Journal of Actuarial Practice, Volume 4, No.1, 1996

Colin Ramsay, Editor

University of Nebraska - Lincoln, cramsay@unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/joap

Part of the Accounting Commons, Business Administration, Management, and Operations Commons, Corporate Finance Commons, Finance and Financial Management Commons, Insurance Commons, and the Management Sciences and Quantitative Methods Commons

http://digitalcommons.unl.edu/joap/64

This Article is brought to you for free and open access by the Finance Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Journal of Actuarial Practice 1993-2006 by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
ARTICLES

An Approach to Estimating Market Value and Duration of Interest-Sensitive Whole Life Contracts
Thomas J. Merfeld .............................................................................................................. 5

Participating GICs: Performance Attribution Analysis
Alec Stais and John P. Toohey III .................................................................................... 27

Asset Allocation in Investing to Meet Liabilities
Anthony Dardis and Vinh Loi Huynh................................................................................. 47

A Possibilistic Linear Programming Method for Asset Allocation
Lijia Guo and Zhen Huang ................................................................................................ 67

Nonmedical Limits in Individual Life Insurance
James B. Ross and Shalini Perumpral ............................................................................. 91

A Proposed New System of Financing Health Care in Singapore
Robert Keng Heong Lian and Loi Soh Loi ...................................................................... 115

Concentration in American Property-Casualty Companies
Edward Nissan ............................................................................................................... 131

Bias of Excluding High and Low Data for Long-Tailed Distributions
Cheng-sheng Peter Wu .................................................................................................. 143

Disclosure and Confidentiality Requirements of Corporate Pension Plan Actuaries
Theodore Konshak .......................................................................................................... 159

Discussion
Richard Daskais ............................................................................................... 169
Brian A. Jones .................................................................................................. 170
Author's Reply .................................................................................................. 172
EDITORIAL POLICY

The aim of this international journal is to publish articles pertaining to the "art" and/or "science" involved in contemporary actuarial practice.

The Journal welcomes articles providing new ideas, strategies or techniques (or articles improving existing ones) that can be used by practicing actuaries. In addition, the Journal of Actuarial Practice provides a forum for the presentation and discussion of issues of interest to actuaries. One of the goals of the Journal of Actuarial Practice is to improve communication between the practicing and the academic actuarial communities.

The Journal publishes articles in a wide variety of formats, including technical papers, commentaries/opinions, discussions, essays, book reviews, and letters. The technical papers published in the Journal are neither abstract nor esoteric; they are practical and readable. Topics suitable for this journal include the following areas:

- AIDS
- annuity products
- asset-liability matching
- cash-flow testing
- casualty ratemaking
- credibility theory
- credit insurance
- disability insurance
- expense analysis
- experience studies
- FASB issues
- financial reporting
- group insurance
- health insurance
- individual risk taking
- insurance regulations
- international issues
- investments
- liability insurance
- loss reserves
- marketing
- pensions
- pricing issues
- product development
- reinsurance
- reserving issues
- risk-based capital
- risk theory
- social insurance
- solvency issues
- taxation
- valuation issues
- workers' compensation

REVIEW PROCESS

A paper submitted first is screened by the editor for suitability. If it is deemed suitable, copies are sent to at least three independent referees. The name of the author(s) of the paper under consideration is anonymous to the referees, and the identities of referees are not revealed to the author(s).

The paper is reviewed for content, originality, and clarity of exposition. On the basis of the referee reports, the editor makes one of the following recommendations: (1) accept subject to minor revisions, (2) accept contingent on substantive revisions, (3) resubmit: return to the author(s) for major revisions and subsequent resubmission, or (4) reject.

The editor communicates the recommendation to the author(s) along with copies of the referee reports. The entire process is expected to take three to four months.

See back cover for instructions to authors.
An Approach to Estimating Market Value and Duration of Interest-Sensitive Whole Life Contracts

Thomas J. Merfeld*

Abstract†

A fixed premium interest-sensitive whole life contract is analyzed in order to estimate its market value. In addition, using various definitions of duration, we determine the duration of the contract for each definition. The results of this analysis have implications for market value accounting of life insurance liabilities and for life company portfolio management.

Key words and phrases: cash flow, accounting, assets, liabilities, surrender benefits, yield curve

1 Introduction

American life insurance companies traditionally have not assessed the market value of their liabilities. Life actuaries compute statutory reserves under strictly defined rules. Neither state regulatory bodies

*Thomas J. Merfeld, Ph.D., C.P.A., F.L.M.I., is vice president of asset/liability management for CIMCO, the registered investment advisor for CUNA Mutual Group. He has asset/liability management experience in mortgage-related securities, financial derivatives, and interest-sensitive products. He is a former chief financial officer and director of Savers Life Insurance Company. Mr. Merfeld was a senior financial economist at Franklin Savings Association and served as assistant research economist at the Federal Reserve Bank of Kansas City. His articles have appeared in Economic Review, Proceedings of the International Actuarial Association, and ARCH. He was a member of the American Academy of Actuaries Fair Value of Liabilities Task Force.

Mr. Merfeld’s address is: CUNA Mutual Group, 5910 Mineral Point Road, Madison WI 53705, USA.

†This paper is based on the author’s earlier article: "Market Value and Duration Estimates of Interest-Sensitive Life Contracts" that appeared in ARCH 1995.2. The author acknowledges useful input from Mr. Dale R. Oldham, F.S.A., and from the referees and the editor of the Journal.
nor the Financial Accounting Standards Board (FASB) have required or allowed market value liability reporting.

Until recently, neither the nature of the life insurance products sold nor the financial markets have made it necessary for companies to assess the market risk or the market value of their liabilities. Market values are sensitive to exogenous factors such as changing interest rates. For much of this century financial markets enjoyed enough stability that risk assessments were not necessary. Life products were relatively simple and margins relatively wide; insolvencies were generally not due to the lack of liability market risk measurements.

But needs have changed. The U.S. Congress is applying pressure on state regulators to tighten statutory reporting requirements. Financial reporting under Generally Accepted Accounting Principles (GAAP) is moving in a market value direction. Complicated new products funding complex new securities are accentuating the effects of the heightened financial market volatility of the last 15 years. The market value of life company surplus has become volatile enough that effective “insolvencies” can result. These insolvencies can be masked by traditional accounting standards, but are revealed by market value methods.

Life company financial management generally has not implemented the technology necessary to measure the risk of surplus volatility. Although statutorily defined cash flow testing procedures adequately reveal certain cases of financial ruin, they also can provide false indications of problems. Most valuation actuary certifications are designed to protect contract holders without regard to stockholders; as a consequence, these certifications constitute another test of solvency rather than of surplus volatility. Even internal portfolio managers remain too far removed from liabilities to apply established market value assessment and risk management procedures to liabilities.

Securities-type valuation and risk management methodologies can be applied to a particularly complex type of life product: the interest-sensitive whole life contract (ISWL). Over the past decade, American life companies have added these and similar universal life contracts with an aggregate face amount of approximately $1.5 trillion. As a result, interest-sensitive products represent a significant portion of life company obligations. It is not clear that either the literature or the industry has assessed the valuation or risk of this particular liability in a thorough way.

This article analyzes a fixed premium interest-sensitive whole life contract with the intent to:

- Develop an approach to estimating its market value; and
• Assess various notions of its duration.

The results of these analyses have implications for market value accounting life insurance liabilities and for life company portfolio management.

The FASB, through Statement of Accounting Standard (SFAS) 115, has taken dramatic steps to implement mark-to-market accounting for securities portfolios. The typical insurance company has responded by holding about 80 percent of its bond portfolio in the "available for sale" category, with annual changes in the market values reflected in GAAP surplus adjustments. The problem for insurance companies is that SFAS 115 does not provide for a parallel market value accounting of liabilities. As a consequence, a company can have no economic exposure to changing interest rates (i.e., be perfectly matched) and yet be reporting enormous changes in GAAP surplus from year to year. This hardly serves investor needs, which is one of the objectives of GAAP accounting. Some analysts call for a return to amortized cost accounting for bonds. A better approach may be to encourage the FASB to adopt market value accounting for insurance liabilities. This paper attempts to further the discipline of liability fair valuation.

Many product managers and portfolio managers attempt to duration-match their assets with their liabilities. These managers believe that interest rate risk is not adequately compensated for by markets. In the past life contracts were generally considered to call for a very long bond portfolio. That is not necessarily the case with interest-sensitive business, due to the resetting feature of the crediting rate. Having developed a methodology to estimate the fair value of liabilities, we can then re-estimate their values after shocking market rates. The relationship can then inform portfolio managers on how to invest bonds to match the value sensitivities.

2 The Contract

2.1 Provisions

ISWL contracts are similar to traditional whole life contracts. Whole life products are structured to produce a level premium at any issue age for a large pool of insureds. This premium can be regarded as the periodic contribution to a fund that, when accumulated at an implicit rate of interest, is expected to produce periodic death benefits and is exhausted at the end of the pool of lives. This condition holds after allowing for loadings for coverage of company issuance and maintenance
expenses and a competitive return to capital. An ISWL contract initially is structured with a level premium.

ISWL contracts differ from traditional whole life contracts in two ways. First, although the contracted premium must be paid in order to avoid nonforfeiture status, ISWL premiums can have a more complex structure. Much of this complexity can be considered as discrete contract owner options to adjust the mix of premium paid and coverage amount within constraints imposed by the issuer and tax code. The contract owner uses private information in exercising these options. The issuer can respond by adjusting the cost of insurance assessment within contractual and legal constraints. In a yet more general type of permanent life contract, universal life (i.e., flexible premium interest-sensitive life), the contract owner effectively has continuous options to adjust the premium/coverage mix.¹

Second, in an ISWL contract the accumulation interest rate is explicit and adjustable. Whatever rate structure the issuer originally used to price the premium structure, the contract fund value is credited with a rate that resets periodically at the option of the issuer, subject to a floor.

ISWL contracts also allow the contract owner to remove the accumulated fund value by:

- Taking a loan (which continues to commit the policy owner to the premium schedule and can expose the loaned portion of the fund value to a lower crediting rate);
- Making a partial withdrawal; or
- Canceling the entire contract (which typically exposes the policy owner to a schedule of surrender assessments).

ISWL contracts also provide the standard array of nonforfeiture options, ignored in this analysis. See the appendix for a summary of terms and in force assumptions of the ISWL contracts analyzed.

2.2 Cash Flow Components

Contract cash flows belong to four classes:

- Premium inflow;
- Surrender benefit;

¹The application of game theoretic modeling to this field could provide rich insights into the optimum use of this option.
• Death benefit; and
• Servicing expense and commission.

Projecting surrender benefit cash flows entails certain complexities. First, the industry comparable product crediting rate must be modeled. Second, the company crediting strategy must be established. Third, a function of contract holder surrender behavior with respect to these differential rates must be integrated into the overall model.

2.2.1 Industry Interest-Sensitive Life Crediting Function

Industry interest-sensitive life crediting rates track fixed income market rates with lags for several institutional reasons:

• Company investment committees often establish crediting rates as a spread to the asset book value earnings rate. Once a particular asset is on the books at a given rate of return for purposes of accruing interest income, its rate of return will be assumed to remain constant until those particular assets are rolled into a new security with a new rate of return. This practice delays the recognition that market rates have changed;

• The institution then needs to declare the new rate;

• Companies are reluctant to change rates too often for marketing purposes; and

• The effective crediting rate can be delayed further due to certain contractual provisions allowing new rates to be credited only at the policy’s anniversary; hence, contracts containing this provision have a built-in expected delay of six months.

The crediting rate series comes from monthly data from January 1985 through December 1992 from the TULAS industry survey published by the actuarial consulting firm Tillinghast. The period covers a substantial fall in rates and changes in yield curve shape.

Let \( ICR_t \) be the industry crediting rate at month \( t \), \( T3M_t \) be the three month risk-free bill rate at month \( t \); and \( T5Y_t \) be the five year risk-free note rate at month \( t \). The subscript \( t - m \) refers to the rate \( m \) months prior to the interest-sensitive life crediting rate. A linear regression with

\[ \text{For more information of the TULAS industry survey, write to: Tillinghast, 245 Park Avenue, 18th Floor, New York NY 10167-0128, USA.} \]
ICR_t as the dependent variable and T3M_{t-6}, T5Y_t and T5Y_{t-18} yields the following equation:

\[
ICR_t = 3.78 + 0.119 \times T3M_{t-6} + 0.342 \times T5Y_t + 0.146 \times T5Y_{t-18} \quad (1)
\]

with \( R^2 = 0.72 \). Equation (1) is not useful in describing the response of the crediting rate to a change in market rate because the nonconstant term coefficients sum to significantly less than one. For example, the model indicates that the crediting rate would never fully respond to a permanent parallel shift in rates.

The following is a similar model, fit without the constant term, and weighs recent observations more heavily. This regression yields the following equation:

\[
ICR_t = 0.424 \times T5Y_t + 0.297 \times T5Y_{t-6} + 0.328 \times T5Y_{t-18} \quad (2)
\]

with \( R^2 = 0.99 \). In this case equation (2) also appears to give a good fit; in addition, because its coefficients sum nearly to one, the equation merely describes the timing by which the crediting rate will reflect changed market rates.

2.2.2 Company Crediting Strategy

The crediting rate structure that a company establishes is the result of myriad marketing, investment, legal, and game theory considerations. Crediting rate rationale is not well understood from a theoretical perspective. Companies attempting to establish market share often credit above-market rates. Companies often engage in "bait and switch tactics" (i.e., crediting above-market rates on new money and below-market rates at subsequent policy years). When a company has closed a block and is running off the reserves, it often credits low rates. Company crediting rates are often not closely related to the company's claims-paying capacity. For this analysis the rate is set at the industry rate less 125 basis points.\(^3\)

2.2.3 Contract Holder Surrender Profile

Contract holder behavior is affected fundamentally by the spread between the rate earned by a particular contract and the rate provided

\(^3\)One basis point is the smallest measure used in quoting yields on mortgages, bonds and notes. It represents 0.01% of yield. Thus a 5.77% rate that drops by 125 basis points becomes 4.52%.
by a similar contract issued by an industry competitor. Surrenders also are relatively higher during the early stages of the contract. Surrender behavior can be volatile at critical points in the life of the contract, such as premium redetermination periods and the end of the surrender charge period. Furthermore, it is affected by critical ages, such as retirement, of the contract holder.

The analysis assumes baseline surrenders of 7.5 percent annually with adjustments for:

- Early policy years due to buyer remorse;
- Later policy years, when surrender charges are lower;
- Higher insured ages, when contract owners typically have a postretirement need for cash and when they find surrender to be tax-efficient; and
- Policy-crediting rates that are different from rates on policies of competitor companies. Over the projection horizon the crediting rate is set at the projected industry rate less 125 basis points. The low crediting rate used in this analysis produces moderately higher surrenders.

To date, little empirical analysis has been done on contract holder surrender profile. There are significant barriers to conducting such research:

- Lack of product homogeneity;
- Differences in sales methods;
- Differences in product management;
- Differences in perceived insurance company risk; and
- Absence of usable historical data maintained by companies.

Contrast this with the extensive analytic work that has been conducted on the prepayment behaviors of single family homeowners. The mortgage is relatively a homogeneous product. Good data are available for at least 15 years, and there are few company-specific issues related to them.

This asymmetric behavior reveals the fundamental interest rate option that the contract holder owns. When market rates rise there is the potential problem of disintermediation, i.e., the insured can surrender
to pursue higher new money rates elsewhere (subject to a declining surrender charge). Consequently, the company needs to choose between raising the crediting rate to keep the policy or to maintain the rate and cash out the policy. In the former case the company pays more interest over time, perhaps even a greater rate than it earns on its investment. In the latter case the company fails to recoup a portion of its acquisition cost and tends to liquidate bonds at depressed prices.

If rates fall, however, the results are not symmetrically bad for the contract holder. The insured merely keeps the policy knowing that the company can reduce its rate only to the guaranteed rate.

The value of this interest rate option is affected by the same two fundamental variables that affect all other interest rate options: interest rate volatility and time to expiry. Greater market rate volatility provides a greater range of opportunity for the insured to exercise. New money rates can become higher, providing more incentive to pursue them. They also can be driven lower, increasing the chance that the insured enjoys a guaranteed rate in excess of new money rates. In addition, these contracts often have long lives, which provides further opportunity for rates to move to the insureds' advantage. The effect of this option on insurance company surplus is exacerbated by the likelihood that the company has written similar options in its mortgage-backed security and callable bond portfolios.

This analysis uses a quadratic function to derive excess surrenders (i.e., surrenders in excess of baseline) stimulated by the contract holders' desire to earn higher rates. Thus

\[ ES_t = (1 + \rho \times TD_t)^2 - 1 \]

where

- \( ES_t \) = Excess surrenders at time \( t \);
- \( \rho \) = Propensity to achieve higher rates through surrender;
- \( TD_t \) = \( \max \left\{ \left[ CR_t - PR_t - 0.0115 - \frac{SC_t(1 + PR_t)^{-SP_t}}{SP_t} \right], 0 \right\} \)
- \( CR_t \) = Identified competitor's credit rate at time \( t \);
- \( PR_t \) = Product credit rate at time \( t \);
- \( SC_t \) = Remaining surrender charge at time \( t \);
- \( SP_t \) = Number of surrender charge periods remaining at time \( t \).

Note that the rate of differential must exceed 115 basis points (0.0115) before any excess surrenders occur, and that the remaining surrender
charge is then amortized off. This paper uses the 125 basis points spread between the competitor rate and the product rate. The effect is to moderate excess surrenders in most forward rate and surrender charge environments.

Surrender cash flows include new policy loans. Policy loans typically do not pay cash interest and, in any case, are usually not repaid. Instead, interest is capitalized, increasing the loan amount. New policy loans are modeled as a premature cash outflow from the fund. When the contract is extinguished by surrender, the surrender benefit is reduced by the amount of the loan; when it is extinguished by ultimate death of the insured, the death benefit is reduced by the amount of the loan.

We assume future loan balances to be a constant proportion of contract cash values. Under older whole life policies the policy loan was a further powerful option enjoyed by the contract holder. In high rate environments the insured withdrew the policy's cash value at a rate fixed at policy inception and invested in high market rates elsewhere. Under newer interest-sensitive contracts the loan rate depends on market rates. Consequently, there may be no further benefit to taking the loan in high or low rates. In high interest rate scenarios, however, cash values increase quickly, loan balances capitalize quickly, and cash paid upon death of the insured is reduced.

In high interest rate scenarios our model shifts cash flows to earlier periods and from the death benefit component to the surrender benefit component. This modeling assumption does not distort aggregate contract duration, but may affect the respective durations of the surrender component and death benefits component.

### 2.2.4 Cash Flow Generation

Various interest rate scenarios are generated in estimating market values and durations. Each scenario is associated with a yield curve. Under each scenario for each projected month premium cash flows into, and benefit and expense cash flows out of, the fund are generated for each contract remaining in force. In addition, the fund accumulates by capitalizing interest at the credited rate and is reduced by the cost of insurance and expense charges deducted for each policy in force. Industry and company crediting rates are functions of medium term forward rates implied by the scenario yield curve. Contract surrenders are triggered, generating surrender cash flows. Finally, mortality cash flows are generated, further reducing the number of policies in force. The process is repeated for each month and each yield curve.
2.2.5 Duration

The market value of financial instruments is sensitive to interest rates at various points on the yield curve. The concept of duration provides a measure of the extent of this sensitivity. The basic duration measure is the Macaulay duration; see, for example, Boyle (1992, Chapter 3). It is defined as follows for an arbitrary series of cash flows $C_k$ paid at time $k$, $k = 1, 2, \ldots, n$ with yield-to-maturity rate of $y_n$

$$\text{Macaulay Duration} = \sum_{k=1}^{n} \frac{k C_k (1 + y_n)^{-k}}{\sum_{k=1}^{n} C_k (1 + y_n)^{-k}}. \tag{3}$$

For example, a one year Treasury bill has only one cash flow; its duration is explained by what happens to the one year rate. A five year investment grade corporate bond's duration is explained largely by the five year rate, but also by the change in rates at the same term as coupon payments. A mortgage-backed security can be affected by medium term rates because its cash flow is centered in that region, and also because refinance rates, which affect the amount and incidence of cash flow itself, are centered there. Interest-sensitive life contracts can have highly complex cash flow magnitude and incidence of cash flow profiles.

In general, the duration of a financial instrument can be calculated as follows: let $P_0$ be the current price of the financial instrument, $P_+$ be the price after a very small increase of $y_+$ in the interest rate, and $P_-$ be the price after a very small decrease of $y_-$ in the interest rate. The duration $D$ is defined as:

$$D = \frac{P_- - P_+}{(y_+ + y_-)P_0}. \tag{4}$$

For example, if a bond currently trading at par (i.e., $P_0 = 100$) is subjected to an instantaneous rate increase of one half of one percent ($y_+ = 0.005$) and its price drops to 98 ($P_+ = 98$), while if it is subjected to an instantaneous rate decrease of one half of one percent ($y_- = 0.005$) and its price increases to 102 ($P_- = 102$), then $D = 4$.

Duration increases as the time to maturity increases, other things being equal. Floating rate instruments usually have quite short durations. Although duration was initially introduced for fixed income instruments, all financial instruments have a duration as defined in equation (2).
3 Market Value Estimate

Consider a comparison to corporate debt whereby the accumulating ISWL fund represents the proceeds from company debt issuance in the same way as an industrial firm can issue debt as part of the financing structure for a capital project. In either case the issuer is accumulating assets with the promise to repay the debt, with interest, at a future date. The issuing firm (in the case of debt issuance) and the life company (in the case of contract issuance) are charged a certain number of basis points as a risk premium by entities providing the funds. Issuers in both cases incur issuance costs and enjoy related tax advantages. Because the ISWL contract provides risk-reducing services to the provider of funds (the contract holders), some or all of such cost of funds is offset. The insurance company is rewarded for providing risk intermediation services. Consider the aggregate of these costs and their offsets to be a cost of funds basis point spread to Treasury.

The spread can be estimated for a line of business or a block of policies at any time in its life. To accomplish this, solve for the spread that when added to the respective risk-free zero coupon rates discounts all future expected contract cash flows to the market value of funds that insureds willingly provide to the insurer. Immediately prior to issuance such assets equal zero; after the initial premium such assets equal the premium reduced by the commission and other acquisition costs incurred in issuance; at any time such assets equal the sum of all net cash flows and their investment earnings at risk-adjusted rates.

Let \( CF_{c,t} \) be the cash flow of component \( c \) at projection month \( t \) where \( c \) is a nominal integer parameter taking possible values from 1 to 4 (1 = premium, 2 = surrender benefit, 3 = death benefit, and 4 = expense), and \( t = 1, 2, \ldots, 480 \) (that is, 40 years). Thus, we are interested in projecting cash flows component by component over 40 years. The cash flows can depend on the level and shape of the yield curve at any projection month, and each month's discount rate also can be different. This component by component approach is consistent with pricing of mortgage-backed securities and derivatives that have been in place for about a decade. See, for example, Roll (1988).

Let \( z_t \) be equal to the risk-free zero coupon rate at projection month \( t \), and \( \xi_{k,t} \) be an instantaneous shift in the zero coupon rate at time \( t \) for key rate \( k \), \( k = 1, 2, \ldots, 9 \). The shift is applied to \( z_t \) at all \( t \) near the term of a Treasury on-the-run instrument. Finally, let \( s \) be the spread. In this analysis the resultant presumed spread is set to 20 basis points for the entire projection period. Using the semi-annual coupon convention of the investment literature, let \( d \) be the market price discounting factor.
such that, at any month $t$,

$$d_t = \left(1 + \frac{Z_t + \xi_{kt} + s}{2}\right)^{-t/6}. \quad (5)$$

A simple procedure provides the market value ($MV$) of ISWL,

$$MV = \sum_{c=1}^{4} \sum_{t=1}^{\infty} CF_{c,t} d_t. \quad (6)$$

This analysis uses 480 months as a practical analog for infinity. Applying the methodology produces Table 1. Because this market value measure is that of a liability, cash flows of the premium component have a negative sign.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Base Market Value ISWL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
<td>Market Value</td>
</tr>
<tr>
<td>Premium</td>
<td>(71,687,488)</td>
</tr>
<tr>
<td>Surrender Benefit</td>
<td>53,262,613</td>
</tr>
<tr>
<td>Death Benefit</td>
<td>34,190,301</td>
</tr>
<tr>
<td>Expense</td>
<td>7,852,738</td>
</tr>
<tr>
<td>Aggregate</td>
<td>23,618,164</td>
</tr>
</tbody>
</table>

A few points are worth noting about these values relative to the various in force values in the appendix:

- The aggregate market value is only about 50 percent of the statutory reserve, reflecting the conservative nature of statutory accounting principles;
- The surrender benefit is even higher than the fund or cash values due primarily to the inclusion of future premiums in this component; and
- The death benefit is only about 5 percent of face amount, indicating the impact of withdrawals and the time value of money.

4 Notions of Duration

In general duration is a measure of the sensitivity of a stream of cash flows to an interest rate change. Different notions of interest rate shifts
are depicted graphically in Figures 1 and 2. Figure 1 shows parallel and nonparallel (i.e., dampened) shifts from a base case yield curve. Figure 2 shows a piecewise dampened shift. Each shift produces a different sensitivity, or duration, of the present value of the product cash flow.

Consider the following four concepts of duration.

**Spread Duration:** Spread duration is the percent sensitivity of market value to a shift in $s$. This is similar to Macaulay duration and modified duration which are measures of sensitivities to changes in yield for fixed and certain cash flows. It is useful as an average life index and as a measure of the capitalized value of a unit of higher cost of funds. Mathematically,

$$ Spread Duration = \frac{1}{MV} \frac{dMV}{ds}. \tag{7} $$

**Parallel Effective Duration:** Parallel effective duration is the percent sensitivity of market value to a parallel shift in $z$ simultaneously over all $k$. That is, $\xi_{k,t}$ are equal at all $k$. This is the type of duration customarily used in asset and liability matching. Mathematically,

$$ Parallel\ Effective\ Duration = \frac{1}{MV} \frac{dMV}{dz}. \tag{8} $$

**Dampened Effective Duration:** Dampened effective duration is the percent sensitivity of market value to a function $f$ that produces nonparallel shift vector in $z$ over all $k$ such that: $\xi_{k,t} > \xi_{k+1,t}$. Rates for months within a region also are shifted according to the nonparallel function $f$. This is similar to effective duration, but recognizes that short rate volatility typically exceeds long rate volatility. In this analysis the average shift over the region $k = 9$ is about 60 percent of the average shift over the region $k = 1$. Mathematically,

$$ Dampened\ Effective\ Duration = \frac{1}{MV} \frac{dMV}{d\tilde{z}} \tag{9} $$

where the vector $\tilde{z}$ is produced according to $f$.

**Key Region Effective Duration:** Key region effective duration is the percent sensitivity of market value to a set of shifts in $z$ sequentially over each $k$. Measures under either of the following two subsets of this notion can be considered partial durations:
Figure 1
Base and Shifted Yield Curves

Parallel: This duration measures shifts by region of the curve. Mathematically,

\[
Parallel\ Duration = \frac{1}{MV} \frac{dMV}{d\xi_{k,t}}. \tag{10}
\]

The methodology developed in this paper provides a duration, termed piecewise duration, to each significant point on the yield curve. It effectively decomposes overall duration to assess the impact of changes of pieces of the yield curve. For example, the duration of a five year bond is about 4.2 percent; about 70 percent of this duration can be attributed to the five year rate. For instruments with fixed cash flows, piecewise durations approximately sum to the aggregate durations. For instruments whose cash flows are interest contingent piecewise durations may sum to a different number. The piecewise (region-by-region) shifts in the yield curve are similar to Ho’s (1992) key rate duration methodology. This analysis shifts entire segments of the yield curve by cutting and excising
these portions. This happens over all $t$ within a given region $k$. Ho uses a series of triangle functions in shifting Treasury on-the-run rates. Both methodologies create kinked undifferentiable curves. The curves used in this paper are also discontinuous. This key region methodology, however, improves the performance of the aggregation of piecewise duration.

**Dampened Duration**: Dampened duration is similar to dampened effective duration, but relates to the effect of dampened shifts by region of the curve. Mathematically,

$$Dampened\ Duration = \frac{1}{MV} \frac{dMV}{d\xi_{k,t}}.$$  \hspace{1cm} (11)

**Figure 2**
Base and Shifted Yield Curves: Key Rates
5 Duration Calculations

Consider point estimates of these measures shown in Table 2. These figures are scaled to a 100 basis point shift. Under dampened measures figures are scaled to reflect an approximately 100 basis point shift over the region \( k = 1 \). We show only the dampened category of key region measurements.

<table>
<thead>
<tr>
<th>Component</th>
<th>Parallel Spread</th>
<th>Effective Spread</th>
<th>Dampered Effective Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premium</td>
<td>-5.94</td>
<td>-5.94</td>
<td>-4.94</td>
</tr>
<tr>
<td>Surrender Benefit</td>
<td>10.86</td>
<td>3.78</td>
<td>3.55</td>
</tr>
<tr>
<td>Death Benefit</td>
<td>14.15</td>
<td>7.93</td>
<td>6.51</td>
</tr>
<tr>
<td>Expense</td>
<td>5.49</td>
<td>5.49</td>
<td>4.59</td>
</tr>
<tr>
<td>Aggregate</td>
<td>28.79</td>
<td>3.82</td>
<td>3.95</td>
</tr>
</tbody>
</table>

Premium duration figures are provided with respect to their effect on the value of the liability. For example, at a higher discount spread of 100 basis points the market value of premium will become less negative, producing a higher overall liability. In this sense premium has a negative duration as well as a negative market value.

Consider also the more intuitive market value change by component shown in Figure 3. See also the aggregate market value changes by effective and dampened effective-type shifts shown in Figure 4. Each of the figures shows a market value change from the base case resulting from the indicated shift.
6 Discussion

Estimating the cost of funds is a complex process. The rate credited to the fund accumulations explains less of such cost than do analogous rates in depository institution or corporate bond environments. A simplistic comparison between a life company's asset earning rate and its crediting rate provides inadequate information with respect to its net spread. Such comparison in a depository context, however, provides much more significant information. Market value approximations for complex insurance liabilities, however, can be made. Actuaries can provide market value estimates using the ideas presented in this paper.
The death benefit component is highly positively convex\(^4\) due to the mere length of the cash flows and to the effect of policy loans. Our model of policy loans causes the death benefit component to have greater absolute cash flows in lower interest rate environments, a combination sure to produce convexity. Liability convexity is not desirable. The surrender benefit component is also highly positively convex due primarily to the presence of the guaranteed rate in the contracts, which acts as an interest rate floor. The effective duration of the surrender benefit component, primarily a floating rate set of cash flows, is long due to the lagged nature of adjustments to crediting rates.

\(^4\)The concept of convexity refers to the extent that the price function changes over greater yield intervals. A financial instrument is said to be positively (negatively) convex when its duration increases (decreases) as the yield decreases. Positive convexity is desirable for assets but not for liabilities.
Financial managers need to decide whether to use the dampening notion of effective duration or whether to use the strictly parallel approach. The alternative methods can give significantly different results. Long-tailed cash benefits and the customary front-loaded premium structure of whole life combine to produce a net asset at the beginning of contract time and a net liability at the end of such time. The present value of this configuration is subject to yield curve twists such as are emphasized in the difference between effective and dampened effective durations.

Spread duration, indicating average lives of cash flow components, are long with respect to the following components:

- Surrender benefits, as interest is capitalized rather than paid currently on the surviving contracts;
- Death benefits, because as the insureds of surviving contracts grow older, cash flow increases; and
- Aggregate, because it is leveraged, in a sense, by negative premium flows in earlier periods. The leveraged characteristic of this component renders its spread duration a less reliable measure of cash flow life. It remains a measure of the value of a basis point change in discounting spread.

Spread and effective durations are equal for fixed cash flow components such as premium and expense, and are different for interest rate-sensitive cash flows such as surrender and death benefits.

Dampened effective durations and aggregate dampened effective partial durations are equal for fixed cash flow components and different for cash flow components. Premium cash flows are fixed over different scenarios even though surrender cash flows change. We have fixed the incidence of surrender in the analysis by leaving the crediting strategy unchanged—hence the constant premium cash flows. On the other hand, the amount of surrender cash flows can change dramatically under different forward rate scenarios.

Because the partial durations with respect to interest-sensitive cash flows are relatively high compared to an equivalent dampened effective duration (in the context of Reitano (1990)), it appears that the interest-sensitive nature of the cash flows has the effect of adding durational leverage to the component. Other elements capable of causing the effect include negative partial durations, portfolios of securities, and short positions. Option-adjusted spread methodology would provide yet additional texture to the duration results. Additionally, analogous
shocks can be made in mortality and surrender scenarios to reveal additional risk measures.

A further consideration is the duration of cash flows distributable to the stockholders. Under life insurance accounting conventions in most states, dividends generally can be paid without restriction to the extent of the prior year statutory earnings. In this case statutorily defined accounting and legal conventions actually drive a cash flow stream. In most cases these conventions have a significant historical cost component that stabilizes earnings. As a result, intracompany durations and the durations of returns to the ownership rights to such company may diverge substantially. Management and stockholders rarely synthesize an optimal duration position.

7 Conclusions

Life insurance companies can adapt market pricing methodologies that customarily are used in the financial area to the cash flow provisions of their particular contracts. Adapting such methodologies can provide rich insight into the market value of liabilities. These modeling methodologies can guide investment strategies and provide management with measures of the market value of life company surplus.

References


Appendix—Hypothetical ISWL Contract Terms and In Force Assumptions

Following is a sample cell of the ISWL model used in the analysis.

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td>Weighted Average Issue Age</td>
<td>32</td>
</tr>
<tr>
<td>Smoking Classification</td>
<td>Nonsmoker</td>
</tr>
<tr>
<td>Premium Mode</td>
<td>Monthly</td>
</tr>
<tr>
<td>Underwriting Class</td>
<td>Nonmedical</td>
</tr>
<tr>
<td>Weighted Average Policy Age</td>
<td>10 Years</td>
</tr>
<tr>
<td>Number of Policies</td>
<td>70</td>
</tr>
<tr>
<td>Basic Annual Premium</td>
<td>$5.60/$1000</td>
</tr>
<tr>
<td>Total Face Amount (Policy yrs 1-10)</td>
<td>$7,605,500</td>
</tr>
<tr>
<td>Total Face Amount (Policy yrs 11+)</td>
<td>$6,844,950</td>
</tr>
<tr>
<td>Total Fund Value</td>
<td>$266,633</td>
</tr>
<tr>
<td>Total Cash Value</td>
<td>$226,639</td>
</tr>
<tr>
<td>Total Statutory Reserve</td>
<td>$250,740</td>
</tr>
<tr>
<td>Total Policy Loan</td>
<td>$22,664</td>
</tr>
</tbody>
</table>

Common to all 102 cells of the analysis are a guaranteed crediting rate of 5.5 percent; a 15 year declining surrender charge schedule as a percentage of existing fund value; a provision whereby interest crediting on loaned balances can be less than on unloaned balances; a specified agent commission schedule; and a standard servicing expense structure. The descriptive statistics for the insured groups, weighted by face amount, are: (i) issue age = 41.5, and (ii) months seasoned = 55.0.

Other beginning characteristics of the model contract are a face amount of $709,585,000; cash value of $36,348,600; fund value of $50,426,020; statutory reserve of $47,150,500; annualized premium in projection month one of $5,886,800; and 6,951 policies. The analysis assumes a simple mortality experience equal to 80 percent of the 1975-80 basic select and ultimate mortality table.
Participating GICs: Performance Attribution Analysis

Alec Stais* and John P. Toohey III†

Abstract

The increasing popularity of participating GICs has created a need for an objective understanding of their performance. The fixed income attribution techniques are not adequate for measuring participating GIC performance because they typically restrict performance measurement to concepts such as duration management, sector rotation, and issue selection. We develop an attribution technique based on four components or effects that are helpful in explaining the changes in credited rates. They are the constant duration effect, the reinvestment effect, the cash flow effect, and the investment effect. The underlying mathematical approach to calculating these effects is presented along with examples.

