study has been to allocate assets in ways which are best suited the liability profile. Therefore, if the mix and size of liabilities are expected to change in the future, it would seem more sensible to consider a dynamic strategy rather than a static one. A situation where this may apply would be a fund which is to be closed to new business at some stage. In this case, the optimal asset allocation at the point of closure will be vastly different from that just before the last policy matures. The main aim of this chapter is to investigate how the optimal asset mix should change over time for a closed fund.

According to immunization theory, the asset mix for a closed fund (with liabilities predominantly fixed in nominal terms) should typically be in long term gilts at the point of closure, moving to shorter term gilts nearer the end of the period of closure. This would seem appropriate for a low risk investment strategy, which is what an immunized position implies. If, on the other hand, a static optimization model were to be used, the asset allocation derived would most probably be a compromise between these two extreme positions, i.e. medium term gilts throughout the period of closure. While this may be unsatisfactory in a closed fund situation, a static strategy should be adequate in the case of a stationary fund. This was why the open funds investigated in Chapters 6 and 7 were assumed to be stationary in real terms.

Another issue which has not been analysed in detail until now is the effect of a horizon date on the decision-making process. Although some arbitrary horizon date had to be assumed in the models used so far (an infinite time horizon is not computationally feasible), the notion of a horizon date is much more intuitive in the case of a closed fund or an individual life insurance contract. In actuarial circles at least, it is widely accepted that for an individual with-profit policy, the investment mix may be more adventurous early on in the policy, as long as the asset mix is gradually allowed to move into a more matched position towards maturity. Similar strategies also seem to be adhered to in other spheres of asset/liability management, such as pension schemes. However, unlike immunization, this strategy does not appear to be based on any firm theoretical footing.

There are at least four possible reasons for suggesting such an investment strategy in the context of a single with-profit endowment contract. The first concerns the level of free assets available over the life of the policy. In the normal course of events, the ratio of free assets to total assets will tend to be highest soon after inception, and least towards the maturity date, because reversionary bonuses are continually being attached to the sum assured. This offers some justification for greater mismatching at early durations.

The second reason may be tactical. If there is a long time period between the inception of the policy and when the assets need to be sold to pay the maturity value, it may seem sensible to invest predominantly in equities initially and then gradually switch into gilts over time. Presumably, this is based on the premise that equities should outperform gilts in the long term. Hence, given a sufficiently long investment period, there is less danger that equities will need to be realized when markets are depressed. A criticism of this, however, is that it assumes that markets will recover given a reasonable cushion of time in which to do so: this would seem to imply market inefficiency. Even if this were true, hardly any tactical advantage could be gained by specifying that the asset mix should evolve over time in a predictable manner. Instead, the process would have to be capable of identifying when the market is or is not depressed and then act accordingly. This form of tactical manoeuvring is well beyond the scope of these investigations.

Two other features that may have some affect on the asset allocation over time are the notions of increasing wealth and time dependency. The former simply refers to when the value of the investment as a proportion of wealth increases over time. This should a feature of most regular premium contracts. Under these circumstances, it is conceivable that investors would behave more cautiously nearer the horizon date as a greater proportion of wealth may be at stake. Time dependent utility functions are best understood in the context of dynamic programming (see section 8.2 below). Due to their more abstract nature, both these features are only of limited interest in the context of the dynamic strategies discussed here.

In view of these arguments, the use of a static optimization model in the open fund situation would seem justified on the whole. The mix of business and the statutory free asset ratio were assumed to remain about constant over time. Assuming that business volumes would on average be stationary in real terms should also have resulted in the discounted mean term of the liabilities remaining constant throughout.

However, in situations where volumes, mix of business or solvency levels may change quite dramatically, it would be necessary to employ dynamic asset allocation strategies, and these will be investigated in this chapter. Initially though, it would be helpful to review the method of dynamic programming before more pragmatic approaches to the problem are discussed.

8.2 Dynamic Programming

8.2.1 Introduction

Dynamic programming is described in Walsh (1975) as an optimization technique applicable to problems involving multistage decisions. It is based on the Principle of Optimality by Bellman (1957), which states that: "*An optimal policy has the property that, whatever the initial state and initial decision, the remaining decisions must constitute an optimal policy with regard to the state resulting from the initial decision.*"

In the field of finance, dynamic programming has been applied to various multiperiod investment-consumption problems (see for example, Samuelson, 1969; Merton, 1971). However, it is only valid for problems where, having obtained the decision variable for a particular stage, the remaining sub-problem retains the exact structural form of the previous sub-problem. This is best illustrated by the two-period investment problem described in Mossin (1968).

8.2.2 Mossin's Two-Period Investment Problem with Quadratic Utility

In many respects, the two-period portfolio selection problem is a very intriguing and difficult problem to solve. Here, the decision involves choosing a portfolio now which maximizes the expected utility of wealth at the end of two consecutive periods, while allowing for the option of revising the portfolio at the end of the first period. Clearly, at the end of the first period, the amount of wealth held then will be known and the process of optimizing the asset mix for the second sub-period is similar to that of a single period problem. However, this decision cannot be determined at the outset as the amount of wealth held at the end of the first period is not known in advance. Similarly, the first period decision will usually depend on the investment policy in the second period.

Despite the apparent circularity of such an argument, Mossin (1968) showed how under certain circumstances, a multiperiod problem may be solved through the use of dynamic programming. The general approach for an n-period problem involves expressing the maximum utility of ultimate wealth with respect to the nth period decision, in terms of the amount held at the start of that period, i.e.:

$$\max_{\mathbf{w}_n} E[U(A_n)] = F_{n-1}(A_{n-1})$$

where $\mathbf{w}_k$ is the decision vector in respect of period k, and A_k is the amount of wealth held at the start of the kth period. Similarly, the maximum of $E[F_{n-1}(A_{n-1})]$ may be expressed as $F_{n-2}(A_{n-2})$, and so on until the ultimate objective may be rewritten as one of maximizing $E[F_1(A_1)]$ with respect to $\mathbf{w}_1$. Because $E[F_1(A_1)]$ and the optimal decision vector, $\mathbf{w}_1$* are functions of A_0 (which is known), the multiperiod problem may then be solved as a series of single period problems.

Mossin demonstrated this technique by using the results from a one-period problem to solve a two-period portfolio selection problem. Consider first the one-period portfolio selection problem involving two risky assets with random rates of return, X_1 and X_2. If A is the initial amount invested and w is the proportion invested in asset 1, then the accumulated amount of assets at the end of the period will be:

$$Y = A[(1 + X_2) + w(X_1 - X_2)]$$

Now let $E_i = E(X_i)$, $V_i = Var(X_i)$, and for convenience assume that $Cov(X_1, X_2) = 0$, i.e. that X_1 and X_2 are uncorrelated. If the investor's risk preference may be represented by the quadratic utility function, $U(Y) = Y - aY^2$, it then follows that:

$$\begin{aligned} E[U(Y)] = A[(1 + E_2) + w(E_1 - E_2)] - aA^2[\, (V_2 + (1 + E_2)^2) \\ - 2w(V_2 - (1 + E_2)(E_1 - E_2)) + w^2((E_1 - E_2)^2 + V_1 + V_2)] \end{aligned} \quad (8.1)$$

Differentiating $E[U(Y)]$ with respect to w and solving for $E'[U(Y)] = 0$, will yield the optimal proportion in asset 1:

$$w^* = \frac{E_1 - E_2 - 2aA[(1+E_2)(E_1 - E_2) - V_2]}{2aA((E_1 - E_2)^2 + V_1 + V_2)} \tag{8.2}$$

Substituting (8.2) into (8.1), it may be shown that the expected utility for the optimal portfolio is given by:

$$E[U(Y)]^* = g(A - hA^2) + c \tag{8.3}$$

where:

$$h = a\left(\frac{V_1(1+E_2)^2 + V_2(1+E_1)^2 + V_1V_2}{V_1(1+E_2) + V_2(1+E_1)}\right) \tag{8.4}$$

$$g = \frac{V_1(1+E_2) + V_2(1+E_1)}{(E_1 - E_2)^2 + V_1 + V_2}$$

$$c = \frac{(E_1 - E_2)^2}{4a[V_1 + V_2 + (E_1 - E_2)^2]}$$

In his two-period example, Mossin assumed that the random rates of return, X_1 and X_2 apply to both periods and that these returns are serially independent. Without the latter assumption, the problem would be much more complicated as ultimate wealth would not only depend on the amount of wealth at the start of the second period, but also on the returns experienced in the previous period. He denoted the amount of initial wealth to be A_0 and defined A_1 and A_2 to be the amounts of wealth at the end of the first and second periods respectively. Similarly, he denoted w_1 and w_2 to be the proportions invested in asset 1 during the first and second periods. Hence, the objective would be to maximize $E[U(A_2)]$.

At the beginning of the second period, the single period case may be applied directly:

$$\max_{w_2} E[U(A_2)] = \max_{w_2} E[A_2 - aA_2^2] = F_1(A_1)$$

where $F_1(A_1)$ is given by $g(A_1 - hA_1^2) + c$ from equation (8.3) earlier. It then follows that the optimal portfolio at the start of the first period will be that which maximizes the expectation of the intermediate return function, $F_1(A_1)$. But:

$$\max_{w_1} E[F_1(A_1)] = \max_{w_1} E[A_1 - hA_1^2]$$

takes exactly the same functional form as that of the one-period case in (8.1). So from (8.2) and (8.4), the optimal proportion in asset 1 at the start of the first period will be:

$$w_1{}^* = \frac{E_1 - E_2 - 2hA[(1+E_2)(E_1 - E_2) - V_2]}{2hA((E_1 - E_2)^2 + V_1 + V_2)}$$

where:

$$h = a\left(\frac{V_1(1+E_2)^2 + V_2(1+E_1)^2 + V_1V_2}{V_1(1+E_2) + V_2(1+E_1)}\right)$$

Hence, the optimal proportion for the first period may be found through maximizing the expectation of a quadratic function in A_1, as one would have done if it were a single period situation. The difference is that the parameter values of the quadratic function in A_1 may only be obtained by considering the situation in the second period. The information derived from optimizing the second period's decision feeds through to the first period in the parameter h. A point worth noting is that the optimal proportion in the second period cannot be determined initially, as it is a function of A_1 and will only be known at the end of the first period. Although such a policy makes intuitive sense, a corollary to this is that the optimal decision may not be obtained solely through the use of numerical optimization routines. Numerical routines only work on the basis that the level of wealth at the start of the period concerned is known. This places quite severe restrictions on the types of problems that may be solved using dynamic programming techniques.

8.2.3 The Exponential and Logarithmic Utility Functions

For instance, consider a one-period portfolio selection problem using the exponential utility function. As general analytical solutions in this situation are not possible without more simplifying assumptions being made, dynamic programming would not be suitable for multiperiod versions of this problem either. But even if the required assumptions are made to enable closed form solutions to be obtained, it does not automatically follow that the identical problem structure will be preserved over successive recursions when applied to the multiperiod case.