Key words and phrases: investment, duration, yield, spread, cash flow, investment manager

* Alec Stais, A.S.A., C.F.A., is the GIC product specialist for Goldman Sachs Asset Management. In addition, he is responsible for client services for their insurance and financial institution clients. Prior to joining Goldman Sachs & Co., Mr. Stais was a managing director at MetLife Investment Management Corporation where his responsibilities included product development and client servicing. He received his B.A. from Haverford College in 1982 and his M.B.A. from New York University in 1986.

Mr. Stais' address is: Goldman Sachs Asset Management, One New York Plaza-41st Floor, New York NY 10004, USA.

† John P. Toohey III, F.S.A., is a senior actuarial associate in the individual business-annuity pricing area at MetLife. He received his B.A. in mathematics, magna cum laude, from Williams College in 1990. Since joining MetLife in 1990, his assignments have included risk management, pricing and product development for various GIC products and analysis of the impact of taxes on pricing and financial results.

Mr. Toohey's address is: MetLife, One Madison Avenue, New York, NY 10010-3690, USA. Internet address: jptoohey@aol.com
1 Introduction

Defined contribution plans often offer a stable value fund as an investment alternative. Stable value funds provide a fixed credited rate and a guarantee of principal to participants. Plan sponsors historically have invested in nonparticipating insurance company contracts called GICs (guaranteed investment contracts) to support the principal plus interest guarantees to participants. The Stable Value Association estimates that aggregate stable value funds currently account for approximately $300 billion in assets.\(^1\) Recent years have seen the advent of participating GIC products (often referred to as synthetic or separate account GICs). These products reflect the investment experience of a specific asset or portfolio of assets. Participating GICs now command over half of all stable value sales, according to a recent Stable Value Association survey.\(^2\)

The increasing use of participating GICs within GIC/stable value portfolios has created a need for an objective understanding of participating GIC performance. Performance measures were not required for traditional GICs because performance essentially was guaranteed (ignoring default risk) in the form of a fixed credited rate. Fixed income attribution techniques\(^3\) are not sufficient for measuring participating GIC performance. Such techniques typically restrict performance measurement to concepts such as duration management, sector rotation, and issue selection. These techniques decompose portfolio experience into components in order to isolate the performance effect of various investment strategies.

Other concepts are often more important, however, in explaining changes in participating GIC credited rates. Rennie (1994) deals with some of these concepts but he does not explain the mechanics of an attribution. As a result we have developed an attribution technique for explaining changes in participating GIC credited rates. In particular, we develop four key components or effects that are helpful in explaining the changes in credited rates:

- The constant duration effect;

---


\(^3\)Fixed income attribution techniques decompose income security total returns into components based on the sources or factors that contributed to the return. See, for example, Dietz, Fogler and Hardy (1980) and Fong, Pearson and Vasicek (1983).
• The reinvestment effect;
• The cash flow effect; and
• The investment effect.

Likewise, the investment effect consists of five subcomponents: investment market, spread; investment manager, time-weighted and nonconstant duration.

This approach described in this paper is not the only way to attribute changes in credited rates, but we believe our approach offers several advantages. It helps plan sponsors to manage their stable value portfolios by explaining, in an intuitive way, why credited rates on participating GICs change. The framework can be integrated with the investment manager’s attribution of investment performance. The framework provides a mechanism to compare the participating GIC products of different providers.

To explain the conceptual framework behind our approach, we discuss the characteristics of participating GICs, the basic resetting formula and introduce our notation in Section 2. In Section 3 we develop our attribution methodology, which is based on the four key components or effects described above that are helpful in explaining the changes in credited rates. Section 4 consists of two detailed hypothetical examples illustrating our attribution methodology.

2 Background

2.1 Participating GIC Characteristics

Participating GIC products typically have these characteristics:

• They combine a market value fixed-income portfolio (bond fund) with a book value guarantee (wrapper) that covers benefit payments to participants. A credited interest rate is established at the inception of the contract and reset periodically (quarterly, semiannually, or annually) to reflect contract experience.

• They are similar to bond funds in that performance is recognized immediately in the market value. In addition, participating GICs possess other characteristics of bond funds: active management, total return, market value, yield, and duration. Unlike bond funds, however, performance is not immediately passed to participants but is reflected prospectively in the form of adjustments to the credited interest rate.
• They possess some of the characteristics of traditional GICs: a credited rate and a guarantee of principal. Unlike traditional GICs that have a fixed credited rate that is guaranteed for the term of the contract, participating GICs have a credited rate that changes to reflect contract experience.

• Most participating GICs are designed as evergreen products with no maturity date and assets managed around a constant duration.

We discuss only this common form of participating GICs, although our framework can be modified to accommodate participating GICs with a target maturity date.

2.2 The Resetting Formula

The formula used for resetting the credited rate for participating GICs at time \( t (t = 0, 1, 2, \ldots) \) typically is:

\[
(mv_t + wc_f_t) \times (1 + \gamma_t - e_t)^d = (cv_t + wc_f_t) \times (1 + cr_t)^d
\]

(1)

where the variables below are defined with respect to portfolio values at time \( t \):

- \( mv_t \): Market value of assets in fixed income portfolio;
- \( cv_t \): Contract value or book value of liability;
- \( wc_f_t \): Cash flow expectation for the period between resets;
- \( \gamma_t \): Actual annualized market yield to maturity of the assets;
- \( e_t \): Contract expenses/fees;
- \( d_t \): Duration in years (usually the portfolio duration); and
- \( cr_t \): Credited rate

with \( mv_0 = cv_0 \) and \( cr_0 = \gamma_0 - e_0 \) at the inception of the contract.

Equation (1) can be solved directly for \( cr_t \) as a function of several variables:

\[
cr_t = cr(mv_t, cv_t, wc_f_t, \gamma_t - e_t, d_t)
\]

\[
= \left( \frac{mv_t + wc_f_t}{cv_t + wc_f_t} \right)^{1/d_t} \times (1 + \gamma_t - e_t) - 1.
\]

(2)

Equation (2) illustrates that the credited rate reflects current investment yields plus the amortized effect of any differences between market value and contract value.
2.3 Notation

Throughout the rest of this paper, the following notation is used in subscripts and superscripts to denote quantities subjected to the particular effect:

\[ \begin{align*}
CD &= \text{Constant duration;} \\
RI &= \text{Reinvestment;} \\
CF &= \text{Cash flow;} \\
IV &= \text{Investment;} \\
IM &= \text{Investment market;} \\
SP &= \text{Spread;} \\
IX &= \text{Investment index;} \\
MG &= \text{Investment manager;} \text{ and} \\
ND &= \text{Nonconstant duration.}
\end{align*} \]

In addition, without loss of generality, we will change our time reference point from the time of the inception of the contract to the time of the prior (previous) reset, i.e.,

- as time \( t - 1 \) represents the prior (previous) reset, events at time \( t - 1 \) are denoted with the subscript 0; and

- as time \( t \) represents the current reset, events at time \( t \) are denoted with the subscript 1.

Also, fees are assumed to be held constant between the prior reset and the current reset, i.e., \( \eta_t = \eta_0 = \eta \) for \( 0 \leq t < 1 \).

3 The Attribution Methodology

3.1 The Four Components

Let \( \Delta c r_0 \) denote the change in the credited rate that occurs from prior reset (time 0) to current reset (time 1), i.e.,

\[ \Delta c r_0 = c r_1 - c r_0. \] (3)

The values of the variables \( m v_0, c v_0, w c f_0, y_0, e_0, \) and \( d_0 \) serve as the starting point for credited rate reset attribution. All effects are measured relative to these values. Each of the intermediate steps involved
in changing the values of these variables at time 0 to their respective values at time 1 corresponds to a component or effect.

The change in the credited rate can be decomposed into four components: a constant duration effect, a reinvestment effect, a cash flow effect and an investment effect. The fundamental equation for the decomposition of the change in the credited rate is as follows:

$$\Delta cr_0 = \Delta cr_0^{(CD)} + \Delta cr_0^{(RI)} + \Delta cr_0^{(CF)} + \Delta cr_0^{(IV)}$$

(4)

where

- $\Delta cr_0^{(CD)}$ = The constant duration effect;
- $\Delta cr_0^{(RI)}$ = The reinvestment effect;
- $\Delta cr_0^{(CF)}$ = The cash flow effect; and
- $\Delta cr_0^{(IV)}$ = The investment effect.

In addition, the investment effect is decomposed further into five subcomponents as follows:

$$\Delta cr_0^{(IV)} = \Delta cr_0^{(IM)} + \Delta cr_0^{(SP)} + \Delta cr_0^{(MG)} + \Delta cr_0^{(TW)} + \Delta cr_0^{(ND)}$$

(5)

where

- $\Delta cr_0^{(IM)}$ = The investment market effect;
- $\Delta cr_0^{(SP)}$ = The spread effect;
- $\Delta cr_0^{(MG)}$ = The investment manager effect
- $\Delta cr_0^{(TW)}$ = The time-weighted effect; and
- $\Delta cr_0^{(ND)}$ = The nonconstant duration effect.

This attribution framework is sufficiently flexible to accommodate a wide variety of methods of measuring yields and durations including:

- Portfolio, index, or model yields and durations;
- Duration-weighted or market value-weighted yields;
- Option-adjusted yields and durations; and
- Macaulay or modified durations.

In our examples in Section 4, we use duration-weighted portfolio yields, Macaulay portfolio durations as the amortization period, and annual resets. We do not use time-weighted cash flows.

Next we develop the formulas needed to compute magnitude of each effect.
3.2 The Constant Duration Effect

Participating GICs have a constant duration product design, i.e., they do not have a maturity date. Assets are managed around a constant (or target) duration, \( d \), which is consistent with plan sponsor objectives. As equation (2) shows, any difference between \( mv_1 \) and \( cv_1 \) is amortized over the constant duration, \( d \). The amortization process acts as a smoothing mechanism, reducing the volatility of returns to participants. It also will cause \( mv_t \) and \( cv_t \) to converge to each other (as \( t \to \infty \)) in a stable interest rate environment.

Assume market value is below contract value. As equation (2) indicates, the credited rate will be lower than the market yield, as the loss is amortized. In a constant duration product the difference grows smaller each year, but is still amortized over the initial duration. This creates a declining drag on the credited rate; see Figure 1. With a fixed maturity product design (declining duration) the credited rate would have remained unchanged until the end of the fourth year, at which time market value would have equaled contract value.

With a constant amortization period the credited rate will change even if all other prior reset assumptions are realized. Over time the credited rate will approach the portfolio's net yield. In summary, the constant duration effect captures the effect of a constant duration product design as opposed to a fixed maturity product design.

The constant duration effect is calculated by assuming that the portfolio returns its net yield and the contract value grows at the credited rate. In mathematical notation,

\[
\Delta cr_0^{(CD)} = cr_1^{(CD)} - cr_0
\]  

where

\[
mv_1^{(CD)} = (mv_0 + wc f_0) \times (1 + y_0 - e) \quad (7)
\]

\[
cv_1^{(CD)} = (cv_0 + wc f_0) \times (1 + cr_0) \quad (8)
\]

\[
cr_1^{(CD)} = cr(mv_1^{(CD)}, cv_1^{(CD)}, 0, y_0 - e, d_0) \quad \text{and} \quad (9)
\]

\[
cr_0 = cr(mv_0, cv_0, wc f_0, y_0 - e, d_0). \quad (10)
\]

3.3 The Reinvestment Effect

As interest rates change, the effect on the bond portfolio is twofold. A drop in interest rates tends to increase the portfolio's market value and decrease the portfolio's yield. In an instantaneous interest rate
As the portfolio moves through time in a lower interest rate environment, however, it tends to earn less coupon income. Thus, at the end of the year the portfolio's market value will reflect the impact of the lower coupon income. Consequently, the credited rate will be reduced. In other words, the overall impact on the credited rate should be in the same direction as the movement in market interest rates. In summary, the reinvestment effect captures the responsiveness of the credited rate to the movement in market interest rates.

The reinvestment effect is calculated by assuming that the yield on the portfolio changes by the change in interest rates and the portfolio return reflects the change in yield. We use Treasury yields as a proxy for interest rates. Using Treasury yields and temporarily ignoring the effect of spread, credit, and prepayment factors, the portfolio return is
given by:

\[ R_{RI} = (1 + y_0 - e + \frac{(T_1 - \bar{T}_{ex})}{2}) \left( \frac{1 + y_0 - e}{1 + y_0 - e + (T_1 - \bar{T}_{ex})} \right)^{d_0} - 1 \]  \hfill (11)

where

\[ T_t = \text{The Treasury rate at time } t \ (0 \leq t \leq 1); \]
\[ \bar{T}_{ex} = \frac{mv_0 \times T_0 + wc_{f0} \times \bar{T}_{av}}{mv_0 + wc_{f0}}; \text{ and} \]
\[ \bar{T}_{av} = \text{A weighted average of Treasury rates over time.} \]

Note that \( \bar{T}_{av} \) is given in the examples that follow. Also, \( \bar{T}_{ex} \) is the weighted average of Treasury rates using the initial market value and the expected cash flows as weights.

In mathematical notation,

\[ \Delta cr_{0}^{(RI)} = cr_{1}^{(RI)} - cr_{1}^{(CD)} \]  \hfill (13)

where

\[ mv_{1}^{(RI)} = (mv_0 + wc_{f0})(1 + R_{RI}) \]  \hfill (14)
\[ y_1^{(RI)} = y_0 + T_1 - \bar{T}_{ex} \text{ and} \]
\[ cr_{1}^{(RI)} = cr (mv_{1}^{(RI)}, cv_{1}^{(CD)}, 0, y_1^{(RI)} - e, d_0). \]  \hfill (16)

### 3.4 The Cash Flow Effect

Many participating GICs accept ongoing contributions and make benefit payments at book value. Cash flow expectations are reflected in the reset formula, i.e., in the prior reset assumption \( wc_{f0} \). If actual cash flow experience (amount and timing) matches expectations, we define the cash flow effect to be zero. Therefore, our cash flow effect incorporates deviations from both the expected amount of cash flow and expected interest rates (at which the cash flow was invested). In summary, the cash flow effect captures the effect of unexpected cash flows on the credited rate.

The cash flow effect is calculated by adjusting the market value and contract value to reflect the actual as opposed to the expected cash flows.

\[ R_{CF} = (1 + y_0 - e + \frac{(T_1 - \bar{T}_{ac})}{2}) \left( \frac{1 + y_0 - e}{1 + y_0 - e + (T_1 - \bar{T}_{ac})} \right)^{d_0} - 1 \]  \hfill (17)
where
\[ \tilde{T}_{ac} = \frac{mv_0 \times T_0 + \sum_{i=0}^{\infty} c_{fi} \times T_{ti}}{mv_0 + \sum_{i=0}^{\infty} c_{fi}} \] (18)

and \( c_{fi} \) is the actual \( i \)-th cash flow at time \( t_i \), with \( 0 = t_0 < t_1 < t_2 < \ldots \). Note that \( \tilde{T}_{ac} \) is the weighted average of Treasury rates using the initial market value and the actual cash flows as weights.

In mathematical notation,
\[ \Delta cr_0^{(CF)} = cr_1^{(CF)} - cr_1^{(RI)} \] (19)

where
\[
\begin{align*}
acf_0 & = \text{The actual cash flow received;} \\
mv_1^{(CF)} & = (mv_0 + acf_0) \times (1 + R_{CF}) \quad (20) \\
cv_1^{(CF)} & = cv_1 \quad \text{and} \\
cr_1^{(CF)} & = \text{cr} (mv_1^{(CF)}, cv_1^{(CF)}, 0, y_1^{(RI)} - e, d_0). \quad (22)
\end{align*}
\]

3.5 The Investment Effect

This investment effect component can be further divided into five subcomponents: investment market, spread, investment manager, time-weighted and nonconstant duration. From equation (4), we know that
\[
\Delta cr_0^{(IV)} = (cr_1 - cr_0) - (cr_1^{(CD)} - cr_0) \\
= (cr_1^{(RI)} - cr_1^{(CD)}) - (cr_1^{(CF)} - cr_1^{(RI)}) \\
= cr_1 - cr_1^{(CF)}. \quad (23)
\]

It follows that the sum of the effects of the five subcomponents must satisfy equation (23). We now review each of these subcomponents individually.

3.5.1 The Investment Market Effect

The investment market effect captures the impact of the chosen investment universe on the credited rate. The investment universe is represented by the index or other benchmark selected to evaluate the portfolio's performance. Participating GIC contracts allow for a variety of portfolio structures and benchmarks, developed by mutual agreement between the plan sponsor and the investment manager.

The investment market effect is designed to provide a basis for evaluating benchmark performance by a manager. (The basis or proxy is
indexed management or a manager whose return exactly matches the chosen benchmark.) The performance of the index is compared to the rate needed to support the initial credited rate. This is expressed in terms of a Treasury-plus-spread benchmark. Treasuries are used because they are the universal underpinning of all fixed income prices and returns.

We calculate the investment market effect by performing a rate reset with a market value calculated using the index return. In mathematical notation,

\[ \Delta cr_{0}^{(IM)} = cr_{1}^{(IM)} - cr_{1}^{(CF)} \]  

(24)

where

\[ R_{IX} = \text{The index return.} \]
\[ mv_{1}^{(IX)} = (mv_{0} + acf_{0})(1 + R_{IX}) \text{ and} \]
\[ cr_{1}^{(IM)} = cr(mv_{1}^{(IX)}, cv_{1}, 0, y_{1}^{(RI)} - e, d_{0}). \]  

(25)

(26)

The comparison of a standard market index, such as the Lehman aggregate or government/corporate index, to a Treasury-plus-spread benchmark implicitly assumes the use of non-Treasury assets. The actual impact on the credited rate is derived from the fact that the total return of the index incorporates initial spread plus changes in spreads, while the Treasury-plus-spread proxy only incorporates the initial spread.

For example, the Lehman MBS index may yield 100 basis points over Treasuries at the start of the year, with a four year index duration. Assume that mortgage spreads widen 50 basis points during the year, with Treasury rates and durations remaining unchanged. The mortgage portfolio will underperform duration-matched Treasuries by roughly 75 basis points during the year (assuming spread change occurs uniformly over the year). The investment market effect will be roughly -44 basis points (-75 basis points actual performance versus +100 basis points assumed performance produces a 175 basis point difference amortized over four years). The spread widening will be reflected in the spread effect (see below) of +50 basis points, more than offsetting the investment market effect.

### 3.5.2 The Spread Effect

Spread widening/tightening may offset the investment market effect results because widening asset spread to Treasuries negatively impacts portfolio return. The impact of spread changes on portfolio return is
recognized in the investment market effect. The impact on the yield to maturity is recognized in the spread effect. These effects therefore should be viewed together rather than individually. An increase in spread translates directly into a like increase in yield which in turn translates directly into a like increase in the credited rate. The opposite holds for a decrease in spread. Thus

\[ \Delta cr_0^{(SP)} = cr_1^{(SP)} - cr_1^{(IM)} \]  

(27)

where

\[ cr_1^{(SP)} = cr(mv_1^{(IX)}, cv_1, 0, y_1 - e, d_0). \]  

(28)

3.5.3 The Investment Manager Effect

The investment manager effect is designed to show the effect of a manager's performance (relative to a benchmark) on the credited rate. The manager's relative performance (also called the excess return) is \( R_{MG} - R_{IX} \) where \( R_{MG} \) is the portfolio return.

We compute the investment manager effect by performing a rate reset with a market value calculated using the actual portfolio return. In mathematical notation,

\[ \Delta cr_0^{(MG)} = cr_1^{(MG)} - cr_1^{(SP)} \]  

(29)

where

\[ mv_1^{(MG)} = (mv_0 + acf_0)(1 + R_{MG}) \] and  

(30)

\[ cr_1^{(MG)} = cr(mv_1^{(MG)}, cv_1, 0, y_1 - e, d_0). \]  

(31)

This effect can be analyzed in further detail using standard portfolio management attribution techniques. These techniques typically allocate the manager's excess return to categories such as duration management, yield curve management, sector and security selection, and other factors relative to the performance benchmark. This analysis can be incorporated easily into the credited rate attribution framework by applying a factor of \((1/d_0)\) to each category in the portfolio attribution analysis. For example, if the manager added 15 basis points of return through duration management for a five year duration portfolio, the credited rate impact would be \((15/5 =) 3\) basis points.
3.5.4 The Time-Weighted Effect

So far we have used the time-weighted rates of return\(^4\) adjusted for the impact of actual cash flows for three previous effects (cash flow effect, the investment market effect, and the investment manager effect). As a result, there is a residual effect, which we call the time-weighted effect that is given by:

\[
\Delta cr_0^{(TW)} = cr_1^{(TW)} - cr_1^{(MG)}
\]  

(32)

where

\[
cr_1^{(TW)} = cr(mv_1, cv_1, 0, y_1 - e, d_0).
\]  

(33)

3.5.5 The Nonconstant Duration Effect

The duration of the portfolio at the current reset (time 1) probably will change from the duration at the prior reset (time 0). Regardless of the reason (rate anticipation strategy, revised client objectives, duration drift), duration changes will affect both the portfolio return (investment manager effect) and the amortization period. For example, duration shortening will accelerate the recognition of any gains or losses.\(^5\)

The duration effect is calculated by substituting the actual portfolio duration at time 1. In mathematical notation,

\[
\Delta cr_0^{(ND)} = cr_1^{(ND)} - cr_1^{(TW)}
\]  

(34)

where

\[
cr_1^{(ND)} = cr(mv_1, cv_1, 0, y_1 - e_1, d_1).
\]  

(35)

3.6 Other Considerations

3.6.1 Investment Risks

Plan sponsors are accustomed to thinking in terms of investment risks including reinvestment risk, credit risk, prepayment risk, liquidity risk, and active management risk. These risks are inherent in one or more of the four components or effects introduced in Section 2 above. In particular,

\(^4\)The dollar-weighted return may have been used instead, but it requires more complicated calculations.

\(^5\)This effect refers to the end-of-year portfolio duration, not the management of the portfolio duration during the year. Duration management has been accounted for in the investment manager effect.
• Reinvestment risk is captured by the reinvestment effect;

• Credit risk, i.e., a decline in credit quality, appears as a negative impact in the investment market effect and a positive impact in the spread effect;

• Prepayment (extension) risk appears as a negative impact in the investment market effect and presumably a positive impact in the spread effect. In addition, prepayments (extensions) reduce (increase) duration and thus the amortization period. The direction of the impact on the duration effect depends on the ratio of market value to book value;

• The realization of liquidity risk appears as a negative impact in the investment market effect and a positive impact in the spread effect; and

• Active management risk affects the investment manager effect.

3.6.2 Treatment of Cash Flows

We ignore future cash flow expectations for the current reset, i.e., we assume $wc_{f1} = 0$. If $wc_{f1}$ does not equal zero, the future cash flow assumption will affect the credited rate because it affects the market-to-contract value ratio. Positive assumed cash flow increases the credited rate when market value is less than book value. Negative assumed cash flow decreases the credited rate when market value is less than book value. The effects are reversed in a situation where market value is above contract value.

Though cash flow is not a significant element for most contracts, some contracts have significant and frequent cash flows. For these contracts cash flows can be handled more precisely by using dollar-weighted portfolio returns instead of time-weighted portfolio returns. Using dollar-weighted returns will eliminate the time-weighted effect described in Section 3.5.4. Otherwise, the change in credited rate may be incorrectly allocated between the cash flow effect, the investment market effect, and the investment manager effect. The allocation problem concerns the timing of any market or manager outperformance during the reset period.

3.6.3 Participant Equity

The smoothing mechanism raises the issue of equity among plan participants. For instance, when the market value is below contact
value, i.e., \( m_{t} < c_{t} \), the credited rate will be lower than current the market yield, i.e., \( c_{r_{t}} < y_{t} \) as indicated by equation (2). A participant who withdraws funds at contract value causes the portfolio to realize a market value loss. As the remaining participants participate in the experience of the portfolio, they will be hurt by the withdrawal in the form of lower future credited rates.

The participant equity issue is inherent in any blended rate stable value plan. The extent of the participant equity issue depends on the overall plan design and operation, not just the characteristics of one of the plan's investments. By helping plan sponsors to better understand the behavior of one possible investment (the participating GIC), the attribution methodology can help plan sponsors better address the equity issue.

4 Two Examples

4.1 Example 1: Prior Reset Assumptions Realized

4.1.1 Prior Reset Assumptions

These assumptions were made one year ago (i.e., at time \( t = 0 \)), the last time the credited interest rate was reset:

- The portfolio had a 2.1% market value loss. This implies that \( m_{v_{0}} = (1 - 0.021) \times c_{v_{0}} = 97.9\% \) of \( c_{v_{0}} \);
- \( m_{v_{0}} = 93.0 \) and \( c_{v_{0}} = 95.0 \);
- A $5 cash flow expected at \( t = 0.5 \), i.e., \( w_{c f_{0}} = 5.0 \);
- Cash flows are not time-weighted;
- The portfolio net yield will be \( y_{0} - e = 8\% \);
- Interest rates will remain unchanged; and
- On average, funds are to mature in four years, making \( d_{0} = 4 \).

From equation (2) we have

\[
c_{r_{0}} = \left( \frac{93.0 + 5.0}{95.0 + 5.0} \right)^{1/4} \times (1 + 0.08) - 1 = 0.0746.
\]
4.1.2 Actual Activity Since Prior Reset

As we prepare to reset the interest rate today (i.e., at time $t = 1$), we must review the actual performance of the contract. For this example, let us assume all prior reset assumptions were realized, meaning:

- Actual cash flow of $5 at $t = 0.5$;
- No change in interest rates;
- The portfolio actually returned a net yield of 8%; and
- On average, funds are now expected to mature in three years, making the duration $d_1 = d_0 - 1 = 3$.

4.1.3 The Calculations

Once the assumptions are exactly realized, the credited rate must be equal to the constant duration credited rate, i.e.,

$$ cr_1 = cr_1^{(CD)}. $$

So, using equations (7) and (8) we see that:

$$ mv_1 = 105.84 \text{ from equation (7)}; $$

$$ cv_1 = 107.46 \text{ from equation (8)}; \text{ and} $$

$$ y_1 - e = y_0 - e = 8%. $$

Equation (9) gives

$$ cr_1 = cr(105.84, 107.46, 0, 8%, 3) = 7.46%. $$

As the assumptions were exactly realized, it is not surprising that the credited rate is unchanged from the initial credited rate established at time 0.

4.2 Example 2: Prior Reset Assumptions Not Realized

In developing this example, we use the same set of prior reset assumptions that are described in Section 4.1.1. As a result $cr_0 = 7.46%$ as before.
4.2.1 Actual Activity Since Last Reset

As we prepare to reset the interest rate today, the actual activities since last reset are listed below:

- Interest rates (using Treasuries as a proxy) declined 1% from time 0 to time 1. Specifically, Treasury rates were $T_0 = 7.50\%$, $T_{0.5} = 7.00\%$, $T_1 = 6.50\%$, and $T_{av} = 7.00\%$;

- The only actual cash flow received was $15 (acf_0 = 15)$ immediately following the prior reset at time 0;\(^6\)

- $R_{ix} = 0.11$, i.e., the index returned 11% from time 0 to time 1 (net of fees $e$);\(^7\)

- The manager returned 11.50% from time 0 to time 1 (net of contract fees $e$), i.e., $R_{mg} = 11.5\%$. (The manager outperformed the index by 50 basis points);

- The portfolio duration at time 1 is 4.25 years, i.e., $d_1 = 4.25$. In other words, the portfolio lengthened by 1/4 year;

- The actual portfolio net yield at time 1 is 7.10%, i.e., $y_1 - e = 7.10\%$;\(^8\)

This gives

\[
\begin{align*}
env_1 &= (93 + 15) \times 1.115 = 120.50 \\
env_1^{CD} &= (93 + 5) \times 1.08 = 105.84 \\
env_1^{CD} &= (95 + 15) \times 1.0746 = 118.21 \\
\end{align*}
\]

\[
\begin{align*}
cr_1 &= Cr(120.50, 118.21, 0, 7.10\%, 4.25) = 7.58\%.
\end{align*}
\]

4.2.2 The Component Effects

**Constant Duration Effect:** From equation (6),

\[
\begin{align*}
env_1^{CD} &= (93 + 5) \times 1.08 = 105.84 \\
env_1^{CD} &= (95 + 15) \times 1.0746 = 107.46 \\
\end{align*}
\]

\[
\begin{align*}
\text{\textit{Cr}(105.84, 107.46, 0, 8\%, 4)} &= 7.59\%. \\
\end{align*}
\]

\(^6\)The additional cash flow was received while the contract was in deficit (market value below contract value). This unexpected cash flow reduced the deficit and thus increased the credited rate. The subsequent fall in interest rates during the year has been captured in the reinvestment effect.

\(^7\)The index returned 58 basis points less than the Treasury-plus-spread bogey.

\(^8\)Thus, yields fell 90 basis points (from 8% to 7.10%) from time 0 to time 1 while interest rates fell 100 basis points. In other words, spreads increased 10 basis points.

giving

\[ \Delta cr_0^{(CD)} = 7.59\% - 7.46\% = 0.13\%. \]

**Reinvestment Effect:** From equation (12), \( T_{ex} = 7.47 \), so \( R_{RI} \) can be calculated from equation (11) to give \( R_{RI} = 11.47\% \). This rate ignores the effect of spread, credit, and prepayment factors. Thus, from equations (14), (15) and (16), we have

\[
\begin{align*}
\nu_1^{(RI)} & = (93 + 5) \times (1 + 0.1147) = 109.24 \\
\gamma_1^{(RI)} - e & = \gamma_0 - e + T_1 - T_{ex} = 7.03\% \\
cr_1^{(RI)} & = cr(109.24, 107.46, 0, 7.03\%, 4) = 7.47\%
\end{align*}
\]

giving

\[ \Delta cr_0^{(RI)} = 7.47\% - 7.59\% = -0.12\%. \]

**Cash Flow Effect:** To isolate the effect of cash flows we adjust the \( mv \) and \( cv \) terms for the amount and timing of actual cash flow, \( acf_0 \). As there was only one cash flow and it occurred at time 0, then \( T_{ac} = T_0 = 7.5\% \). Accounting for unexpected cash flow, the portfolio returned 11.58\%, i.e., \( R_{CF} = 0.1158 \) (after using equation (17)).

\[
\begin{align*}
\nu_1^{(CF)} & = (93.0 + 15.0) \times (1 + 0.1158) = 120.51 \\
\gamma_1^{(CF)} & = cv_1 = 118.21 \\
cr_1^{(CF)} & = cr(120.51, 118.21, 0, 7.03\%, 4) = 7.55\%
\end{align*}
\]

giving

\[ \Delta cr_0^{(CF)} = 7.55\% - 7.47\% = 0.08\%. \]

**Investment Market Effect:** Using equations (2), (25) and (26) to get:

\[
\begin{align*}
\nu_1^{(MX)} & = (93.0 + 15.0) \times (1 + 0.11) = 119.88 \\
cr_1^{(IM)} & = cr(119.88, 118.21, 0, 7.03\%, 4) = 7.41\%
\end{align*}
\]

giving

\[ \Delta cr_0^{(IM)} = 7.41\% - 7.59\% = -0.14\%. \]
**Spread Effect:** Using equation (28) we find:

\[ cr^{(SP)}_1 = cr(119.88, 118.21, 0, 7.10\%, 4) = 7.48\% \]

giving

\[ \Delta cr^{(SP)}_0 = 7.48\% - 7.41\% = 0.07\%. \]

**Investment Manager Effect:** Using equations (30) and (31), we find:

\[
\begin{align*}
mv^{(MG)}_1 &= (93.0 + 15.0) \times (1 + 0.115) = 120.42 \\
cr^{(MG)}_1 &= cr(120.42, 118.21, 0, 7.10\%, 4) = 7.60\% \\
\end{align*}
\]

giving

\[ \Delta cr^{(MG)}_0 = 7.60\% - 7.48\% = 0.12\%. \]

**Time-Weighted Effect:** Using equations (32) and (33), we find:

\[
\begin{align*}
ctw^{(TW)}_1 &= cr(120.42, 118.21, 0, 7.10\%, 4) = 7.60\% \\
\end{align*}
\]

giving

\[ \Delta cr^{(TW)}_0 = 7.60\% - 7.60\% = 0.00\%. \]

Note that the time-weighted return equals the dollar weighted return since the only cash flow occurs at time 0.

**Duration Effect:** Using equations (34) and (35) we find:

\[
\begin{align*}
c^{(ND)}_1 &= cr(120.42, 118.21, 0, 7.10\%, 4.25) = 7.57\% \\
\end{align*}
\]

giving

\[ \Delta cr^{(ND)}_0 = 7.57\% - 7.60\% = -0.03\%. \]

### 4.3 Summary of Attribution

The credited interest rate in our example changed from 7.46 percent to 7.57 percent, a net increase of 11 basis points in a declining interest rate environment. The credited rate benefited from the amortization of a previous market value loss, unexpected cash flow when market rates were favorable, manager outperformance, and widening spreads. Adverse factors included a declining interest rate environment, a longer portfolio duration, and underperformance by the portfolio’s investment universe relative to initial assumptions.
The increasing use of participating GICs within GIC/stable value portfolios has created a need for objective measures of performance attribution. When a credited rate is reset, providers should expect to provide detailed attribution analyses that explain how the new credited rate was derived.

Current explanations discuss, in general terms, such concepts as issue selection and yield curve positioning. These explanations ignore, however, other important concepts such as product design (constant duration effect), credited rate responsiveness (reinvestment effect), cash flows (cash flow effect), initial return assumption (investment market effect), and amortization period (duration effect). We have presented a framework to explain the performance of participating GICs by isolating the effects of all relevant factors.

Our approach analyzes these effects in static fashion and in a particular order. We recognize that in actuality these components are likely to affect performance in a dynamic way.
Asset Allocation in Investing to Meet Liabilities
Anthony Dardis* and Vinh Loi Huynh†

Abstract

We present some rudimentary concepts on asset/liability management and describe an approach to asset allocation modeling for institutions that invest to meet liabilities. The traditional risk/reward framework of financial economics is used as a starting point. The definitions of risk and reward are then refined with regard to the institution under consideration. A simple model of a U.S. life office is examined. We assume that the only investments available are domestic stocks and long-dated government bonds. Stochastic simulation is used to create a large number of future investment scenarios using historical total return data for these asset classes. The ability of the institution to meet its liabilities under each simulated scenario is examined. We construct optimal risk/reward profiles, and hence the optimal asset allocation strategy, and show that they can vary considerably by liability profile.

Key words and phrases: asset/liability management, Monte Carlo simulation, risk, reward, solvency

*Anthony Dardis, F.I.A., A.SA, M.AAA., is a consulting actuary with Tillinghast-Towers Perrin. He received a B.A. in economics from Leeds University, England. Mr. Dardis has over 15 years actuarial experience covering a wide range of fields including merchant banking, reinsurance, and life consulting.

Mr. Dardis is a regular participant at Society of Actuaries' meetings and has published articles in the Record of the Society of Actuaries and in various Society of Actuaries' newsletters. He is a member of the editorial board of the Society's Investment Section newsletter, and is a member of the Society of Actuaries' Finance Practice Area Education Committee.

Mr. Dardis' address is: Tillinghast—Towers Perrin, 12377 Merit Drive, Suite 1200, Dallas TX 75251. Internet address: Dardist@Tillinghast.com

†Vinh Loi Huynh, M.A., is a senior actuarial assistant with Employers Reassurance International Limited, an international reinsurance company. He received his B.A. and M.A. in mathematics from Cambridge University, England. Mr. Huynh's main work involves pricing life and health reinsurance products. He also designs computer programs for actuarial pricing and cash flow testing. He has presented papers at the 1994 AFIR in Orlando and the Staple Inn Society of Actuaries in London.