In Chapter 4, it was shown that if returns are normally distributed, then the expected utility for an exponential utility function may be expressed in a form which can be solved analytically: $E[U(Y)] = -E[\exp(-Y/r)] = -\exp(-E(Y)/r + Var(Y)/2r^2)$. However, having solved for the optimal weight in the two asset case as given by equation (4.1), it may be shown (see Appendix B) that inserting this expression for w^* back into $E[U(Y)]$ yields a rather more complicated expression:

$$E[U(Y)]^* = -q\exp\left\{\frac{A^2(V_1V_2 - C_{12}{}^2) - 2Ar[(V_2 - C_{12})(E_1 - E_2) + E_2(V_1 + V_2 - 2C_{12})]}{2r^2(V_1 + V_2 - 2C_{12})}\right\}$$

where: $$q = \exp\left\{\frac{-(E_1 - E_2)^2}{2(V_1 + V_2 - 2C_{12})}\right\}$$

be maintained. Furthermore, the return function $E[U(Y)]^*$ is the exponent of a quadratic in A, which means that it can only be maximized numerically. So although it would be possible for dynamic programming techniques to be applied to a two-period problem for the exponential utility function with normally distributed returns, the process may break down for cases involving more than two periods.

An interesting situation arises, though, if the logarithmic utility function is used. While it may be the case that single-period portfolio selection problems cannot be solved analytically for the logarithmic utility function, this fact does not exclude the function from multiperiod problems. It has been shown in Mossin (1968) and Samuelson (1969) that the logarithmic utility function leads to optimal policies which are *myopic*. This simply means that the optimal decision in any period is not dependent on future events. In a multiperiod selection problem, an investor with a logarithmic utility function will only need to be concerned with maximizing the expected utility of wealth at the end of the first period.

In relation to the asset fund analysed in Chapter 4, this means that an investor with a logarithmic utility function should make the same optimal decision over the first year of investment regardless of whether the horizon date is in one or twenty years time. However, as the distributions of simulated returns are only approximately the same from year to year, the one-period optimal decisions in subsequent years will only be roughly the same as that derived in the first year.

When long-term liabilities are involved, the situation is less simple: the amount of wealth (surplus) at the end of each period is only an estimate, as surplus is a function of the valuation basis used. The true amount of wealth may only be determined when all business is run off the companies books or if the business is wound up at or before some finite horizon date. Hence, even if a logarithmic utility function is used in the context of a model office, it would seem inappropriate just to consider the expected utility of surplus at the end of the first period, as this would place excessive importance on the valuation basis assumed.

So apart from exceptional circumstances, there generally remains a need to consider investment decisions in a multiperiod framework. Although dynamic programming has been shown to work under certain conditions, its range of applications is very limited indeed. Therefore, in the following section, a brief review is given on some of the more pragmatic approaches which have been suggested for dynamically reallocating assets.

8.3 Pragmatic Decision Rules

8.3.1 Resilience Testing

A well known dynamic reallocation rule, though rarely thought of as a decision rule (probably due to the fact that it features as a statutory requirement), is resilience testing (see Purchase *et al*, 1989). Rather than specifying a percentage of funds which needs to be matched to liabilities or invested in gilts, resilience testing simply ensures that the free assets held by an office are adequately resilient to fluctuations in key economic variables within a specified range. For example, consider the rule that free assets should be resilient to falls in equity values of up to x%. An office operating on this basis would be allowed to invest in any asset mix, providing that this requirement is satisfied. Should the free asset ratio fall beyond the level needed to maintain this degree of resilience, then the x% rule imposes the required switch from equities into gilts. This rule gives insurers the choice of paying the cost of higher capital, which may permit an investment that leads to a higher expected return, or holding less capital with a more matched investment strategy.

Resilience testing may be quite useful in optimization models, acting as a partial constraint on the optimal asset allocation. If a sufficiently risk tolerant utility function is used, the model should still permit adventurous asset mixes when the solvency position is strong, but curtail excessive ruin probabilities by limiting the proportion in equities when the solvency position is weak. Static optimization models may be criticized on the grounds that the optimal asset mixes are based on the assumption that no intermediate reallocation of assets is possible. Incorporating this type of rule may help to make these models more realistic.

There are, however, a number of points which should be made regarding this approach. Firstly, the rule is a practical method for controlling insolvency probabilities and is not intended to lead to an optimal investment strategy in any sense. Secondly, if the rule is convoluted (including mis-matching tests for various other asset classes say), then the procedure for determining the appropriate switches may be extremely complex. Thirdly, the approach may not produce meaningful optimal asset mixes if the rule actually happens to take effect most of the time. This may occur if the rule is too stringent relative to the initial solvency position. Lastly, the rule should also reflect how the life office would be expected to react in those circumstances when they arise. It would defeat the point of imposing the rule in the first place if the optimization process assumed a switching rule that would not actually be implemented in practice.

8.3.2 Solvency Based Reallocations

A technique related to resilience testing involves a switching rule that is dependent on some measure of solvency. For example, a rule may be imposed whereby no less than $100(x - \rho)/(x - y)\%$ of the assets must be invested in gilts if the ratio of assets to liabilities, ρ falls below x but remains above y. If ρ falls beyond y, then all assets must be invested in gilts. In Ross (1989) and Hardy (1993), $x = 1.25$ and $y = 1.05$. This not only encourages a more cautious investment strategy when solvency is onerous, but also aims to improve the statutory solvency margin by permitting a higher valuation rate of interest to be used in calculating the statutory reserves.

However, the considerations outlined with regards to resilience testing also apply here. In general, the rule is not optimal and only works well if the solvency position is strong to begin with. It also assumes the initial asset mix is predominantly in equities and other risky asset classes.

8.3.3 Tactical Switching

Although it is not the intention of this thesis to investigate tactical allocations, such methods tend to be dynamic by nature and would be worth mentioning in the context of dynamic optimization models. An example of a tactical decision rule has been discussed in MacDonald (1991). Using the investment model developed by Wilkie (1986), MacDonald attempted to determine which asset class (gilts or equities) would be more likely to perform better following a recent spell of good performance relative to the other asset class. As expected, equities were found to be better following poorer performances relative to gilts and vice-versa. MacDonald acknowledged that this was not really surprising given the autoregressive nature of Wilkie's model.

Regardless of whether models such as this may be useful in practice (particularly due to their excessive reliance on the more predictive characteristics of the investment model), tactical switching does not really deal with concerns about changing solvency positions, as tackled to a certain extent in resilience testing or the solvency based switching rule described in Sections 8.3.1 and 8.3.2 respectively. All three models, however, do not address directly the problem of a changing liability profile, which would be a key issue in the asset allocation strategy for a closed fund.

8.4 The Closed Fund

8.4.1 Duration Based Models

The closed fund is an attractive situation to consider because it allows ultimate surplus to be accounted for in a more realistic manner than in an open fund. Furthermore, there is naturally a dynamic element involved in the asset allocation of a closed fund. The main reason for this relates to the fact that the unexpired duration of the liabilities is reducing over time. From an immunization perspective, the average term to redemption of fixed-interest assets held should also be reducing in tandem with the duration of the fixed liabilities.

For a fairly risk averse investor, a suitable strategy could be to immunize (as far as this is possible) an amount of assets equal to the value of the liabilities and to invest the rest in assets which have higher expected returns. But should there be ample free assets available, it may be unnecessary to be as cautious as this. In any case, it may be more useful to investigate the optimal strategies for a range of risk tolerance levels, as done in previous chapters, rather than just a single level of risk tolerance.

Ideally, the strategy used for the closed fund should at least be dynamic enough to take account of changes in duration of its liabilities and be able to accommodate a range of risk preferences. A possible approach could be to optimize a separate vector of weights for each time interval. If there are H years to the time when the last liability payment is due, then the objective may be to optimize the ultimate surplus at this horizon date in respect of the decision vectors, $\mathbf{w}_1$, ..., $\mathbf{w}_H$, where $\mathbf{w}_t$ is the asset mix in the tth year. However, as well as having to optimize over $H \times (M-1)$ parameters, M being the number of asset classes available, there would also be an element of tactical switching involved if the distributions of returns differ from year to year. The latter problem would almost be inevitable with simulated returns.

In view of this, it would be desirable if fewer parameters could be used, both for reasons of computational efficiency and to reduce the effect of tactical switches on the optimal strategy. A simple remedy could be to split the whole period into fewer sub-periods of five year intervals, for example. However, even with a twenty year time horizon, such an approach could produce fairly crude changes in asset mix whilst involving as many as $4 \times (M-1)$ parameters.

Before an alternative approach is considered, it would be worth briefly returning to the numerical routine used to derive the optimal portfolios in a static context. In Chapters 4 and 6, the basic procedure had been to set the proportion in the kth asset class, $w_k = a_k^2$, for $k = 1, ..., M-1$, leaving the balancing proportion, $w_M = 1 - (w_1 + ... + w_{M-1})$. The optimal asset mix could then be found by unconstrained optimization of the objective function with respect to the parameters $a_1, ..., a_{M-1}$. However, this is not the only transformation that results in asset proportions that are non-negative and sum to unity. An equivalent formulation could be:

$$w_k = \frac{a_k^2}{\sum_{j=1}^{M} a_j^2} \tag{8.5}$$

while setting $a_M = c$ (an arbitrary constant) to avoid over-parameterization.

Returning to the dynamic case, consider the following function:

$$w_k(s,t) = \frac{a_k^2 + b_k^2 \phi_{s,t}}{\sum_{j=1}^{M} (a_j^2 + b_j^2 \phi_{s,t})} \tag{8.6}$$

where $\phi_{s,t}$ is a random component dependent on simulation s at time t; $a_1, ..., a_{M-1}$ and $b_1, ..., b_M$ are the parameters over which the objective function is to be optimized, with

$a_M = c$. If ϕ represents the discounted mean term of liabilities, then for a given set of parameter values, the asset proportions will vary as the discounted mean term changes. In this way, the asset mix can be made a function of the duration of the liabilities. Hence, the problem may be approached from the perspective of optimizing the expected utility of payouts in relation to the parameters $a_1, ..., a_{M-1}$ and $b_1, ..., b_M$.

A feature of this formulation is that it effectively reduces to the static case if the liability profile is constant over s and t, or if $b_k = 0, \forall k$. This means that it will always yield an expected utility which is at least as high as that of a static strategy i.e. expression (8.5). It also has relatively few parameters, making it computationally efficient and should result in smoothed asset mixes over time. In so doing, however, there is a possibility that this function may not be sufficiently flexible in accommodating the necessary switches in the asset proportions over time. Nevertheless, it should be worthwhile investigating how the dynamic reallocation method given by expression (8.6) performs in the closed fund situation.

8.4.2 Optimal Strategies with Nominal Payouts

Unless otherwise stated, Model A will be assumed throughout. This means that negative dividends are permitted when the guarantee fund is breached. Therefore, the fund will always remain in operation throughout the period of closure. At the start of the projection period, the fund is assumed to be stationary in real terms. The liabilities in force are assumed to be twenty year endowments, as in Chapter 6. Closure to new business is assumed to take immediate effect, which means that the last contract is issued a year before the start of the projection period. This implies that the last maturity payment is due at time 19. The algorithm used for distributing dividends is identical to that used in the open fund (which had been aimed at maintaining the same ratio of assets to published liabilities throughout the projection period).

The discounted mean term of liabilities is calculated on the same basis as the realistic reserves, i.e. gross premium valuation at a rate of interest equal to the current yields on Consols and index-linked gilts for nominal and real liabilities respectively. As these yields are stochastic from time 1 onwards, only the asset mix at time 0 will be known for a given set of parameter values. In subsequent periods, the parameter values just determine the allocation rule, rather than the unconditional asset mixes. However, as a particular set of parameter values may be difficult to interpret in isolation, all optimal strategies obtained will be graphically illustrated in relation to a 'typical' scenario. This typical scenario assumes that the relevant yields remain at their neutral values, i.e. 8.5% for Consols and 3.5% for index-linked gilts. As the discounted mean terms do not vary extensively between scenarios, the trends in asset proportions shown by means of this neutral scenario should be fairly representative of the general trends that would be observed from each of the one thousand simulations used.

Figure 8.1 shows how the optimal proportion in each asset class changes as the duration of the liabilities decreases. The initial solvency margin is assumed to be sm = 15%. Payouts are measured in nominal terms and the level of risk tolerance used is $r = 2$. The probability of ruin together with the mean and standard deviation of the nominal payouts are also indicated at the top of the graph.

Overall, the asset allocation strategy over time seems broadly intuitive, with the optimal proportions in cash and Consols rising and falling respectively as the discounted mean term of the liabilities decreases. Holdings in index-linked gilts and equities remain small throughout, which is to be expected given the low level of risk tolerance assumed. In fact, the asset mix seems to be quite well immunized, with about 25% in cash and over 70% in Consols when the duration is highest, evolving to about 80% in cash and 20% in Consols at the shortest duration. The immunized strategy also appears to be reflected in the low ruin probability of 1% observed for the nineteen year projection period.

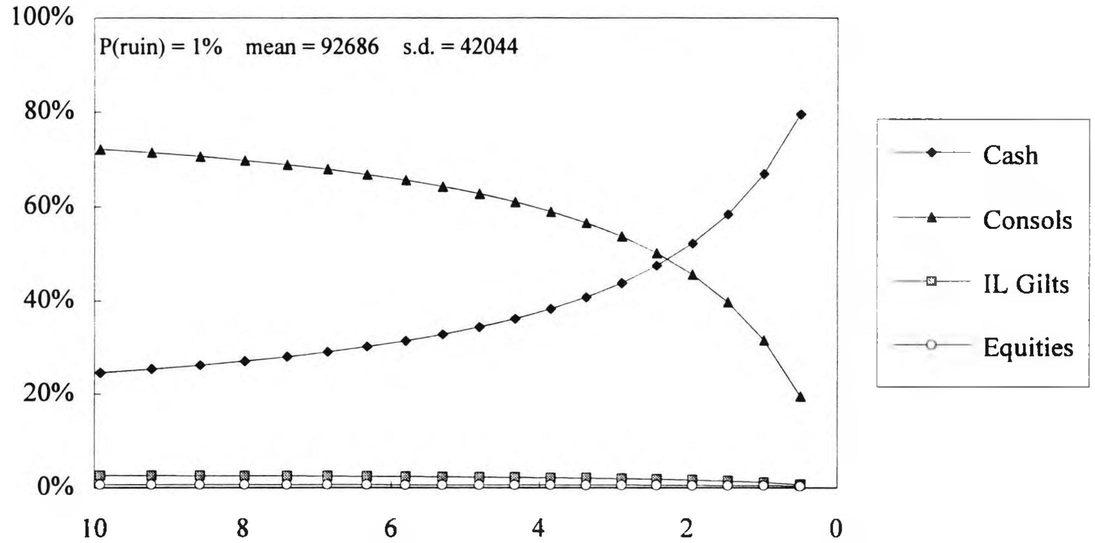


Figure 8.1 Graph of Optimal Asset Proportions vs. Discounted Mean Term of liabilities: sm = 15%, in nominal terms, at r = 2.

In Figure 8.2, the same situation applies except that a 'medium risk' strategy with r = 8 is considered. The mean and standard deviation of payouts shown here are higher than those observed in Figure 8.1, which is facilitated by investing greater proportions of assets in the higher yielding real asset classes. Given the small fraction of inflationary liabilities incurred throughout the projection period, this strategy appears to be more mis-matched in relation to the liabilities, particularly at the shorter durations.

Although it is intuitive that the fund would be mis-matched to a greater extent (with the aim of achieving higher average payouts), it is less obvious how this mis-matching should evolve across time. As the ratio of assets to published liabilities should, on average, remain constant over time, it would be reasonable to expect the optimal proportions invested in real assets to also remain the same throughout the projection period. This does not appear to be the case here, with the percentage invested in fixed income assets falling steadily from 73% in the first year to 23% in the last year. Despite this peculiar trend, the observed probability of ruin still remains at 1%.

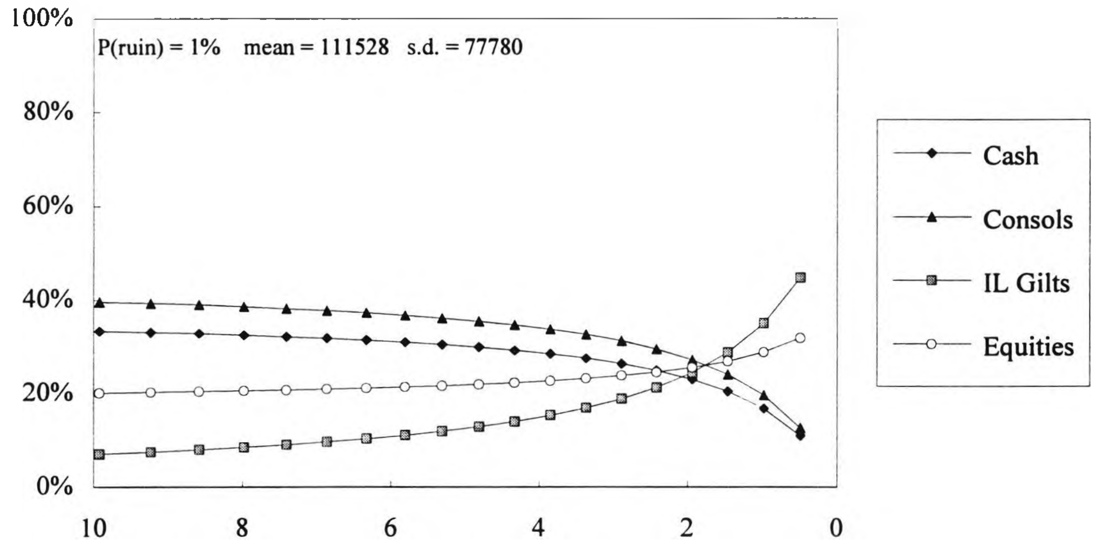


Figure 8.2 Graph of Optimal Asset Proportions vs. Discounted Mean Term of liabilities: sm = 15%, in nominal terms, at $r = 8$.

In view of these results, it may be helpful to examine the free asset ratio (FAR) over time, both on a realistic and a published basis. Define the free asset ratio at time t to be:

$$FAR(t) = \frac{Acc(t) - Liab(t)}{Acc(t)},$$

where $Acc(t)$ is the total amount of assets immediately following the distribution of dividends at time t (see Section 5.3) and $Liab(t)$ is the value placed on liabilities at time t. The realistic free asset ratio uses $Liab(t) = Real\_Liab(t)$ whereas the published free asset ratio uses $Liab(t) = Stat\_Liab(t)$, as defined in Section 5.3 earlier.

Figure 8.3 below shows the average realistic and published free asset ratios at the start of each projection year, assuming the optimal strategy at $r = 8$ is adopted. In many respects, the free asset ratios shown are not unreasonable. The published FAR at $t = 0$ of 13.04% is consistent with the initial solvency margin of 15% as a proportion of published liabilities, i.e. $0.1304 \approx 1 - 1/1.15$. The average published FAR also remains below the average realistic FAR throughout the projection period, which reflects the fact that the published basis is more stringent than the realistic basis. As the horizon date approaches, the two ratios must converge because the reserves will become less sensitive to interest rate assumptions as the duration of the liabilities shortens.

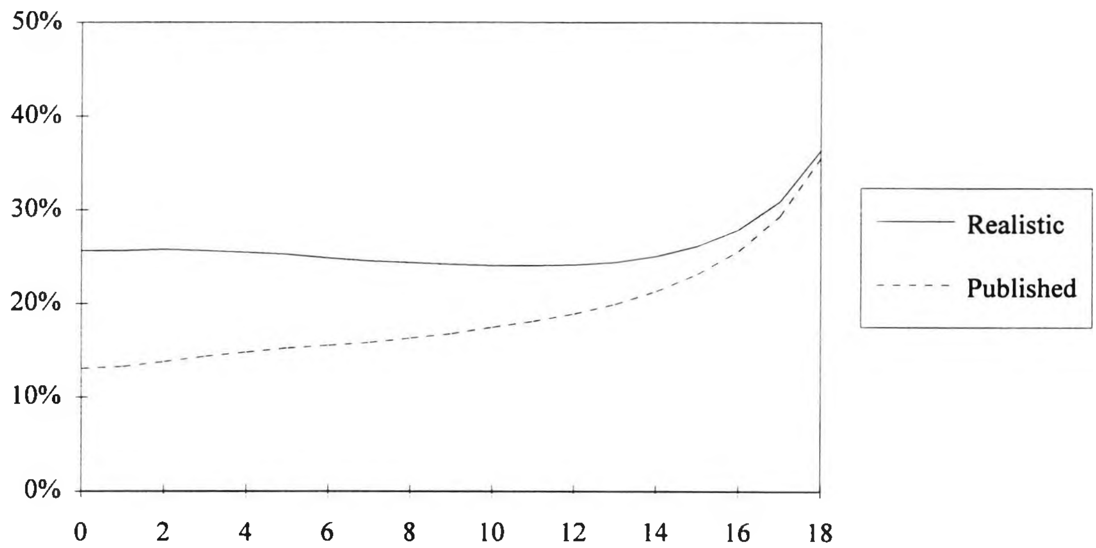


Figure 8.3 Graph of Mean Realistic Free Asset Ratio and Mean Published Free Asset Ratio over time: sm = 15%, in nominal terms, at r = 8.

A striking feature of Figure 8.3 is the steep rise in free asset ratios towards the horizon date. The intended distribution policy had been to declare dividends in a manner which would maintain the published solvency margin held initially throughout the projection period. While the formula for doing this had worked well in the case of the open fund, the same formula appears to be inadequate for the closed fund in this respect. One may recall that the smoothing parameters k, c and e had been specifically chosen so as to satisfy this objective in the open fund situation. It therefore seems probable that an alternative set of smoothing parameters may be more appropriate for the closed fund situation if constant published free asset ratios are to be maintained. Nevertheless, the scenario being considered here is very much a hypothetical one and it is debatable whether striving towards such an objective would be ideal in practice. A positive aspect of the resulting distribution policy is that it errs on the side of prudence by allowing the published FARs to build up over time, thus making insolvency less likely.