Mr. Huynh's address is: Employers Reassurance International Limited, Portsoken House, 155-157 Monories, London EC3 N1BU, ENGLAND. Internet address: dkerivh0@ibmmail.com
1 Historical Overview

Both financial economists and actuaries have been involved in the development of quantitative asset allocation techniques for many years. The two major asset allocation techniques examined are immunization and mean/variance analysis.

1.1 Immunization

In 1952 Redington introduced the theory of immunization. Since then, the theory of immunization has had a profound influence on the way actuaries approach the valuation of insurance companies and their assessment of solvency. As a practical asset/liability management (ALM) model for insurance companies, immunization has had little competition to date. Tilley (1988) remarked that “a whole investment advisory business has grown up in the United States around immunization concepts.”

The idea of equating the duration$^1$ of assets with the duration of liabilities has been used widely by insurance companies worldwide. Recently the notion of convexity (which is similar to duration, but with second derivatives replacing first derivatives) has given immunization new life. But immunization does have its limitations: (i) it has little relevance to interest-sensitive and performance-linked products; and (ii) immunization immunizes against profits as well as against loss.

Redington's ideas today may be viewed as the classical actuarial approach to ALM. The success of the Redington model as an accepted ALM tool lies in its relative simplicity and the ease with which the calculations necessary to test immunization can be made. As Buff (1989) states “if you can’t compute it, you can’t compute it.” It is not possible to use theoretical advances unless it is feasible to execute the calculations necessary in these advances.

Actuarial research into ALM modeling was muted for many years after Redington, but the actuarial profession recently has found a new interest in the subject. Some of the most interesting work has been in the United Kingdom where pioneering stochastic investment modeling has been done by Wilkie (1986). The focus of the U.K. work, however, is in the area of solvency testing. See Hardy (1993) for an excellent example of the usage of stochastic modeling in assessing solvency.

A new concept is beginning to appear in the ALM literature in the U.S. concerning solvency in connection with ALM; this is the idea that ALM$^1$

---

$^1$Duration of a financial instrument is a measure of its sensitivity to interest rates at various points on the yield curve.
focuses on asset/liability surplus management (ALSM). ALSM refers to ALM that focuses on the NAIC risk-based capital standards. These standards require certain minimum surplus amounts to be maintained in respect of various classifications of risk. An ALSM model might assess how well the required minimum surplus levels are likely to hold using the potential investment strategies under consideration. Hepokoski (1994) gives an excellent introduction to ALSM as an extension of ALM. In practice, an asset allocation model might view the risk-based capital standards as constraints rather than defining risk purely in terms of those standards.

1.2 Mean/Variance Analysis

The same year that Redington published his ideas on immunization, one of the most important papers of modern financial economics also was published. In 1952 Markowitz introduced the idea of asset allocation within a risk/reward tradeoff framework.

Markowitz notes that a reduction in risk, measured by the standard deviation of return on assets, could be achieved by diversification (into assets whose returns are uncorrelated) without any reduction in return. Markowitz also introduced the idea of an efficient frontier, which is a curve joining the risk/reward combinations of asset mixes that give the highest reward for any given level of risk.

At the time the financial world was not ready for the concept of an efficient frontier—to return to Buff’s truism, computer power had not reached the stage where Markowitz’s ideas could be implemented. A practical adaptation of these ideas had to wait over a decade, when Sharpe (1963) introduced the diagonal model that suggests that the future price of a security depends on its alpha, the market return through its beta, and a random error term, the values based on simple linear regression on historical data. This marked the birth of the now widely used capital asset pricing model.

Sharpe (1970) suggests that mean and variance alone “may suppress too much reality” and that a different utility curve may be needed to compare different portfolios of different riskiness. Risk is not necessarily the same for all investors; in Arthur's words (1989), “risk is in the eye of the beholder.”

Many of Sharpe’s ideas in the area of asset/liability management are summarized in Managing Investment Portfolios (1990). In this volume he presents the concepts of risk/reward indifference curves and states that “the optimal asset mix lies at the point at which an indifference curve is tangent to (i.e., touches but does not intersect) the curve along
which the efficient investment opportunities lie." He also presents a complete ALM model for a defined benefit pension scheme in which reward is defined in terms of surplus return (equal to the change in value of surplus divided by the initial asset value) and risk is the standard deviation of the surplus return. Such a model is close to the model we use in this paper.

1.3 Other Approaches to ALM

Other variations on the efficient frontier idea experiment with constraints that can be used to narrow acceptable portfolio mixes on the efficient frontier. They also attempt to be dynamic in the sense that acceptable portfolio mixes change and reflect the particular market conditions present at any particular time.

A good example of such a model is developed by Leibowitz, Kogelman, Bader, and Dravid (1994). Looking at a one year time horizon, their model updates the asset allocation strategy whenever interest rates move. Their model does not just look at portfolios on the efficient frontier, but also introduces the constraint that portfolios must have no more than a specified probability of generating one year returns that fall below a certain level. This is incorporated by the introduction of a *shortfall line*, such that all portfolios above the line of constraint meet the maximum probability criterion.

If interest rates fall (with the equity risk premium, stock and bond volatilities, and stock/bond correlations all held constant) the entire risk/reward curve will shift down, decreasing the expected returns of all potential portfolio mixes. With the shortfall line unchanged, market conditions make all portfolios riskier in shortfall terms, and few portfolios will fall above the shortfall line. This requires revision of the bond/equity mixes previously deemed acceptable.

1.4 Objectives of this Paper

We will develop a simple model of an insurance company and use it to explore some of the basic concepts of ALM. The model is a true ALM model where the two sides of an institution's balance sheet are considered equally in setting or appraising long-term investment policy.

The nature of financial risk is briefly explored. Risk is related to the chance of not meeting the rate of return that is required to support a life insurer's liabilities, rather than the typical risk measures postulated by mean/variance models. A specific measure of risk, based on
the probability of continued solvency rather than asset volatility, is introduced.

The company is assumed to invest only in domestic stocks and long-dated government bonds. Stochastic simulation is used to create a large number of future investment scenarios using historical total return data for domestic stocks and long-dated government bonds. The ability of the institution to meet its liabilities under each simulated scenario is examined for each possible mix of assets and risk characteristic. Thus, for each asset mix, the model produces a certain level of risk and a certain level of reward.

We then assess the minimum level of risk for any particular level of reward; the asset mix that produces such a level of risk is retained, and all such retained risk/reward points are plotted to create an optimal risk/reward profile. The paper demonstrates that such optimal risk/reward profiles—and hence the optimal asset allocation strategy—can vary considerably by liability profile.

2 What is Return?

Although the meaning of the term return on assets usually can be taken for granted in financial modeling—it is based on market value changes after allowing for positive and negative cash flows—this is not the case in an ALM model. This extra consideration arises because definitions of return and risk must be consistent.

In our paper risk is viewed as the ability of the financial institution to demonstrate, from time to time, that it is in a financially stable situation. This requires an assessment of the solvency of the institution by comparing the actual value of assets with the value of assets required for the institution to meet future liabilities.

For a U.S. life office a solvency valuation is required by regulation, and asset values in such a solvency test are prescribed by state law or by the National Association of Insurance Commissioners (NAIC). This valuation generally requires carrying assets at market values, although there are important exceptions such as amply secured bonds not in default that are written up or down in order that the value at maturity will equal the maturity value. To be consistent with the risk/solvency assessment, return must be defined in terms of return on the actuarial value of assets as carried in the solvency valuation, and these values may or may not be market values.

Our highly simplified model office is in a financially unstable condition if the office becomes insolvent, which will arise if the actual return
on assets falls below the expected return on assets used in pricing the liabilities. (The expected return is the terms on which the business was sold.) In other words, risk is defined in terms of underperforming the pricing assumptions over the lifetime of the policies. No reference is made to valuation margins and capital (the reserves and, therefore, the value of assets required from year to year are based purely on the assumptions used in setting the premium rates), and the values of assets are based on market returns. In effect, risk is in terms of actual market returns underperforming expected market returns, so both risk and return need to be defined in terms of market values. The actuarial value of assets is defined as the market value of assets.

In practice, the risk of insolvency should not be judged against pricing assumptions with no valuation margins and no capital; if valuation margins and capital were incorporated, it would be more appropriate for the risk measures to be based on statutory results rather than simply on market values. Also, risk ideally should be measured not just by the probability of underperforming, but by the amount of underperformance.

3 The Model

There are four important stages in the development and exploration of the model:

- An assessment has to be made of the probability distribution of the returns on assets available to the financial institution;
- An accurate cash flow projection must be made of the future liability outgo of the financial institution;
- Using the information about the probability distribution of asset returns, large numbers of possible investment scenarios must be derived. The performance of the fund in meeting the liabilities under each scenario must be examined; and
- A large number of runs will enable an assessment of how a particular mix of the various asset classes will meet the liabilities. This assessment forms the basis for the construction of a risk/reward profile from which possible optimal asset mixes can be considered for investment policy.

The means used to explore the model is via simulation, where the simulated variable is the return on assets. Simulation is necessary because
a mathematical solution to the model is too complicated. This compi-
lcation is due not to the intractability of the return on assets, but is
due to the other variable in the model—the risk variable—which is not
necessarily a straightforward variable to handle mathematically. Even
the risk variable that we adopted for this paper, although simple in
concept, is difficult to express mathematically. More sophisticated de-
definitions of risk that also incorporate constraints would pose even more
of a challenge.

As a result, the simulation process starts by generating random ob-
servations for the random variable with a known distribution (or at
least a distribution for which a reasonable assessment can be made)
that can be used to calculate random observations for the complicated
random variable. From these observations it is possible to make infer-
ences about the distribution of the complicated variable.

We assume that the financial institution being assessed is a life office
that issues a large number of level annual premium whole life policies
on male lives age 50 at entry, and these policies all begin today. The
only decrement is mortality, and this is assumed to accord with the
Society of Actuaries (SOA) 75-80 15 Year Select and Ultimate Table (age
nearest birthday). All expenses and commissions are assumed to be
zero.

Let $F_t$ be the fund at end of policy year $t$. The model tracks forward
for each of the years for which the whole life contracts are expected to
be in force and computes the following for $t = 1, 2, \ldots$:

$$F_t = (F_{t-1} + P_t)(1 + i) - C_t(1 + i)^{1/2}$$

where

- $i$ = Interest rate used to determine the net premium;
- $P_t$ = Net premium received at start of policy year $t$; and
- $C_t$ = Aggregate claims in policy year $t$.

Claims are assumed to occur on average in the middle of the year.

Thus, the sequence $\{F_t\}$ represents the target fund level to which
the office should strive, based on an investment return equal to that
assumed in the premium basis. If the actual fund falls persistently
below this target fund in practice, the office is heading toward financial
difficulties. It is therefore appropriate to examine the success of any
particular investment policy in generating a fund size consistently at
least as great as the target fund.

Let $N_t$ be the simulated fund at the end of policy year $t$, then

$$N_t = (N_{t-1} + P_t)(1 + s_t) - C_t(1 + s_t)^{1/2}$$
where \( s_t \) is the simulated annual rate of return in policy year \( t \); and \( P_t \) and \( C_t \) are as previously defined. Mathematically, \( s_t \) is defined as the weighted average of the simulated annual rates of return from stocks and bond, i.e.,

\[
s_t = \rho_t \times \text{Simulated Annual Return on Stocks} + (1 - \rho_t) \times \text{Simulated Annual Return on Bonds} \tag{3}
\]

where \( \rho_t \) is the proportion of assets invested in stocks during policy year \( t \).

Thus, for example, suppose that in year 1 we have the following information:

- There is a mix of 60 percent in stocks and 40 percent in bonds;
- The annual rates of return on stocks and bonds are 0.1507 and -0.0014, respectively;
- The level net annual premium for a whole life policy covering a male age 50, face amount of $1,000, using the SOA 75-80 15 Year Select and Ultimate Tables and assuming a rate of interest of 6 percent, is \( P_1 = 16.38 \); and
- The expected claims cost for year 1 is \( C_1 = 1.7 \).

We can determine \( F_1 \) and \( N_1 \) as follows:

\[
F_1 = (0 + 16.38) \times 1.06 - 1.7 \times (1.06)^{1/2} = 15.61
\]

\( \rho_1 = 0.6 \) and

\[
s_1 = 0.6 \times 0.1507 + 0.4 \times -0.0014 = 0.8986\%
\]

yielding

\[
N_1 = (0 + 16.38) \times (1 + 0.08986) - 1.7 \times (1 + 0.08986)^{1/2} = 16.08.
\]

As the simulated fund is in excess of the target fund, the office may be off to a good start.

4 The Probability Distribution of Asset Classes

The most difficult aspect of the construction of the model is determining the probability distribution of the available asset classes. To
avoid complicated analysis, we consider exclusively common stocks and long-dated government bonds. In our opinion this is a reasonable starting point for any discussion of the basic asset allocation decision process for a U.S. financial institution.

We use the annual total returns for common stocks and long-term bonds compiled by Ibbotson Associates of Chicago. The data for these returns go back as far as 1926 and are shown in Table 1.

The first, and most critical, step in using this historical data is to establish the framework in which the data set can be used as a forecasting tool. The objective is to use the historical data as a basis for saying something about future returns. This raises three questions:

- How much emphasis do we place on old data?
- Do returns move randomly over time? And
- Is there a relationship between stock and bond returns?

Indeed, the question of whether the past is any indicator of the future is contentious in itself. This last assumption is not justified by either intuition or empirical evidence—it is a simplifying assumption for the purposes of the example presented in this paper. The decision on how to model returns could affect the results of the model materially.

Considerable evidence exists to justify that stock prices, like bond prices, vary inversely with interest rate movements—see Solnik (1983) and Peavy (1992) for good discussions on the subject—so that some correlation should be recognized between stock and bond returns. In order to keep the model simple and to concentrate on illustrating ideas outside those of modeling stock and bond returns, however, we employ the assumption that both stock and bond returns move randomly and independently of each other. Re-running the model to incorporate an approach that correlates successive returns in some fashion or recognizes a relationship between stock and bond returns would introduce a major layer of complexity to the modeling process.

To test the success of a fund in meeting its liabilities using any particular mix of stocks and bonds (assuming random returns), we derive a large number of potential individual investment scenarios by creating a set of random rates of return for each year for which the projection is made, where these random rates of return are based on cumulative probability distributions constructed from the historical data. The projection period extends to the year in which all policyholders are expected to have died, in this case 52 years on the basis of the SOA 75-80 table for a portfolio comprising exclusively 50 year old males.
### Table 1

**Annual Returns for Common Stocks and Long-Term U.S. Government Bonds**

<table>
<thead>
<tr>
<th>Year</th>
<th>Stocks (%)</th>
<th>Bonds (%)</th>
<th>Year</th>
<th>Stocks (%)</th>
<th>Bonds (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1926</td>
<td>11.62</td>
<td>7.77</td>
<td>1959</td>
<td>11.96</td>
<td>2.26</td>
</tr>
<tr>
<td>1927</td>
<td>37.49</td>
<td>8.93</td>
<td>1960</td>
<td>0.47</td>
<td>13.78</td>
</tr>
<tr>
<td>1928</td>
<td>43.61</td>
<td>0.10</td>
<td>1961</td>
<td>26.89</td>
<td>0.97</td>
</tr>
<tr>
<td>1929</td>
<td>(8.42)</td>
<td>3.42</td>
<td>1962</td>
<td>(8.73)</td>
<td>6.89</td>
</tr>
<tr>
<td>1930</td>
<td>(24.90)</td>
<td>4.66</td>
<td>1963</td>
<td>22.80</td>
<td>1.21</td>
</tr>
<tr>
<td>1931</td>
<td>(43.34)</td>
<td>(5.31)</td>
<td>1964</td>
<td>16.48</td>
<td>3.51</td>
</tr>
<tr>
<td>1932</td>
<td>(8.19)</td>
<td>16.84</td>
<td>1965</td>
<td>12.45</td>
<td>0.71</td>
</tr>
<tr>
<td>1933</td>
<td>53.99</td>
<td>(0.07)</td>
<td>1966</td>
<td>(10.06)</td>
<td>3.65</td>
</tr>
<tr>
<td>1934</td>
<td>(1.44)</td>
<td>10.03</td>
<td>1967</td>
<td>23.98</td>
<td>(9.18)</td>
</tr>
<tr>
<td>1935</td>
<td>47.67</td>
<td>4.98</td>
<td>1968</td>
<td>11.06</td>
<td>(0.26)</td>
</tr>
<tr>
<td>1936</td>
<td>33.92</td>
<td>7.52</td>
<td>1969</td>
<td>(8.50)</td>
<td>(5.07)</td>
</tr>
<tr>
<td>1937</td>
<td>(35.03)</td>
<td>0.23</td>
<td>1970</td>
<td>4.01</td>
<td>12.11</td>
</tr>
<tr>
<td>1938</td>
<td>31.12</td>
<td>5.53</td>
<td>1971</td>
<td>14.31</td>
<td>13.23</td>
</tr>
<tr>
<td>1939</td>
<td>(0.41)</td>
<td>5.94</td>
<td>1972</td>
<td>18.98</td>
<td>5.69</td>
</tr>
<tr>
<td>1940</td>
<td>(9.78)</td>
<td>6.09</td>
<td>1973</td>
<td>(14.66)</td>
<td>(1.11)</td>
</tr>
<tr>
<td>1941</td>
<td>(11.59)</td>
<td>0.93</td>
<td>1974</td>
<td>(26.47)</td>
<td>4.35</td>
</tr>
<tr>
<td>1942</td>
<td>20.34</td>
<td>3.22</td>
<td>1975</td>
<td>37.20</td>
<td>9.20</td>
</tr>
<tr>
<td>1943</td>
<td>25.90</td>
<td>2.08</td>
<td>1976</td>
<td>23.84</td>
<td>16.75</td>
</tr>
<tr>
<td>1944</td>
<td>19.75</td>
<td>2.81</td>
<td>1977</td>
<td>(7.18)</td>
<td>(0.69)</td>
</tr>
<tr>
<td>1945</td>
<td>36.44</td>
<td>10.73</td>
<td>1978</td>
<td>6.56</td>
<td>(1.18)</td>
</tr>
<tr>
<td>1946</td>
<td>(8.07)</td>
<td>(0.10)</td>
<td>1979</td>
<td>18.44</td>
<td>(1.23)</td>
</tr>
<tr>
<td>1947</td>
<td>5.71</td>
<td>(2.62)</td>
<td>1980</td>
<td>32.42</td>
<td>(3.95)</td>
</tr>
<tr>
<td>1948</td>
<td>5.50</td>
<td>3.40</td>
<td>1981</td>
<td>(4.91)</td>
<td>1.86</td>
</tr>
<tr>
<td>1949</td>
<td>18.79</td>
<td>6.45</td>
<td>1982</td>
<td>21.41</td>
<td>40.36</td>
</tr>
<tr>
<td>1950</td>
<td>31.71</td>
<td>0.06</td>
<td>1983</td>
<td>22.51</td>
<td>0.65</td>
</tr>
<tr>
<td>1951</td>
<td>24.02</td>
<td>(3.93)</td>
<td>1984</td>
<td>6.27</td>
<td>15.48</td>
</tr>
<tr>
<td>1952</td>
<td>18.37</td>
<td>1.16</td>
<td>1985</td>
<td>32.16</td>
<td>30.97</td>
</tr>
<tr>
<td>1953</td>
<td>(0.99)</td>
<td>3.64</td>
<td>1986</td>
<td>18.47</td>
<td>24.53</td>
</tr>
<tr>
<td>1954</td>
<td>52.62</td>
<td>7.19</td>
<td>1987</td>
<td>5.23</td>
<td>(2.71)</td>
</tr>
<tr>
<td>1955</td>
<td>31.56</td>
<td>(1.29)</td>
<td>1988</td>
<td>16.81</td>
<td>9.67</td>
</tr>
<tr>
<td>1956</td>
<td>6.58</td>
<td>(5.59)</td>
<td>1989</td>
<td>31.49</td>
<td>18.11</td>
</tr>
<tr>
<td>1957</td>
<td>(10.78)</td>
<td>7.46</td>
<td>1990</td>
<td>(3.17)</td>
<td>6.18</td>
</tr>
<tr>
<td>1958</td>
<td>43.36</td>
<td>(6.09)</td>
<td>1991</td>
<td>30.55</td>
<td>19.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1992</td>
<td>7.67</td>
<td>8.05</td>
</tr>
</tbody>
</table>

5 A Note on the Number of Simulations

We need a sufficiently large number of simulations to ensure that a smooth curve can be drawn between any set of risk/reward points at a particular interest rate assumption and to accurately estimate the probability of insolvency, \( p \). Suppose we perform \( n \) simulations and \( X \) is the number of times insolvency results, then we can estimate the probability of insolvency by \( \hat{p} = X/n \). However, to ensure that our estimate has a high probability, say \( 1 - \epsilon \) (for small \( \epsilon > 0 \)), of being within, say, a margin of 100\( \alpha \)% of the true value \( p \), we must ensure that \( n \) satisfies

\[
\Pr[(1 - \alpha)p \leq X/n \leq (1 + \alpha)p] \geq 1 - \epsilon.
\]

As \( n \) is large, we can appeal to the central limit theorem and show that the smallest value of \( n \) is given by:

\[
n \geq \left( \frac{Z_y}{\alpha} \right)^2 \left( \frac{1 - p}{p} \right)
\]

where \( Z_y \) is the 100(1 - \( y \))% percentage point of the standard normal distribution. Suppose \( p \) is calculated using \( n = 1,000 \) simulations and the result is \( p = 0.1 \). Then if \( \epsilon = 0.05 \), equation (4) gives the minimum value of \( \alpha = 0.186 \), derived from

\[
1000 = \left( \frac{1.96}{\alpha} \right)^2 \left( \frac{1 - 0.1}{0.1} \right).
\]

This is not small enough to make a credible graphical presentation of a smooth risk profile. For example, if the values of \( p \) were plotted, a smooth curve (shape) would not be achieved, but a sample of random points would result. If 25,000 simulations were used for \( p = 0.1 \), however, the margin \( \alpha \) would be 0.037, which is accurate enough for graphical purpose. As a result, throughout this paper we use 25,000 simulations.

6 The Simulation Process

The life office's simulation process must be built into its liability cash flow framework. This means the target fund needs to be compared with the simulated fund in each year of projection, as derived under each simulated investment scenario. The progress of the target and simulated funds is tracked for the full expected future term of the business in force. This is repeated for various simulated stock
and bond returns—as discussed in Section 5, the model has been run using 25,000 simulations—using all possible combinations of stocks and bonds in steps of 1 percent and using liability profiles based on actuarial interest rate assumptions of 0 percent, 2 percent, 4 percent, 6 percent, and 8 percent.

For each simulated investment scenario, the internal rate of return for each mix of stocks and bonds is calculated as:

$$r_n = \left[ \prod_{t=1}^{n} (1 + s_t) \right]^{1/n} - 1$$

where $n$ is the projection period (in years). This return ($r_n$) is averaged over the 25,000 simulated scenarios to derive an expected rate of return on the fund for any particular mix of stocks and bonds. This expected rate remains the same regardless of the liability profile under consideration.

The next step is to determine how risk should be specified within the framework of the cash flow projections for any particular liability profile. This assessment is critical to the modeling process. For the purposes of this paper, our measure of risk is defined as the probability of the simulated fund being less than the target fund for three consecutive years during the full projection period. A three year period (rather than a one year time horizon) is chosen on the premise that if the fund has gone this amount of time in an unbalanced financial position, it may have long-term financial problems.

This definition of risk is a convenient way of assessing the real risk in the example used (the real risk is underperforming the pricing assumptions over the lifetime of the policies) because of the way in which solvency is defined (i.e., in terms of market value of assets versus liabilities valued on the basis of original pricing assumptions). This would not be the case in a more sophisticated model that incorporates valuation margins and capital in its computations. In such instances an alternative measure of underperformance against pricing assumptions may be more satisfactory, with the risk of insolvency incorporated as a constraint. In addition, risk should be measured not just by the probability of underperforming, but by the amount of underperformance.

We now formally define our measure of risk, $R$, algebraically as follows: For $t = 3, 4, \ldots, n$, let

$$R_t = \begin{cases} 
1 & \text{if } N_{t-k} < F_{t-k} \text{ for } k = 0, 1, 2; \text{ and} \\
0 & \text{otherwise.}
\end{cases}$$
The measure of risk for any particular mix of stocks and bonds is the sum of all values of \( R \) over the 25,000 simulations, divided by 25,000 to give an average probability of insolvency.

7 The Monte Carlo Sampling Method Used

In order to use Monte Carlo simulation, we first need to specify a distribution function of asset returns. In our case this function is an empirical function. If the rate of return on a particular asset class is defined as a random variable, \( S \), then the empirical probability distribution function (pdf), \( f(s) \) and the empirical cumulative probability distribution (cdf), \( F(s) \), of \( S \) must be determined.

Suppose we have data on \( S \) and we construct a relative frequency histogram with \( m \) (a positive integer) distinct intervals such that a return of \( s_{k-1} < S \leq s_k \) occurs with relative frequency \( f_k \geq 0 \), for \( k = 1, 2, \ldots, m \) with \( \sum_{i=1}^{m} f_k = 1 \). Following Hogg and Klugman (1984, Chapter 3), we can construct a continuous cdf using a piecewise linear approximation. First we choose a sequence of points \( \{c_k\} \) such that \( s_{k-1} < c_k < s_k \) for \( k = 1, 2, \ldots, m - 1 \), and \( c_0 = s_0 \) and \( c_m = s_m \). The \( c_k \)'s do not have to be equidistant. It can easily be verified that the cdf is given by

\[
F(s) = \begin{cases} 
0 & \text{for } s < c_0 \\
(s - c_0) f_1 / (c_1 - c_0) & c_0 \leq s \leq c_1 \\
\vdots & \vdots \\
(s - c_{k-1}) f_k / (c_k - c_{k-1}) + \sum_{j=1}^{k-1} f_j & c_{k-1} \leq s \leq c_k \\
\vdots & \vdots \\
(s - c_{m-1}) f_m / (c_m - c_{m-1}) + \sum_{j=1}^{m-1} f_j & c_{m-1} \leq s \leq c_m \\
1 & s > c_m.
\end{cases}
\] (7)

Having defined the cumulative distribution function of \( S \), it is now possible to demonstrate how the random variable \( S \) is simulated (i.e., how samples of the observation of the variable \( S \) are generated). Let \( U \) be a uniform distribution on \([0,1]\). The standard approach to generating a random variable \( S \) is as follows (see, for example, Bratley, Fox, and Schrage (1983, Chapter 5.2.2)): Suppose \( U_i \) is a random observation from \( U \). We must determine the \( c_j \) be such that \( c_j \leq U_i \leq c_{j+1} \). It
follows that the corresponding observation from $S$ is $S_t$ where

$$S_t = \frac{[F(c_{j+1}) - U_t] \times c_j + [U_t - F(c_j)] \times c_{j+1}}{F(c_{j+1}) - F(c_j)}.$$

(8)

Table 2 shows the cumulative distributions for the two asset classes (common stocks and long-term bonds) used in the model. Table 2 is derived from the basic data of Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Common Stocks</th>
<th>Long-Term Bonds</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>$c_k$</td>
<td>$c_k$</td>
</tr>
<tr>
<td>0</td>
<td>-45</td>
<td>-10</td>
</tr>
<tr>
<td>1</td>
<td>-40</td>
<td>-5</td>
</tr>
<tr>
<td>2</td>
<td>-35</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>-30</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>-25</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>-20</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>-15</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>-10</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>-5</td>
<td>30</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>40</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>45</td>
</tr>
<tr>
<td>12</td>
<td>15</td>
<td>50.75%</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>62.69%</td>
</tr>
<tr>
<td>14</td>
<td>25</td>
<td>73.13%</td>
</tr>
<tr>
<td>15</td>
<td>30</td>
<td>76.12%</td>
</tr>
<tr>
<td>16</td>
<td>35</td>
<td>88.06%</td>
</tr>
<tr>
<td>17</td>
<td>40</td>
<td>92.54%</td>
</tr>
<tr>
<td>18</td>
<td>45</td>
<td>95.52%</td>
</tr>
<tr>
<td>19</td>
<td>50</td>
<td>97.01%</td>
</tr>
<tr>
<td>20</td>
<td>55</td>
<td>100.00%</td>
</tr>
</tbody>
</table>
8 Results of the Simulation

The results for each liability profile under consideration (i.e., for each rate of interest assumption) using various combinations of stocks and bonds are summarized in Figure 1. In particular, Figure 1 shows the optimal risk/reward points for various possible combinations of stocks and bonds under the various interest rate assumptions used in pricing the underlying liabilities. A curve is drawn through the points to create a risk/reward profile for each rate of interest. All points on any line to the left of the minimum risk point can be ignored, because it is possible to achieve simultaneously a higher return and a lower risk by altering the mix of stocks and bonds.

The point at which the minimum level of risk is achieved depends on the liability structure under consideration (i.e., assumed interest rate used for pricing). At a rate of interest of 2 percent the minimum risk is
achieved where 32 percent of the fund is held in stocks and 68 percent of the fund is held in bonds. This minimum risk point shifts toward a heavier weighting in stocks as the rate of interest rises; at high rates of interest the minimum risk point is not achieved until 100 percent is held in stocks.

Our results are intuitive, i.e., if there is a high minimum guarantee, the office will be driven to more volatile assets (with higher potential upside and downside) because they are the only assets with a chance of outperforming the guarantee. An easy target is associated with a conservative strategy, and a difficult target is associated with a not-so-conservative strategy. Although the not-so-conservative strategy often will miss the difficult target, it nevertheless has a higher probability of exceeding the target than a more conservative strategy that never hits the target. On the other hand, if a company is adequately capitalized, a 2 percent minimum guarantee will not result in a low stock holding. Further, there is a business risk associated with not being able to offer competitive guarantees.

The curves based on the low rate of interest assumptions look like traditional efficient frontiers. This is not surprising. At relatively low rates of interest the nature of the liabilities becomes relatively unimportant, so the model reverts to the conventional asset-only model. But at relatively high rates of interest the concept of an efficient frontier collapses, and at a rate of interest of 10 percent there is only one efficient point (where 100 percent is held in stocks).

In practice a life office may be required to hold certain asset categories. For example, there may be an investment policy constraint within the office that at least 50 percent of the portfolio must be held in long-term bonds. Moreover, in many countries there are legal restrictions on the extent to which certain categories of asset may be held by life offices. Thus, this paper does not concentrate on analyzing the efficient combinations of the various asset classes.

The final part of the exercise is to determine an acceptable level of risk; having decided this, it is possible to derive a uniquely defined optimal asset mix. For example, for the fund that has used a rate of interest assumption of 2 percent in its pricing assumptions, it may be appropriate to go 100 percent into stocks (and therefore go for the maximum possible return) if a probability of insolvency level of around 30 percent were deemed acceptable.

Setting an acceptable level of risk is a largely subjective decision, and in practice the usefulness of this model is in assessing the relative riskiness of various portfolio mixes rather than in making any sense of
the absolute values generated for the risk and reward of any particular investment policy in isolation. The absolute values for the probability of insolvency in the model look extremely high across the board, the result of the relatively large probability of a market crash in any one investment scenario. See Hardy (1993) for similar findings when using a stochastic model.

A highly artificial liability profile is being considered, and the fact that the results drive toward a higher stock allocation than most companies hold in practice indicate the failings of the simplified model. A significant shortfall of the model is that the RBC implications of any particular asset allocation recommendation have not been considered; in practice, this would be a major constraint on the asset allocation decision.

Because the example in this paper has been highly simplified, the liability structure is expressed entirely in terms of a pricing interest rate. Similar efficient frontiers could be created by merely comparing the simulated portfolio returns to the 0 percent to 8 percent pricing rates.

Finally, the complete asset allocation model should incorporate the full range of assets available to the financial institution, which for a life office should include cash, property, and overseas stock and bond investment in addition to domestic stocks and bonds.

9 Summary and Conclusion

This paper describes an approach to asset allocation modeling for institutions that invest to meet liabilities. The model is consistent with conventional financial economics. Traditional risk/reward profiles become apparent where the nature of the liabilities is not considered or is relatively unimportant, but such traditional risk/reward profiles may or may not become apparent once the nature of the liabilities is introduced. Thus, the traditional ideas of financial economics have been shown to be a special case of the more general asset allocation system using a true ALM model.

We have concentrated exclusively on the applications of an ALM model in the context of a highly simplified life office issuing purely nonparticipating whole life assurance. The principles can be applied equally, however, to any financial institution that is concerned with investing to meet liabilities.
The critical element of the model is the definition of risk. It is not important that risk is taken as some measure of exposure to insolvency, but that it incorporates the liabilities.

Refinement of the model to incorporate the features of participating business should not be problematic; this would be akin to lowering the rate of interest assumption used in pricing the liabilities which implicitly means a general reduction in the risk profile and, hence, potentially greater freedom in investment policy.

The application of our model to a pension fund poses some interesting issues, although these issues are specific to the particular country under consideration. Although it is recognized that pension funds can overcome deficit situations by increasing contribution rates from time to time, pension fund trustees may be interested in knowing whether a particular investment policy is more likely to lead to persistent deficits. Alternatively, if a primary objective of pension fund investment policy is to avoid unduly fluctuating contribution rates, then the ALM model could use a refined definition of risk (say, the probability of the fund falling outside a certain surplus or deficit range).

Incorporating the inflationary aspects of a pension fund model is problematic. Possible approaches include linking inflation to the yield curve or stochastically modeling inflation as an independent variable. It is not obvious which of the two approaches may be more appropriate; perhaps, given the major uncertainties associated with inflation, there is no definitive model. Some innovative research in this area has been done by Wilkie (1986).

There remains much exciting asset/liability modeling work to be done. Dramatic developments can be expected as microcomputer processing power becomes more widely appreciated and makes the type of stochastic model described in this paper a standard tool of financial analysis.

References


A Possibilistic Linear Programming Method for Asset Allocation

Lijia Guo* and Zhen Huang†

Abstract‡

The mean-variance method has been one of the popular methods used by most financial institutions in making the decision of asset allocation since the 1950s. This paper presents an alternative method for asset allocation. Instead of minimizing risk for a given expected return or maximizing expected return for a fixed level of risk, our approach considers simultaneously maximizing the rate of return of portfolio, minimizing the risk of obtaining lower return, and maximizing the possibility of reaching higher return. By using a triangular possibilistic distribution to describe the uncertainty of the return, we introduce a possibilistic linear programming model which we solve by a multiple objective linear programming technique with two control constraints. We present a solution algorithm that provides maximal flexibility for decision makers to effectively balance the portfolio's return and risk. Numerical examples show the efficiency of the algorithm.

Key words and phrases: Mean-variance method; possibility distribution; multiple-objective, fuzzy sets.

*Lijia Guo, Ph.D., A.S.A., is an assistant professor of actuarial science at Ball State University.

Dr. Guo's address is: Department of Mathematical Science, Ball State University, Muncie, IN 47304 U.S.A. Internet address: lguo@math.bsu.edu

†Zhen Huang is an assistant professor of mathematics at Otterbein College.

Dr. Huang's address is: Department of Mathematical Sciences, Otterbein College, Westerville, OH 43081. Internet address: zhuang@otterbein.edu

‡This research was completed while Dr. Guo was at Ohio State University. Dr. Guo would like to acknowledge the financial support from the Ohio State University Actuarial Faculty Fund. The authors would like to thank the anonymous referees for their helpful suggestions and comments.
1 Introduction

Asset allocation decisions are often reached through a three-step process: first, the risk and return characteristics of available and relevant investment opportunities are identified; second, the investor's risk tolerances or expected returns are parameterized; and finally, the risk-return trade-offs of the investor are combined with those observed in the market to produce an optimal asset allocation. A frequently used tool for asset allocation problems is the mean-variance optimization technique developed by Markowitz (1952).