In each projection year, the office is retaining a modest amount of surplus over and above the intended level. As the published reserve tends to remain above its initial level for the first half of the projection period, the published free asset ratio is only seen to rise very gradually for the first 10-15 years. However, in the last five or so years, the published reserve begins to decline much more rapidly. Although the absolute amount of free assets held is also decreasing, it is doing so at a much slower rate (due to the lag in the dividend distribution policy); so towards the horizon date, the free assets may be seen to rise more sharply as a proportion of total assets.

Given the free asset ratios shown in Figure 8.3, the optimal strategy shown in Figure 8.2 now appears more intuitive. The office is quite well matched initially, with both the amount invested in fixed income assets and the realistic value of liabilities forming roughly three quarters of the total fund at the start. Even though the proportion invested in fixed income assets decreases with time, the fund does not on average deviate too far from a matched position over the next ten or so years. For example, at time 10 when the mean realistic FAR is 24%, only 22% and 14% of assets are invested in equities and index-linked gilts respectively. This makes it unlikely that free asset ratios will fall to an onerous level during this period.

Now consider the more extreme position at the start of the final year when 76% of the fund (32% and 44% in equities and index-linked gilts respectively) is invested in real assets. At first this asset mix may appear much more prone to lead to insolvency with both the mean realistic and published FARs standing at about 36%. However, even if the returns on equities and index-linked gilts were to fall by two standard deviations in one year, i.e. returns of about -31.8% and -2.8% respectively (see Table 3.8), this would *ceteris paribus* result in a reduction in total fund value of about 11%. So as well as requiring the free asset ratio to fall to about a third of its mean value by $t = 18$, the office would have to experience another year of very poor returns for insolvency to occur then. This explains the very low ruin probability observed in Figure 8.2.

A useful comparison to study may be the 'high risk' strategy with $r = 32$ shown in Figure 8.4. From this graph, there appears to be a definite preference for more volatile assets nearer the horizon date, as had been noted in Figure 8.2 earlier. This time though, the asset allocation strategy is much more extreme, with none of the fund being held in fixed interest assets at any time. The resulting 91% ruin probability would seem consistent with such a strategy, as would the relatively high mean and standard deviation of payouts shown.

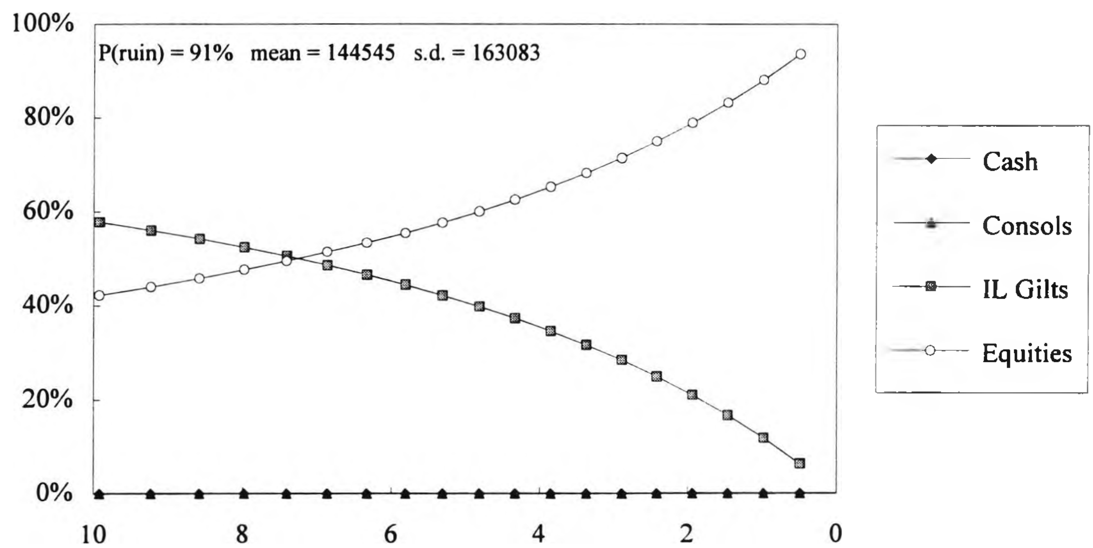


Figure 8.4 Graph of Optimal Asset Proportions vs. Discounted Mean Term of liabilities: $sm = 15\%$, in nominal terms, at $r = 32$.

Looking at the mean published free asset ratio shown in Figure 8.5, there should be little difficulty in explaining why the observed ruin probability for such a strategy is so high. Being invested entirely in real assets, the valuation rate of interest that may be used for published reserves is generally very low. This is clear from the published FAR at $t = 0$. As the office is barely able to cover the guarantee fund at this early stage and is heavily mis-matched throughout, ruin would seem the most probable outcome.

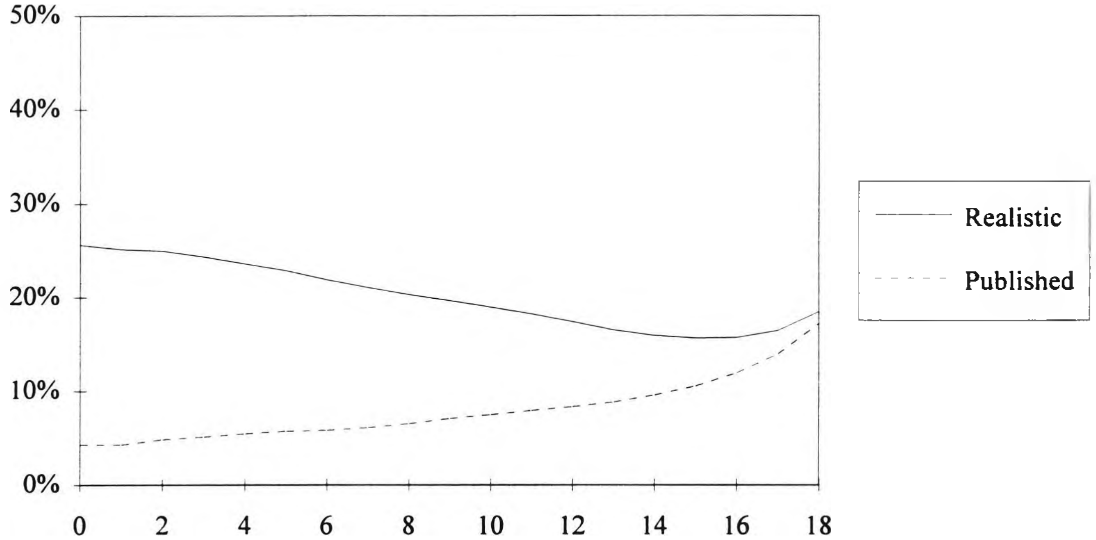


Figure 8.5 Graph of Mean Realistic Free Asset Ratio and Mean Published Free Asset Ratio over time: sm = 15%, in nominal terms, at r = 32.

A far more challenging task though would be to explain the trend towards the more volatile asset categories over time. Unlike the optimal strategy obtained when $r = 8$, the strategy shown in Figure 8.4 does not appear to be greatly influenced by the office's free asset ratios. In Figure 8.2 earlier, the optimal proportions in real assets remained low for the majority of the projection period, rising sharply only when both the realistic and published FARs rose sharply as well. In contrast to this, the rate at which index-linked gilts are being switched into equities when $r = 32$ is much more uniform over the entire period. While this may reflect the gradually increasing published FAR seen in Figure 8.5, intuition would tend to suggest the realistic FAR as being a more influential factor in determining the optimal asset allocation strategy. However, the realistic FAR in this case appears to exhibit more of a downward trend over time.

An alternative explanation of this result may relate to the way in which the objective function has been defined. In order to allow for time preference, dividends that have been declared are rolled up to the horizon date at the rate of return earned on equities during that accumulation period. This not only increases the significance of early dividends in terms of expected payouts, but also enhances their contribution to the variability of these payouts. With $r = 32$, relatively more heed is given to the average payouts than to the variance of payouts. Hence, it is possible that the contribution to the variance of payouts from later dividends is not sufficiently material relative to that of earlier dividends to require asset mixes later in the projection period to be as cautious as those preferred early on. This could explain the trend towards increasing equity investment over time, as seen in Figure 8.4.

8.4.3 Optimal Strategies with Real Payouts

The above investigations were also carried out in real amounts, the results of which are discussed in this section. Figure 8.6 below shows the optimal trend when $r = 2$ and exhibits many of the features seen in Figure 8.2 earlier. The initial asset mix seems quite reasonable with about 72% of the fund being invested in fixed income assets, although the optimal proportion gradually falls to 57% after 10 years. This is not too dissimilar to the trend shown in Figure 8.2, with the corresponding proportions being 73% at $t = 0$ and 64% at $t = 10$. As the outstanding duration of the liabilities decreases further, the optimal proportion in the real asset classes rises quite sharply, eventually reaching about 85% in the final year of operation. This compares with the figure of 76% as implied by Figure 8.2.

Although the optimal strategies in Figure 8.2 (nominal payouts with $r = 8$) and Figure 8.6 (real payouts with $r = 2$) should not be directly comparable, there are reasons for the similarities between the two. Both require a reasonable degree of matching initially as their free asset ratios are at their lowest levels during this time (see Figures 8.3 and 8.7). As their free asset ratios rise towards the horizon date a greater degree of mis-matching is acceptable. However, the two strategies tend to mis-match for different reasons.

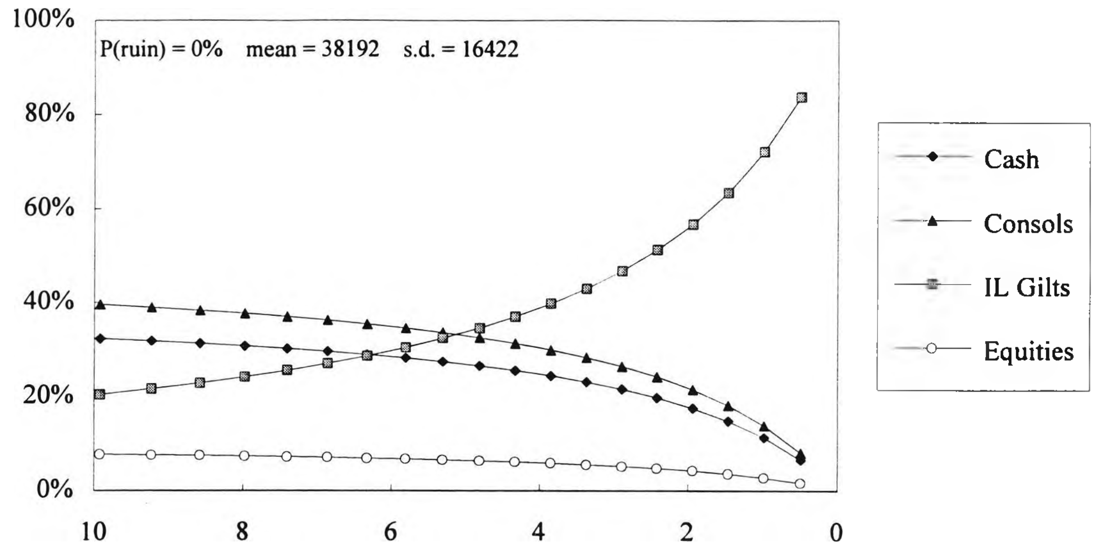


Figure 8.6 Graph of Optimal Asset Proportions vs. Discounted Mean Term of liabilities: sm = 15%, in real terms, at r = 2.

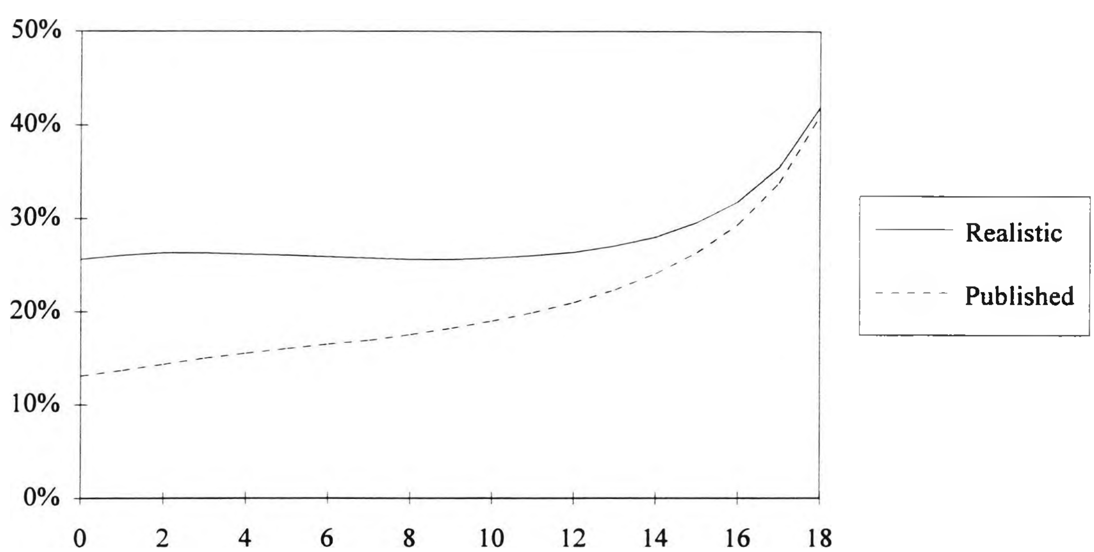


Figure 8.7 Graph of Mean Realistic Free Asset Ratio and Mean Published Free Asset Ratio over time: sm = 15%, in real terms, at r = 2.

The strategy in Figure 8.6 aims to keep the variability of real payouts down to a very low level, thus strongly favouring index-linked gilts at shorter durations. The strategy in Figure 8.2, though, aims to increase the mean nominal payout while constraining the variability of nominal payouts to a suitable level. Hence, this strategy favours a more balanced mix between equities and index-linked gilts. As with Figure 8.2, the strategy in Figure 8.6 still manages to maintain a negligible ruin probability despite appearing mis-matched at the shorter durations. Given the high free asset ratios observed towards the end of the projection period (see Figure 8.7) and the low proportion invested in equities throughout, this result is hardly surprising.

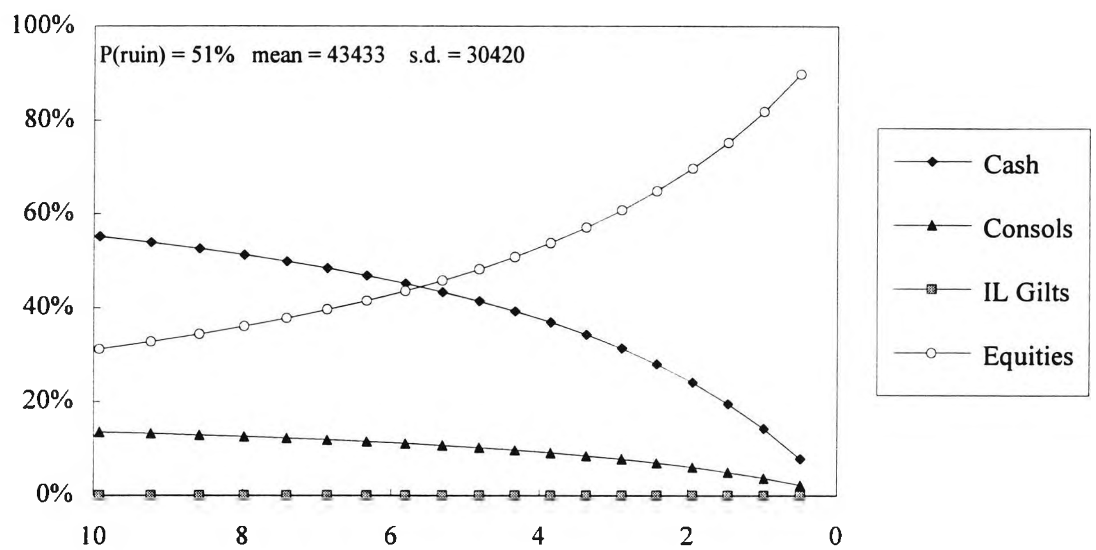


Figure 8.8 Graph of Optimal Asset Proportions vs. Discounted Mean Term of liabilities: sm = 15%, in real terms, at r = 8.

At the higher risk tolerance level of $r = 8$, the optimal strategy for real payouts is shown in Figure 8.8. As expected, the ruin probability, mean and standard deviation of real payouts are all greater than for $r = 2$. Despite having substantial proportions in nominal assets at the earlier stages, the fund still appears mis-matched as less than 15% of assets are invested in Consols at all times. Investment in index-linked gilts is negligible throughout, being the asset class with the lowest mean real accumulation over this period. The high proportions held in equities in the later years contribute to the high ruin probability of 51%, as do the more moderate free asset ratios seen throughout (see Figure 8.9). The optimal strategy for $r = 32$ is not shown here as it simply comprises 100% investment in equities throughout.

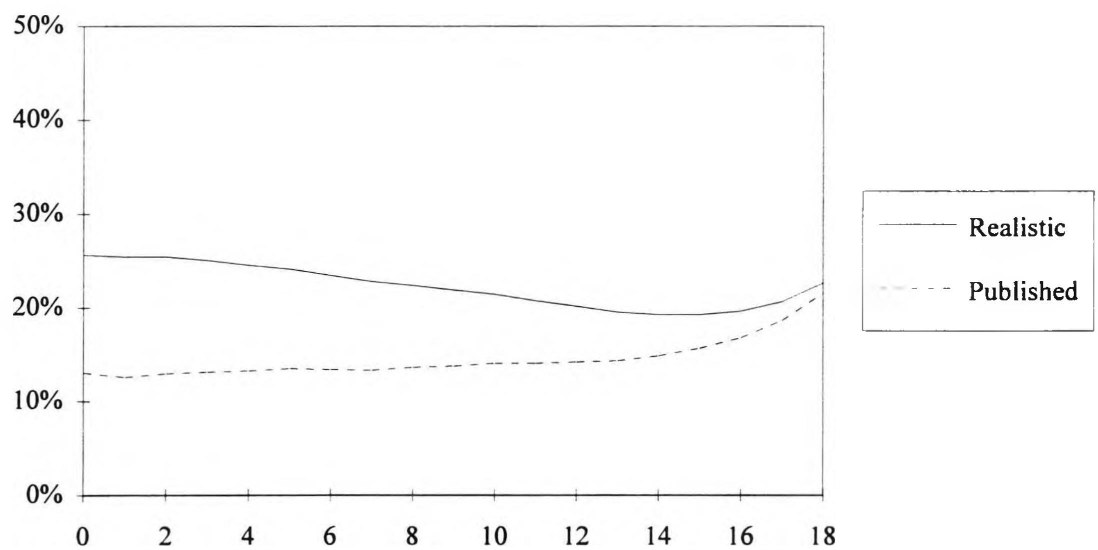


Figure 8.9 Graph of Mean Realistic Free Asset Ratio and Mean Published Free Asset Ratio over time: sm = 15%, in real terms, at r = 8.

8.4.4 Optimal Strategies for a Simple Closed Fund

Although the distribution policy used in Sections 8.4.2 and 8.4.3 above failed to fulfil the aim of maintaining a constant free asset ratio, the study did demonstrate the ability of duration based models to deal with changing liability profiles. Nevertheless, it also seemed possible that factors other than those relating to increasing free asset ratios may have been affecting the optimal strategies obtained. In order to investigate this further, it may be more appropriate to analyse a much simpler version of the closed fund.

Consider a fund at time 0, $F_0 = 38000$. At the start of each year, the fund, F_t is invested in asset mix, $\mathbf{w}_t$ to yield a random amount, $[1 + i(\mathbf{w}_t)].F_t$ in one year's time. At the end of each year for the next 19 years, a fixed amount of 2000 is paid out from the fund. Immediately following each payment made at time t, a reserve of $V_t = 2000.(19 - t)$ is set up, i.e. the reserve is equal to the sum of the remaining payments due in the future. The difference between the remaining fund and the reserve at time t is:

$$d_t = [1 + i(\mathbf{w}_{t-1})].F_{t-1} - 2000 - V_t$$

where d_t is the dividend declared at time t, which is assumed to be invested in equities until the horizon date. It then follows that $F_t = V_t$.

If this simple fund is then optimized, with the objective of minimizing the variance of real accumulated dividends (payouts), the optimal strategy is similar to that shown in Figure 8.6. Initially, the fund favours fixed income assets, although towards the later projection years the optimal asset mix mostly comprises index-linked gilts. In a way this was to be expected as the minimum variance strategy should be very similar to a utility maximizing strategy at low levels of risk tolerance. However, this result also shows that the trend seen in Figure 8.6 is not unique to the complex liability structure and distribution policy assumed in the model office. In fact, the same trend emerged even when dividends from the simple fund were accumulated using a constant rate of interest, thus reinforcing this view.

A useful feature of the simple closed fund is that the fund value at the start of each year is known with certainty. Hence, the variance of dividends arising from any one year may be computed independently of the outcome in any other year. In this sense, each year of operation may be treated as a separate single period problem. So when the objective was to minimize the variance of $d_{19}/Q(19)$ with respect to $\mathbf{w}_{19}$, where $Q(t)$ is the retail price index at time t, the optimal asset mix in the final year was shown to consist mostly of cash, which is intuitive. However, this contrasts with the result of the optimal multiperiod strategy which favours more index-linked gilts in the 19th year.

In order to resolve this paradox, consider a variant of the simple closed fund. As before, the objective will be to minimize the variance of real payouts. However, instead of optimizing the asset proportions dynamically, let the investment strategy in the first 18 years be fixed and for the sake of argument let this be assumed to follow the optimal multiperiod strategy derived for the simple closed fund earlier. It then follows that the only decision variables involved are the asset proportions in year 19. Under these circumstances, the optimal asset mix was found to be entirely in index-linked gilts. Therefore, it must be deduced that the anomaly is not caused by any limitation of the dynamic optimization procedure but is the result of some other influence.

Assuming that these results have been derived accurately, the only other explanation for this inconsistency must lie with the different objective functions used. In other words, the aim of minimizing the variance of real dividends in the final year may not be consistent with the aim of minimizing the variance of real accumulated dividends at the horizon date. As an immunized strategy is closely related to one which minimizes the variance of dividends over one time interval, this would imply that a strategy which minimizes the variance of payouts may also be inconsistent with an immunized strategy.