Mean-variance optimization refers to a mathematical process to determine the security (or asset class) weights that provide a portfolio with the minimum risk for a given expected return or, conversely, the maximum expected return for a given level of risk. The inputs needed to conduct mean-variance optimization are security expected returns, expected standard deviations, and expected cross-security correlation. The Markowitz model has been one of the methods widely used by institutional investors, retail brokerage houses, and pension fund managers.

Another type of asset allocation strategy is dynamic asset allocation, which continually adjusts a portfolio's allocation in response to changing market conditions. The most popular use of these strategies is portfolio insurance, which attempts to remove the downside risk faced by a portfolio. A popular means of implementing portfolio insurance is to engage in a series of transactions that give the portfolio the return distribution of a call option. Black and Scholes (1973) show that under certain assumptions, the payoff of an option can be duplicated through a continuously revised combination of the underlying asset and a risk-free bond. Rubenstein and Leland (1981) extend this insight by showing that a dynamic strategy that increases the stock allocation of a portfolio in rising markets and reinvested the remaining portion in cash would replicate the payoffs to a call option on an index of stocks.

Portfolio insurance concentrates on only two assets, both of which are carefully predetermined. To the extent that its assumptions about the behavior of uninsured investors turn out to be less than 100% correct, however, the increasing volatility of risky assets could drive insured portfolios to sell or buy even more aggressively than they would have in the first place; see Sharpe (1992).

On the other hand, tactical asset allocation (active asset allocation) is the process of diverging from the strategic asset allocation when an investor's short-term forecasts deviate from the long-term forecasts used to formulate the strategic allocation. If the investor can make accurate short-term forecasts, tactical asset allocation has the potential to en-
hance returns. In practice, tactical asset allocators are the investors providing portfolio insurance; see Sharp and Perrold (1988).

In this study, instead of the traditional mean-variance approach, we describe the uncertainty of the rate of return by a triangular possibilistic distribution. A possibilistic linear programming model (see, Lai and Hwang (1992, Chapter 5)) is formulated and then solved by introducing two control constraints to the auxiliary multi-objective linear programming model. By selecting different values for the parameters in control constraints, our method can be applied in solving the following problems:

- Maximizing the most possible return and minimizing the risk of obtaining lower return as well as maximizing the possibility of obtaining higher return.
- Minimizing the risk of obtaining lower return and maximizing the possibility of obtaining higher return for a specified most possible return.
- Maximizing the most possible return and maximizing the possibility of obtaining higher return for a given risk tolerance.

2 Models

Let us consider the problem of allocating capital $C$ among $N$ asset classes, $S_1, S_2, \ldots, S_N$. In the mean-variance optimization method [Fong and Fabozzi (1992)], the rate of return, $R_i$, of asset $S_i$ is assumed to be a random variable with $\mu_i$ and $\sigma_i$ denoting the mean and standard deviation, respectively, of $R_i$ for $i = 1, 2, \ldots, N$, and $\rho_{ij}$ denoting the correlation between $R_i$ and $R_j$ for $i, j = 1, 2, \ldots, N$.

If the $N$ assets are combined linearly to form a portfolio, where the allocation weight, $x_i$, for asset class $S_i$, is equal to the dollar value of the asset class relative to the dollar value of the portfolio, then the rate of return of the portfolio is

$$R_p = \sum_{i=1}^{N} x_i R_i$$

which is also a random variable. The expected return of the portfolio is

$$\mu_p = \sum_{i=1}^{N} x_i \mu_i$$
and the variance of the portfolio is
\[ \sigma_p^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} x_i x_j \sigma_i \sigma_j \rho_{ij}. \]

2.1 Mean-Variance Analysis

The mean-variance method for determining weights \( x_1, x_2, \ldots, x_N \) is to fix the expected portfolio return \( \mu_p \) to a desirable level \( \mu \) and determine the allocation weights \( x_1, x_2, \ldots, x_N \) that minimize the risk level \( \sigma_p^2 \) of the portfolio for the fixed \( \mu \). The following quadratic programming model (1) is employed to accomplish this goal:

**Model 1:**

\[
\begin{align*}
\text{min} & \quad \sigma_p^2 = \sum_{i,j=1}^{N} x_i x_j \sigma_i \sigma_j \rho_{ij} \\
\text{subject to} & \quad \sum_{i=1}^{N} x_i \mu_i = \mu \\
& \quad \sum_{i=1}^{N} x_i = 1 \\
& \quad l_i \leq x_i \leq u_i, \quad i = 1, 2, \ldots, N
\end{align*}
\]

where \( l_i \) is the lower bound and \( u_i \) is the upper bound on funds allocated to the \( i \)-th asset class, \( i = 1, 2, \ldots, N \).

An equivalent approach is to fix the risk level \( \sigma_p^2 \) of a portfolio to a tolerable level \( \sigma^2 \) and determine the weights \( x_1, x_2, \ldots, x_N \) that maximize the expected portfolio return \( \mu_p \) for the fixed \( \sigma^2 \). The following quadratic programming model (2) is employed to accomplish this goal:

**Model 2:**

\[
\begin{align*}
\text{max} & \quad \mu_p = \sum_{i=1}^{N} x_i \mu_i \\
\text{subject to} & \quad \sum_{i,j=1}^{N} x_i x_j \sigma_i \sigma_j \rho_{ij} = \sigma^2 \\
& \quad \sum_{i=1}^{N} x_i = 1 \\
& \quad l_i \leq x_i \leq \mu_i, \quad i = 1, 2, \ldots, N.
\end{align*}
\]
2.2 Mean-Variance-Skewness Analysis

The mean-variance method, which does not consider the skewness of the return random variable $R_p$, is frequently used in practice for asset allocation. In a continuous time model with asset prices following a diffusion process, Ito's differentiation rule\(^1\) implies the higher moments are irrelevant to asset allocation decisions. In this case, the mean-variance method provides optimal portfolio selection. In a discrete model, however, Samuelson (1970) shows that the mean-variance efficiency becomes inadequate and the higher moments become relevant to the portfolio selection.

It has been shown empirically by Simkowitz and Beedles (1978) and Singleton and Wingender (1986) that stock return distributions are often positively skewed. Under asymmetrically distributed asset returns, it is important to take skewness into consideration in discrete models of portfolio selection. Arditti and Levy (1975) have illustrated the important role of skewness in the pricing of stocks. As shown by Arditti (1967), the investor’s preference for more skewness is consistent with the notion of decreasing absolute risk aversion, because a positive-skewness asset return refers to a right-hand elongated tail of density function of asset return.

If the skewness of $R_p$, defined as $E[(R_p - \mu_p)^3]/\sigma_p^3$, is incorporated in the mean-variance method, model (1) then becomes a multiple objective nonlinear programming model:

\[
\text{Model 3:} \\
\begin{align*}
\min & \quad \sigma_p^2 = \sum_{i,j=1}^{N} x_i x_j \sigma_i \sigma_j \rho_{ij} \\
\max & \quad E[(R_p - \mu_p)^3]/\sigma_p^3 \\
\text{subject to} & \quad \sum_{i=1}^{N} x_i \mu_i = \mu
\end{align*}
\]

\(^1\)Ito's differentiation rule states that if $f = f(X, t)$ and $X$ follows
\[
dX = \alpha dt + \sigma dW.
\]

Then $df$ can be written:
\[
df = (\alpha f_x + \frac{1}{2} \sigma^2 f_{xx} + f_t) dt + \sigma f_x dW.
\]

while model (2) becomes the following:

Model 4:

\[
\begin{align*}
\text{max} & \quad \mu_p = \sum_{i=1}^{N} x_i \mu_i \\
\text{max} & \quad \frac{E[(R_p - \mu_p)^3]}{\sigma_p^3} \\
\text{subject to} & \quad \sum_{i,j=1}^{N} x_i x_j \sigma_i \sigma_j \rho_{ij} = \sigma^2 \\
& \quad \sum_{i=1}^{N} x_i = 1 \\
& \quad l_i \leq x_i \leq \mu_i, \quad i = 1, 2, \ldots, N
\end{align*}
\]

where

\[
E[(R_p - \mu_p)^3] = \sum_{i,j,k=1}^{N} x_i x_j x_k \sigma_{ijk}
\]

and \(\sigma_{ijk}\) is defined as

\[
E[(R_i - \mu_i)(R_j - \mu_j)(R_k - \mu_k)].
\]

The skewness of a portfolio of securities is not simply a weighted average of the skewness of the component securities. Like variance, it depends on the joint movement of securities. This means that to measure the skewness on a portfolio, a great number of estimates of joint movement must be made. As indicated by Elton and Gruber (1995), for these estimates to be feasible, it requires the type of single indexed model or multiple indexed model development to calculate the correlation structure of security returns and some simple techniques for determining the three dimensional efficient frontier. This developmental work has not been done.

In this paper, instead of dealing with models (3) or (4) directly, we present a possibilistic linear programming model to implement the idea of maximizing the expected return, minimizing the risk, and maximize skewness simultaneously without estimating the third moments.
of securities. Possibility theory studies primarily imprecise phenomena. Possibilistic decision making models handle practical decision making problems where input data are imprecise. Applications of possibility theory to linear programming problems with imprecise coefficients have been discussed by Lai and Hwang (1992).

3 The Possibilistic Model

The possibility distribution \( \pi_X(u) \) of an event \( X \) states the degree possibility of the occurrence of the event. To illustrate the difference between the possibility distribution and the probability distribution, we consider the following simple example due to Zadeh (1978, p. 8): Consider the statement "Hans ate \( X \) eggs for breakfast," where \( X = \{1, 2, \ldots\} \). A possibility distribution as well as a probability distribution may be associated with \( X \). The possibility distribution \( \pi_X(u) \) can be interpreted as the degree of ease with which Hans can eat \( u \) eggs while the probability distribution \( P_X(u) \) might have been determined by observing Hans at breakfast for 100 days. The values of \( \pi_X(u) \) and \( P_X(u) \) might be as shown in the following table: We observe that a high degree of possibility does not imply a high degree of probability. If, however, an event is not possible, it is also improbable. Thus, in a way the possibility is an upper bound for the probability. For a more detailed discussion of possibility theory, readers are referred to Zimmermann (1991, Chapter 8) or Dubois and Prade (1988).

<table>
<thead>
<tr>
<th>( u )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi_X(u) )</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0.8</td>
<td>0.6</td>
<td>0.4</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>( P_X(u) )</td>
<td>0.1</td>
<td>0.8</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

For our model, we use the possibility distribution to describe the uncertainty of the rate of return. Because uncertainty from the return of assets can be regarded as the nature of imprecision, possibility distributions are suitable for characterizing such kinds of uncertainty. Moreover, using the possibility distribution may also reduce the impact of the underlying structure of the asset market shifts.

For the \( i \)-th asset \( S_i, i = 1, 2, \ldots, N \), we describe the imprecise rate of return by \( \bar{r}_i = (r_{i}^p, r_{i}^m, r_{i}^o) \), where \( r_{i}^p \), \( r_{i}^m \), and \( r_{i}^o \) are the most pessimistic value, the most possible value, and the most optimistic value for the rate of return, respectively. Assume that the imprecise rate of
return, $\tilde{r}_i = (r_i^p, r_i^m, r_i^o)$, with $r_i^p < r_i^m < r_i^o$, has the triangular possibility distribution $\pi_{\tilde{r}_i}$ defined as:

$$
\pi_{\tilde{r}_i}(r) = \begin{cases} 
0 & \text{for } r < r_i^p \text{ or } r > r_i^o \\
1 & \text{for } r = r_i^m \\
(r - r_i^p)/(r_i^m - r_i^p) & \text{for } r_i^p \leq r < r_i^m \\
(r_i^o - r)/(r_i^o - r_i^m) & \text{for } r_i^m \leq r < r_i^o
\end{cases}
$$

(1)

and is displayed in Figure 1.

Figure 1

The Triangular Possibility Distribution of $\tilde{r}_i$

As shown in Figure 1, possibility distribution of the rate of return describes the possibility degree of occurrence of each possible rate of return. For example, if for $i$-th asset $S_i$, $\pi_{\tilde{r}_i}(0.10) = 0.8$, then the possibility degree of occurrence of $n = 10\%$ is 0.8.

Next, let $x_i$ denote the allocation weight and $\tilde{r}_i = (r_i^p, r_i^m, r_i^o)$ denote the imprecise rate of return to asset $S_i$ for $i = 1, 2, \ldots, N$. Then the imprecise rate of return for the portfolio is

$$
\tilde{r} = \sum_{i=1}^{N} \tilde{r}_i x_i.
$$
Linear combinations of triangular possibility distributions are also triangular possibility distribution, and $\bar{r} = (r^p, r^m, r^o)$ is given by

$$r^p = \sum_{i=1}^{N} r^p_i x_i, \quad r^m = \sum_{i=1}^{N} r^m_i x_i, \quad r^o = \sum_{i=1}^{N} r^o_i x_i.$$  

Notice that the triangular possibility distribution $\pi_{\bar{r}}$ for $\bar{r}$, as indicated by a bold triangle in Figure 2, is determined by $r^p$, $r^m$, and $r^o$ according to the definition of $\pi_{\bar{r}}$.

**Figure 2**

The Triangular Possibility Distribution of $\bar{r}$

We now select the optimal portfolio that maximizes the portfolio return by solving the following possibilistic linear programming model:

Model 5:

$$\max \sum_{i=1}^{N} \tilde{r}_i x_i$$

subject to

$$\sum_{i=1}^{N} x_i = 1$$

$$l_i \leq x_i \leq u_i, \quad i = 1, 2, \ldots, N.$$
4 Solution Procedures

From Figure 2, we observe that $r^m$ has the highest degree of possibility to be the rate of return for the portfolio; we therefore define portfolio return as $r^m$. We also notice that, in Figure 2, the larger the area of region (I) is, the more possible it is for the portfolio to obtain lower return. As the area of region (I) is $(r^m - r^p)/2$, we define portfolio risk as $(r^m - r^p)$. Similarly, $(r^o - r^m)/2$ is the area of region (II) in Figure 2. Larger values of $(r^o - r^m)/2$ indicate higher degrees of possibility for the portfolio to reach higher return. We define portfolio skewness as $(r^o - r^m)$.

In order to maximize the imprecise rate of portfolio return, $\tilde{r}$, we select the optimal portfolio in the sense of maximizing portfolio return, minimizing portfolio risk, and maximizing portfolio skewness. Therefore model (5) can be approximated by the multiple objective linear programming model (6):

**Model 6:**

\[
\begin{align*}
\text{max} & \quad z^{(1)} = \sum_{i=1}^{N} r^m_i x_i \\
\text{min} & \quad z^{(2)} = \sum_{i=1}^{N} (r^m_i - r^p_i) x_i \\
\text{max} & \quad z^{(3)} = \sum_{i=1}^{N} (r^o_i - r^m_i) x_i \\
\text{subject to} & \quad \beta_i \leq \sum_{i=1}^{N} r^m_i x_i \leq \beta_u \\
& \quad \gamma_i \leq \sum_{i=1}^{N} (r^m_i - r^p_i) x_i \leq \gamma_u \\
& \quad \sum_{i=1}^{N} x_i = 1 \\
& \quad l_i \leq x_i \leq \mu_i, \quad i = 1, 2, \ldots, N.
\end{align*}
\]

There are three objectives in model (6):

- The first objective is to maximize portfolio return;
- The second objective is minimizing portfolio risk; and
- The third objective is maximizing portfolio skewness.
This strategy is essentially analogous to maximizing mean return, minimizing variance, and maximizing skewness for a random rate of return. In Figure 2, the triangle made by the thin lines denotes the optimal triangular possibility distribution for the imprecise rate of return for the portfolio.

4.1 Selecting the Parameters

Model (6) has two control constrains:

\[ \beta_l \leq \sum_{i=1}^{N} r_{i}^{m} x_{i} \leq \beta_u \]

and

\[ \gamma_l \leq \sum_{i=1}^{N} ( r_{i}^{m} - r_{i}^{p} ) x_{i} \leq \gamma_u . \]

By selecting parameters \( \beta_l \) and \( \beta_u \), the decision makers could use the first control constraint to assure portfolio return within the desirable range. On the other hand, by selecting parameters \( \gamma_l \) and \( \gamma_u \), the second control constraint can be used to adjust portfolio risk to a tolerable range. In the following, we discuss three special cases for selecting parameters \( \beta_l, \beta_u, \gamma_l, \) and \( \gamma_u \).

Case 1: If we set

\[ \beta_l = \min_{i=1,2,...,N} \{ r_{i}^{m} \}, \quad \beta_u = \max_{i=1,2,...,N} \{ r_{i}^{m} \}, \]

and

\[ \gamma_l = \min_{i=1,2,...,N} \{ r_{i}^{m} - r_{i}^{p} \}, \quad \gamma_u = \max_{i=1,2,...,N} \{ r_{i}^{m} - r_{i}^{p} \}, \]

both control constraints become inactive and model (6) is reduced to model (7), as proposed by Lia and Hwang (1992):

Model 7:

\[
\begin{align*}
\max \quad & z^{(1)} = \sum_{i=1}^{N} r_{i}^{m} x_{i} \\
\min \quad & z^{(2)} = \sum_{i=1}^{N} ( r_{i}^{m} - r_{i}^{p} ) x_{i} \\
\max \quad & z^{(3)} = \sum_{i=1}^{N} ( r_{i}^{o} - r_{i}^{m} ) x_{i}
\end{align*}
\]
subject to \[ \sum_{i=1}^{N} x_i = 1 \]
\[ l_i \leq x_i \leq \mu_i, \quad i = 1, 2, \ldots, N. \]

**Case 2:** If we set \( \beta_l = \beta_u = \beta \), a constant,
\[ y_l = \min_{i=1,2,\ldots,N} \{ r_i^m - r_i^p \}, \]
and
\[ y_u = \max_{i=1,2,\ldots,N} \{ r_i^m - r_i^p \}, \]
then the first objective and the second control constraint in model (6) become inactive. In this case, model (6) becomes the following:

**Model 8:**
\[
\begin{align*}
\min & \quad z^{(2)} = \sum_{i=1}^{N} (r_i^m - r_i^p) x_i \\
\max & \quad z^{(3)} = \sum_{i=1}^{N} (r_i^o - r_i^m) x_i \\
\text{subject to} & \quad \sum_{i=1}^{N} r_i^m x_i = \beta \\
& \quad \sum_{i=1}^{N} x_i = 1 \\
& \quad l_i \leq x_i \leq \mu_i, \quad i = 1, 2, \ldots, N.
\end{align*}
\]

It is easy to notice the similarity between models (3) and (8).

**Case 3:** If we set \( y_l = y_u = y \), a constant,
\[ \beta_l = \min_{i=1,2,\ldots,N} \{ r_i^m \} \quad \text{and} \quad \beta_u = \max_{i=1,2,\ldots,N} \{ r_i^m \}, \]
then the second objective and the first control constraint in model (6) become inactive. In this case, model (6) becomes:

**Model 9:**
\[
\begin{align*}
\max & \quad z^{(1)} = \sum_{i=1}^{N} r_i^m x_i \\
\max & \quad z^{(3)} = \sum_{i=1}^{N} (r_i^o - r_i^m) x_i
\end{align*}
\]
subject to
\[ \sum_{i=1}^{N} (r_i^m - r_i^p) x_i = y \]
\[ \sum_{i=1}^{N} x_i = 1 \]
\[ l_i \leq x_i \leq \mu_i, \quad i = 1, 2, \ldots, N. \]

Notice that model (9) is also analogous to model (4).

The selection of parameters $\beta_l, \beta_u, \gamma_l,$ and $\gamma_u$ may be based on either experience or managerial judgment. The examples given in Section 5 show the significance of our control constraints.

4.2 The Solution

In trying to find the solution to model (6), we must remember that model (6) has three simultaneous objectives: (i) maximizing portfolio return, (ii) maximizing portfolio skewness, and (iii) minimizing portfolio risk. With these multiple conflicting and competing objectives we cannot expect to achieve the best values for all objectives simultaneously. Therefore trade-offs among conflicting objectives are necessary.

There are various techniques to handle these trade-offs. Examples of such techniques include utility theory, goal programming, fuzzy programming, or iterative approaches. In this paper, we use Zimmermann's fuzzy programming method (1978) with a normalization process to solve the multiple objective linear programming model (6).

Let $X$ denote the set of feasible solutions satisfying all the constraints in programming model (6). Next, for the objective function $z^{(1)}$ defined in model (6), we first calculate

\[ z_{\text{min}}^{(1)} = \min_{x \in X} \sum_{i=1}^{N} r_i^m x_i \]

and

\[ z_{\text{max}}^{(1)} = \max_{x \in X} \sum_{i=1}^{N} r_i^m x_i. \]

Then we define the linear membership function $\mu_{z^{(1)}}$ as

\[ \mu_{z^{(1)}}(z) = \begin{cases} 
1 & \text{if } z \geq z_{\text{max}}^{(1)}; \\
(z - z_{\text{min}}^{(1)})/(z_{\text{max}}^{(1)} - z_{\text{min}}^{(1)}) & \text{if } z_{\text{min}}^{(1)} < z < z_{\text{max}}^{(1)}; \\
0 & \text{if } z \leq z_{\text{min}}^{(1)}. 
\end{cases} \]
Now, for the second objective function $z^{(2)}$ of model (6), we calculate

$$z_{\text{min}}^{(2)} = \min_{x \in X} \sum_{i=1}^{N} (r_i^m - r_i^p) x_i$$

and

$$z_{\text{max}}^{(2)} = \max_{x \in X} \sum_{i=1}^{N} (r_i^m - r_i^p) x_i.$$

The corresponding linear membership function $\mu_{z^{(2)}}(z)$ is:

$$\mu_{z^{(2)}}(z) = \begin{cases} 
1 & \text{if } z \leq z_{\text{min}}^{(2)}; \\
\frac{z_{\text{max}}^{(2)} - z}{z_{\text{max}}^{(2)} - z_{\text{min}}^{(2)}} & \text{if } z_{\text{min}}^{(2)} < z < z_{\text{max}}^{(2)}; \\
0 & \text{if } z \geq z_{\text{max}}^{(2)}. 
\end{cases}$$

Similarly, for the objective function $z^{(3)}$ of model (6), we compute

$$z_{\text{min}}^{(3)} = \min_{x \in X} \sum_{i=1}^{N} (r_i^m - r_i^p) x_i$$

and

$$z_{\text{max}}^{(3)} = \max_{x \in X} \sum_{i=1}^{N} (r_i^m - r_i^p) x_i$$

and the corresponding linear membership function $\mu_{z^{(3)}}(z)$

$$\mu_{z^{(3)}}(z) = \begin{cases} 
1 & \text{if } z \geq z_{\text{max}}^{(3)}; \\
\frac{z - z_{\text{min}}^{(3)}}{z_{\text{max}}^{(3)} - z_{\text{min}}^{(3)}} & \text{if } z_{\text{min}}^{(3)} < z < z_{\text{max}}^{(3)}; \\
0 & \text{if } z \leq z_{\text{min}}^{(3)}. 
\end{cases}$$

Finally, we solve the following max-min problem

$$Y = \max_{x \in X} \{ \min_{x \in X} (\mu_{z^{(1)}}(x), \mu_{z^{(2)}}(x), \mu_{z^{(3)}}(x)) \}$$

(5)

to obtain the optimal allocation weights $x_1, x_2, \ldots, x_N$.

By introducing a variable $y$, equation (5) is then equivalent to a single-objective linear programming problem:
Model 10:

\[
\begin{align*}
\text{max} & \quad y \\
\text{subject to} & \quad \mu_{z^{(1)}}(x) \geq y \\
& \quad \mu_{z^{(2)}}(x) \geq y \\
& \quad \mu_{z^{(3)}}(x) \geq y \\
& \quad x \in X.
\end{align*}
\]

The optimal solution of model (10) provides a satisfying solution under the strategy of maximizing portfolio return \( r^p \), minimizing portfolio risk \( (r^m - r^p) \), and maximizing portfolio skewness \( (r^o - r^m) \).

Our algorithm for asset allocation is now summarized as follows:

**Step 1:** For each available asset \( S_i \), estimate the most possible return rate \( r_i^m \), the most pessimistic return \( r_i^p \), and the most possible return rate \( r_i^o \), \( i = 1, 2, \ldots, N. \)

**Step 2:** Determine the initial values for parameters \( \beta_l, \beta_u, \gamma_l, \) and \( \gamma_u \) calculated by:

\[
\beta_l = \min_{i=1,2,\ldots,N} \{ r_i^m \}, \quad \beta_u = \max_{i=1,2,\ldots,N} \{ r_i^m \}
\]

and

\[
\gamma_l = \min_{i=1,2,\ldots,N} \{ r_i^m - r_i^p \}, \quad \gamma_u = \max_{i=1,2,\ldots,N} \{ r_i^m - r_i^p \}.
\]

The parameters \( \beta_l, \beta_u, \gamma_l, \) and \( \gamma_u \) can also be determined by experience and managerial judgment.

**Step 3:** For each objective function \( z^{(j)} (j = 1, 2, 3) \) in model (6), use linear programming techniques to find its maximal value \( z_{max}^{(j)} \) and its minimal value \( z_{min}^{(j)} \) subjected to the four constraints in model (6).

**Step 4:** Solve the following linear programming model with \( N + 1 \) variables to determine allocation weights \( x_1, x_2, \ldots, x_N \):

\[
\begin{align*}
\text{max} & \quad y \\
\text{subject to} & \quad \sum_{i=1}^{N} r_i^m x_i - (z_{max}^{(1)} - z_{min}^{(1)}) y \geq z_{min}^{(1)}
\end{align*}
\]
\[ \sum_{i=1}^{N} (r_i^m - r_i^p) x_i + (z_{max}^{(2)} - z_{min}^{(2)}) y \leq z_{max}^{(2)} \]
\[ \sum_{i=1}^{N} (r_i^o - r_i^m) x_i + (z_{max}^{(3)} - z_{min}^{(3)}) y \geq z_{min}^{(3)} \]

**Step 5:** For the optimal solution \( x_1^*, x_2^*, \ldots, x_N^* \), calculate

\[ (r^p)^* = \sum_{i=1}^{N} r_i^p x_i^*, \quad (r^m)^* = \sum_{i=1}^{N} r_i^m x_i^* \]

and

\[ (r^o)^* = \sum_{i=1}^{N} r_i^o x_i^*. \]

If \( (r^m)^* - (r^p)^* \geq \xi \), where \( \xi \) is the risk tolerance bound, then decrease the value of \( y_u \) and goto Step 2;

Else if \( (r^m)^* \leq \eta \), where \( \eta \) is the lower bound for the most possible rate of return, then increase \( \beta_l \) and goto Step 2;

else STOP! \( x_1^*, x_2^*, \ldots, x_N^* \) is the optimal solution.

5 Numerical Examples

5.1 Data Used to Construct Examples

Assume there are six asset classes in the market and the \( i \)-th asset class has mean and standard deviation of \( \mu_i \) and \( \sigma_i \) respectively, \( i = 1, 2, \ldots, 6 \). The values of \( \mu_i \) and \( \sigma_i \) are taken from Fong and Fabozzi (1992, p. 145) and Lederman and Klein (1994, Chapter 2, p. 27). Next we define \( r_i^p, r_i^m, r_i^o \) by setting \( r_i^p = \mu_i - 2\sigma_i, r_i^m = \mu_i, \) and \( r_i^o = \mu_i + 3\sigma_i \), \( i = 1, 2, \ldots, 6 \), with some adjustment. Table 1 displays the basic data used in the examples. The data are summarized in Table 1.

**Example 1:** We solve model (7) by setting \( \beta_l = 0.05, \beta_u = 0.17, y_l = 0.008, \) and \( y_u = 0.4 \). (See Case 1 of model (6).) The optimal allocation weights are \( x_1 = 0.0061, x_2 = 0.5, x_3 = 0.0354, x_5 = 0.4584, \) and \( (r^p, r^m, r^o) = (-0.0881, 0.1078, 0.4307) \). The optimal allocation is almost a combination of the second most risky asset and the second most conservative asset.
Table 1

Data For Examples

<table>
<thead>
<tr>
<th></th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\gamma_p$</th>
<th>$\gamma_m$</th>
<th>$\gamma^0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock 1</td>
<td>0.17</td>
<td>0.200</td>
<td>-0.230</td>
<td>0.17</td>
<td>0.800</td>
</tr>
<tr>
<td>Stock 2</td>
<td>0.15</td>
<td>0.185</td>
<td>-0.220</td>
<td>0.15</td>
<td>0.750</td>
</tr>
<tr>
<td>Bound 1</td>
<td>0.12</td>
<td>0.055</td>
<td>0.010</td>
<td>0.12</td>
<td>0.270</td>
</tr>
<tr>
<td>Bound 2</td>
<td>0.08</td>
<td>0.050</td>
<td>-0.020</td>
<td>0.08</td>
<td>0.200</td>
</tr>
<tr>
<td>Cash</td>
<td>0.06</td>
<td>0.005</td>
<td>0.050</td>
<td>0.06</td>
<td>0.090</td>
</tr>
<tr>
<td>T-bill</td>
<td>0.05</td>
<td>0.004</td>
<td>0.042</td>
<td>0.05</td>
<td>0.075</td>
</tr>
</tbody>
</table>

**Example 2:** We solve model (3.3) by fixing portfolio return at 22 different values. (See Case 2 of model (6).) The computational results are summarized in the following Tables 2 and 3.

The fifth column in Table 3 gives the set $\{r \mid \pi_{\gamma}(r) \geq 0.85\}$, which contains all the possible values of the return rate whose degree of occurrence is at least 0.85. This interval is called the acceptable event with degree of occurrence at least 0.85. Similarly, the last column in Table 3 gives the acceptable event with degree of occurrence at least 0.95.

We observe that both portfolio risk $(\gamma_m - \gamma_p)$ and portfolio skewness $(\gamma^0 - \gamma_m)$ increase as portfolio return $\gamma_m$ increases, which is consistent with the fact that as $\gamma_m$ is pushed higher, more weight should be allocated to higher risk assets. We also observe that when $\gamma_m$ increases gradually, the weights are adjusted gradually, showing that our numerical results are stable.

**Example 3:** We solve model (9) by fixing portfolio risk for 22 different values. (See Case 3 of model (6).) The computational results are summarized in Tables 4 and 5.