8.4.5 Multiperiod Consumption and Ultimate Surplus

Regardless of whether the deductions made in Section 8.4.4 had been justified adequately, there remains a fundamental problem with the results obtained for the closed fund. When accumulated dividends were expressed in real terms, the optimal strategy was to invest most of the fund in fixed income assets initially, while gradually switching more and more of the fund into index-linked gilts over time. This resulted in the optimal mix for the final year being predominantly invested in index-linked gilts. Nevertheless, the optimal mix for the first year had been well matched, which is consistent with a minimum variance strategy. Therefore, at least the decision at time 0 seemed reasonable, even if it had been based on the supposition that less intuitive decisions would be made at a later stage. However, the main limitation of this strategy is that it is incompatible with the strategies which will be optimal in the future.

For example, with just one year remaining, the optimal strategy for the simple closed fund was to invest a significant proportion of assets in cash. However, this seemed to contradict the optimal dynamic strategy derived at time 0, which found index-linked gilts to be the optimal asset class in the final year of operation. So not only has the optimal decision at time 0 been reliant on less intuitive strategies being adopted later on, but it has also been based on the assumption that certain decisions will be made in the future, decisions which may not actually be made at such a time.

In effect, the objective function itself is evolving as time progresses. The objective changes from being one of minimizing (at $t = 0$) the variance of accumulated dividends to one of minimizing (at $t = 18$) the variance of the final year's dividends. Hence, the concept of ultimate surplus is perhaps less robust than had been thought in Section 5.4. In this particular regard, time additive multiattribute utility functions may have been more suitable, although these would still have suffered from one or more practical limitations relating to non-constant levels of relative risk aversion, inconsistencies with present values and the inability to deal with negative valued outcomes.

8.5 Summary

It has been seen in this chapter how the decision-making process can be complicated considerably when analysed in a multiperiod setting. Under simplistic conditions, the powerful optimization technique of dynamic programming may be used to solve such problems. This process essentially reduces the multiperiod problem to a sequence of single period problems through backward induction. However, these single period problems need to have certain analytical properties in order for the whole procedure to work, placing restrictions on the types of problems that may be dealt with in this manner.

When faced with more complex situations, particularly in cases involving life office models, pragmatic approaches to dynamic asset allocation strategies may be more appropriate. These have usually focused on the reallocation of assets when the solvency position is relatively poor, either by switching into a more resilient or matched position, or by investing more heavily in gilts to increase the statutory solvency margin. However, these approaches have tended to be predetermined rather than optimal and do not tackle the issue of changing liability profiles over time.

Hence, the objective of the final part of this chapter had been to consider a dramatically changing liability profile by looking at the closed fund situation. A possible approach was to extend the static strategy, making the asset mix a function of some key variable such as the discounted mean term of the liabilities. Despite being very much at its developmental stage, the method appeared to be capable of yielding some quite intuitive results, especially by maintaining a roughly immunized position for a low risk strategy when payouts were expressed in nominal terms. When payouts were expressed in real terms, there was a strong tendency for the low risk strategy to be mis-matched towards the horizon date which could partly be attributed to the steeply rising free asset ratios observed towards the end of the projection period. This trend in free asset ratios was the result of the dividend policy used.

Investigations into a simple closed fund threw further light onto the issue. It was shown here that similar results could be obtained without many of the complexities built into the model office. From these investigations, it also became apparent that the objective of minimizing the variance of accumulated dividends was not necessarily consistent with one of minimizing the variance of one-period dividends and hence not necessarily consistent with one of achieving an immunized position.

9. CONCLUSIONS

9.1 Overview

In this thesis, it has been suggested that the optimal asset allocation strategies for a life insurance company may be determined in a rational decision-making framework. The approach used builds on methods which have been developed in financial economics by introducing complex liability structures into the analysis. It also enhances many of the studies carried out in the field of asset/liability modelling by attempting to optimize the investment strategies, from a static as well as a dynamic perspective.

As decisions were considered in the context of utility theory, the approach is general, and has been shown to be broadly compatible with many other existing techniques for investment decision-making, such as mean-variance analysis and immunization theory. In order to accommodate greater realism in these decision models, the results were computed using stochastic simulation techniques and numerical optimization routines. Hence, many of the restrictive assumptions usually associated with portfolio selection models could be relaxed.

9.2 Main Results

In relation to the simulations results produced by the investment model, the optimal asset mixes for a pure asset fund seemed intuitive. At low levels of risk tolerance over a one year time horizon, much of the optimal fund was in cash, while over a twenty year time horizon, the optimal fund was predominantly in Consols. When the fund had been expressed in real terms, the optimal proportions in index-linked gilts were seen to increase. At higher levels of risk tolerance, the optimal asset mixes were dominated by equities, which is reasonable given that this is the asset class with the highest expected return over any period. The results were generally consistent with a mean-variance approach, although the optimal portfolios at low risk tolerance levels were occasionally shown to lie away from the *E-V* efficient frontier. This was attributed to the inability of simulations to represent adequately the tails of the underlying distributions.

In the case of Model A, it was assumed that shareholders would provide additional capital to make up the guarantee fund when this had been breached, meaning that the objective function would be smooth. Consequently, the optimal portfolios were quite indifferent to the ruin probabilities produced. Initially though, it seemed less intuitive why the optimal portfolios at low levels of risk tolerance, such as 100% Consols, also had very high ruin probabilities. This was primarily due to the upper limit imposed on the valuation rate of interest used for calculating published reserves.

The results also showed that it was possible for the optimal portfolios at low risk tolerance levels to deviate from an immunized position, as the objective function had been defined in terms of accumulated dividends over a twenty year time horizon. Immunization, on the other hand, is only valid in respect of surpluses which arise as a result of a single movement in interest rates. In this sense, ruin probabilities were felt to be superior in assessing mis-matching risk: they are not restricted to situations involving a once and for all change in interest rates and reflect the risk position better when the valuation basis is complex. However, they cannot distinguish between the different extents to which an office may be insolvent. A suitable compromise between dividend performance and solvency may be achieved by placing ruin constraints on the utility maximizing portfolios.

With Model B, hardly any mis-matching was noticeable in the optimal portfolios at all levels of risk tolerance because winding up the business in the event of insolvency had a penalizing effect on the payouts. This, in turn, was due to the loss of future profits and the possibility that the office could be wound up under unfavourable terms when insolvent. Therefore, Model B often resulted in better matched and more diverse optimal portfolios than Model A. This pattern was also maintained when different liability profiles were assumed in Model B. In situations where the office was assumed to be writing mainly index-linked annuity contracts, the optimal portfolios were heavily biased in favour of index-linked gilts. Similarly, the discounted mean terms of the optimal portfolios tended to be shorter for offices issuing ten year endowments than for those issuing twenty year endowments.

When Model A was considered in a closed fund situation, the optimal asset proportions were defined as functions of the discounted mean terms of the liabilities. This was done as a means for allowing the asset mixes to change dynamically with the ever-maturing liability profile. At low levels of risk tolerance for nominal payouts, the optimal asset mix was largely in Consols at the start of the period and largely in cash at the end of the period of closure, which is consistent with immunization theory. Increasing the risk tolerance level encouraged higher proportions in index-linked gilts and equities which is also reasonable, although the proportions in these more risky asset classes were seen to increase as the horizon date approached. This affinity for higher yielding assets towards the horizon date was partly due to the rising free asset ratios seen during this period.

A similar trend was also observed when payouts had been expressed in real terms. Here, the optimal proportions in index-linked gilts for low levels of risk tolerance increased quite rapidly towards the end of the period of closure, even though cash would have seemed a better match. From simpler investigations performed in parallel with this, it appeared that the results were in fact accurate, in that these strategies did genuinely produce the most stable real payouts. However, it was also shown that investment in cash would result in more stable real dividends in the final year. Therefore, the anomaly was attributed to the difference between an immunized strategy and a strategy which sought to maximize the expected utility of real payouts at very low levels of risk tolerance.

9.3 Review of the Methodology

A logical step towards concluding this research would be to analyse the three stages of rational decision-making (assumptions, decision criteria and optimization) in the light of the investigations performed. In addition, it would also seem appropriate to include 'interpretation' as a final stage in the decision-making process. Where appropriate, suggestions will be made regarding possible extensions to this research.

9.3.1 Assumptions

The investment model used in this research has been based on the model developed by Wilkie (1986). Despite it being the most widely accepted stochastic investment model in the actuarial profession to date, many independent reviews of the model have drawn attention to levels of uncertainty associated with both its parameter estimates and linear structure. The instability of linear models over time was also evident when a model for cash yields had been fitted over the period from 1955 to 1993: this model implied an average yield of nearly 2% greater than when the same model structure had been fitted using data from 1923 onwards (see Wilkie, 1995a). The existence of such discrepancies illustrated the importance of interpretation when using any stochastic investment model.

This concern was heightened when the optimal portfolios obtained for the open fund were found to be very sensitive to the assumptions made in the investment model. Even when assuming the model structure to be correct, changes in parameter values caused by using different data periods resulted in significant alterations in the optimal portfolios obtained. This brought into question the credibility of the optimal proportions produced. However, as many of the deficiencies in Wilkie's model relate to problems which have yet to be resolved in the field of econometrics, uncertainty is likely to remain a crucial barrier in most practical applications of stochastic investment models.

In addition to the investment model, consideration should also be given the sensitivity of the results to the assumptions made in respect of the liability model. For example, the difference between Model A and Model B is in the assumed course of action when the fund is technically insolvent. From these investigations, the implications of the two models were seen to have a significant influence on the optimal strategies derived. For Model B, the discontinuity in the payout at insolvency produced more intuitive asset allocation strategies and appeared to override some of the deficiencies in the decision criteria. This implicit penalty on the objective function also took account of the *severity* as well as the incidence of insolvency, which would not have been achieved by simply imposing ruin constraints on the objective function. As a result, Model B was shown to be superior to Model A.

Similarly, it may be reasonable to expect that the terms on which the liabilities may be transferred, either at the point of ruin or at the horizon date, will have some impact on the results. The same could be said for assumptions relating to the distribution policy or the valuation bases employed. Hence, this research could be enriched by studying the possible effects which these and other assumptions may have on the optimal portfolios.

9.3.2 Decision Criteria

An advantage which utility theory has over mean-variance analysis is that it provides a means of assessing *which E-V* efficient portfolio should be chosen. However, the theory does not suggest how the decision-maker's utility function may be ascertained, which implies that the utility maximization approach could be quite subjective in practice. Nevertheless, utility theory can be a useful method of ensuring that consistent decisions are made in all situations. Having selected a particular level of risk tolerance, a life company which maintains this risk tolerance level by maximizing the expected utility of dividends should, hypothetically, be fulfilling its shareholders' reasonable expectations.

While it appears sensible that a utility function used in this context should exhibit constant relative risk aversion, the only utility functions with this property are the log function and the power function. These utility functions only operate over positive real numbers, which is consistent with the notion of utility being a function of wealth. However, there may be occasions, as in Model A, where it is expedient to assume otherwise. In these situations, utility functions without this property of constant relative risk aversion would have to be used.