**Example 4:** We solve model (6) by adjusting $\beta_i$ to control portfolio return. The computational results are summarized in Tables 6 and 7. From these tables we observe that $\beta_i$ controls portfolio return effectively.
<table>
<thead>
<tr>
<th>No.</th>
<th>$(\beta_1, \beta_2, \gamma_1, \gamma_2)$</th>
<th>Optimal Solution $X^*$</th>
<th>$\pi = (\rho^p, \rho^m, \rho^o)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(0.055, 0.055, 0.008, 0.4)$</td>
<td>$x_5 = 0.5, x_6 = 0.5$</td>
<td>$(0.046, 0.055, 0.0825)$</td>
</tr>
<tr>
<td>2</td>
<td>$(0.060, 0.060, 0.008, 0.4)$</td>
<td>$x_2 = 0.0454, x_3 = 0.0152, x_5 = 0.4394, x_6 = 0.5$</td>
<td>$(0.0331, 0.6, 0.1152)$</td>
</tr>
<tr>
<td>3</td>
<td>$(0.065, 0.065, 0.008, 0.4)$</td>
<td>$x_2 = 0.0713, x_3 = 0.0597, x_5 = 0.3690, x_6 = 0.5$</td>
<td>$(0.0244, 0.065, 0.1403)$</td>
</tr>
<tr>
<td>4</td>
<td>$(0.070, 0.070, 0.008, 0.4)$</td>
<td>$x_2 = 0.0935, x_3 = 0.1098, x_5 = 0.2967, x_6 = 0.5$</td>
<td>$(0.0164, 0.07, 0.164)$</td>
</tr>
<tr>
<td>5</td>
<td>$(0.075, 0.075, 0.008, 0.4)$</td>
<td>$x_2 = 0.1157, x_3 = 0.1598, x_5 = 0.2245, x_6 = 0.5$</td>
<td>$(0.0084, 0.075, 0.1876)$</td>
</tr>
<tr>
<td>6</td>
<td>$(0.080, 0.080, 0.008, 0.4)$</td>
<td>$x_2 = 0.1381, x_3 = 0.2096, x_5 = 0.1524, x_6 = 0.5$</td>
<td>$(0.0003, 0.08, 0.2113)$</td>
</tr>
<tr>
<td>7</td>
<td>$(0.085, 0.085, 0.008, 0.4)$</td>
<td>$x_2 = 0.1605, x_3 = 0.2593, x_5 = 0.0802, x_6 = 0.5$</td>
<td>$(-0.0077, 0.085, 0.2351)$</td>
</tr>
<tr>
<td>8</td>
<td>$(0.090, 0.090, 0.008, 0.4)$</td>
<td>$x_2 = 0.1829, x_3 = 0.3089, x_5 = 0.0081, x_6 = 0.5$</td>
<td>$(-0.0157, 0.09, 0.2588)$</td>
</tr>
<tr>
<td>9</td>
<td>$(0.095, 0.095, 0.008, 0.4)$</td>
<td>$x_2 = 0.2310, x_3 = 0.3128, x_6 = 0.4562$</td>
<td>$(-0.0285, 0.095, 0.2919)$</td>
</tr>
<tr>
<td>10</td>
<td>$(0.100, 0.100, 0.008, 0.4)$</td>
<td>$x_2 = 0.2801, x_3 = 0.3142, x_6 = 0.4057$</td>
<td>$(-0.0414, 0.1, 0.3253)$</td>
</tr>
<tr>
<td>11</td>
<td>$(0.105, 0.105, 0.008, 0.4)$</td>
<td>$x_2 = 0.3250, x_3 = 0.3214, x_6 = 0.3536$</td>
<td>$(-0.0534, 0.105, 0.357)$</td>
</tr>
<tr>
<td>12</td>
<td>$(0.110, 0.110, 0.008, 0.4)$</td>
<td>$x_2 = 0.3701, x_3 = 0.3284, x_6 = 0.3015$</td>
<td>$(-0.0655, 0.11, 0.3889)$</td>
</tr>
<tr>
<td>13</td>
<td>$(0.115, 0.115, 0.008, 0.4)$</td>
<td>$x_2 = 0.4150, x_3 = 0.3357, x_6 = 0.2493$</td>
<td>$(-0.0775, 0.115, 0.4206)$</td>
</tr>
<tr>
<td>14</td>
<td>$(0.120, 0.120, 0.008, 0.4)$</td>
<td>$x_2 = 0.4601, x_3 = 0.3428, x_6 = 0.1972$</td>
<td>$(-0.0895, 0.12, 0.4524)$</td>
</tr>
<tr>
<td>15</td>
<td>$(0.125, 0.125, 0.008, 0.4)$</td>
<td>$x_1 = 0.0092, x_2 = 0.5, x_5 = 0.3415, x_6 = 0.1494$</td>
<td>$(-0.1024, 0.125, 0.4858)$</td>
</tr>
<tr>
<td>16</td>
<td>$(0.130, 0.130, 0.008, 0.4)$</td>
<td>$x_1 = 0.0504, x_2 = 0.5, x_5 = 0.3421, x_6 = 0.1074$</td>
<td>$(-0.1137, 0.13, 0.5157)$</td>
</tr>
<tr>
<td>17</td>
<td>$(0.135, 0.135, 0.008, 0.4)$</td>
<td>$x_1 = 0.0956, x_2 = 0.5, x_5 = 0.3361, x_6 = 0.0683$</td>
<td>$(-0.1258, 0.135, 0.5473)$</td>
</tr>
<tr>
<td>18</td>
<td>$(0.140, 0.140, 0.008, 0.4)$</td>
<td>$x_1 = 0.1410, x_2 = 0.5, x_5 = 0.3296, x_6 = 0.0293$</td>
<td>$(-0.1379, 0.14, 0.579)$</td>
</tr>
<tr>
<td>19</td>
<td>$(0.145, 0.145, 0.008, 0.4)$</td>
<td>$x_1 = 0.2204, x_2 = 0.4660, x_3 = 0.3136$</td>
<td>$(-0.1501, 0.145, 0.6105)$</td>
</tr>
<tr>
<td>20</td>
<td>$(0.150, 0.150, 0.008, 0.4)$</td>
<td>$x_1 = 0.3137, x_2 = 0.4771, x_3 = 0.2092$</td>
<td>$(-0.175, 0.15, 0.6653)$</td>
</tr>
<tr>
<td>21</td>
<td>$(0.155, 0.155, 0.008, 0.4)$</td>
<td>$x_1 = 0.4067, x_2 = 0.4888, x_3 = 0.1045$</td>
<td>$(-0.2, 0.155, 0.7202)$</td>
</tr>
<tr>
<td>22</td>
<td>$(0.160, 0.160, 0.008, 0.4)$</td>
<td>$x_1 = 0.5, x_2 = 0.5$</td>
<td>$(-0.225, 0.16, 0.775)$</td>
</tr>
<tr>
<td>No.</td>
<td>$r^m$</td>
<td>$r^m - r^p$</td>
<td>$r^o - r^m$</td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>1</td>
<td>0.055</td>
<td>0.0090</td>
<td>0.0275</td>
</tr>
<tr>
<td>2</td>
<td>0.060</td>
<td>0.0269</td>
<td>0.0552</td>
</tr>
<tr>
<td>3</td>
<td>0.065</td>
<td>0.0406</td>
<td>0.0753</td>
</tr>
<tr>
<td>4</td>
<td>0.070</td>
<td>0.0536</td>
<td>0.0940</td>
</tr>
<tr>
<td>5</td>
<td>0.075</td>
<td>0.0666</td>
<td>0.1126</td>
</tr>
<tr>
<td>6</td>
<td>0.080</td>
<td>0.0797</td>
<td>0.1313</td>
</tr>
<tr>
<td>7</td>
<td>0.085</td>
<td>0.0927</td>
<td>0.1501</td>
</tr>
<tr>
<td>8</td>
<td>0.090</td>
<td>0.1057</td>
<td>0.1688</td>
</tr>
<tr>
<td>9</td>
<td>0.095</td>
<td>0.1235</td>
<td>0.1969</td>
</tr>
<tr>
<td>10</td>
<td>0.100</td>
<td>0.1414</td>
<td>0.2253</td>
</tr>
<tr>
<td>11</td>
<td>0.105</td>
<td>0.1584</td>
<td>0.2520</td>
</tr>
<tr>
<td>12</td>
<td>0.110</td>
<td>0.1755</td>
<td>0.2789</td>
</tr>
<tr>
<td>13</td>
<td>0.115</td>
<td>0.1925</td>
<td>0.3056</td>
</tr>
<tr>
<td>14</td>
<td>0.120</td>
<td>0.2095</td>
<td>0.3324</td>
</tr>
<tr>
<td>15</td>
<td>0.125</td>
<td>0.2274</td>
<td>0.3608</td>
</tr>
<tr>
<td>16</td>
<td>0.130</td>
<td>0.2437</td>
<td>0.3857</td>
</tr>
<tr>
<td>17</td>
<td>0.135</td>
<td>0.2608</td>
<td>0.4123</td>
</tr>
<tr>
<td>18</td>
<td>0.140</td>
<td>0.2779</td>
<td>0.4390</td>
</tr>
<tr>
<td>19</td>
<td>0.145</td>
<td>0.2951</td>
<td>0.4655</td>
</tr>
<tr>
<td>20</td>
<td>0.150</td>
<td>0.3250</td>
<td>0.5153</td>
</tr>
<tr>
<td>21</td>
<td>0.155</td>
<td>0.3550</td>
<td>0.5647</td>
</tr>
<tr>
<td>22</td>
<td>0.160</td>
<td>0.3850</td>
<td>0.6150</td>
</tr>
<tr>
<td>No.</td>
<td>$\left(\beta_p, \beta_u, \gamma_p, \gamma_u\right)$</td>
<td>Optimal Solution $X^*$</td>
<td>$\tilde{r} = (r^p, r^u, r^o)$</td>
</tr>
<tr>
<td>-----</td>
<td>-----------------------------------</td>
<td>------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>1</td>
<td>$(0.05, 0.17, 0.009, 0.009)$</td>
<td>$x_5 = 0.5, x_6 = 0.5$</td>
<td>$(0.046, 0.055, 0.0825)$</td>
</tr>
<tr>
<td>2</td>
<td>$(0.05, 0.17, 0.0269, 0.0269)$</td>
<td>$x_2 = 0.0202, x_3 = 0.1038, x_5 = 0.5, x_6 = 0.376$</td>
<td>$(0.0408, 0.0643, 0.1164)$</td>
</tr>
<tr>
<td>3</td>
<td>$(0.05, 0.17, 0.0406, 0.0406)$</td>
<td>$x_2 = 0.0359, x_3 = 0.1824, x_5 = 0.5, x_6 = 0.2817$</td>
<td>$(0.0308, 0.0714, 0.1423)$</td>
</tr>
<tr>
<td>4</td>
<td>$(0.05, 0.17, 0.0536, 0.0536)$</td>
<td>$x_2 = 0.0506, x_3 = 0.2577, x_5 = 0.5, x_6 = 0.1917$</td>
<td>$(0.0245, 0.0781, 0.1669)$</td>
</tr>
<tr>
<td>5</td>
<td>$(0.05, 0.17, 0.0666, 0.0666)$</td>
<td>$x_2 = 0.0746, x_3 = 0.3001, x_5 = 0.5, x_6 = 0.1253$</td>
<td>$(0.0169, 0.0835, 0.1914)$</td>
</tr>
<tr>
<td>6</td>
<td>$(0.05, 0.17, 0.0927, 0.0927)$</td>
<td>$x_2 = 0.1061, x_3 = 0.3167, x_5 = 0.5, x_6 = 0.0773$</td>
<td>$(0.0081, 0.0878, 0.2159)$</td>
</tr>
<tr>
<td>7</td>
<td>$(0.05, 0.17, 0.1057, 0.1057)$</td>
<td>$x_2 = 0.1382, x_3 = 0.3301, x_5 = 0.5, x_6 = 0.0317$</td>
<td>$(-0.0008, 0.0919, 0.2402)$</td>
</tr>
<tr>
<td>8</td>
<td>$(0.05, 0.17, 0.1235, 0.1235)$</td>
<td>$x_2 = 0.1704, x_3 = 0.3435, x_5 = 0.4861$</td>
<td>$(-0.0097, 0.0959, 0.2643)$</td>
</tr>
<tr>
<td>9</td>
<td>$(0.05, 0.17, 0.1414, 0.1414)$</td>
<td>$x_2 = 0.2189, x_3 = 0.3469, x_5 = 0.4342$</td>
<td>$(-0.0230, 0.1005, 0.2969)$</td>
</tr>
<tr>
<td>10</td>
<td>$(0.05, 0.17, 0.1584, 0.1584)$</td>
<td>$x_2 = 0.2674, x_3 = 0.3514, x_5 = 0.3812$</td>
<td>$(-0.0363, 0.1052, 0.3297)$</td>
</tr>
<tr>
<td>11</td>
<td>$(0.05, 0.17, 0.1755, 0.1755)$</td>
<td>$x_2 = 0.3139, x_3 = 0.3538, x_5 = 0.3322$</td>
<td>$(-0.0489, 0.1095, 0.3608)$</td>
</tr>
<tr>
<td>12</td>
<td>$(0.05, 0.17, 0.1925, 0.1925)$</td>
<td>$x_2 = 0.3610, x_3 = 0.3552, x_5 = 0.2837$</td>
<td>$(-0.0617, 0.1138, 0.3922)$</td>
</tr>
<tr>
<td>13</td>
<td>$(0.05, 0.17, 0.2095, 0.2095)$</td>
<td>$x_2 = 0.4071, x_3 = 0.3596, x_5 = 0.2333$</td>
<td>$(-0.0743, 0.1182, 0.4234)$</td>
</tr>
<tr>
<td>14</td>
<td>$(0.05, 0.17, 0.2274, 0.2274)$</td>
<td>$x_2 = 0.4511, x_3 = 0.3710, x_5 = 0.1779$</td>
<td>$(-0.0867, 0.1229, 0.4545)$</td>
</tr>
<tr>
<td>15</td>
<td>$(0.05, 0.17, 0.2437, 0.2437)$</td>
<td>$x_2 = 0.4975, x_3 = 0.3831, x_5 = 0.1194$</td>
<td>$(-0.0996, 0.1278, 0.4873)$</td>
</tr>
<tr>
<td>16</td>
<td>$(0.05, 0.17, 0.2608, 0.2608)$</td>
<td>$x_1 = 0.0411, x_2 = 0.5, x_3 = 0.3765, x_5 = 0.0823$</td>
<td>$(-0.1116, 0.1321, 0.5169)$</td>
</tr>
<tr>
<td>17</td>
<td>$(0.05, 0.17, 0.2779, 0.2779)$</td>
<td>$x_1 = 0.0911, x_2 = 0.5, x_3 = 0.3529, x_5 = 0.0561$</td>
<td>$(-0.1246, 0.1362, 0.5482)$</td>
</tr>
<tr>
<td>18</td>
<td>$(0.05, 0.17, 0.2791, 0.2791)$</td>
<td>$x_1 = 0.1476, x_2 = 0.5, x_3 = 0.3034, x_5 = 0.049$</td>
<td>$(-0.1385, 0.1394, 0.5794)$</td>
</tr>
<tr>
<td>19</td>
<td>$(0.05, 0.17, 0.2951, 0.2951)$</td>
<td>$x_1 = 0.2042, x_2 = 0.5, x_3 = 0.2546, x_5 = 0.0412$</td>
<td>$(-0.1524, 0.1427, 0.6108)$</td>
</tr>
<tr>
<td>20</td>
<td>$(0.05, 0.17, 0.325, 0.325)$</td>
<td>$x_1 = 0.3025, x_2 = 0.5, x_3 = 0.1703, x_5 = 0.0272$</td>
<td>$(-0.1765, 0.1485, 0.6654)$</td>
</tr>
<tr>
<td>21</td>
<td>$(0.05, 0.17, 0.355, 0.355)$</td>
<td>$x_1 = 0.4013, x_2 = 0.5, x_3 = 0.0848, x_5 = 0.0139$</td>
<td>$(-0.2008, 0.1542, 0.7202)$</td>
</tr>
<tr>
<td>22</td>
<td>$(0.05, 0.17, 0.385, 0.385)$</td>
<td>$x_1 = 0.5, x_2 = 0.5$</td>
<td>$(-0.225, 0.16, 0.775)$</td>
</tr>
</tbody>
</table>
Table 5
Optimal Portfolio Return and Risk Analysis Using Table 4

<table>
<thead>
<tr>
<th>No.</th>
<th>$r^m$</th>
<th>$r^m - r^P$</th>
<th>$r^o - r^m$</th>
<th>85%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0550</td>
<td>0.0090</td>
<td>0.0275</td>
<td>(0.0537, 0.0541)</td>
<td>(0.0500, 0.0564)</td>
</tr>
<tr>
<td>2</td>
<td>0.0643</td>
<td>0.0235</td>
<td>0.0521</td>
<td>(0.0608, 0.0721)</td>
<td>(0.0631, 0.0669)</td>
</tr>
<tr>
<td>3</td>
<td>0.0714</td>
<td>0.0406</td>
<td>0.0709</td>
<td>(0.0653, 0.0820)</td>
<td>(0.0694, 0.0749)</td>
</tr>
<tr>
<td>4</td>
<td>0.0781</td>
<td>0.0536</td>
<td>0.0888</td>
<td>(0.0701, 0.0914)</td>
<td>(0.0754, 0.0825)</td>
</tr>
<tr>
<td>5</td>
<td>0.0835</td>
<td>0.0666</td>
<td>0.1079</td>
<td>(0.0735, 0.0997)</td>
<td>(0.0802, 0.0889)</td>
</tr>
<tr>
<td>6</td>
<td>0.0878</td>
<td>0.0797</td>
<td>0.1281</td>
<td>(0.0758, 0.1070)</td>
<td>(0.0838, 0.0942)</td>
</tr>
<tr>
<td>7</td>
<td>0.0919</td>
<td>0.0927</td>
<td>0.1483</td>
<td>(0.0780, 0.1142)</td>
<td>(0.0873, 0.0993)</td>
</tr>
<tr>
<td>8</td>
<td>0.0959</td>
<td>0.1056</td>
<td>0.1684</td>
<td>(0.0801, 0.1212)</td>
<td>(0.0906, 0.1043)</td>
</tr>
<tr>
<td>9</td>
<td>0.1005</td>
<td>0.1235</td>
<td>0.1964</td>
<td>(0.0820, 0.1300)</td>
<td>(0.0943, 0.1103)</td>
</tr>
<tr>
<td>10</td>
<td>0.1052</td>
<td>0.1414</td>
<td>0.2245</td>
<td>(0.0840, 0.1389)</td>
<td>(0.0981, 0.1164)</td>
</tr>
<tr>
<td>11</td>
<td>0.1095</td>
<td>0.1584</td>
<td>0.2513</td>
<td>(0.0857, 0.1472)</td>
<td>(0.1016, 0.1221)</td>
</tr>
<tr>
<td>12</td>
<td>0.1138</td>
<td>0.1755</td>
<td>0.2784</td>
<td>(0.0875, 0.1556)</td>
<td>(0.1050, 0.1277)</td>
</tr>
<tr>
<td>13</td>
<td>0.1182</td>
<td>0.1925</td>
<td>0.3052</td>
<td>(0.0893, 0.1640)</td>
<td>(0.1086, 0.1335)</td>
</tr>
<tr>
<td>14</td>
<td>0.1229</td>
<td>0.2096</td>
<td>0.3316</td>
<td>(0.0915, 0.1726)</td>
<td>(0.1124, 0.1395)</td>
</tr>
<tr>
<td>15</td>
<td>0.1278</td>
<td>0.2274</td>
<td>0.3595</td>
<td>(0.0937, 0.1817)</td>
<td>(0.1165, 0.1458)</td>
</tr>
<tr>
<td>16</td>
<td>0.1321</td>
<td>0.2437</td>
<td>0.3848</td>
<td>(0.0958, 0.1898)</td>
<td>(0.1199, 0.1513)</td>
</tr>
<tr>
<td>17</td>
<td>0.1362</td>
<td>0.2608</td>
<td>0.4120</td>
<td>(0.0971, 0.1980)</td>
<td>(0.1232, 0.1568)</td>
</tr>
<tr>
<td>18</td>
<td>0.1394</td>
<td>0.2779</td>
<td>0.4400</td>
<td>(0.0977, 0.2054)</td>
<td>(0.1255, 0.1614)</td>
</tr>
<tr>
<td>19</td>
<td>0.1427</td>
<td>0.2951</td>
<td>0.4681</td>
<td>(0.0984, 0.2129)</td>
<td>(0.1279, 0.1661)</td>
</tr>
<tr>
<td>20</td>
<td>0.1485</td>
<td>0.3250</td>
<td>0.5169</td>
<td>(0.0998, 0.2260)</td>
<td>(0.1323, 0.1743)</td>
</tr>
<tr>
<td>21</td>
<td>0.1542</td>
<td>0.3550</td>
<td>0.5660</td>
<td>(0.1100, 0.2391)</td>
<td>(0.1365, 0.1825)</td>
</tr>
<tr>
<td>22</td>
<td>0.160</td>
<td>0.3850</td>
<td>0.6150</td>
<td>(0.1023, 0.2523)</td>
<td>(0.1408, 0.1908)</td>
</tr>
</tbody>
</table>

6 Summary

We presented an asset allocation method using possibilistic programming techniques to characterize the imprecise nature of the rate of return. Unlike the traditional mean-variance method, our asset allocation method takes the portfolio's skewness into consideration. It provides two control constraints that permit maximal flexibility for decision makers to effectively balance the portfolio's return and the portfolio's risk. The optimal allocation decision is made by solving several linear programming problems. Software packages are available that can efficiently solve linear programming problems.
Table 6
Solutions for Different Values of $\beta_i$

<table>
<thead>
<tr>
<th>No.</th>
<th>$(\beta_p, \beta_u, \gamma_p, \gamma_u)$</th>
<th>Optimal Solution $X^*$</th>
<th>$\tilde{r} = (r_p, r^m, r^o)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\phi = (0.08, 0.17, 0.008, 0.4)$</td>
<td>$x_2 = 0.4977, x_3 = 0.2554, x_4 = 0.2469$</td>
<td>$(-0.0946, 0.1201, 0.4645)$</td>
</tr>
<tr>
<td>2</td>
<td>$\phi = (0.11, 0.17, 0.008, 0.4)$</td>
<td>$x_1 = 0.0788, x_2 = 0.5, x_5 = 0.3596, x_6 = 0.0616$</td>
<td>$(-0.1214, 0.1352, 0.5407)$</td>
</tr>
<tr>
<td>3</td>
<td>$\phi = (0.14, 0.17, 0.008, 0.4)$</td>
<td>$x_1 = 0.4137, x_2 = 0.3113, x_3 = 0.275$</td>
<td>$(-0.1609, 0.15, 0.6387)$</td>
</tr>
<tr>
<td>4</td>
<td>$\phi = (0.16, 0.17, 0.008, 0.4)$</td>
<td>$x_1 = 0.5, x_2 = 0.5$</td>
<td>$(-0.225, 0.16, 0.775)$</td>
</tr>
</tbody>
</table>

Table 7
Optimal Portfolio Return and Portfolio Risk Analysis

<table>
<thead>
<tr>
<th>No.</th>
<th>$r^m$</th>
<th>$r^m - r_p$</th>
<th>$r^o - r^m$</th>
<th>85%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1201</td>
<td>0.2147</td>
<td>0.3444</td>
<td>(0.0879, 0.1718)</td>
<td>(0.0879, 0.1373)</td>
</tr>
<tr>
<td>2</td>
<td>0.1352</td>
<td>0.2566</td>
<td>0.4055</td>
<td>(0.0967, 0.196)</td>
<td>(0.1224, 0.1555)</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
<td>0.3009</td>
<td>0.4887</td>
<td>(0.1049, 0.2233)</td>
<td>(0.135, 0.1744)</td>
</tr>
<tr>
<td>4</td>
<td>0.16</td>
<td>0.385</td>
<td>0.615</td>
<td>(0.1023, 0.2523)</td>
<td>(0.1408, 0.1908)</td>
</tr>
</tbody>
</table>
Some problems still remain to be solved. For example, instead of obtaining the most pessimistic value, the most probable value, and the most optimistic value for the rate of return from mean and standard deviation as shown in the examples, we could use simulation to generate the data directly from the historical resources. We could also use the possibility programming method to solve multistage asset allocation problems and asset/liability management problems.

References


Nonmedical Limits in Individual Life Insurance

James B. Ross* and Shalini E. Perumpral†

Abstract

This paper shows data that illustrate the substantial variation among non­
medical schedules and the dramatic increase in their amount limits from 1972
through 1992. Coefficients of variation are analyzed for several data subsets.
We find that the variation of schedules in the sample of all firms has increased
throughout the 1972-1992 period for issue ages up to 30, but has declined for
issue ages beyond 30 during the 1982-1992 period. For the non-New York and
stock companies our statistical tests indicate an increase in the variability of
schedules over the full period 1972 to 1992.

Key words and phrases: mortality, underwriting, medical examinations, sched­
ules, coefficient of variation

1 Introduction

The practice of granting life insurance without a medical examina­
tion began in England when underwriting evidence consisted of per­
sonal interviews, opinions of associates and friends, and/or attending

* James B. Ross, F.S.A., earned his A.B. from Harvard College, his M.B.A. from the
University of Rhode Island, and his Ph.D. from the University of Connecticut. A fellow
of the Society of Actuaries since 1954, Dr. Ross served as a senior officer of several
financial services companies before becoming associate professor of finance at Radford
University in 1992. He has published in the Journal of Actuarial Practice, Transactions
of the Society of Actuaries, and Journal of Financial Services Research.

Dr. Ross’s address is: Radford University, Department of Finance, P.O. Box 6951,
Radford VA 24142, USA. Internet address: jross@runet.edu

† Shalini E. Perumpral earned her B.A. (honors) from Delhi University, India, her M.S.
in economics from Purdue University, and her M.B.A. and Ph.D. in finance from Virginia
Tech. Dr. Perumpral is currently a professor of finance at Radford University. She has

Dr. Perumpral’s address is: Radford University, Department of Finance, P.O. Box
6951, Radford VA 24142, USA. Internet address: sperump@runet.edu
physicians' statements. Medical evidence began to be required in 1850, and a medical examination was considered essential until 1885. In 1886 cautious experiments to remove the medical examination on smaller policies began, albeit with substantial restrictions that were gradually lifted in view of favorable results.

The rationale for nonmedical limits\(^1\) for insurance policies had been that the savings in medical exam expenses were sufficient to offset the additional mortality experienced in the absence of underwriting information from medical exams. A shortage of medical examiners in rural areas following World War I led a group of Canadian companies to begin nonmedical programs with restrictions on issue ages and amounts. The practice was well received in the field, the early experience was favorable, and the Canadian program was liberalized and expanded. Beginning in 1925 nonmedical underwriting spread rapidly through the American life insurance industry, and by 1935 86 percent of the 129 members of the American Life Convention had adopted nonmedical programs. Today nearly every life insurer in the United States and Canada accepts some nonmedically underwritten business, and it is estimated that 67 percent of new ordinary policies and 33 percent of new ordinary amounts are written nonmedically (Black and Skipper, 1994, Chapter 24, p. 671).

Because the insurer pays for medical evidence it uses in underwriting the application, there are initial expense savings when no medical examination is required. The actuarial mechanics of the construction of such schedules are well established: the present value over the policy life of the excess mortality experienced under nonmedical underwriting is equated to the expense savings at issue, and the equation is solved for the face amount that balances it.

Nonmedical limit schedules theoretically should depend on the cost of medical exams and the additional mortality experienced in their absence, suggesting that the schedules for different companies should not vary much. In practice, however, variation among companies enters via differing attitudes in areas such as mortality selection standards, persistency rates, returns on investments, target markets, degrees of accommodation to the writing agent, safety/profit margins in the premium structure, and stock versus mutual forms of insurer organization. This paper addresses questions raised by the existence of a large number of nonmedical limit schedules that exhibit substantial variation.

\(^1\)A nonmedical limit for a new life insurance policy is the maximum amount of insurance that can be issued without the benefit of a medical or paramedical examination.
Changes in nonmedical limits over the last two decades have been characterized in Black and Skipper (1994, Chapter 24, p. 672) as "nonmedical limits exploded." Great increases in nonmedical limits represent the responses by companies to large increases in the cost of medical examinations over the period of this study. Companies have dealt with the cost increases in medical examinations by using less expensive paramedical exams and by making cost-effective use of blood and urine testing.

The extent to which nonmedical limit schedules vary is an empirical question. This paper seeks to determine both the degree of the current variation and the trend in variation over time: Is competition driving the schedules together, or are individual company differences forcing them apart? We show how nonmedical limits have developed, summarize the current situation, and explore the variations of schedules of different insurers.

2 Factors Impacting Nonmedical Limits

While this paper focuses on nonmedical limits, there is a continuum of underwriting approaches of which medically examined business and nonmedical business are the extremes. All variations are driven by the trade-off between expense savings and differential mortality costs. This dynamic trade-off is a function of the increase in the cost-effectiveness of underwriting tools, increases in medical exam costs, and continuing improvements in insured mortality. Paramedical underwriting provides the best example (Woodman, 1992). Paramedical underwriting has advanced to the point where separate mortality experiences are maintained for this approach. Blood and urine testing also offer protective values that are cost-effective at levels less than full nonmedical limits. Additionally, companies review periodically their use of other underwriting tools such as inspection reports, attending physicians' statements (APSs), personal health interviews (PHIs), and motor vehicle records (MVRs). These reviews may cause companies to revise the issue amounts at which they order such tools.

Inflation is one of the major forces that drew attention to the nonmedical area. The chairman of an extended discussion in 1970 on the impact of inflation on underwriting remarked: "There is evidence that the cost of medical underwriting has increased more rapidly than the health care index, so we can conclude that the major components of underwriting costs have increased more rapidly than the Consumer Price Index" (Taylor 1970). The Statistical Abstracts of the United States
provides the data for the percentage increases in related price indices shown below for the periods 1972-1982, 1982-1992, and 1972-1992. Data from the *Life Insurance Fact Book* show that the percentage increases in the average size policy issued have more than kept pace with these inflationary increases in the several price indices.

The onset of AIDS as a significant factor in underwriting occurred during the period 1982-1992. During this period AIDS dominated discussions of underwriting in the actuarial literature. Company responses have included blood testing at much lower face amount levels in applicant cohorts where AIDS is a concern. Prior to 1985 blood testing generally was not requested until face amounts applied for exceeded $1 million. HIV/AIDS changed that dramatically. Blood/urine/saliva testing for HIV now begins at $25,000 to $100,000. Additionally, some observers feel that companies may have slowed increases in nonmedical limits and conformed their nonmedical schedules by issue ages to those of competitors to avoid being selected against by the HIV-infected.

### 3 Literature Review

This literature review concentrates on papers and discussions dealing with the factors impacting nonmedical limits. Outside the actuarial literature there is substantial additional underwriting material relevant to this subject, particularly in the publications of the Home Office Life Underwriters Association and the Institute of Home Office Underwriters.

---

For the period up to December 31, 1991, during which information could affect company decisions on nonmedical limits for 1992, there were several papers and task force reports on AIDS (though not all focused on underwriting) that were published by the Society of Actuaries. These include the *Guide for Practicing Actuaries* (1988), Panjer (1989), Plumley (1989), Ramsay (1989 and 1990), the *Report of the Society of Actuaries Committee on HIV Research* (1990), and the *Report of the Task Force on the Financial Implications of AIDS* (1990).
The literature contains three themes. The first theme consists of historical examinations of nonmedical limits in ordinary (and industrial) life insurance. Parker (1921) reviews the Canadian experiment. Auden (1938) gives a brief history, an update on the practice of 114 companies, a review of the reasons for writing nonmedical business, and a report on the generally favorable mortality. Morton (1977) discusses nonmedical and paramedical underwriting in his review of underwriting principles and practices. Sankey (1990) and Black and Skipper (1994, Chapter 24, pp. 671-672) provide historic treatments for more recent periods.

The second theme, review and liberalization, consists of a long series of discussions in the actuarial and underwriting journals responding to questions by editors. Smith (1924) emphasizes the early success of the Canadian nonmedical program. Larus (1925) cautions against competition on nonmedical limits, while Parker (1925) feels that companies doing a nonmedical business contribute meaningfully to the information maintained by the Medical Impairment Bureau.

As liberalizations develop, the discussions focus on nonmedical mortality experience relative to that of medically examined business. Smith (1930) uses Canadian male select mortality as a benchmark; Shepherd (1930) benchmarks against American male select mortality. Both find the ratios of actual-to-expected mortality (A/E ratios) for nonmedical issues higher than the ratios for medically examined business; both find the A/E ratios for nonmedical issues in age groups beyond age 45 substantially higher than their medically examined counterparts. Smith and Cross (1930) indicate higher lapse rates on the nonmedical issues. Marshall (1932) provides data showing the favorable mortality experience of Connecticut Mutual. Discussions in Record of the American Institute of Actuaries (1934) identify issue age 40 as the supportable upper age for nonmedical schedules, providing several examples at older issue ages of substantially higher A/E ratios (relative to American male select mortality) for nonmedical issues than for those medically examined.

Auden (1938) cites reductions from upper age 45 to age 40 as the trend of the day, with nonmedical persistency still poor but nonmedical mortality satisfactory. He discusses the value of the forgone expense of the medical exam in offsetting additional mortality. Hunter (1940) inventories mortality studies (up to 1931 for three Canadian companies and five American companies) and adds New York Life data through 1939 to show generally favorable nonmedical experience. Discussions in Record of the American Institute of Actuaries (1942) center around the problems of obtaining medical examiners during World War II and the nonmedical liberalizations that would help reduce the load on examin-
ers (the consensus was “yes” to amounts, “no” to age extensions). The increase in the percentage of applications on a nonmedical basis that accompanied nonmedical schedule liberalizations is discussed, with one large company’s percentage in 1942 going from 9 percent in July to 30 percent in October!

The central issue in Record of the American Institute of Actuaries (1946) is wartime mortality; all commentators on nonmedical limits come to the same general conclusion, viz. that nonmedical business still could be written satisfactorily at issue ages under 40 for amounts up to $5,000. The discussions in the Transactions of the Society of Actuaries (1950) indicate that the triggering incident for the announcement of nonmedical limit increases is a specific increase in medical examiner fees.

Merriam (1951) describes an increase in medical examiner fees of about one-third, with resulting extensions of nonmedical limits in the Metropolitan Life to the age groups 41–45 and 46–50. Mathews (1953) provides survey evidence from 108 companies that such extensions are not common—only 5 percent of the companies issue nonmedically above age 40. Morton (1954) reports that most Canadian companies continue some nonmedical issue amount to age 45, but provides discounted extra mortality costs that suggest only nominal amounts are feasible. Van Keuren (1956) indicates that Metropolitan Life, which introduced nonmedical issues above age 40 in 1951, has discontinued them because of unsatisfactory mortality experience and the necessity to obtain medical exams on 25 percent of nonmedical applicants.

Jacoby and Tookey (1959) both indicate pressure from physicians to increase the medical examination fees. They attribute this to doctors’ aversion to paper work, the lagging of fees behind price levels, and resentment that insurers would attempt to fix doctors’ fees. All discussants (Transactions of the Society of Actuaries, 1960) note increases of $25,000 to $30,000 up to age 30, but few increases thereafter.

Lew (1966) predicts increased use of bodily fluids testing to extend the use of nonmedical limits to older age groups. Gauer and van Keuren (1967) explore the use of technicians and early paramedical techniques. The difficulty of finding physicians willing to serve as medical examiners is noted. Many discussants note the use of medical information phoned-in and recorded. Keltie (1969) attributes the slowdown in mortality improvement on medically examined business to the spread of paramedical exams and alludes to reductions in the use of inspection reports and attending physicians’ reports.

The third theme consists of the readings gathered by the Society of Actuaries under the rubric of cost implications in the Professional Actu-
Ormsby (1963) first examines the economics of underwriting in a paper that addresses the considerations involved in ordering inspection reports. He provides formulas for "... converting changes in underwriting action attributable to information in the APS (attending physician's statement) into equivalent 'net' single premiums at issue so that a comparison can be made of these 'net' single premiums with the total cost of obtaining and processing the statement itself ..." The techniques outlined are applicable to the construction of nonmedical limit schedules.

Mast (1978) discusses each element of the nonmedical limit question. His paper determines the break-even amount as "... the policy size at which the increased mortality costs resulting from the lack of a medical examination are approximately counterbalanced by the consequent savings in underwriting expenses." He mentions an asset share approach, and discusses the net single premium technique used by Ormsby: "... the relationship between the expenses associated with obtaining a medical examination and the present value of the increased mortality cost per $1,000 is used to determine the break-even amount."

Reitano (1979) provides a consistent theory for evaluating the interplay between the cost of underwriting tools and the resulting mortality. He discusses two cases:

- The actuarial approach typically used in setting nonmedical limits, using the present value of the difference between medical and nonmedical mortality experience (the two table technique); and

- The underwriting approach for valuing underwriting tools (as in Ormsby), under which the value of the tool is the present value of the extra mortality costs that are saved by removing certain lives from the standard issue class (the single table method).

Bergstrom (1989, 1991) discusses the assumptions and calculations that provide estimates for the protective values of blood chemistry profile and urinalysis testing. The earlier study gives protective values for life insurance, the latter for major medical insurance. The reports show the techniques for expressing the results in terms of amount levels above which the testing is cost-justified and in terms of return on the investment (ROI) in the testing.

Mills (1991) provides a general model for such protective value studies, utilizing the axiom that "... a particular underwriting procedure has positive economic value if its cost is less than the savings in mortality (or morbidity) made possible by its use." Mills provides an example for valuing the attending physician's statement in connection with disability income.
Woodman (1992) assesses the value of the paramedical examination using the tools and approach specified by Bergstrom. He provides comparisons between medical, paramedical, and nonmedical mortality experience; his further analysis indicates the age-at-issue groups and amount levels for which the several underwriting approaches are most appropriate.

4 Data Sources

The data used in the statistical analyses are the nonmedical limit amounts published in Best’s Flitcraft Compend (Life-Health) for the editions dated 1973, 1983, and 1993. The data collection procedures used by A.M. Best Co. are such that the data relate to the years 1972, 1982, and 1992. It is these latter years that are used in the table headings and the text.

The nonmedical limit information, when available, is given in the policy analysis section (preceding the statistical sections) of the Flitcraft Compend. The availability of nonmedical schedules is shown in Table 1, which gives in the panel headings the number of companies contributing nonmedical limit schedules to each year of the study. The material available for analysis grew substantially from 1972 to 1982, then shrank in 1992 because the A.M. Best Company split the Flitcraft Compend into two sections, only one of which preserved the nonmedical data. As a result there are data on 113 companies for 1972, 164 companies for 1982, and 119 companies for 1992. Forty-eight companies provided data for all three years.

The basic data (not shown) consist of values for the nonmedical limits across each of the 15 issue age groups for each company plus additional values for independent variables representing specific characteristics of individual companies. The issue age groups used by different life insurers in practice are so similar that less than 20 forcings were needed to put the nonmedical schedules into the common format of 15 groups by age at issue.
Table 1

Characteristics of the Sample of All Firms Nonmedical Limits (000s)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Companies*</td>
<td></td>
<td>115</td>
<td>115</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
<td>118</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>222</td>
<td>225</td>
<td>224</td>
<td>220</td>
<td>208</td>
<td>209</td>
<td>190</td>
<td>140</td>
<td>86.6</td>
<td>66.3</td>
<td>25.4</td>
<td>17.5</td>
<td>5.50</td>
<td>1.63</td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td></td>
<td>153</td>
<td>155</td>
<td>137</td>
<td>138</td>
<td>134</td>
<td>134</td>
<td>99.8</td>
<td>89.2</td>
<td>65.1</td>
<td>65.3</td>
<td>53.6</td>
<td>50.1</td>
<td>18.3</td>
<td>7.42</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>150</td>
<td>100</td>
<td>75</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Mode</td>
<td></td>
<td>250</td>
<td>250</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td></td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>500</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Companies*</td>
<td></td>
<td>162</td>
<td>162</td>
<td>163</td>
<td>164</td>
<td>164</td>
<td>164</td>
<td>164</td>
<td>164</td>
<td>162</td>
<td>150</td>
<td>82</td>
<td>37</td>
<td>22</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>134</td>
<td>139</td>
<td>139</td>
<td>141</td>
<td>141</td>
<td>145</td>
<td>144</td>
<td>108</td>
<td>70.6</td>
<td>40.1</td>
<td>25.2</td>
<td>15.8</td>
<td>14.4</td>
<td>8.50</td>
<td>5.57</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td></td>
<td>84.5</td>
<td>82.5</td>
<td>81.3</td>
<td>81.0</td>
<td>80.9</td>
<td>86.4</td>
<td>86.9</td>
<td>85.2</td>
<td>82.5</td>
<td>75.4</td>
<td>75.9</td>
<td>72.2</td>
<td>72.3</td>
<td>47.5</td>
<td>42.4</td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td>50</td>
<td>20</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mode</td>
<td></td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td></td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>Minimum</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>15</td>
<td>25</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: 1972 Sample (N = 113)</th>
<th>Age at Issue</th>
<th>0-4</th>
<th>5-14</th>
<th>15</th>
<th>16-17</th>
<th>18-20</th>
<th>21-25</th>
<th>26-30</th>
<th>31-35</th>
<th>36-40</th>
<th>41-45</th>
<th>46-50</th>
<th>51-55</th>
<th>56-60</th>
<th>61-65</th>
<th>66-70</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Companies*</td>
<td></td>
<td>113</td>
<td>113</td>
<td>113</td>
<td>113</td>
<td>113</td>
<td>112</td>
<td>111</td>
<td>111</td>
<td>71</td>
<td>18</td>
<td>9</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>30.2</td>
<td>31.9</td>
<td>32.7</td>
<td>33.0</td>
<td>33.1</td>
<td>32.9</td>
<td>32.0</td>
<td>20.9</td>
<td>11.0</td>
<td>3.61</td>
<td>0.50</td>
<td>0.15</td>
<td>0.07</td>
<td>0.02</td>
<td>0</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td></td>
<td>9.18</td>
<td>7.83</td>
<td>6.61</td>
<td>6.46</td>
<td>6.39</td>
<td>6.19</td>
<td>6.77</td>
<td>6.04</td>
<td>5.22</td>
<td>3.62</td>
<td>1.49</td>
<td>0.58</td>
<td>0.34</td>
<td>0.17</td>
<td>0</td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>35</td>
<td>35</td>
<td>30</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mode</td>
<td></td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td></td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>25</td>
<td>10</td>
<td>4</td>
<td>2.5</td>
<td>1.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Minimum</td>
<td></td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* Number of companies with nonzero nonmedical limits
5 Methodology

For each age group for each of the years 1972, 1982, and 1992 these univariate statistics for the nonmedical limits are calculated: mean, median, mode, maximum, minimum, and standard deviation. We also count and display the number of companies that provide nonzero nonmedical limits to a particular age group. These characteristics are displayed in Table 1. The same statistics are provided in Table 2 for the 48 companies with data for all three years.

Because our interest is to determine the extent of current variation among issuers and the trend in variation over time, a test for stationarity of variance seems logical. Given the tremendous increase in nonmedical limits in the decade from 1972 to 1982, however, stationarity tests of the variance do not provide any insight as to the real divergences in behavior within the industry. Therefore, coefficients of variation are calculated for each age group for the years 1972, 1982, and 1992, and a series of nonparametric tests is performed on this statistic.

Statistical tests are used to determine: (i) whether variation within the industry has remained consistent for the two decades—this test was suggested in 1937 by Friedman (1991); and (ii) whether the variation has consistently increased or decreased over the two decades—this test was suggested in 1963 by Page (1991). Appendix A describes these tests for the entire sample, giving the null and alternative hypotheses, the calculated coefficients of variation, formulas for the test statistics, and the cut-off points for rejection at selected confidence levels. The Friedman and Page tests are performed on the entire sample and repeated again for those 48 companies for which data are available for both decades. This approach allows us to isolate any bias that may have been introduced by outliers or new entrants into the full sample.

The 48 firms for which data are available for 1972, 1982 and 1992 are also split into stock (22) and mutual (26) companies and New York (22) and non-New York (26) insurers. Similar tests are performed on these samples to determine whether there are any identifiable differences in behavior among these subgroups. The stock/mutual split is chosen to explore whether the philosophy or practices inherent in the form of company organization may influence the nonmedical limits. The non-New York/New York split is chosen to test whether the New York expense and commission limitations (and perhaps the extraterritoriality) would impact the nonmedical limits.
Table 2  
Characteristics of the Sample of 48 Firms Nonmedical Limits (000s)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Companies*</td>
<td>47</td>
<td>47</td>
<td>47</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>47</td>
<td>47</td>
<td>47</td>
<td>47</td>
<td>21</td>
<td>11</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Mean</td>
<td>253</td>
<td>259</td>
<td>261</td>
<td>264</td>
<td>252</td>
<td>252</td>
<td>252</td>
<td>204</td>
<td>149</td>
<td>83.5</td>
<td>66.4</td>
<td>29.1</td>
<td>21.5</td>
<td>7.04</td>
<td>1.63</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>163</td>
<td>167</td>
<td>166</td>
<td>163</td>
<td>162</td>
<td>162</td>
<td>162</td>
<td>105</td>
<td>93.4</td>
<td>73.9</td>
<td>72.9</td>
<td>71.9</td>
<td>71.8</td>
<td>22.3</td>
<td>8.00</td>
</tr>
<tr>
<td>Median</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>200</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mode</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>250</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>500</td>
<td>500</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>100</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>Minimum</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Age at Issue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Companies*</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>47</td>
<td>47</td>
<td>42</td>
<td>16</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Mean</td>
<td>137</td>
<td>136</td>
<td>135</td>
<td>135</td>
<td>136</td>
<td>135</td>
<td>134</td>
<td>94.0</td>
<td>50.7</td>
<td>23.1</td>
<td>7.42</td>
<td>1.28</td>
<td>0.11</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>70.0</td>
<td>68.3</td>
<td>65.1</td>
<td>65.1</td>
<td>64.9</td>
<td>64.3</td>
<td>65.0</td>
<td>60.3</td>
<td>40.0</td>
<td>24.2</td>
<td>18.0</td>
<td>4.43</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>Median</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>125</td>
<td>125</td>
<td>120</td>
<td>75</td>
<td>50</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mode</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>200</td>
<td>100</td>
<td>25</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Minimum</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Number of Companies*</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>47</td>
<td>47</td>
<td>27</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mean</td>
<td>29.7</td>
<td>31.6</td>
<td>32.8</td>
<td>33.2</td>
<td>33.2</td>
<td>33.1</td>
<td>32.1</td>
<td>19.8</td>
<td>9.96</td>
<td>2.96</td>
<td>0.21</td>
<td>0.16</td>
<td>0.03</td>
<td>0.03</td>
<td>0</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>9.31</td>
<td>7.52</td>
<td>5.92</td>
<td>5.60</td>
<td>5.60</td>
<td>5.61</td>
<td>6.09</td>
<td>5.62</td>
<td>3.43</td>
<td>3.20</td>
<td>0.92</td>
<td>0.69</td>
<td>0.22</td>
<td>0.22</td>
<td>0</td>
</tr>
<tr>
<td>Median</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mode</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>40</td>
<td>25</td>
<td>12</td>
<td>5</td>
<td>4</td>
<td>1.5</td>
<td>1.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Minimum</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* Number of companies with nonzero nonmedical limits
6 Analysis and Findings

Panels A, B, and C of Table 1 show the descriptive statistics for all firms for the set of 15 age groups over 1972-1992. The data show the stunning increases in nonmedical limits, particularly over 1972-1982. The mean is consistently higher than the median and the mode, with few exceptions, suggesting that some companies offer significantly larger nonmedical limits than their competitors.

Panels A, B, and C of Table 2 show the descriptive statistics for the 48 companies. The same patterns of skewness, with the mean being higher than the median and the mode, emerge for 1992 and 1982, while the 1972 figures emulate a normal distribution.

Table 3 shows the percentage increase in the mean nonmedical limits for the periods 1972-1982, 1982-1992, and from 1972-1992. Percentage increases for 1972-1982 are substantial in every age category, especially beyond issue age 40. There are further increases in the mean nonmedical limits for every issue age category in the second decade. These increases are much smaller than those in the earlier decade but more evenly distributed along the age range.

Table 4 shows that the percentage of companies offering nonzero nonmedical limits at issue ages beyond age 40 has risen dramatically since 1972. This may reflect the lower mortality rates due to improved health care and the reduction of death rates from diseases significant to the elderly. The percentage of companies offering nonzero nonmedical insurance to groups below the age of 15 dropped slightly.

A comparison of the various coefficients of variation suggests that the differences among companies increased over both decades for the first seven age groups (0-30) in the total sample, particularly in the decade from 1972 to 1982 (Table 5). For the next five age groups (31-55) the variation increased from 1972 to 1982, but the differences in nonmedical limits among companies decline markedly. For the last three age groups the variation among companies from 1972 to 1992 consistently declined. Much of the reduction in variation at the older ages can be attributed to those companies which went from zero to positive nonmedical limits in that age range. The data further suggest that positive socioeconomic factors for the older age groups in the decade from 1982 to 1992 may have overridden any differences in individual company underwriting costs. The greater variability in practice for the lower age groups, however, suggests that company policies differ more in targeting this age group.

---

3 The coefficient of variation is the ratio of the standard deviation to the (nonzero) mean.
Table 3  
Percentage Increases in Mean  
Non-Medical Limits: All Firms (in %)

<table>
<thead>
<tr>
<th>Age Range</th>
<th>1972-82</th>
<th>1982-92</th>
<th>1972-92</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>347</td>
<td>66</td>
<td>640</td>
</tr>
<tr>
<td>5-14</td>
<td>336</td>
<td>62</td>
<td>605</td>
</tr>
<tr>
<td>15-15</td>
<td>325</td>
<td>61</td>
<td>585</td>
</tr>
<tr>
<td>16-17</td>
<td>327</td>
<td>56</td>
<td>567</td>
</tr>
<tr>
<td>18-20</td>
<td>326</td>
<td>48</td>
<td>528</td>
</tr>
<tr>
<td>21-25</td>
<td>344</td>
<td>43</td>
<td>535</td>
</tr>
<tr>
<td>26-30</td>
<td>353</td>
<td>44</td>
<td>553</td>
</tr>
<tr>
<td>31-35</td>
<td>417</td>
<td>67</td>
<td>761</td>
</tr>
<tr>
<td>36-40</td>
<td>545</td>
<td>97</td>
<td>1,173</td>
</tr>
<tr>
<td>41-45</td>
<td>1,118</td>
<td>115</td>
<td>2,516</td>
</tr>
<tr>
<td>46-50</td>
<td>4,980</td>
<td>161</td>
<td>13,160</td>
</tr>
<tr>
<td>51-55</td>
<td>10,500</td>
<td>60</td>
<td>16,833</td>
</tr>
<tr>
<td>56-60</td>
<td>21,186</td>
<td>17</td>
<td>24,900</td>
</tr>
<tr>
<td>61-65</td>
<td>42,900</td>
<td>(36)</td>
<td>27,400</td>
</tr>
<tr>
<td>66-70</td>
<td>∞</td>
<td>(71)</td>
<td>∞</td>
</tr>
</tbody>
</table>

For the entire sample the null hypothesis that variation among firms did not change from decade to decade is rejected at the 5 percent level using Friedman's nonparametric test (Table 6). The alternate hypothesis that the variation increased over time could neither be accepted nor rejected using Page's ordered test, while a second alternate hypothesis that the variation decreased over time failed to be accepted (Table 7). The analysis suggests that the divergent pattern in nonmedical limits for the younger age groups more than offsets the convergent patterns for the older age groups, but only to a small extent. There is no ordered pattern to this variation, however; neither the highest nor the lowest nonmedical limits fall in the same issue age category for the years 1972, 1982, and 1992.

The results are similar when the tests are performed only on the 48 firms for which data are available for both decades. The null hypothesis that variations among firms did not change from decade to decade fails to be rejected (Table 6 and Table 7). This is true even though the pattern
Table 4
Percentage of Companies Offering Nonzero Nonmedical Privileges

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>100.0</td>
<td>98.8</td>
<td>96.6</td>
<td>(1.2)</td>
<td>(2.2)</td>
<td>(3.4)</td>
</tr>
<tr>
<td>5-14</td>
<td>100.0</td>
<td>98.8</td>
<td>96.6</td>
<td>(1.2)</td>
<td>(2.2)</td>
<td>(3.4)</td>
</tr>
<tr>
<td>15-15</td>
<td>100.0</td>
<td>99.4</td>
<td>99.2</td>
<td>(0.6)</td>
<td>(0.2)</td>
<td>(0.8)</td>
</tr>
<tr>
<td>16-17</td>
<td>100.0</td>
<td>100.0</td>
<td>99.2</td>
<td>0</td>
<td>(0.8)</td>
<td>(0.8)</td>
</tr>
<tr>
<td>18-20</td>
<td>100.0</td>
<td>100.0</td>
<td>99.2</td>
<td>0</td>
<td>(0.8)</td>
<td>(0.8)</td>
</tr>
<tr>
<td>21-25</td>
<td>100.0</td>
<td>100.0</td>
<td>99.2</td>
<td>0</td>
<td>(0.8)</td>
<td>(0.8)</td>
</tr>
<tr>
<td>26-30</td>
<td>99.1</td>
<td>100.0</td>
<td>99.2</td>
<td>0.9</td>
<td>(0.8)</td>
<td>0.1</td>
</tr>
<tr>
<td>31-35</td>
<td>98.2</td>
<td>98.8</td>
<td>99.2</td>
<td>0.6</td>
<td>0.4</td>
<td>1.0</td>
</tr>
<tr>
<td>36-40</td>
<td>98.2</td>
<td>98.8</td>
<td>99.2</td>
<td>0.6</td>
<td>0.4</td>
<td>1.0</td>
</tr>
<tr>
<td>41-45</td>
<td>62.8</td>
<td>91.5</td>
<td>96.6</td>
<td>45.7</td>
<td>5.6</td>
<td>53.8</td>
</tr>
<tr>
<td>46-50</td>
<td>15.9</td>
<td>50.0</td>
<td>89.9</td>
<td>214.5</td>
<td>79.8</td>
<td>465.4</td>
</tr>
<tr>
<td>51-55</td>
<td>8.0</td>
<td>22.6</td>
<td>43.7</td>
<td>182.5</td>
<td>93.4</td>
<td>446.3</td>
</tr>
<tr>
<td>56-60</td>
<td>4.4</td>
<td>13.4</td>
<td>30.3</td>
<td>204.5</td>
<td>126.1</td>
<td>588.6</td>
</tr>
<tr>
<td>61-65</td>
<td>1.7</td>
<td>9.8</td>
<td>15.1</td>
<td>444.4</td>
<td>54.1</td>
<td>738.9</td>
</tr>
<tr>
<td>66-70</td>
<td>0</td>
<td>6.1</td>
<td>8.4</td>
<td>∞</td>
<td>37.7</td>
<td>∞</td>
</tr>
</tbody>
</table>

The null hypothesis fails to be rejected because there is a strong pattern of convergence in company practices in the age groups extending from 41 to 70. The increasing similarity in the behavior of these companies may have allowed other more independent firms to carve niches in these target markets, which would explain the ambivalence in the results for the entire sample.
Table 5

Coefficients of Variation for All Age Groups

<table>
<thead>
<tr>
<th>Sample: All Firms</th>
<th>Age</th>
<th>0-4</th>
<th>5-14</th>
<th>15</th>
<th>16-17</th>
<th>18-20</th>
<th>21-25</th>
<th>26-30</th>
<th>31-35</th>
<th>36-40</th>
<th>41-45</th>
<th>46-50</th>
<th>51-55</th>
<th>56-60</th>
<th>61-65</th>
<th>66-70</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td></td>
<td>69.02</td>
<td>69.17</td>
<td>61.34</td>
<td>62.82</td>
<td>64.32</td>
<td>64.19</td>
<td>64.19</td>
<td>55.31</td>
<td>63.51</td>
<td>75.13</td>
<td>98.50</td>
<td>211.04</td>
<td>286.21</td>
<td>332.91</td>
<td>453.86</td>
</tr>
<tr>
<td>1982</td>
<td></td>
<td>63.00</td>
<td>59.59</td>
<td>58.35</td>
<td>57.51</td>
<td>57.35</td>
<td>59.38</td>
<td>60.10</td>
<td>78.89</td>
<td>116.88</td>
<td>188.10</td>
<td>300.59</td>
<td>457.71</td>
<td>500.98</td>
<td>558.50</td>
<td>762.51</td>
</tr>
<tr>
<td>1972</td>
<td></td>
<td>30.43</td>
<td>24.54</td>
<td>20.20</td>
<td>19.58</td>
<td>19.33</td>
<td>18.38</td>
<td>21.15</td>
<td>28.95</td>
<td>47.30</td>
<td>109.30</td>
<td>297.31</td>
<td>386.89</td>
<td>518.27</td>
<td>763.39</td>
<td>0</td>
</tr>
<tr>
<td>Sample: 48 Firms</td>
<td>Age</td>
<td>0-4</td>
<td>5-14</td>
<td>15</td>
<td>16-17</td>
<td>18-20</td>
<td>21-25</td>
<td>26-30</td>
<td>31-35</td>
<td>36-40</td>
<td>41-45</td>
<td>46-50</td>
<td>51-55</td>
<td>56-60</td>
<td>61-65</td>
<td>66-70</td>
</tr>
<tr>
<td>1992</td>
<td></td>
<td>64.55</td>
<td>64.40</td>
<td>63.59</td>
<td>61.83</td>
<td>64.23</td>
<td>64.23</td>
<td>64.23</td>
<td>51.23</td>
<td>62.82</td>
<td>88.49</td>
<td>109.89</td>
<td>247.35</td>
<td>334.40</td>
<td>316.56</td>
<td>492.25</td>
</tr>
<tr>
<td>1982</td>
<td></td>
<td>51.29</td>
<td>50.33</td>
<td>48.15</td>
<td>48.15</td>
<td>47.65</td>
<td>47.54</td>
<td>48.60</td>
<td>64.20</td>
<td>78.86</td>
<td>104.66</td>
<td>242.14</td>
<td>345.96</td>
<td>631.64</td>
<td>631.64</td>
<td>631.64</td>
</tr>
<tr>
<td>1972</td>
<td></td>
<td>31.35</td>
<td>23.83</td>
<td>18.03</td>
<td>16.86</td>
<td>16.86</td>
<td>16.95</td>
<td>18.97</td>
<td>28.40</td>
<td>34.47</td>
<td>108.35</td>
<td>442.35</td>
<td>363.82</td>
<td>692.82</td>
<td>692.82</td>
<td>0</td>
</tr>
<tr>
<td>Sample: Mutual Companies - 26 Firms</td>
<td>Age</td>
<td>0-4</td>
<td>5-14</td>
<td>15</td>
<td>16-17</td>
<td>18-20</td>
<td>21-25</td>
<td>26-30</td>
<td>31-35</td>
<td>36-40</td>
<td>41-45</td>
<td>46-50</td>
<td>51-55</td>
<td>56-60</td>
<td>61-65</td>
<td>66-70</td>
</tr>
<tr>
<td>1992</td>
<td></td>
<td>63.28</td>
<td>64.94</td>
<td>62.95</td>
<td>58.90</td>
<td>66.06</td>
<td>66.06</td>
<td>66.06</td>
<td>18.19</td>
<td>46.72</td>
<td>45.81</td>
<td>53.78</td>
<td>129.79</td>
<td>216.15</td>
<td>429.98</td>
<td>509.90</td>
</tr>
<tr>
<td>1982</td>
<td></td>
<td>44.72</td>
<td>45.03</td>
<td>45.03</td>
<td>45.03</td>
<td>45.03</td>
<td>74.42</td>
<td>45.98</td>
<td>65.86</td>
<td>57.50</td>
<td>90.57</td>
<td>261.28</td>
<td>269.03</td>
<td>509.90</td>
<td>509.90</td>
<td>509.90</td>
</tr>
<tr>
<td>1972</td>
<td></td>
<td>25.44</td>
<td>15.61</td>
<td>15.61</td>
<td>15.10</td>
<td>15.10</td>
<td>15.29</td>
<td>17.23</td>
<td>19.22</td>
<td>23.83</td>
<td>97.30</td>
<td>355.62</td>
<td>363.32</td>
<td>509.90</td>
<td>509.90</td>
<td>509.90</td>
</tr>
<tr>
<td>Sample: Stock Companies - 22 Firms</td>
<td>Age</td>
<td>0-4</td>
<td>5-14</td>
<td>15</td>
<td>16-17</td>
<td>18-20</td>
<td>21-25</td>
<td>26-30</td>
<td>31-35</td>
<td>36-40</td>
<td>41-45</td>
<td>46-50</td>
<td>51-55</td>
<td>56-60</td>
<td>61-65</td>
<td>66-70</td>
</tr>
<tr>
<td>1992</td>
<td></td>
<td>56.21</td>
<td>56.21</td>
<td>53.23</td>
<td>54.73</td>
<td>53.23</td>
<td>53.23</td>
<td>54.95</td>
<td>77.74</td>
<td>102.65</td>
<td>133.59</td>
<td>216.90</td>
<td>267.07</td>
<td>235.59</td>
<td>409.65</td>
<td>0</td>
</tr>
<tr>
<td>1982</td>
<td></td>
<td>60.07</td>
<td>56.91</td>
<td>52.80</td>
<td>52.80</td>
<td>51.54</td>
<td>51.54</td>
<td>52.15</td>
<td>65.28</td>
<td>97.72</td>
<td>119.65</td>
<td>210.28</td>
<td>431.80</td>
<td>458.26</td>
<td>458.26</td>
<td>458.26</td>
</tr>
<tr>
<td>Sample: New York Companies - 22 Firms</td>
<td>Age</td>
<td>0-4</td>
<td>5-14</td>
<td>15</td>
<td>16-17</td>
<td>18-20</td>
<td>21-25</td>
<td>26-30</td>
<td>31-35</td>
<td>36-40</td>
<td>41-45</td>
<td>46-50</td>
<td>51-55</td>
<td>56-60</td>
<td>61-65</td>
<td>66-70</td>
</tr>
<tr>
<td>1992</td>
<td></td>
<td>65.02</td>
<td>63.02</td>
<td>63.02</td>
<td>56.59</td>
<td>62.48</td>
<td>62.48</td>
<td>62.48</td>
<td>51.75</td>
<td>67.61</td>
<td>73.49</td>
<td>88.90</td>
<td>150.74</td>
<td>191.76</td>
<td>248.84</td>
<td>342.88</td>
</tr>
<tr>
<td>1982</td>
<td></td>
<td>54.97</td>
<td>56.56</td>
<td>52.83</td>
<td>52.83</td>
<td>52.83</td>
<td>52.94</td>
<td>54.77</td>
<td>75.99</td>
<td>79.04</td>
<td>104.48</td>
<td>218.55</td>
<td>231.40</td>
<td>469.04</td>
<td>469.04</td>
<td>469.04</td>
</tr>
<tr>
<td>1972</td>
<td></td>
<td>23.14</td>
<td>17.01</td>
<td>17.01</td>
<td>17.01</td>
<td>17.01</td>
<td>17.27</td>
<td>19.73</td>
<td>19.35</td>
<td>36.54</td>
<td>120.77</td>
<td>325.86</td>
<td>333.01</td>
<td>469.04</td>
<td>469.04</td>
<td>0</td>
</tr>
<tr>
<td>Sample: Non-New York Companies - 26 Firms</td>
<td>Age</td>
<td>0-4</td>
<td>5-14</td>
<td>15</td>
<td>16-17</td>
<td>18-20</td>
<td>21-25</td>
<td>26-30</td>
<td>31-35</td>
<td>36-40</td>
<td>41-45</td>
<td>46-50</td>
<td>51-55</td>
<td>56-60</td>
<td>61-65</td>
<td>66-70</td>
</tr>
<tr>
<td>1992</td>
<td></td>
<td>49.92</td>
<td>49.92</td>
<td>48.78</td>
<td>51.20</td>
<td>52.47</td>
<td>52.47</td>
<td>52.47</td>
<td>49.53</td>
<td>60.06</td>
<td>93.39</td>
<td>124.77</td>
<td>236.58</td>
<td>331.10</td>
<td>340.63</td>
<td>500.00</td>
</tr>
<tr>
<td>1982</td>
<td></td>
<td>46.45</td>
<td>43.44</td>
<td>43.44</td>
<td>42.71</td>
<td>42.71</td>
<td>42.97</td>
<td>48.82</td>
<td>80.75</td>
<td>97.89</td>
<td>233.97</td>
<td>500.00</td>
<td>500.00</td>
<td>500.00</td>
<td>500.00</td>
<td>500.00</td>
</tr>
<tr>
<td>1972</td>
<td></td>
<td>36.00</td>
<td>28.82</td>
<td>19.14</td>
<td>17.09</td>
<td>17.09</td>
<td>17.09</td>
<td>18.96</td>
<td>34.13</td>
<td>33.35</td>
<td>95.74</td>
<td>500.00</td>
<td>500.00</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 6
The Friedman Test

<table>
<thead>
<tr>
<th>Sample</th>
<th>Friedman Test Statistic</th>
<th>Test Indication on Null at 5% Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Firms</td>
<td>6.93</td>
<td>Reject</td>
</tr>
<tr>
<td>48 Firms</td>
<td>2.53</td>
<td>Fail to reject</td>
</tr>
<tr>
<td>Mutual Companies (26)</td>
<td>3.63</td>
<td>Fail to reject</td>
</tr>
<tr>
<td>Stock Companies (22)</td>
<td>19.07</td>
<td>Reject</td>
</tr>
<tr>
<td>New York Companies (22)</td>
<td>3.73</td>
<td>Fail to reject</td>
</tr>
<tr>
<td>Non-New York Companies (26)</td>
<td>11.03</td>
<td>Reject</td>
</tr>
</tbody>
</table>

Note: \( H_0: t_1 = t_2 = t_3; \)
\[ H_1: \text{At least one of the } t_is \text{ is different.} \]
The 5% critical value is for this test is 5.99.

Table 7
The Page Test

<table>
<thead>
<tr>
<th>Sample</th>
<th>Page Test Statistic</th>
<th>Test Indication on Null at 5% Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Firms</td>
<td>( H_1: 190 )</td>
<td>Unclear</td>
</tr>
<tr>
<td></td>
<td>( H_2: 170 )</td>
<td>Fail to reject</td>
</tr>
<tr>
<td>48 Firms</td>
<td>( H_1: 187 )</td>
<td>Fail to reject</td>
</tr>
<tr>
<td></td>
<td>( H_2: 173 )</td>
<td>Fail to reject</td>
</tr>
<tr>
<td>Mutual Companies (26)</td>
<td>( H_1: 188.5 )</td>
<td>Fail to reject</td>
</tr>
<tr>
<td></td>
<td>( H_2: 171.5 )</td>
<td>Fail to reject</td>
</tr>
<tr>
<td>Stock Companies (22)</td>
<td>( H_1: 195 )</td>
<td>Reject</td>
</tr>
<tr>
<td></td>
<td>( H_2: 170 )</td>
<td>Fail to reject</td>
</tr>
<tr>
<td>New York Companies (22)</td>
<td>( H_1: 188 )</td>
<td>Fail to reject</td>
</tr>
<tr>
<td></td>
<td>( H_2: 170 )</td>
<td>Fail to reject</td>
</tr>
<tr>
<td>Non-New York Companies (26)</td>
<td>( H_1: 196 )</td>
<td>Reject</td>
</tr>
<tr>
<td></td>
<td>( H_2: 164 )</td>
<td>Fail to reject</td>
</tr>
</tbody>
</table>

Note: \( H_0: t_1 = t_2 = t_3; \)
\[ H_1: t_1 < t_2 < t_3; \]
\[ H_2: t_1 > t_2 > t_3. \]
The 5% critical value is for this test is 190.
When the 48 firms are divided into New York carriers and non-New York insurers, the statistical tests provide interesting results. The tests indicate that the variation among New York carriers did not change over the two decades, while the null (no change) is strongly rejected for non-New York insurers (Table 6). Furthermore, Page's ordered test rejects the null in favor of the alternate that the variation among firms is increasing over time for the non-New York carriers (Table 7). The pattern in the coefficient of variation for the New York insurers remains similar to that for the sample of 48 firms.

When the sample of 48 firms is split on the basis of organization into stock and mutual firms, we again find interesting differences. For the stock companies, the null hypothesis that the variation in company practices did not change over time is strongly rejected in favor of the alternate (Table 6). Furthermore, Page's test rejects the null in favor of the alternate that the variation in company practices is increasing over time (Table 7). These variations are preponderant in the issue age groups from 0 to 30. Although the pattern of increasingly divergent practices exists at the lower age groups for the mutual companies, there seems to be convergence at the higher age groups. As a result, Friedman's test fails to reject the null of no changes. This result is further confirmed by Page's test—the null fails to be rejected in favor of either increasing or decreasing divergence in mutual company practices over time.

7 Conclusions

This study examines nonmedical limits for a sample of life insurance companies over a 20 year period to determine the extent of variability in company practices at several points in time and the change in variability over time. The study shows a greater variability in company practices for the lower age groups than for higher age groups. Part of this variability could be attributed to the fact that almost all companies offer nonmedical insurance in the lower age brackets. The number of companies offering nonmedical insurance at higher age brackets decreases sharply, particularly after age 50.

Analysis of data over time shows that the percentage of companies offering insurance at the higher age brackets has risen while the percentage at lower age brackets has dropped slightly. The number of companies offering nonmedical insurance to those below age 45 increased substantially in the first decade of our study, but decreased
slightly in the second decade. There is a continuous increase, however, in the number of firms offering nonmedical insurance at the higher age brackets. This fact could be attributed to improved mortality rates for the older population and to companies’ increased interest in the senior citizen market.

When the entire sample is examined, statistical tests suggest an increase in variability of company nonmedical limit schedules. When the subsample of 48 firms for which data are available over both decades is examined, however, there appears to be no substantive change in the variability of nonmedical limits. One possible explanation for these results is that new firms entering or leaving the market attempt to carve special niches that contribute to the greater variability in nonmedical limits.

Interesting questions about nonmedical limits in practice abound. Do companies construct new nonmedical limit schedules analytically along the lines suggested earlier in this paper? Or do they forego such calculations and base their decisions in part on the schedules of other companies—particularly competitors? How do companies manage their agency operations with nonmedical limits less liberal than competitors? And where is the industry headed with respect to limits for nonmedical and paramedical acceptances and for blood/urine testing? Qualitative data are required to provide useful answers to these questions. Perhaps these data are best secured through a survey instrument addressed to the companies. The survey approach would have the additional benefit of providing a larger sample by avoiding the data limitations that a source such as the *Best’s Flitcraft Compend (Life-Health)* necessarily imposes.

References


“Informal Discussion-NonMedical Business.” Record of the American Institute of Actuaries 14 (June 1925): 116-140.


“Underwriting.” Record of the American Institute of Actuaries 31 (October 1942): 544-555.

“Informal Discussion-Underwriting.” Record of the American Institute of Actuaries 33 (June 1944): 135-149.

“Underwriting.” Record of the American Institute of Actuaries 35 (June 1946): 134-149.


Appendix

The Friedman and Page statistics are explained below for the sample of all firms; for more details on these statistics see Hettmansperger (1991). They are nonparametric tests and are performed on the coefficient of variation for the sample of all firms and for all the subsamples.

The first column of Table A1 recognizes that there are 15 issue age groups in the sample. In the remaining three columns the values of the coefficient of variation (CV) and the respective ranking of each year based on the CVs are provided. A value of three is given to the year with the highest value of the CV, and the other years are rank-ordered accordingly for each age group. The years 1972, 1982, and 1992 are represented by \( t_1 \), \( t_2 \), and \( t_3 \), respectively, in the tests below.

For the Friedman test, the null hypothesis and the alternative hypothesis are:

\[
H_0 : \quad t_1 = t_2 = t_3;
\]
\[
H_1 : \quad \text{At least one of the } t_i \text{s is different.}
\]

The test statistic is:

\[
K^* = \frac{12}{nk(k+1)} \sum_{j=1}^{k} (R_{ij})^2 - 3n(k + 1)
\]
\[
= \frac{12}{15 \times 3 \times 4} \times [(22)^2 + (36)^2 + (32)^2] - 3 \times 15 \times 4
\]
\[
= 6.93
\]
Table A1
Coefficient of Variation
And Rank (in parentheses)

<table>
<thead>
<tr>
<th>Age Group</th>
<th>1972</th>
<th>1982</th>
<th>1992</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30.4 (1)</td>
<td>63.0 (2)</td>
<td>69.0 (3)</td>
</tr>
<tr>
<td>2</td>
<td>24.5 (1)</td>
<td>59.6 (2)</td>
<td>69.2 (3)</td>
</tr>
<tr>
<td>3</td>
<td>20.2 (1)</td>
<td>58.3 (2)</td>
<td>61.3 (3)</td>
</tr>
<tr>
<td>4</td>
<td>19.6 (1)</td>
<td>57.5 (2)</td>
<td>62.8 (3)</td>
</tr>
<tr>
<td>5</td>
<td>19.3 (1)</td>
<td>57.3 (2)</td>
<td>64.3 (3)</td>
</tr>
<tr>
<td>6</td>
<td>18.8 (1)</td>
<td>59.4 (2)</td>
<td>64.2 (3)</td>
</tr>
<tr>
<td>7</td>
<td>21.1 (1)</td>
<td>60.1 (2)</td>
<td>64.2 (3)</td>
</tr>
<tr>
<td>8</td>
<td>28.9 (1)</td>
<td>78.9 (3)</td>
<td>55.3 (2)</td>
</tr>
<tr>
<td>9</td>
<td>47.3 (1)</td>
<td>116.9 (3)</td>
<td>63.5 (2)</td>
</tr>
<tr>
<td>10</td>
<td>109.3 (2)</td>
<td>188.1 (3)</td>
<td>75.1 (1)</td>
</tr>
<tr>
<td>11</td>
<td>297.3 (2)</td>
<td>300.6 (3)</td>
<td>98.5 (1)</td>
</tr>
<tr>
<td>12</td>
<td>386.9 (2)</td>
<td>457.7 (3)</td>
<td>211.0 (1)</td>
</tr>
<tr>
<td>13</td>
<td>518.3 (3)</td>
<td>501.0 (2)</td>
<td>286.2 (1)</td>
</tr>
<tr>
<td>14</td>
<td>767.4 (3)</td>
<td>558.5 (2)</td>
<td>332.9 (1)</td>
</tr>
<tr>
<td>15</td>
<td>0.0 (1)</td>
<td>762.5 (3)</td>
<td>453.9 (2)</td>
</tr>
</tbody>
</table>

\[ R_j = \sum_{i=1}^{n} R_{ij} \quad j = 1, 2, \ldots, k. \]

where \( k \) is the total number of years \((k = 3)\); \( n \) is the number of issue age groups \((n = 15)\); \( R_{ij} \) is the rank of the \( i \)-th observation in year \( j \) relative to the other \( k - 1 \) years; and

The calculated value of \( K^* \) has a chi-square distribution with two degrees of freedom. The critical values at the 5 percent and 10 percent levels are 5.99 and 4.61, respectively. Thus, the hypothesis is rejected in favor of the alternative that the variations in the years 1972, 1982, and 1992 are not the same. (The hypothesis, however, fails to be rejected at the 1 percent level.)

Page's test for ordered alternatives asks whether the variable (in this case the coefficient of variation) is increasing over time or is decreasing
over time. The null hypothesis and the alternatives are:

\[ H_0 : \quad t_1 = t_2 = t_3; \]
\[ H_1 : \quad t_1 < t_2 < t_3; \]
\[ H_2 : \quad t_1 > t_2 > t_3. \]

The test statistic for \( H_1 \) is:

\[
L = \sum_{j=1}^{k} j \times R_j \\
= 1 \times 22 + 2 \times 36 + 3 \times 32 \\
= 190.
\]

The value of the test statistic is equal to the critical value of 190 at the 5 percent confidence level. Therefore the hypothesis is neither accepted nor rejected in favor of the alternative \( H_1 \) that the coefficient of variation is increasing over time.

The test statistic for \( H_2 \) is:

\[
L = \sum_{j=1}^{k} (k - j + 1) \times R_j \\
= 3 \times 22 + 2 \times 36 + 1 \times 32 \\
= 170.
\]

Because the calculated value of 170 is less than the critical value of 190, the hypothesis that the coefficient of variation remains constant over time fails to be rejected.
Abstract

Like many other countries, including the United States, Singapore faces the dual problems of rising health care costs and an aging population. To cope with these problems, the Singapore government introduced the Medishield scheme in 1989 that provides low cost catastrophic medical insurance coverage. The scheme suffers from a serious deficiency, however: coverage ceases at age 70. This deficiency is exacerbated by Medishield's premium payment structure which is akin to the premium structure of a one year renewable term policy so no reserves are developed. As a result, coverage beyond age 70 requires exorbitant premiums that are beyond the reach of the average Singaporean.

* Robert Keng Heong Lian, M.Sc., A.S.A., is a senior lecturer in actuarial science and insurance at Nanyang Technological University, Singapore. He received his masters degree in actuarial science from the Northeastern University. Mr. Lian was the general manager/principal officer of British American Life Assurance before becoming a lecturer.