There is, though, another practical limitation with both the logarithmic and the power utility function. In the case of Model B, it has been shown how the log function tends to yield optimal portfolios which appear similar to those using the exponential function at fairly high levels of risk tolerance. However, it was also shown in Section 2.3.2 that the power utility function was in fact more risk tolerant than the log utility function. Hence, the range of risk tolerance levels available in the exponential utility function would not generally be possible with either of these constant relative risk averse utility functions. In using the latter, one appears to be implicitly assuming high levels of risk tolerance.

A further difficulty arises in the case of a life fund because the multiperiod consumption of dividends needs to be taken into account. In much of finance literature, multiperiod consumption is usually dealt with by assuming the utility at each epoch to be time additive. Although the problem is reduced to one of discounting expected utilities, the results are generally inconsistent with the notion of present values. Furthermore, if the utility function does not have the property of constant relative risk aversion, the effective level of risk tolerance at each epoch will depend on the general level of surplus available. This would be a particularly serious implication in a closed fund situation.

Therefore, it was felt that the most suitable approach to the problem of multiperiod consumption would be to accumulate dividends to the horizon date and to treat these payouts as a form of ultimate surplus. As long as these accumulations were separated from the life fund, there would not be any danger of artificially high levels of surplus building up over time, thus giving a false picture of the solvency position. However, this approach led to some less intuitive results for Model A.

In an open fund, with payouts measured in nominal terms, the optimal low risk portfolios had durations which were too long in relation to the liabilities. Although the variability of the dividends in each year were more stable in portfolios with slightly shorter durations than this, the accumulated dividends were found to be more volatile. Analogous to this is the effect seen in the simulations produced by the investment model: while the annual return from Consols is more volatile than cash over one year, the opposite is true over twenty years. This feature is probably linked to the assumption that the Consols yield in the investment model is mean reverting, outlining the potential for misinterpreting the results.

With a closed fund situation, the optimal low risk portfolios when payouts were measured in real terms also seemed mis-matched in relation to the nominal liabilities. Significant proportions were invested in index-linked gilts towards the horizon date, even though large proportions in cash would have produced the most stable dividends in the final few years. Again, this highlighted the difference between the stability of dividends in each year and the stability of their accumulated amounts. But in the closed fund situation, this result is less acceptable than it is with the open fund. In minimizing the variability of real payouts, the apparent purpose of investing in more index-linked gilts near the horizon date is to reduce the contribution to this variability of real payouts from dividends arising in earlier years. However, when these earlier dividends are then paid, the investor is left with the decision criterion of minimizing the variability of accumulated dividends over the last few years, for which the optimal strategy is to invest mainly in cash rather than index-linked gilts. Hence, the optimal portfolios in earlier years are conditional upon a strategy at a later stage that may not be pursued when the time actually comes, which seems illogical.

The most important thing to recognize with these results is that the use of payouts in decision criteria may not be as robust as previously expected. Having said this, the incorporation of ruin criteria, through the use of constraints or discontinuities in the objective function, have been shown to alleviate this problem in the open fund situation. However, with a closed fund, the low ruin probabilities observed for the low risk strategies imply that such measures would have little impact on these strategies. Hence, the development of more pragmatically and theoretically sound approaches to multiperiod consumption would seem to be a fruitful area for further research.

9.3.3 Optimization

In most of the situations considered in this research, analytical methods were found to be inappropriate in dealing with more realistic circumstances. Therefore, numerical optimization algorithms had been applied instead. As far as numerical methods were concerned, optimization algorithms seemed to be the only feasible method for producing solutions to a high degree of precision. However, given how sensitive the optimization results were to the investment model parameter values, the need to obtain optimal portfolios with high levels of accuracy was deemed to be spurious.

With just four asset classes being considered in these investigations, the approximate method of testing a representative sample of all possible combinations of portfolios (referred to as the grid approach) was often shown to be just as useful. The approach is not dependent on the behaviour of the objective function, unlike gradient methods used in most optimization algorithms. Hence, discontinuities in the objective function and ruin constraints may be introduced with no additional effort when the grid approach is used. However, the grid approach would have been impractical in the case of a closed fund, where as many as seven decision variables were required to be optimized.

9.3.4 Interpretation

In the traditional approach to portfolio optimization suggested by Markowitz (1952), the selection procedure ended at the point where the optimum solution was obtained. The simplicity of mean-variance analysis in respect of one-period returns meant that detailed examination of the results was unnecessary. Moreover, the enormous number of securities involved would have made interpreting these results virtually impossible. The approach was treated as an objective means of making investment decisions.

However, with just four asset classes involved in these investigations, it has been feasible to take the decision-making process beyond the stage of computing the optimal strategies to one of actually analysing the results. Where liabilities were present, it was also possible to interpret the results against the background of existing theories such as immunization, and a conceptual understanding of the economic quantities of different assets. On the few occasions where these decisions were felt to be less intuitive, attempts were made at determining the assumptions which led to these results: thus providing feedback on the weaknesses of the model. However, the application of simulation techniques did make the task of interpreting the results quite difficult.

9.4 Theoretical and Practical Implications

One of the contributions of this research has been to establish a more unified framework in which investment decisions may be made. It links the approaches of utility theory, mean-variance analysis, downside risk measures and immunization theory more closely, and demonstrates how analytical and numerical solutions may complement each other. The work also furthers the theoretical developments of others by actually implementing these ideas in a simulated environment. This has led to the discussion of practical issues such as those relating to multiperiod consumption and dynamic strategies. Although the proposed solutions to these problems have generally been found to work satisfactorily, these approaches do leave some room for improvement. In particular, it would be worth extending the scope of dynamic strategies beyond duration based reallocations and developing an alternative to ultimate surplus which would lead to more consistent decisions over time.

From a practical perspective, portfolio selection models are often criticized because the results produced can be counterintuitive and highly sensitive to the assumption set used. Nevertheless, the majority of the results obtained in this research are perfectly intuitive, which seems to refute the first criticism. Furthermore, those which were less intuitive (as were all the results) had been a direct consequence of the assumptions and decision criteria adopted: the optimal decisions in any decision model will only be as meaningful as the set of inputs used. However, the sensitivity of the results to assumptions such as the investment model also means that the ultimate aim of obtaining the optimal solutions will not be possible.

Notwithstanding these limitations, investment decisions still need to be made in practice. If these decisions are to be made rationally, they will have to be based on all the information known about the variables concerned and the criteria which the investors must satisfy. This is precisely the approach adopted throughout this thesis.

Having said this it is vital that the extent of the uncertainty in the optimal strategies obtained is understood; and where the results are believed to be less intuitive, it is important that the factors which have led to these solutions are isolated. Although decision models may not be able to point the investor at the true optimum, they should at least be capable of increasing the investor to understand better the implications of the assumptions being made and the decision criteria used. Thus, this approach essentially epitomizes the actuarial methodology. Mathematical and statistical techniques may be used to solve financial problems but mathematical and statistical models can never describe or represent the financial problem in its full complexity.

REFERENCES

ALLAIS, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole americaine. *Econometrica*, **21**: 503-546.

BAILEY, A. (1862). On the Principles on which the Funds of Life Assurance Societies should be Invested. *Journal of the Institute of Actuaries*, **10**: 142-147.

BANK OF ENGLAND. (1995). *Bank of England Quarterly Bulletin*, **35**: 219-309.

BEALE, E. (1988). *Introduction to Optimization* - edited by Lynne Mackley (Chichester: Wiley)

BELLMAN, R. (1957). Dynamic Programming. (Princeton University Press).

BLAKE, D. (1990). *Financial Market Analysis*. (London: McGraw-Hill).

BOOTH, P. (1995a). The management of investment risk for defined contribution pension schemes. *Transactions of the 25th International Congress of Actuaries, Brussels, 1995*, **3**: 105-122.

BOOTH, P. (1995b). Utility theory and actuarial investment risk. City University Actuarial Research Report, (to appear).

BOOTH, P. and ONG, A. (1994). A Simulation Based Approach to Asset Allocation Decisions. *Proceedings of the 4th Actuarial Approach to Financial Risks International Colloquium, Orlando, 1994*, **1**: 217-240.

BOX, M. (1966). A comparison of several current optimization methods and the use of transformations in constrained problems. *The Computer Journal*, **9**: 67-77.

CHATFIELD, C. (1989). *The Analysis of Time Series - an introduction*, Fourth Edition. (London: Chapman & Hall).

CHATFIELD, C. (1995). Model Uncertainty, Data Mining and Statistical Inference. *Journal of the Royal Statistical Society*, A, **158**: 363-417.

CLARKE, H. (1954). A broad analysis of the problem of the investment of life funds. *Journal of the Institute of Actuaries*, **80**: 335-389.

CLARKSON, R. (1989). The Measurement of Investment Risk. *Journal of the Institute of Actuaries*, **116**: 127-178.

DAY, J. (1966). The Actuarial Contribution to Investment Thought. *Journal of the Institute of Actuaries*, **92**: 253-295.

DAYKIN, C. and HEY, G. (1990). Managing Uncertainty in a General Insurance Company. *Journal of the Institute of Actuaries*, **117**: 173-277.

EFRON, B. and TIBSHIRANI, R. (1993). *An introduction to the Bootstrap*. (London: Chapman & Hall).

ELTON, E. and GRUBER, M. (1991). *Modern Portfolio Theory and investment analysis*, Fourth Edition. (Wiley).

FELLOWS, D. (1981). Pension Fund Liabilities and Asset Matching. *Journal of the Institute of Actuaries*, **108**: 211-228.

FISHBURN, P. (1988). *Nonlinear preference and utility theory*. (Wheatsheaf).

GALLEN, M. and KIPLING, M. (1995). Recent developments in life office financial reporting. Presented to the Staple Inn Actuarial Society, March, 1995.

GEOGHEGAN, T., CLARKSON, R., FELDMAN, K., GREEN, S., KITTS, A., LAVECKY, J., ROSS, F., SMITH, W. and TOUTOUNCHI, A. (1992). Report on the Wilkie Stochastic Investment Model. *Journal of the Institute of Actuaries*, **119**: 173-228.

GILL, P., MURRAY, W. and WRIGHT, M. (1981). *Practical Optimization*. (London: Academic Press).

HANOCH, G. and LEVY, H. (1969). The Efficiency Analysis of Choices Involving Risk. *The Review of Economic Studies*, **36**: 335-346.

HARDY, M. (1993). Stochastic Simulation in Life Office Solvency Assesment. *Journal of the Institute of Actuaries*, **120**: 131-152.

HARDY, M. (1994a). Incorporating individual life company variation in simulated equity returns. *Proceedings of the 4th Actuarial Approach to Financial Risks International Colloquium, Orlando, 1994*, **3**: 1147-1162.

HARDY, M. (1994b). Simulating the Relative Insolvency Probability of Life Insurers. Presented to the 4th International Conference on Insurance Solvency and Finance, Philadephia, April, 1994.

HARVEY, A. (1989). *Forecasting, structural time series models and the Kalman filter*. (Cambridge University Press).