Mr. Lian's address is: Nanyang Technological University, School of Accountancy and Business, Nanyang Avenue, Singapore 639-798. Internet address: akhlian@ntuvax.ntu.ac.sg

† Loi Soh Loi, Ph.D., is a lecturer at Nanyang Technological University, Singapore. She received her Ph.D., in mathematics from the University of Canterbury. Before joining the Nanyang Technological University, she was a member of the faculty of the National University of Singapore, having previously worked for the Department of Statistics, Christchurch, New Zealand.

Dr. Loi's address is: Nanyang Technological University, School of Accountancy and Business, Nanyang Avenue, Singapore 639-798. Internet address: aslo@ntuvax.ntu.ac.sg

‡ The authors wish to acknowledge the many helpful comments from Professor J.H. Pollard of Macquarie University, Australia, and the many helpful comments from their colleague Dr. Ponniah Elancheran. They are also grateful to the anonymous referees for their many helpful comments and suggestions that have improved considerably the quality and scope of the paper. The authors are particularly indebted to the editor for his assistance in developing several of the equations in this paper.
We propose a premium payment structure under which the annual premium payable remains level in real terms throughout the lifetime of the insured. This makes the premium structure similar to one that is a level percentage of salary. The model uses such key variables as the rate of return, the rates of inflation of general costs and of medical costs, and the rate of increase in the morbidity rate.

Key words and phrases: *premiums, inflation, reserves, gains and losses, surplus*

1 Introduction

Singapore soon will face the problem of rising health care costs as its population ages. To cope with this anticipated problem, the Singapore government introduced the Medisave scheme in 1984 under which every working person must contribute a certain percent of his or her income to meet personal or immediate family hospitalization expenses, particularly medical expenses incurred during old age.\(^1\)

1.1 The Medisave Scheme

Medisave is compulsory and is administered by the Central Provident Funds (CPF) Board, a statutory board of the Singapore government. All employed persons are required to be members of the CPF and must contribute 40 percent of their income\(^2\) to meet their retirement, housing, education, investment, and health care needs. This 40 percent contribution is jointly and equally shared by the employer and the employee and is allocated as follows:

- 6 to 8 percent of the total contribution goes to the Medisave account;
- 4 percent of the total contribution goes to the special account;
- The balance (28 to 30 percent) of the total contribution remains in the ordinary account.

The operation of the Medisave special account and ordinary account are described below.

**The Medisave Account:** Funds in the Medisave accounts may be used to meet personal or immediate family's hospitalization expenses,

\(^{1}\)Medisave coverage ceases at age 70.

\(^{2}\)Income is subject to a maximum monthly contribution of $2,400.
especially after retirement. For example, these funds can be used to pay for the hospital bills of the member's spouse, children, parents, or grandparents up to:

- $300 per day for daily hospital charges; and
- A fixed limit per table of surgical operation according to the complexity of the operation.

If the hospital bill exceeds Medisave limits, the member is obligated to pay the part of the hospital bill not covered by Medisave. In addition, Medisave can be used to pay for the hospital stay for the delivery of the first three children.

Employee contribution rates progressively increase with age as shown in Table 1: For self-employed workers, contribution rates are lower. Medisave contributions and savings earn interest at prevailing market rates and are tax-deductible. The Medisave balance can be accumulated up to $17,000; any amount in excess of this limit is automatically transferred into the ordinary account.

**Table 1**

<table>
<thead>
<tr>
<th>Age</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 35</td>
<td>6%</td>
</tr>
<tr>
<td>35 to 44</td>
<td>7%</td>
</tr>
<tr>
<td>45 and above</td>
<td>8%</td>
</tr>
</tbody>
</table>

The **Special Account**: Funds in the special account can be used to finance the Minimum Sum scheme, which is a compulsory national retirement scheme to help members support a modest standard living during retirement. Starting from $40,000 in 1995, the minimum sum will be raised by $5,000 a year until it reaches $80,000 in 2003. At least half the minimum sum must be in cash, the other half may consist of tangible assets such as property. The cash portion ensures members of a monthly income in retirement. Prior to retirement, members have three options to invest their minimum sum: (i) buy a life annuity from an approved insurance company, (ii) keep it with an approved bank, or (iii) leave it with the CPF Board. If the income in the special account is less than...
the minimum sum, the balance of the minimum sum is covered by the income in the ordinary account.

**The Ordinary Account:** Funds in the ordinary accounts can be used for several different purposes including: (i) retirement (together with the special account to meet the minimum sum requirement); (ii) housing (can be government or private houses, for the purpose of owner occupation or for investment in one or more houses); (iii) education (restricted to local tertiary education), and (iv) investment (in common stocks and bonds, government bonds, fixed deposit, unit trust, gold, as well as endowment insurance policy).

1.2 The Medishield Scheme

CPF members' Medisave accounts, however, may be strained in the event of a prolonged illness that requires long-term medical treatment. To protect the Medisave account, the Singapore government introduced the Medishield scheme in 1990, a low cost medical insurance that adds more value to Medisave. Medishield is major medical insurance with participation built in by way of deductible and co-insurance as measures to contain cost. All CPF members automatically are covered under Medishield unless they elect to opt out. The Medishield premium and any costs not covered by Medishield, such as deductibles, co-insurances, and amounts in excess of the maximum claimable amount, can be paid from the Medisave account. Table 2 shows the benefits and claim limits under Medishield. Those limits are per policy. Each of the family member will have his/her own policy. Therefore, for example the deductible $4,000 is for the particular family member and not the entire family. Likewise for all of the rest.

Let $AP(x, z)$ be the actual annual premium charged for one year health insurance coverage in calendar year $z$ payable under Medishield. The actual premium charged is determined from expected claims by applying the percentage loading according to the formula:

$$AP(x, z) = (1 + \theta)EC(x, z)$$

where $EC(x, z)$ is the expected claim cost in the calendar year $z$ for insured age $x$ last birthday on 1/1/z, and $\theta$ is the expense and other loading applied to the net premium to get the actual annual premium charged.

Table 3 shows how these premiums increase with age.

---

3 As these premiums are only for coverage for one year, they are similar to the premiums for yearly renewable term policy in life insurance.
Table 2

Medishield Benefits and Claim Limits (in Singapore $s)

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deductible</td>
<td>$4,000 per policy year</td>
</tr>
<tr>
<td>Claim Limits</td>
<td>$70,000 per policy year; $200,000 lifetime</td>
</tr>
<tr>
<td>Hospital Stay*</td>
<td>$500 per day</td>
</tr>
<tr>
<td>Intensive Care Unit*</td>
<td>$800 per day</td>
</tr>
<tr>
<td>Surgical Operations</td>
<td>$400 to $5,500 per policy year</td>
</tr>
<tr>
<td>Implants</td>
<td>$3,500 per policy year</td>
</tr>
<tr>
<td>Outpatient Treatment</td>
<td>Limits vary according to treatment.</td>
</tr>
</tbody>
</table>

*Includes meals, prescriptions, investigations, and other miscellaneous charges.

The annual premiums shown in Table 1 are the 1995 published rates charged by the CPF's Medishield program on an annual basis for an insured of age \( x \) last birthday. Because we do not have access to original data such as the morbidity rate and other assumptions used by Medishield in determining \( AP(x, z) \), it is difficult to estimate the annual premiums for ages beyond age 70. A quick estimate of the projected premiums beyond age 70 can be obtained by assuming that health care costs continue to increase exponentially at advanced ages. The formula

\[
AP(x, 1995) = 13.17(1.06)^x
\]

for \( x = 25, 35, 45, 55, 62.5, \) and 67.5 gives a good least squares fit to the data in Table 1. Equation (1) is used to provide estimates for \( AP(x, z) \) for higher ages as tabulated in Table 4. Table 4 shows that if coverage is extended beyond age 70, the annual premium for ages over 70 will be more than most Singaporeans can afford.

1.3 Objectives

The objective of this paper is to develop a model of the Medishield program without the restriction that coverage ceases after age 70, where it is needed most. To pay for this extended benefit, we develop a premium structure that remains level in real terms and is approximately a constant percentage of a worker's salary. A method for performing a detailed analysis of the annual gains and losses and the sources of surplus of the Medishield program is provided.
Table 3
Medishield
1995 Annual Premiums*

<table>
<thead>
<tr>
<th>Age (x)</th>
<th>AP(x, 1995)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 30</td>
<td>$ 60</td>
</tr>
<tr>
<td>31 - 40</td>
<td>$ 90</td>
</tr>
<tr>
<td>41 - 50</td>
<td>$180</td>
</tr>
<tr>
<td>51 - 60</td>
<td>$360</td>
</tr>
<tr>
<td>61 - 65</td>
<td>$480</td>
</tr>
<tr>
<td>66 - 70</td>
<td>$660</td>
</tr>
</tbody>
</table>

*In Singapore dollars.

2 Notation

The analysis presented in this paper is based on population information that may be most readily available at the start of each calendar year. If, however, information is gathered on a different basis, such as is the case for fiscal years, then our results can easily be modified. This requires changing 1/1/z to the date at the start of the fiscal year. All events are then defined over the fiscal year in an obvious manner.

The following notation is used throughout this paper:

\[
\begin{align*}
\text{lbd} & = \text{Last birthday;} \\
b & = \text{Annual morbidity rate of increase due to age; } \\
g & = \text{Annual rate of inflation (general cost of living);} \\
m & = \text{Annual rate of medical inflation;} \\
i & = \text{Annual valuation rate of interest;} \\
\nu & = (1 + i)^{-1}; \\
z_0 & = \text{Calendar year of issue on the policy;} \\
x_0 & = \text{Age lbd on 1/1/z}; \\
z & = \text{Current calendar year;} \\
x & = \text{Current age last birthday on 1/1/z;} \\
kP_x^{(r)} & = \text{Probability that a Medshield insured age } x \\
& \text{lbd survives to age } x + k \text{ lbd;} \\
q_x^{(d)} & = \text{Probability that a Medshield insured age } x
\end{align*}
\]
Table 4
Medishield Extrapolated
Annual Premium*

<table>
<thead>
<tr>
<th>Age</th>
<th>Premium</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>$778</td>
</tr>
<tr>
<td>75</td>
<td>$1,041</td>
</tr>
<tr>
<td>80</td>
<td>$1,393</td>
</tr>
<tr>
<td>85</td>
<td>$1,865</td>
</tr>
<tr>
<td>90</td>
<td>$2,495</td>
</tr>
<tr>
<td>95</td>
<td>$3,339</td>
</tr>
</tbody>
</table>

*In Singapore dollars.

\[
q_x^{(w)} = \text{Probability that a Medshield insured age } x \text{ lbd withdraws before age } x + 1 \text{ lbd;}
\]

\[
q_x^{(r)} = q_x^{(d)} + q_x^{(w)};
\]

\[r_x^z = \text{Morbidity rate in calendar year } z \text{ for insured age } x \text{ lbd on } 1/1/;\]

\[ECS(x, z) = \text{Expected Medshield claim cost in calendar year } z \text{ for an insured age } x \text{ lbd on } 1/1/ \text{ who survives to } 1/1/z + 1;\]

\[ECD(x, z) = \text{Expected Medshield claim cost in calendar year } z \text{ for an insured age } x \text{ lbd on } 1/1/ \text{ who dies before } 1/1/z + 1;\]

\[ECW(x, z) = \text{Expected Medshield claim cost in calendar year } z \text{ for an insured age } x \text{ lbd on } 1/1/ \text{ who withdraws before } 1/1/z + 1;\]

\[1 + f_s = \text{Annual extra claim inflation factor for survivors;}\]

\[1 + f_d = \text{Annual extra claim inflation factor for deaths;}\]

\[1 + f_w = \text{Annual extra claim inflation factor for withdrawals;}\]

\[LP(x_0, z_0) = \text{Level (in real terms) annual premium payable upto age 65 for Medshield coverage sold in calendar year } z_0 \text{ to a person age } x_0 \text{ lbd on } 1/1/z_0.\]
Note that for all calculations we will use the Singapore male mortality table for the minimum reserve calculation for annuity (MAS 309)\(^4\) with \(q_{99} = 1\). This table is deemed suitable, as it is based on the \(a(90)\) table with improvement in mortality over time. There is no service table, with decrements for disability and withdrawals, available for use in Singapore's Medishield program as of May 1996.

3 The Individual Model

We develop a model in which the charged premium remains level in real terms throughout the working life of the member (Black & Skipper, 1994, Chapter 22). Level premium in real terms means that the premiums in subsequent years are adjusted by the rate of inflation of the general cost of living. Such level premiums also decrease the rate of growth of premiums at the older ages and tend to make lifetime coverage more affordable. Another advantage of this approach is that it makes financial planning easier in an environment where wages and salaries keep pace with the general cost of living increases; these premiums are approximately a constant percentage of wages and salaries.

A key feature of our model is that it removes the restriction that coverage ceases after age 70, as is the case with the Medishield program. Thus, coverage is available in the advanced ages where people may need it most.

3.1 Benefits

For increased flexibility, we develop three separate expected claim costs in any calendar year: (i) for those who continue their Medishield policy during the next calendar year \((ECS(x, z))\), (ii) for those who die within the calendar year \((ECD(x, z))\), and (iii) for those who withdraw from the Medishield program during the calendar year \((ECW(x, z))\). Using anti-selection arguments, one may expect those who withdraw to be healthier than those who remain, while those who die may tend to have larger than average costs. Thus anti-selection arguments suggest that we adjust the expected claim costs by the appropriate extra claim inflation factor \((1 + f)\) to reflect the different expected experiences among those who die, withdraw, or continue with the policy.

\(^4\)The MAS 309 is the notice No. 309 issued by the Monetary Authority of Singapore (equivalent to the Federal Reserve Board in the United States) to life insurers. In MAS 309, the mortality tables for calculating the minimum reserves for annuities are given.
We will assume that expected claim costs are paid at the end of the calendar year. As a result, we include the expected interest accrued to the end of the calendar year in our definition of the expected claims.

Our model assumes that the expected claim cost for an insured increases for two reasons: (i) increases in the underlying rate morbidity rate, and (ii) increases due to medical inflation. Motivated by equation (1), we model the morbidity rate due to age (within the a calendar year) as:

\[ r_{x+1}^{z} = (1 + b) r_x^{z}, \]

which gives

\[ ECS(x + 1, z) = (1 + b) ECS(x, z). \]

On the other hand, the impact of annual medical inflation on expected claim costs is assumed to affect only the size of claims not the morbidity rate. So, \( r_{x+1}^{z+1} = r_x^{z} \), and

\[ ECS(x, z + 1) = (1 + m)(1 + f_s) ECS(x, z). \]

The final factor to be applied is the extra claim inflation factor. The combined effect on expected claim costs is as follows:

\[
\begin{align*}
ECS(x + k, z + k) & = [(1 + b)(1 + m)(1 + f_s)]^k ECS(x, z) \\
ECD(x + k, z + k) & = [(1 + b)(1 + m)(1 + f_d)]^k ECD(x, z) \\
ECW(x + k, z + k) & = [(1 + b)(1 + m)(1 + f_w)]^k ECW(x, z).
\end{align*}
\]

Note that we expect \( f_w \leq f_s \leq f_d \).

The actuarial present value of the future health claims on 1/1/\( z \) for individual age \( x \) on 1/1/\( z \) is denoted by \( AHC(x, z) \) where

\[
\begin{align*}
AHC(x, z) & = \sum_{k=0}^{\infty} \nu^{(k+1)} p_x^{(\tau)} ECS(x + k, z + k) \\
& \quad + \sum_{k=0}^{\infty} \nu^{(k+1)} p_x^{(\tau)} q_x^{(d)} ECD(x + k, z + k) \\
& \quad + \sum_{k=0}^{\infty} \nu^{(k+1)} p_x^{(\tau)} q_x^{(w)} ECW(x + k, z + k).
\end{align*}
\]

Notice that equation (6) includes costs incurred after age 70.\(^5\) Using equations (3), (4), and (5) the expression for \( AHC(x, z) \) can be rewritten

\(^5\) Recall that the Medishield program ceases coverage at age 70 while the proposed Medishield program provides coverage beyond age 70.
as follows:

\[
AHC(x, z) = ECS(x, z) \sum_{k=0}^{\infty} (1 + i_1)^{-(k+1)} k^{(\tau)} p_x^\tau + ECD(x, z) \sum_{k=0}^{\infty} (1 + i_2)^{-(k+1)} k^{(d)} p_x^\tau q_x^{(d)} + ECW(x, z) \sum_{k=0}^{\infty} (1 + i_3)^{-(k+1)} k^{(w)} p_x^\tau q_x^{(w)}
\]

\[
= ECS(x, z) a_x + ECD(x, z) A_x^{(d)} + ECW(x, z) A_x^{(w)}
\]

where

\[
a_x = \sum_{k=0}^{\infty} (1 + i_1)^{-(k+1)} k^{(\tau)} p_x^\tau
\]

\[
A_x^{(d)} = \sum_{k=0}^{\infty} (1 + i_2)^{-(k+1)} k^{(d)} p_x^\tau q_x^{(d)}
\]

\[
A_x^{(w)} = \sum_{k=0}^{\infty} (1 + i_3)^{-(k+1)} k^{(w)} p_x^\tau q_x^{(w)}
\]

\[
1 + i_1 = \frac{1 + i}{(1 + b)(1 + m)(1 + f_s)}
\]

\[
1 + i_2 = \frac{1 + i}{(1 + b)(1 + m)(1 + f_d)}
\]

\[
1 + i_3 = \frac{1 + i}{(1 + b)(1 + m)(1 + f_w)}
\]

### 3.2 Level Premiums in Real Terms

By assuming a constant rate of inflation, \( g \), we can determine the individual net premium (payable to age 65) that is level in real terms. That is, if \( LP(x, z) \) is the net level (in real terms) annual premium payable up to age 65 for health insurance coverage for a newly insured age \( x \) last birthday on 1/1/\( z \), then the premiums increase annually by a factor of \( 1 + g \). This implies that the premium for policy year \( (k + 1) \) is \( (1 + g)^k LP(x, z) \).
The actuarial present value of the future premiums on 1/1/z₀ for individual age x₀ on 1/1/z₀ is denoted by $AFP(x₀, z₀)$ where

$$AFP(x₀, z₀) = \sum_{k=0}^{65-x₀-1} v^k p_{x₀}^{(x)} (1 + g)^k LP(x₀, z₀)$$

$$= LP(x₀, z₀) s \bar{a}_{x₀; 65-x₀}$$

where

$$s \bar{a}_{x; 65-x} = \sum_{k=0}^{65-x-1} (1 + g)^k v^k p_{x}^{(x)}$$

$$= \sum_{k=0}^{65-x-1} (1 + i_4)^{-k} p_{x}^{(x)}$$

$$= \bar{a}_{x; 65-x} \text{ at } i_4,$$  \hspace{1cm} (15)

and

$$i_4 = (i - g)/(1 + g).$$  \hspace{1cm} (16)

The net premium is then given by

$$LP(x₀, z₀) = [ECS(x₀, z₀)a_{x₀} + ECD(x₀, z₀)A^{(d)}_{x₀} + ECW(x₀, z₀)A^{(w)}_{x₀}] / s \bar{a}_{x₀; 65-x₀}.$$  \hspace{1cm} (17)

### 4 Analysis of Gains and Losses

Once expected claim costs have been determined and the premiums have been set, the actual experience of the fund must be evaluated annually to see how it compares with what was expected. Large surpluses or large deficits may signal problems inherent in the pricing or benefit structure. To get a clear picture of the year's experience, we need to develop a formal system for determining reserves, surpluses, and gains.

Let $V(x, z|x₀, z₀)$ be the actuarial reserve on 1/1/z for an insured active worker who is age x last birthday on 1/1/z but was age x₀ last birthday on 1/1/z₀ at which time the policy was issued.

$$V(x, z|x₀, z₀) = ECS(x, z)a_{x} + ECD(x, z)A^{(d)}_{x} + ECW(x, z)A^{(w)}_{x} - LP(x₀, z₀)(1 + g)^{(z-z₀)} s \bar{a}_{x; 65-x}$$  \hspace{1cm} (18)
for \( x \geq x_0 \). But as

\[
a_x = \frac{p_x^{(\tau)}}{1 + i_1} (1 + a_{x+1}) \quad \text{and} \quad \hat{s}_x = \frac{p_x^{(\tau)}}{1 + i_4} \hat{s}_x^{(1:64-x)}
\]

it follows that

\[
V(x, z|x_0, z_0) = nECS(x, z)p_x^{(\tau)} + nECD(x, z)q_x^{(d)} + nECW(x, z)q_x^{(w)} + nV(x + 1, z + 1|x_0, z_0) - LP(x_0, z_0)(1 + g)^{z-z_0}.
\]

Equation (19) can be rearranged to give

\[
V(x + 1, z + 1|x_0, z_0) = (1 + i)(V(x, z|x_0, z_0)
+ (1 + i)LP(x_0, z_0)(1 + g)^{z-z_0})
- ECS(x, z)p_x^{(\tau)}
- ECD(x, z)q_x^{(d)}
- ECW(x, z)q_x^{(w)}
+ q_x^{(\tau)}V(x + 1, z + 1|x_0, z_0). \quad (20)
\]

Given this recursive relationship between successive years' reserves, we can develop an analysis of gains and losses in a manner similar to Anderson (1990, Chapter 2). To this end we introduce the following notation:

- \( \mathcal{A}_z \) = Set of the Medishield insureds alive on 1/1/z
- \( \mathcal{W}_z \) = Set of withdrawals from \( \mathcal{A}_z \) during calendar year z;
- \( \mathcal{D}_z \) = Set of deaths from \( \mathcal{A}_z \) during calendar year z;
- \( \mathcal{N}_z \) = Set of new Medishield insureds who joined the program during calendar year z.

In addition, let \( j \) denote a member of the set \( \mathcal{A}_z \), and we append the subscript \( j \) to our previous notation to refer to values that depend on the insured \( j \)'s policy experience.

We define \( TV_z \) to be the total actual Medishield reserve as of 1/1/z, i.e.,

\[
TV_z = \sum_{j \in \mathcal{A}_z} V_j(x, z|x_0, z_0)
\]
where $V_j(x, z | x_0, z_0)$ is the reserve for insured $j$. Following Anderson, we decompose $TV_{z+1}$ to give

$$TV_{z+1} = \sum_{j \in A_{z+1}} V_j(x + 1, z + 1 | x_0, z_0)$$

$$= \sum_{j \in A_z} V_j(x + 1, z + 1 | x_0, z_0)$$

$$- \sum_{j \in D_z} V_j(x + 1, z + 1 | x_0, z_0)$$

$$- \sum_{j \in W_z} V_j(x + 1, z + 1 | x_0, z_0)$$

$$+ \sum_{j \in N_z} V_j(x + 1, z + 1 | x_0, z_0).$$

(22)

In order to simplify our mathematical expressions and reduce clutter, we use the following abbreviation:

$$V_j(z + 1) = V_j(x + 1, z + 1 | x_0, z_0).$$

(23)

Substituting equation (20) into equation (22) yields

$$TV_{z+1} = (1 + i)TV_z$$

$$+ (1 + i) \sum_{j \in A_z} LP(x_0, z_0)(1 + g)^{z-z_0}$$

$$- \sum_{j \in A_z} ECS_j(x, z)p_x^{(\tau)}(x)$$

$$- \sum_{j \in A_z} ECD(x, z)q_x^{(d)}(x)$$

$$- \sum_{j \in A_z} ECW(x, z)q_x^{(w)}(x)$$

$$- \left[ \sum_{j \in D_z} V_j(z + 1) - \sum_{j \in A_z} q_x^{(d)}V_j(z + 1) \right]$$

$$- \left[ \sum_{j \in W_z} V_j(z + 1) - \sum_{j \in A_z} q_x^{(w)}V_j(z + 1) \right]$$

$$+ \sum_{j \in N_z} V_j(z + 1).$$

(24)

To develop an expression for the surplus, we must now consider the fund balance at the start of each calendar year. Let $F_z$ be the actual fund balance on 1/1/z.

$$F_{z+1} = F_z + I_z + TP_z - TC_z$$

(25)
where $I_z$ is the actual interest earned during calendar year $z$; $TP_z$ is the total premiums actually received during calendar year $z$; and $TC_z$ is the total claim costs actually paid during calendar year $z$. The surplus in hand on $1/1/z$, $SUR_z$ is defined to be

$$SUR_z = F_z - TV_z. \quad (26)$$

Using equations (24) and (26) yields

$$SUR_{z+1} = (1 + i)SUR_z +$$

$$+ [I_z - iF_z - ITP_z + ITC_z]$$

$$+ [(TP_z + ITP_z) - (1 + i) \sum_{j \in A_z} Lp_j(x_0, z_0) (1 + g)^{z - z_0}]$$

$$+ \left[ \sum_{j \in A_z} ECS_j(x, z) p_x^{(r)} - (TC_z^{(s)} + ITC_z^{(s)}) \right]$$

$$+ \left[ \sum_{j \in A_z} ECD(x, z) q_x^{(d)} - (TC_z^{(d)} + ITC_z^{(d)}) \right]$$

$$+ \left[ \sum_{j \in A_z} ECW(x, z) q_x^{(w)} - (TC_z^{(w)} + ITC_z^{(w)}) \right]$$

$$+ \left[ \sum_{j \in D_z} V_j(z + 1) - \sum_{j \in A_z} q_x^{(d)} V_j(z + 1) \right]$$

$$+ \left[ \sum_{j \in W_z} V_j(z + 1) - \sum_{j \in A_z} q_x^{(w)} V_j(z + 1) \right]$$

$$- \sum_{j \in N_z} V_j(z + 1) \quad (27)$$

where $ITP_z$ is the expected interest earned on $TP_z$; $TC_z^{(s)}$ is the total claims costs generated by those who survive to the end of calendar year $z$; $ITC_z^{(s)}$ is the expected interest earned on $TC_z^{(s)}$; $TC_z^{(d)}$ is the total claims costs generated by those who died during calendar year $z$; $ITC_z^{(d)}$ is the expected interest earned on $TC_z^{(d)}$; $TC_z^{(w)}$ is the total claims costs generated by those who withdrew during calendar year $z$; $ITC_z^{(w)}$ is the expected interest earned on $TC_z^{(w)}$;

$$TC_z = TC_z^{(s)} + TC_z^{(d)} + TC_z^{(w)}$$

$$ITC_z = ITC_z^{(s)} + ITC_z^{(d)} + ITC_z^{(w)}.$$ 

The actuarial gain for calendar year $z$ is defined to be

$$GAIN_z = SUR_{z+1} - (1 + i)SUR_z. \quad (28)$$

The various components of the gain can be identified from equation (27).
5 Closing Comments

The model we have developed is simply a first step in the process of developing a more comprehensive model of the Medishield program. We need actual Singapore data on mortality and morbidity to test the model.

There are several known limitations, however, inherent in our current model:

- We failed to separate morbidity by sex. Females, in general, have different morbidity patterns than males. They tend to have higher morbidity during their child-bearing years.

- We have simplified the morbidity pattern over calendar year as well as over age. A more appropriate model would use very general rates: $r_{x,z}^{(m)}$ for males and $r_{x,z}^{(f)}$ for females. These rates must be developed from population data.

- Insureds who die during a calendar year are likely to have very high medical expenses during the preceding five years of life. This should be explicitly included in the model. The current model accounts for these high end-of-life expenses through experience losses attributable to deaths.

- Persons who withdraw are likely to reenter if they expect to be ill or are ill. This poses severe anti-selection problems.

Our proposed model is only a first attempt at extending the current Medishield program. Much more work needs to be done before this proposal can be implemented in practice.

References


Central Provident Funds Board. Publications on the Medishield Scheme.


Monetary Authority of Singapore's Notices to Life Insurers, MAS 309, *Mortality Tables for Minimum Reserves Calculation for Annuity*. 
Concentration in American Property-Casualty Companies
Edward Nissan*

Abstract†

A Theil's entropy index utilizing premiums written as units is employed to measure trends in concentration of the largest 200 property-casualty companies in the United States between 1985 and 1993 based on Best's Insurance Report data. Each of the indexes confirms that concentration trends experienced no increase for the whole period for all 200 firms, the top 20, and subsets of lower ranked companies. Significant differences are observed, however, between groups of companies for the same period.

Key words and phrases: mergers and acquisitions, Theil's entropy, insurance

1 Introduction

Throughout its history the United States economy has experienced cycles of mergers and acquisitions. The most recent cycle, according to Shleifer and Vishny (1991) and Sikora (1995), occurred during the 1980s. Significant factors contributing to mergers and acquisitions in the 1980s included laxity in antitrust enforcement policies and improvements in takeover technology (such as leveraged buyouts and junk bonds).

The property and casualty insurance industry was not exempt from this merger wave. Farinella (1996) reports that between 1985 and 1995...
some 660 corporate retirements, acquisitions, and mergers occurred in response to what Gart (1994) describes as tremendous changes within the insurance industry. Increasing competition drives the change and forces more company mergers, creating conglomerates of multiple function companies. Multiple functions allow a company to take advantage of the opportunities in the emerging financial services field. Such activities may lead insurance companies to engage in anticompetitive practices, resulting in allegations of collusion, restriction of output, and favorable terms from consumers. Anticompetitive practice may cause the repeal of the McCarran-Ferguson Act which exempts insurance companies from federal antitrust enforcements.

Roughly 3,000 companies constitute the property-liability insurance industry. Due to numerous affiliations, however, there are really only 800 independent decision-making units or groups. In 1993, according to Huebner, Black, and Webb (1996), the net premiums written, the combined admitted assets, and policyholders' surplus totaled $253.8 billion, $571.5 billion, and $182.3 billion, respectively. Policyholders' surplus serves as a cushion so that larger-than-expected losses can be paid. The abundance of cash, access to cheap capital, and low interest rates helped boost the recent trend in mergers and acquisitions (Farinella 1996).

There are two major conflicting arguments regarding mergers. One, according to Gilbert (1989), is that mergers enhance efficiency by promoting consumer welfare through a superior allocation of productive, financial, and managerial resources. Potential competition serves as a control on monopoly power. Salop (1987) and Adams and Brock (1996) note that the rationale for this argument is provided by economists at the University of Chicago. Simply stated, this theory focuses on consumer welfare as the sole concern. The other argument, supported by economists at Harvard University, is that mergers damage the overall working of the economy, lessen competition, and increase concentration of sales and thereby create monopoly power in a given industry. The process of concentration is defined as the increase in the extent economic activity is controlled by large firms.

Clarke (1985) distinguishes market and aggregate concentration and absolute and relative concentration. Market concentration concerns a specific industry under the control of a few large firms which may lead to the exercise of monopoly power. Aggregate concentration occurs when a few large firms control broad segments of the economy (such as manufacturing, financial, and insurance sectors) or when the power of conglomerates extend beyond a particular industry. Changes in aggregate concentration may signal a change in the distribution of economic,
political, or social power. Another distinction is between relative and absolute concentration. If all firms grow the same proportion, concentration increases absolutely but relatively remains the same. Relative concentration is concerned with the share of output held by large firms among a fixed number of firms. Accordingly, various indexes are proposed to measure such concentrations.

The most widely used indexes are the \( k \)-firms concentration ratio, whereby the share of sales of the \( k \) largest firms (out of a total of \( n \) firms) are combined; the Herfindahl, defined as the sum of the shares of all \( n \) firms with the share of each firm weighted by itself; the Gini, which measures the extent to which firms in the industry are unequal in size; the coefficient of variation, which is the ratio of the standard deviation to the mean; and the Theil’s entropy,\(^1\) which is used in physics as a measure of disorder. Hannah and Kay (1977) find Theil’s entropy to be one of the most satisfactory indexes. As a result, Theil’s entropy is used throughout this paper as the measure of concentration.

2 Past Studies

In the two decades spanning 1973 to 1991 three studies provide excellent information on the structure of the property insurance industry. The common denominator of these articles is an assessment of market concentration by product line category or by ownership category. Joskow (1973) informs us that the 1206 property-liability insurance firms as a whole in 1971 held assets of $68 billion and premiums of $35 billion. Over half of the latter was written by the top 20 firms, resulting in a slight increase in concentration since 1961. Joskow also examines concentration within two individual lines, automobile and fire. Again, the top 20 firms accounted for concentration levels of approximately 56 percent in each line, increasing from 45 percent for automobile and from 49 percent for fire in 1954. Joskow believes that as a result of effective competition, consumers moved their business from high cost firms to low cost firms and thereby caused concentration to increase.

\(^1\)Entropy in used in information theory a measure of disorder. In the field of economics, entropy is conveniently translated as a measure of the concentration of firms in an industry. Sawyer (1981, pp. 29) explains that this use of entropy in economics is justified along the lines that an industry will be more competitive the greater the uncertainty as to which of a given number of firms will secure the business of a buyer chosen at random, and entropy is a measure of this uncertainty. Note that a rise in entropy indicates an increase in competitiveness and hence a decrease in concentration. Brockett (1991), Brockett and Song (1995), and references therein provide examples of other applications of information theory to actuarial science.
Mayers and Smith (1988) use geographical concentration, line-of-business concentration, and specialization as indicators of success for the ownership type structure (Lloyds, common stock, mutual, reciprocal) of the insurance companies. Mayers and Smith find stock companies are less concentrated geographically than the other three forms of ownership. When not controlling for size, mutual companies were the least concentrated by line of business. When controlling for size, however, reciprocals were the most concentrated.

The third significant paper, by Cummins and Weiss (1991), assesses concentration for the personal and commercial categories for four, ten, and 50 largest firms. They discover that for the personal lines in 1989 the top four firms controlled 43.2 percent and 41.8 percent of private passenger auto liability and private passenger auto physical damage, respectively. For homeowners, 39.5 percent was controlled by the top four firms. On the commercial side, 26.7 percent of workers' compensation was controlled by the top four firms. Cummins and Weiss echo the assessment of Joskow that concentration in some insurance lines results from the efficiency advantage some companies have in dealing with clients and from gains in market share that accompany lower prices for insureds.

3 Aim and Purpose

Most studies on concentration of property insurance have been based on a product line category or on ownership category, the two important classifications according to Gart (1984). Due to the two interrelated categories, the noticeable reorganization of the insurance industry where there is increased interest in financial services, and general corporate mergers, many insurance companies operate on many lines and thus have become conglomerates. Specifically, mergers among competing firms, who already occupy substantial positions as pointed out by Manne (1965), are viewed with suspicion.

Utton (1970) explains that where overall (aggregate) concentration exists, it is most likely that some individual markets will be highly concentrated. There are also arguments that stress the significance of aggregate concentration. Among these arguments is the notion that when a large proportion of economic activity is held by a relatively few firms, it constitutes a threat to democratic government directly through pressure groups and indirectly through advertising. A follow-up to this notion is the concern that basic policy decisions such as future investment, price, and product policies (which are functions associated with
entrepreneurship) are made by a small number of individuals, perhaps one or two members of the board of directors. Furthermore, large diversified firms can affect the market conduct even though their relative shares do not constitute a monopoly. If one large firm has a stronger position in a specific product line and another firm has a stronger position in another line, it is unlikely that any of the firms compete in the market where it has the advantage for fear of retaliation in the market where its position is not strongly established. The focus of attention, therefore, is the overall concentration of economic power controlled by a small number of large firms, typically the largest 100 or 200 enterprises.

The purpose of this paper is to quantify the effects of reorganization and mergers of recent years on aggregate concentration. Such an assessment will add further insight into the structure of the property-casualty insurance industry with data from Best's Insurance Reports on the largest 200 American property casualty companies. Similar studies concerned with aggregate concentration in the industrial sector have been conducted using the Fortune 500, a popular source of data for such studies. Notable among these are the works by Hexter and Snow (1970), Nissan and Caveny (1985, 1988), Attaran and Saghafi (1988), Saghafi and Attaran (1990), and Deutsch and Silber (1995). The data supplied by Best's Insurance Reports since 1985, whereby the 200 largest American property-casualty companies are ranked, provide a similar opportunity for measuring aggregate concentration for the insurance industry.