HUANG, C. and LITZENBERGER R. (1988). *Foundations for Financial Economics*. (New Jersey: Prentice Hall).

HUBER, P. (1995). A Review of Wilkie's Stochastic Investment Model. *Transactions of the 25th International Congress of Actuaries, Brussels, 1995*, **3**: 333-364.

KAHNEMAN, D. and TVERSKY, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica*, **47**: 263-292.

KITTS, A. (1988). Applications of stochastic financial models: a review. Department of Social Statistics, University of Southampton.

KITTS, A. (1990). Comments on a model of Retail Price Inflation. *Journal of the Institute of Actuaries*, **117**: 407-413.

KROLL, Y., LEVY, H. and MARKOWITZ, H. (1984). Mean-Variance Versus Direct Utility Maximization. *Journal of Finance*, **39**: 47-62.

LEVY, H. and MARKOWITZ, H. (1979). Approximating Expected Utility by a Function of Mean and Variance. *American Economic Review*, **69**: 308-317.

LINTNER, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, **47**: 13-37.

MACDONALD, A. (1991). On investment strategies using the Wilkie model. *Proceedings of the 2nd Actuarial Approach to Financial Risks International Colloquium, Brighton, 1991*, **3**: 413-427.

MACDONALD, A. (1993). What is the value of a valuation? *Proceedings of the 3rd Actuarial Approach to Financial Risks International Colloquium, Rome, 1993*, **2**: 723-743.

MARKOWITZ, H. (1952). Portfolio Selection. *Journal of Finance*, **7**: 77-91

MARKOWITZ, H. (1991). *Portfolio Selection - efficient diversification of investments*, Second Edition. (Cambridge, Mass.: Basil Blackwell).

MCKENNA, F. and KIM, Y. (1986). Managerial Risk Preferences, Real Pension Costs and Long-Run Corporate Pension Fund Investment Policy. *Journal of Risk and Insurance*, **53**: 29-48.

MERTON, R. (1971). Optimal Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory*, **3**: 373-413.

MILLER, J. and NEWBOLD, P. (1995). Uncertainty about Functions of Arima Model Parameters and Estimation of the Persistence of Economic Shocks. *Journal of Business and Economic Statistics*, (to appear).

MOORE, P. (1972). Mathematical Models in Portfolio Selection. *Journal of the Institute of Actuaries*, **98**: 103-130.

MOSSIN, J. (1968). Optimal Multiperiod Portfolio Policies. *Journal of Business*, **41**: 215-229.

NASH, J. (1979). *Compact Numerical Methods for Computers - linear algebra and function minimization.* (Bristol: Hilger).

ONG, A. (1994). A Stochastic Model for Treasury-Bills: An extension to Wilkie's Model. City University Actuarial Research Report 68.

PEGLER, J. (1948). The actuarial principles of investment. *Journal of the Institute of Actuaries*, **74**: 179-211.

PORTER, R. (1974). Semivariance and Stochastic Dominance: A comparison. American Economic Review, **64**: 200-204.

PRATT, J. (1964). Risk Aversion in small and large. *Econometrica*, **32**: 122-136.

PRELEC, D. and LOEWESTEIN, G. (1991). Decision making over time and uncertainty: a common approach. *Management Science*, **37**: 770-786.

PRESS, W., TEUKOLSKY, S., VETTERLING, W. and FLANNERY, B. (1992). *Numerical Recipes in C: The Art of Scientific Computing*, Second Edition. (Cambrige University Press).

PULLEY, L. (1983). Mean Variance Approximations to Expected Logarithmic Utility. *Operations Research*, **31**: 685-696.

PURCHASE, D., FINE, A., HEADDON, C., HEWITSON, T., JOHNSON, C., LUMSDEN, I., MAPLE, M., O'KEEFE, P., POOK, P. and ROBINSON, D. (1989). Reflections on Resilience: Some Consideration of Mismatching Tests, with Particular Reference to Long-term Insurance Business. *Journal of the Institute of Actuaries*, **116**: 241-268.

PYLE, D and TURNOVSKY, J. (1970). Safety-first and expected utility maximization in mean-standard deviation portfolio analysis. *The Review of Economics and Statistics*, **52**: 75-81.

REDINGTON, F. (1952). Review of the principles of life office valuations. *Journal of the Institute of Actuaries*, **78**: 286-340.

REID, D. and TEW, B. (1986). Mean-Variance Versus Direct Utility Maximization: a comment. *Journal of Finance*, **41**: 1177-1180.

REITANO, R. (1994). Non-Parallel Yield Curve Shifts and Immunization. *Proceedings of the 4th Actuarial Approach to Financial Risks International Colloquium, Orlando, 1994*, **1**: 427-456.

ROFF, T. (1992). Asset and Liability Studies on a With Profit Fund. Presented to the Staple Inn Actuarial Society, November 1992.

ROSS, M. (1989). Modelling a With-profits Life Office. *Journal of the Institute of Actuaries*, **116**: 691-716.

SAMUELSON, P. (1969). Lifetime Portfolio Selection by Dynamic Programming. *The Review of Economics and Statistics*, **51**: 239-246.

SAS Institute Inc. (1988). *SAS/ETS® User's Guide, Version 6*, First Edition. (Cary, NC: SAS Institute Inc.).

SCALES, L. (1985). *Introduction to Non-Linear Optimization.* (London: MacMillan).

SHARPE, W. (1963). A Simplified Model for Portfolio Selection. *Management Science*, **9**: 277-293.

SHARPE, W. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, **19**: 425-442.

SHERRIS, M. (1992). Portfolio Selection and Matching: A Synthesis. *Journal of the Institute of Actuaries*, **119**: 87-105.

TOBIN, J. (1958). Liquidity Preference as Behaviour Toward Risk. *Review of Economic Studies*, **26**: 65-86.

TSE, K., UPPAL, J. and WHITE, M. (1993). Downside Risk and Investment Choice. *The Financial Review*, **28**: 585-605.

TVERSKY, A. and KAHNEMANN, D. (1981). The framing of decisions and the psychology of choice. *Science*, **211**: 453-458.

VON NEUMANN, J. and MORGENSTERN, O. (1944). *Theory of Games and Economic Behaviour*, First Edition. (Priceton University Press).

WALSH, G. (1975). *Methods of Optimization.* (London: Wiley).

WILKIE, A. (1984). Steps towards a Stochastic Investment Model for Actuarial Use. Occasional Actuarial Research Discussion Paper 36.

WILKIE, A. (1985). Portfolio Selection in the Presence of Fixed Liabilities: A Comment on "The Matching of Assets to Liabilities". *Journal of the Institute of Actuaries*, **112**: 229-277.

WILKIE, A. (1986). A Stochastic Investment Model for Actuarial Use. *Transactions of the Faculty of Actuaries*, **39**: 341-403.

WILKIE, A. (1992). Stochastic Investment Models for XXIst Century Actuaries. *In*: *Transactions of the 24th International Congress of Actuaries, Montreal, 1992*, **5**: 119-137.

WILKIE, A. (1995a). The risk premium on ordinary shares. *British Actuarial Journal*, **1**: 251-293.

WILKIE, A. (1995b). More on stochastic asset models for actuarial use. *British Actuarial Journal*, (to appear).

WISE, A. (1984a). A Theoretical Analysis of Matching Assets to Liabilities. *Journal of the Institute of Actuaries*, **111**: 375-402.

WISE, A. (1984b). The Matching of Assets to Liabilities. *Journal of the Institute of Actuaries*, **111**: 445-485.

WISE, A. (1987a). Matching and Portfolio Selection: Part 1. *Journal of the Institute of Actuaries*, **114**: 113-133.

WISE, A. (1984b). Matching and Portfolio Selection: Part 2. *Journal of the Institute of Actuaries*, **114**: 551-568.

ZIEMBA, W. and VICKSON, R. (1975). *Stochastic Optimization Models in Finance*. (London: Academic Press).

APPENDIX A. DATA SOURCES

Retail Price Index

General Index of Retail Prices for December in each year:

1955-1993: *Employment Gazette*

Consols Yield

Yield on 2.5% Consolidated Stock at last day of each year:

1955-1993: *Financial Times*

Cash Yield

Average discount on 91 Day Treasury Bills calculated on last Friday of each year:

1955-1958: *The Bankers' Magazine*
1959-1977: *Bank of England Quarterly Bulletin*
1978-1993: *Financial Statistics*

Index-Linked Gilt Yield

Gross Redemption Yield on British Government Index-linked Stock with more than five years to maturity and an inflation assumption of 5% (monthly):

1986-1995: *Datastream*

APPENDIX B. DERIVATIONS

Expression for w^*:

Let w and $(1 - w)$ be the proportions invested in assets 1 and 2 respectively.

Let E_i and V_i be the respective means and variances of the accumulations from asset i.

Hence, the portfolio mean and variance will be:

$$\mu = A(w(E_1 - E_2) + E_2) \tag{B.1}$$

and

$$\sigma^2 = A^2(w^2(V_1 + V_2 - 2C_{12}) - 2w(V_2 - C_{12}) + V_2) \tag{B.2}$$

respectively.

Need to maximize: $E[U(.)] = -\exp(-\mu/r + \sigma^2/2r^2)$.

If $L = \log(E[U(.)]$, then:

$$\frac{dL}{dw} = -\frac{A}{r}(E_1 - E_2) + \frac{A^2}{2r^2}[2w(V_1 + V_2 - 2C_{12}) - 2(V_2 - C_{12})]$$

when $\dfrac{dL}{dw} = 0$,

$$w^* = \frac{(r / A)(E_1 - E_2) + V_2 - C_{12}}{V_1 + V_2 - 2C_{12}} \tag{4.1}$$

Expression for $E[U(Y)]^*$:

Substituting (4.1) into (B.1) and (B.2) yields:

$$\mu^* = \frac{r(E_1 - E_2)^2}{V_1 + V_2 - 2C_{12}} + A\left[\frac{(V_1 - C_{12})(E_1 - E_2)}{V_1 + V_2 - 2C_{12}} + E_2\right]$$

and

$$\sigma^2{}^* = \frac{r^2(E_1 - E_2)^2 + A^2(V_1V_2 - C_{12}{}^2)}{V_1 + V_2 - 2C_{12}}$$

respectively.

The expected utility at w^* is then:

$$E[U(.)]^* = -\exp\left(\frac{-(E_1 - E_2)^2 - (A/r)(V_2 - C_{12})(E_1 - E_2)}{V_1 + V_2 - 2C_{12}} - \frac{A}{r}E_2\right)$$

$$\times \exp\left(\frac{(E_1 - E_2)^2 + (A/r)^2(V_1V_2 - C_{12}{}^2)}{2(V_1 + V_2 - 2C_{12})}\right)$$

Therefore:

$$E[U(Y)]^* = -q\exp\left\{\frac{A^2(V_1V_2 - C_{12}{}^2) - 2Ar[(V_2 - C_{12})(E_1 - E_2) + E_2(V_1 + V_2 - 2C_{12})]}{2r^2(V_1 + V_2 - 2C_{12})}\right\}$$

where: $$q = \exp\left\{\frac{-(E_1 - E_2)^2}{2(V_1 + V_2 - 2C_{12})}\right\}$$