4 Measurement

Using the data on shares of premiums reported in Best's Insurance Reports for the largest 200 firms or groups in 1985, 1989, and 1993, Theil's entropy is used to quantify the degree of concentration. The three periods 1985, 1989, and 1993 are spaced in time to show if any significant changes occurred. Between 1985 and 1989 approximately 250 mergers and acquisitions occurred. Between 1990 and 1993 approximately 300 occurred. The question is whether these mergers have resulted in an increase in concentration.

Theil's (1967) entropy, $E$, is defined as

$$E = - \sum_{i=1}^{n} p_i \log p_i, \quad 0 \leq E \leq \log n$$

where $p_i \geq 0$ is the $i$-th firm's proportional share of premiums; $n$ is the number of firms, and $\sum_i p_i = 1$. 
If all \( n \) firms have an equal share, then \( E = \log n \), and concentration is at a minimum, in contrast to \( E = 0 \) when one firm controls all shares. Therefore, a decline in \( E \) corresponds to an increase in concentration. For a given level of entropy \( E^* \), the numbers equivalent, \( (n^*) \), is the number of equally sized firms it would take to produce the same level of entropy \( E^* \), i.e.,

\[
n^* = e^{E^*}.
\]

(2)

5 Empirical Results

The largest 200 property-casualty insurance companies or groups of companies accounted in 1993 for 73 percent of net premiums written ($189 billion of $259 billion), 78 percent of admitted assets ($527 billion of $672 billion), and 84 percent of holders' surplus ($153 billion of $182 billion). These huge sums indicate that this comparably small number of firms held a significant control of the market.

The three panels of Table 1 report for 1985, 1989, and 1993 the total net premiums, the mean, the standard deviation, the minimum, the maximum, and the coefficient of variation. The information in Table 1 is supplied for the largest 200 companies as well as by smaller sets of four groups of companies 001-020; 021-050; 051-100; 101-200. Total net premiums written by the 200 companies increased from approximately $117 billion in 1985 to $189 billion in 1993, an increase of 62 percent. The largest 20 companies accounted on average for 50 percent of total premiums written throughout the period, followed by the next 30 companies of lesser rank accounting for approximately 20 percent. The lesser ranked sets of 50 companies and 100 companies, ranked 51 to 100 and 101 to 200, accounted for approximately 16 percent and 15 percent, respectively.

In 1993 the average net premiums written for all the 200 companies reached almost $1 billion, with the largest 20 companies writing on average $4.7 billion. The smallest company among the 200 in 1993 wrote $200 million worth of premiums, while the largest wrote well over $22 billion. For all 200 companies the coefficient of variation shows an increase from 1.81 in 1985 to 2.01 in 1989 to 2.18 in 1993. For the respective three periods the most noticeable increase in the coefficient of variation occurred for the top 20 companies, moving from 0.80 to 0.94 to 1.12. The coefficient of variation for the other groups remained virtually the same throughout the three periods.

Table 2 shows the results for the computation of the concentration index, the Theil's entropy \( E \) of equation (1). There are slight changes


Table 1
Summary Information of Net Premiums Written of Largest Property-Casualty Companies

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Companies</td>
<td>Total</td>
<td>Percent</td>
<td>Mean</td>
<td></td>
</tr>
<tr>
<td>All 200</td>
<td>117,103</td>
<td>100</td>
<td>586</td>
<td>1,058</td>
</tr>
<tr>
<td>001-020</td>
<td>56,840</td>
<td>49</td>
<td>2,842</td>
<td>2,281</td>
</tr>
<tr>
<td>021-050</td>
<td>23,460</td>
<td>20</td>
<td>782</td>
<td>183</td>
</tr>
<tr>
<td>051-100</td>
<td>19,005</td>
<td>16</td>
<td>380</td>
<td>79</td>
</tr>
<tr>
<td>101-200</td>
<td>17,798</td>
<td>15</td>
<td>178</td>
<td>46</td>
</tr>
<tr>
<td>All 200</td>
<td>166,645</td>
<td>100</td>
<td>833</td>
<td>1,677</td>
</tr>
<tr>
<td>001-020</td>
<td>83,380</td>
<td>50</td>
<td>4,169</td>
<td>3,920</td>
</tr>
<tr>
<td>021-050</td>
<td>32,706</td>
<td>20</td>
<td>1,090</td>
<td>272</td>
</tr>
<tr>
<td>051-100</td>
<td>27,010</td>
<td>16</td>
<td>540</td>
<td>118</td>
</tr>
<tr>
<td>101-200</td>
<td>23,549</td>
<td>14</td>
<td>236</td>
<td>54</td>
</tr>
<tr>
<td>All 200</td>
<td>189,030</td>
<td>100</td>
<td>945</td>
<td>2,060</td>
</tr>
<tr>
<td>001-020</td>
<td>93,100</td>
<td>49</td>
<td>4,655</td>
<td>5,202</td>
</tr>
<tr>
<td>021-050</td>
<td>36,735</td>
<td>20</td>
<td>1,225</td>
<td>264</td>
</tr>
<tr>
<td>051-100</td>
<td>30,730</td>
<td>16</td>
<td>615</td>
<td>139</td>
</tr>
<tr>
<td>101-200</td>
<td>28,465</td>
<td>15</td>
<td>285</td>
<td>60</td>
</tr>
</tbody>
</table>

Source: Best's Insurance Reports: Property-Casualty United States (Oldwick, New Jersey: A.M. Best Company, 1985, 1989, and 1993) and calculations by the author. Total, mean, standard deviation, minimum, and maximum are in $ millions.

in the magnitudes of $E$ over time among the 200 companies, as well as the smaller subsets of 20, 30, 50, and 100 companies. This is also obvious from the numbers equivalent $n^*$ of equation (2). The numbers equivalent for all the 200 firms was reduced from 99 in 1985 to 92 in 1989 to 91 in 1993. For the top 20 the sequence is 16, 15, 14. Hardly any change is visible for the lower ranked companies.

Next we need to ascertain whether these apparent differences in entropy over time are statistically significant. Let $E_{ij}$ denote the entropy associated with the proportion of premiums ($p_{ij}$) written for firm $i$ in time period $j$, be denoted by

$$E_{ij} = -p_{ij} \log p_{ij}$$

(3)
Table 2  
Concentration of Premiums Written  
of Largest Property-Casualty Companies

<table>
<thead>
<tr>
<th>Companies</th>
<th>Theil's Entropy (E)</th>
<th>Standard Deviation</th>
<th>Numbers Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: 1985</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All 200</td>
<td>1.9976</td>
<td>.0112</td>
<td>99</td>
</tr>
<tr>
<td>001-020</td>
<td>1.2022</td>
<td>.0265</td>
<td>16</td>
</tr>
<tr>
<td>021-050</td>
<td>1.4661</td>
<td>.0079</td>
<td>29</td>
</tr>
<tr>
<td>051-100</td>
<td>1.6901</td>
<td>.0052</td>
<td>49</td>
</tr>
<tr>
<td>101-200</td>
<td>1.9859</td>
<td>.0040</td>
<td>97</td>
</tr>
<tr>
<td>Panel B: 1989</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All 200</td>
<td>1.9661</td>
<td>.0119</td>
<td>92</td>
</tr>
<tr>
<td>001-020</td>
<td>1.1720</td>
<td>.0293</td>
<td>15</td>
</tr>
<tr>
<td>021-050</td>
<td>1.4646</td>
<td>.0084</td>
<td>29</td>
</tr>
<tr>
<td>051-100</td>
<td>1.6890</td>
<td>.0055</td>
<td>49</td>
</tr>
<tr>
<td>101-200</td>
<td>1.9887</td>
<td>.0036</td>
<td>97</td>
</tr>
<tr>
<td>Panel C: 1993</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All 200</td>
<td>1.9594</td>
<td>.0119</td>
<td>91</td>
</tr>
<tr>
<td>001-020</td>
<td>1.1333</td>
<td>.0317</td>
<td>14</td>
</tr>
<tr>
<td>021-050</td>
<td>1.4678</td>
<td>.0073</td>
<td>29</td>
</tr>
<tr>
<td>051-100</td>
<td>1.6881</td>
<td>.0057</td>
<td>49</td>
</tr>
<tr>
<td>101-200</td>
<td>1.9908</td>
<td>.0033</td>
<td>98</td>
</tr>
</tbody>
</table>

Source: Best’s Insurance Reports: Property-Casualty United States (Oldwick, New Jersey: A.M. Best Company, 1985, 1989, and 1993) and calculations from equation (1).

for \( i = 1, 2, \ldots, n, \ j = 1985, 1986, \ldots, 1993 \) and \( \sum_{i=1}^{n} p_{ij} = 1 \). From equation (1), it is obvious that

\[
E_j = \sum_{i=1}^{n} E_{ij} = nE_j.
\]

To test the hypothesis of equality of a pair of total entropies \( E_j \) and \( E_k \), for \( j, k = 1985, 1989 \) and 1993, the appropriate test statistic (assuming that a large sample approximation is appropriate) is

\[
Z = \frac{E_j - E_k}{\sqrt{n(S_j^2 + S_k^2)}} \tag{4}
\]

where

\[
S_j^2 = \frac{\sum_i E_{ij}^2 - E_j^2}{n - 1}.
\]
Under the null hypothesis, the statistic $Z$ has a standard normal distribution.

The results indicate that differences in total entropy are not statistically significant at $\alpha = 0.05$, in which case $|Z| < 1.96$. None of the computed $|Z|$ values for the groups of companies 20, 30, 50, and 100 exceeds 1.96. The important conclusion from these results is that the levels of concentration among these groups remained virtually the same throughout the period 1985-1993. In other words, no substantial shares of premiums written were transferred from one company to another.

The picture looks a bit different when the comparisons are made between the groups at each period. Note in Table 1 that $E$ increases in magnitude, indicating a decrease in concentration as one moves down the hierarchy from the top 20 companies to the bottom 100 companies for every period, thus pointing to the existence of larger concentration among the top 20 than among the next 30. In turn, there is more concentration among this group of 30 than among the smaller group of 50 companies, which has higher concentration than the next group of the smaller 100 companies. The patterns, however, remain the same for every period. Concentration does exist, especially among the largest firms, yet the level of concentration has remained stable.

The reorganization and mergers of recent years have not resulted in a perceptible increase in concentration in the property-liability insurance industry, unlike what has happened in other services such as retail trade, electric and utilities, and the transportation sectors, as shown by O'Neill (1996). These services experienced large increases in concentration in recent years.

6 Summary

This paper focuses on measuring aggregate concentration using as units net premiums written of the 200 largest property casualty companies, an important sector in the U.S. economy. Theil's entropy index is employed for the period 1985 to 1993. The index is not sensitive to measuring an increase in concentration among the 200 companies or by groups of 20, 30, 50, and 100 companies. Concentration between these groups of companies remained stable for every period under consideration. During the period under consideration which was marked by a substantial activity of mergers and takeovers the property-liability insurance industry cannot be accused of increasing its overall economic power in spite of the large number of mergers.
The findings that no perceptible increase is detected in aggregate concentration in the property-liability insurance industry do not preclude the possibility that some lines of insurance have become concentrated as a result of recent mergers and acquisitions. Tests at the aggregate level may mask increasing trends in concentration on a by-line basis. Past studies were conducted using data prior to 1990. Because mergers and acquisitions have been relatively high in the past five years, documenting changes in by-line concentration since 1989 would be useful.

The debate whether industry concentration is due to growth of efficient firms that manage to maintain low cost operations through economies of scale or whether concentration is due to collusion and suppression of competition continues. In the mean time, efforts must be made to provide empirical evidence as to whether concentration exists and, if so, whether it is increasing or decreasing over time.

References


Nissan: Concentration in American Property-Casualty Companies


Bias of Excluding High and Low Data for Long-Tailed Distributions

Cheng-Sheng Peter Wu*

Abstract

Property and casualty actuaries frequently employ a technique of averaging (called high-low averages) that excludes the same amount of data at both ends. For example, (i) in selecting loss development factors, the middle three of the latest five years or the middle eight of latest 12 quarters sometimes are used, or (ii) in calculating average expense ratios, the largest expense ratios and the smallest expense ratios may be removed from the sample. Although high-low averages can reduce the impact of influential data on analyzed results, the averages will result in downward bias when they are applied to pricing or reserving data that exhibit a long-tailed property. We derive the bias for two commonly used distributions: lognormal and Pareto. An example is provided using chain-ladder reserving where loss development factors are assumed to be lognormally distributed.

Key words and phrases: lognormal distribution, Pareto distribution, percentile, influential data, long-tailed distributions, chain-ladder

1 High-Low Percentile Average

Property and casualty (P&C) actuaries often encounter a wide variety of large loss risks. These large loss risks sometimes result in pricing

*Cheng-Sheng Peter Wu, F.C.A.S., A.S.A., M.A.A.A., is a manager with Deloitte & Touche LLP. He received his masters degrees in chemical engineering and in statistics from the Pennsylvania State University. Mr. Wu has published several papers in automotive engineering, tribology (lubrication engineering), statistics, and actuarial science.

Mr. Wu's address is: Deloitte & Touche LLP, 1000 Wilshire Blvd., Los Angeles, CA 90017, U.S.A. Internet address: pwu@dttus.com
or reserving data that have disproportional influence on analyzed results. Long-tailed distributions, such as lognormal or Pareto distributions, have been used to describe the insurance data with the large loss risks.

One remedy to minimize the effect of influential data (i.e., large losses) on analyzed results is to exclude only the influential data. This requires a great deal of caution, however. According to Neter, Wasserman, and Kutner (1989):

... an outlying influential case should not be automatically discarded, because it may be entirely correct and simply represents an unlikely event. Discarding such an outlying case could lead to the undesirable consequences of increased variances of some of the estimated regression coefficients.

P&C actuaries frequently employ a technique of averaging that excludes the same amount of data at both ends. This averaging technique will be called high-low averaging in this paper. High-low averages, such as the averages of the middle three among five years, middle eight among ten quarters, etc., can be used in many different situations including the calculation of loss development factors, to select underwriting expense ratios or to determine loss adjustment expense ratios. This paper, however, focuses on the particular type of high-low averaging called high-low percentile averaging where the data are sampled from a single distribution and the average excludes data lying outside a specified lower and upper pair of percentile points.

Applying high-low percentile averages to data sampled from a long-tailed distribution results in a systematic downward bias. The downward bias is the percentage difference between the mean and the conditional mean given that the data lie between a specified lower and upper pair of percentile points, i.e.,

\[
\text{Downward Bias} = \frac{E_p[X] - E[X]}{E[X]} \times 100\%
\] (1)

with

\[
E[X] = \text{Expected value for random variable } X;
\]

\[
E_p[X] = \text{Expected value of } X \text{ given that } X \text{ lies between its upper and lower } p \text{ percentile points, i.e.,}
\]

\[
= \frac{1}{1 - 2p} \int_{x_1(p)}^{x_2(p)} x \, dF(x)
\]
where $F(x)$ is cumulative probability function (cdf) of $X$, and $x_2(p)$ and $x_1(p)$ are the upper and lower percentile points respectively of $F(x)$, i.e.,

$$F(x_2(p)) = 1 - p \quad F(x_1(p)) = p, \quad x_1(p) \leq x_2(p).$$

Once the downward bias is calculated, the sample high-low average can be modified as follows:

Modified Sample High-Low Average = \frac{\text{Sample High-Low Average}}{1 + \text{Downward Bias}}. \quad (2)

Figure 1 illustrates such downward bias when the upper and the lower deciles ($p = 0.10$) of the lognormal distribution are excluded. The more data excluded or the more skewed the distribution, the higher the downward bias is. Results in Figure 1 can be extended to high-low averages used by P&C actuaries. For example, a middle 8-of-10 average also excludes the upper and lower 10 percent of the data. The only difference is that the high-low average is based on the sample or empirical cumulative distribution function, while Figure 1 is based on the theoretical cumulative distribution function.

In Section 2, the downward bias is derived using the theoretical cumulative distribution function for two commonly used long-tailed distributions: lognormal and Pareto. The downward bias when the sample (empirical) cumulative distribution function is used is not easily derived. Section 3 provides a chain-ladder reserving example in which high-low percentile averages are used to select loss development factors, and the lognormal distribution is assumed for the loss development factors.

2 Downward Bias for the Lognormal and Pareto

In this section two long-tailed distributions, the lognormal and the Pareto, are used for illustration. Other long-tailed distributions should follow the general results given for these two distributions. Many of the results in this section can be found in Hogg and Klugman (1984) or other statistical texts.

2.1 Lognormal Distribution

If $Z$ is a standard normal distribution, then $X$ has a lognormal distribution if and only if

$$\ln X = \mu + \sigma Z, \quad \sigma > 0.$$
Figure 1: High-Low Average for Lognormal Distribution
(μ = 7.0, σ = 0.7, p = 0.1)

- Lower Decile of the Distribution
- E(X) = 1257, Excluding Upper and Lower Deciles of the Distribution
- E(X) = 1510
Let $\Phi$ be the cumulative distribution of the standard normal distribution, i.e.,

$$\Phi(x) = \int_{-\infty}^{x} \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy,$$

it follows that the theoretical cumulative distribution function of $X$ is

$$F(x) = \Phi(\frac{\ln x - \mu}{\sigma}).$$

In addition, we have

$$F(x_1(p)) = p \Rightarrow \ln x_1(p) = \mu + \sigma \Phi^{-1}(p), \quad \text{and}$$

$$F(x_2(p)) = 1 - p \Rightarrow \ln x_2(p) = \mu + \sigma \Phi^{-1}(1 - p).$$

It can be easily proved that

$$E[X] = \exp\{\mu + \frac{1}{2} \sigma^2\}, \quad \text{and}$$

$$E_p[X] = \frac{E[X]}{(1 - 2p)} \left[ \Phi\left(\frac{\ln x_2(p) - \mu - \sigma^2}{\sigma}\right) - \Phi\left(\frac{\ln x_1(p) - \mu - \sigma^2}{\sigma}\right) \right].$$

This gives the downward bias as:

$$\text{Bias} = \frac{1}{(1 - 2p)} \left[ \Phi(\Phi^{-1}(1 - p) - \sigma) - \Phi(\Phi^{-1}(p) - \sigma) \right] - 1. \quad (3)$$

The above result indicates that the degree of bias only depends on $p$, the percentage of data being excluded, and $\sigma$, the shape factor. The bias does not depend on $\mu$, the location parameter. Table 1 above shows the bias for several combinations of $\sigma$ and $p$. 

### Table 1

<table>
<thead>
<tr>
<th>$p$</th>
<th>$\sigma = 0.2$</th>
<th>$\sigma = 1.0$</th>
<th>$\sigma = 1.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>-0.8%</td>
<td>-18.2%</td>
<td>-38.1%</td>
</tr>
<tr>
<td>0.10</td>
<td>-1.1%</td>
<td>-25.1%</td>
<td>-48.6%</td>
</tr>
<tr>
<td>0.20</td>
<td>-1.6%</td>
<td>-32.6%</td>
<td>-59.1%</td>
</tr>
</tbody>
</table>
Table 2  

<table>
<thead>
<tr>
<th>$p$</th>
<th>$\alpha = 3$</th>
<th>$\alpha = 10$</th>
<th>$\alpha = 50$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>-23.1%</td>
<td>-14.0%</td>
<td>-11.7%</td>
</tr>
<tr>
<td>0.10</td>
<td>-31.2%</td>
<td>-20.4%</td>
<td>-17.6%</td>
</tr>
<tr>
<td>0.20</td>
<td>-40.1%</td>
<td>-28.1%</td>
<td>-24.7%</td>
</tr>
</tbody>
</table>

2.2 Pareto Distribution

Here the theoretical cumulative distribution function is more tractable:

$$F(x) = 1 - \left( \frac{\lambda}{\lambda + x} \right)^\alpha$$

where $\lambda > 0$ and $\alpha > 1$ are the location and shape parameters. The upper and lower percentile points are easily seen to be given by:

$$F(x_1(p)) = p \Rightarrow x_1(p) = \lambda \times \left( (1 - p)^{-1/\alpha} - 1 \right), \quad \text{and}$$
$$F(x_2(p)) = 1 - p \Rightarrow x_2(p) = \lambda \times \left( p^{-1/\alpha} - 1 \right).$$

In addition,

$$E[X] = \frac{\lambda}{(\alpha - 1)} \quad \text{and}$$
$$E_p[X] = \frac{E[X]}{(1 - 2p)} \left[ \alpha((1 - p)^{\beta} - p^{\beta}) - (\alpha - 1)(1 - 2p) \right].$$

where $\beta = (\alpha - 1)/\alpha$. Finally, the downward bias is:

$$\text{Bias} = \frac{\alpha}{(1 - 2p)} \left[ (1 - p)^{\beta} - p^{\beta} - (1 - 2p) \right]. \quad (4)$$

The degree of bias for Pareto distribution depends only on the percentage of excluded data ($p$) and the shape factor ($\alpha$), but not on the location parameter ($\lambda$). Table 2 above shows the bias for different combinations of $\alpha$ and $p$.

3 A Case Study: Chain-Ladder Reserving Example

The chain-ladder technique is a loss development technique that assumes that past loss development patterns reflect future loss developments. For more information regarding loss development techniques, see Wiser (1990, Chapter 2).
Let $L_{i,j}$ denote the loss for accident year $i$ developed to age $j$ and $D_{j,j+1}$ denote the age-to-age development factor from development age $j$ to $j + 1$. The estimate of the ultimate loss for year $i$ developed to age $j$, $UL_i(j)$, is:

$$UL_i(j) = L_{i,j} \times UD_i(j)$$  \hspace{1cm} (5)

where $UD_i(j)$ is the age-to-ultimate development factor for year $i$ developed to age $j$ and

$$UD_i(j) = \prod_{k=0}^{\infty} D_{j+k,j+k+1}. \hspace{1cm} (6)$$

Table 3 shows an automobile bodily injury loss triangle and the associated age-to-age development factor triangle. These data were introduced by Zehnwirth (1989) and analyzed by Kelly (1992), and others.

Table 3 also shows three types of averages for selecting the age-to-age and age-to-ultimate development factors: all year straight averages, all year averages excluding one high and one low data, and all year averages excluding two high data and two low data. These averages are factor averages, not volume-weighted averages. When the number of data points for older accident years is not enough to calculate one high-one low (two high-two low) averages, straight averages (one high-one low averages) will be used.

Results in Table 3 show that the straight averages result in the highest estimates, while the two high-two low averages result in the lowest estimates. This suggests that age-to-age loss development factors may have a long-tailed property.

The possibility that age-to-age development factors may have a long tail has not gone without notice. Hayne's study (1986) assumes in quantifying variability of loss reserves, that the age-to-age development factors are lognormally distributed:

$$\ln(D_{j,j+1}) \sim N(\mu_j, \sigma_j^2)$$

where $\mu_j$ and $\sigma_j^2$ are the mean and variance of the normal distribution.

The main advantage of assuming lognormal distributions for age-to-age development factors is that the age-to-ultimate factors and, consequently, the ultimate loss estimates are also lognormally distributed, i.e.,

$$\ln UD_i(j) \sim N(\sum_{k=j}^{\infty} \mu_k, \sum_{k=j}^{\infty} \sigma_k^2).$$
<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Losses:</th>
<th>Development Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1972</td>
<td>$428,753</td>
<td>$1,399,393</td>
</tr>
<tr>
<td>1973</td>
<td>$458,252</td>
<td>$1,447,324</td>
</tr>
<tr>
<td>1974</td>
<td>$355,229</td>
<td>$1,304,036</td>
</tr>
<tr>
<td>1975</td>
<td>$282,419</td>
<td>$970,751</td>
</tr>
<tr>
<td>1976</td>
<td>$267,600</td>
<td>$1,312,390</td>
</tr>
<tr>
<td>1977</td>
<td>$560,307</td>
<td>$1,500,309</td>
</tr>
<tr>
<td>1978</td>
<td>$360,171</td>
<td>$1,371,944</td>
</tr>
<tr>
<td>1979</td>
<td>$445,545</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Age-to-Age Factors:</th>
<th>Development Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-2</td>
<td>2-3</td>
</tr>
<tr>
<td>1971</td>
<td>3.7759</td>
<td>1.5949</td>
</tr>
<tr>
<td>1972</td>
<td>3.2639</td>
<td>1.6831</td>
</tr>
<tr>
<td>1973</td>
<td>3.1584</td>
<td>1.9795</td>
</tr>
<tr>
<td>1974</td>
<td>3.6710</td>
<td>1.9915</td>
</tr>
<tr>
<td>1975</td>
<td>3.4373</td>
<td>2.1937</td>
</tr>
<tr>
<td>1976</td>
<td>4.9043</td>
<td>1.9269</td>
</tr>
<tr>
<td>1977</td>
<td>2.6777</td>
<td>1.7904</td>
</tr>
<tr>
<td>1978</td>
<td>3.8091</td>
<td></td>
</tr>
<tr>
<td>1979</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3 (continued)
Loss and Loss Development Factor Triangles

<table>
<thead>
<tr>
<th>Development Age</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Age-to-Age Development Factors:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Year Average - Straight Average</td>
<td>3.5872</td>
<td>1.8800</td>
<td>1.3427</td>
<td>1.1827</td>
<td>1.0747</td>
<td>1.0411</td>
<td>1.0374</td>
<td>1.0123</td>
<td></td>
</tr>
<tr>
<td>All Year Average - 1 High and 1 Low</td>
<td>3.5192</td>
<td>1.8743</td>
<td>1.3442</td>
<td>1.1783</td>
<td>1.0649</td>
<td>1.0396</td>
<td>1.0374</td>
<td>1.0123</td>
<td></td>
</tr>
<tr>
<td>All Year Average - 2 High and 2 Low</td>
<td>3.5370</td>
<td>1.8989</td>
<td>1.3322</td>
<td>1.1636</td>
<td>1.0649</td>
<td>1.0396</td>
<td>1.0374</td>
<td>1.0123</td>
<td></td>
</tr>
</tbody>
</table>

Age-to-Ultimate Development Factors:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Year Average - Straight Average</td>
<td>12.5843</td>
<td>3.5081</td>
<td>1.8660</td>
<td>1.3897</td>
<td>1.1751</td>
<td>1.0934</td>
<td>1.0502</td>
<td>1.0123</td>
<td>1.0000</td>
</tr>
<tr>
<td>All Year Average - 1 High and 1 Low</td>
<td>12.1459</td>
<td>3.4513</td>
<td>1.8414</td>
<td>1.3699</td>
<td>1.1626</td>
<td>1.0918</td>
<td>1.0502</td>
<td>1.0123</td>
<td>1.0000</td>
</tr>
<tr>
<td>All Year Average - 2 High and 2 Low</td>
<td>12.1050</td>
<td>3.4224</td>
<td>1.8023</td>
<td>1.3529</td>
<td>1.1626</td>
<td>1.0918</td>
<td>1.0502</td>
<td>1.0123</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Notes: 1. These are automobile bodily injury data that were studied by Kelly (1992).
2. No tail development is assumed.
3. If the number of data points is not enough, straight averages (1 high - 1 low) will be used for 1 high - 1 low (2 high - 2 low).
Kelly (1992) and McNichols (1992) also conclude that a lognormal assumption is better in describing age-to-age development factors than a normal assumption because lognormal distributions can take only positive values and their long-tailed property reflects no upper boundary for the development factors.

Fitting lognormal distributions to the age-to-age development factors in Table 3 produces the parameter estimates in Table 4. First, $\mu_j$ and $\sigma_j^2$ are calculated for each development age. These parameters are then summed from that age to ultimate to obtain age-to-ultimate development factors for a development age. No tail development is assumed in Tables 3 or 4.

There are noted differences between the sample variances given in Table 4 and the sample variances given by Kelly (1992). In Table 4 the sample variances are equal to the sum of squares divided by a factor of $n-1$ (to yield the traditional unbiased estimate of the sample variance), while Kelly divides the sum of squares by $n$ (to yield the maximum likelihood estimate of the sample variance).

Given these lognormal parameter estimates, the high-low averages in Table 3 can be modified to correct the downward bias for the averages. The modified averages are given in Table 5. For example, the one high-one low all-year average (middle six of eight) for the one year-to-two year development factor excludes the upper and the lower 12.5 percent of the sample data. According to the results given in Section 2.1, with $P = 12.5\%$, $\mu_1 = 1.2636$, and $\sigma_1^2 = 0.2155$, a bias of -0.97 percent is indicated for the lognormal assumption.

Table 5 shows the indicated biases for each development age and the modified high-low averages. Table 6 compares the estimated ultimate losses and reserves among the straight averages, the high-low averages, and the modified high-low averages.

Three issues for dealing with limited volume data should be noted. First, for limited volume data additional parameter variation is introduced because sample parameters are assumed for true parameters. Second, this is a bootstrap procedure because excluded data are used to calculate the sample parameters, which in turn are used to calculate the degree of bias to modify the high-low averages. Last, even though the true parameters are known, the indicated bias when sample size is small will not be the same as the indicated bias when sample size is large. In the chain-ladder reserving example, the indicated bias is -0.97 percent for the one high-one low all-year average (middle six of eight) for the one year-to-two year development factor.
<table>
<thead>
<tr>
<th>Accident Year</th>
<th>1-2</th>
<th>2-3</th>
<th>3-4</th>
<th>4-5</th>
<th>5-6</th>
<th>6-7</th>
<th>7-8</th>
<th>8-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>1.3286</td>
<td>0.4668</td>
<td>0.1943</td>
<td>0.1515</td>
<td>0.0439</td>
<td>0.0288</td>
<td>0.0109</td>
<td>0.0122</td>
</tr>
<tr>
<td>1972</td>
<td>1.1829</td>
<td>0.5206</td>
<td>0.3820</td>
<td>0.1378</td>
<td>0.1170</td>
<td>0.0531</td>
<td>0.0619</td>
<td></td>
</tr>
<tr>
<td>1973</td>
<td>1.1501</td>
<td>0.6828</td>
<td>0.2872</td>
<td>0.2076</td>
<td>0.0576</td>
<td>0.0388</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1974</td>
<td>1.3005</td>
<td>0.6889</td>
<td>0.2531</td>
<td>0.1515</td>
<td>0.0681</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>1.2347</td>
<td>0.7856</td>
<td>0.3536</td>
<td>0.1887</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1976</td>
<td>1.5901</td>
<td>0.6559</td>
<td>0.2864</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1977</td>
<td>0.9849</td>
<td>0.5825</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1978</td>
<td>1.3374</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Age-to-Age Development Factors:**

<table>
<thead>
<tr>
<th></th>
<th>Lognormal Mean - All Year Average</th>
<th>Lognormal Variance - All Year Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.2636</td>
<td>0.0308</td>
</tr>
<tr>
<td></td>
<td>0.6262</td>
<td>0.0120</td>
</tr>
<tr>
<td></td>
<td>0.2928</td>
<td>0.0046</td>
</tr>
<tr>
<td></td>
<td>0.1674</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>0.0717</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>0.0403</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>0.0364</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>0.0122</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

**Age-to-Ultimate Development Factors:**

<table>
<thead>
<tr>
<th></th>
<th>Lognormal Mean - All Year Average</th>
<th>Lognormal Variance - All Year Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.5106</td>
<td>0.0507</td>
</tr>
<tr>
<td></td>
<td>1.2470</td>
<td>0.0199</td>
</tr>
<tr>
<td></td>
<td>0.6208</td>
<td>0.0079</td>
</tr>
<tr>
<td></td>
<td>0.3280</td>
<td>0.0033</td>
</tr>
<tr>
<td></td>
<td>0.1606</td>
<td>0.0025</td>
</tr>
<tr>
<td></td>
<td>0.0889</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>0.0486</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>0.0122</td>
<td>0.0000</td>
</tr>
<tr>
<td>Accident Year</td>
<td>Development Age</td>
<td>1-2</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>------</td>
</tr>
<tr>
<td>1971</td>
<td></td>
<td>3.7759</td>
</tr>
<tr>
<td>1972</td>
<td></td>
<td>3.2639</td>
</tr>
<tr>
<td>1973</td>
<td></td>
<td>3.1584</td>
</tr>
<tr>
<td>1974</td>
<td></td>
<td>3.6710</td>
</tr>
<tr>
<td>1975</td>
<td></td>
<td>3.4373</td>
</tr>
<tr>
<td>1976</td>
<td></td>
<td>4.9043</td>
</tr>
<tr>
<td>1977</td>
<td></td>
<td>2.6777</td>
</tr>
<tr>
<td>1978</td>
<td></td>
<td>3.8091</td>
</tr>
</tbody>
</table>

Age-to-Age Development Factors:

| All Year Average - Straight Average | 3.5872 | 1.8800 | 1.3427 | 1.1827 | 1.0747 | 1.0411 | 1.0374 | 1.0123 |

Lognormal Parameters:

| Lognormal Mean - All Year Average | 1.2636 | 0.6262 | 0.2928 | 0.1674 | 0.0717 | 0.0403 | 0.0364 | 0.0122 |
| Lognormal Variance - All Year Average | 0.0308 | 0.0120 | 0.0046 | 0.0009 | 0.0010 | 0.0001 | 0.0013 | 0.0000 |

All Year Average - 1 High and 1 Low

| % of High and Low Data Excluded | 12.5% | 14.3% | 16.7% | 20.0% | 25.0% | 33.3% | 0.0% | 0.0% |
| Downward Bias | -0.97% | -0.40% | -0.17% | -0.03% | -0.04% | -0.01% |
| Modified 1 High and 1 Low | 3.5536 | 1.8819 | 1.3464 | 1.1787 | 1.0654 | 1.0397 | 1.0374 | 1.0123 |
Table 5 (continued)

Modified High-Low Averages for Loss Development Factors

<table>
<thead>
<tr>
<th>Age-to-Age Development Factors:</th>
<th>1-2</th>
<th>2-3</th>
<th>3-4</th>
<th>4-5</th>
<th>5-6</th>
<th>6-7</th>
<th>7-8</th>
<th>8-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Year Average - 2 High and 2 Low</td>
<td>3.5370</td>
<td>1.8989</td>
<td>1.3322</td>
<td>1.1636</td>
<td>1.0649</td>
<td>1.0396</td>
<td>1.0374</td>
<td>1.0123</td>
</tr>
<tr>
<td>p% of Data Excluded</td>
<td>25.0%</td>
<td>28.6%</td>
<td>33.3%</td>
<td>40.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Downward Bias</td>
<td>-1.31%</td>
<td>-0.54%</td>
<td>-0.22%</td>
<td>-0.04%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modified 2 High and 2 Low</td>
<td>3.5840</td>
<td>1.9092</td>
<td>1.3351</td>
<td>1.1641</td>
<td>1.0654</td>
<td>1.0397</td>
<td>1.0374</td>
<td>1.0123</td>
</tr>
</tbody>
</table>

Age-to-Ultimate Development Factors:

<table>
<thead>
<tr>
<th></th>
<th>1-2</th>
<th>2-3</th>
<th>3-4</th>
<th>4-5</th>
<th>5-6</th>
<th>6-7</th>
<th>7-8</th>
<th>8-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Straight Average</td>
<td>12.5843</td>
<td>3.5081</td>
<td>1.8660</td>
<td>1.3897</td>
<td>1.1751</td>
<td>1.0934</td>
<td>1.0502</td>
<td>1.0123</td>
</tr>
<tr>
<td>Modified 1 High and 1 Low</td>
<td>12.3453</td>
<td>3.4740</td>
<td>1.8460</td>
<td>1.3711</td>
<td>1.1632</td>
<td>1.0919</td>
<td>1.0502</td>
<td>1.0123</td>
</tr>
<tr>
<td>Modified 2 High and 2 Low</td>
<td>12.3702</td>
<td>3.4515</td>
<td>1.8079</td>
<td>1.3541</td>
<td>1.1632</td>
<td>1.0919</td>
<td>1.0502</td>
<td>1.0123</td>
</tr>
</tbody>
</table>
### Table 6

Comparison of Ultimate Losses and Reserves Between Different Averaging Techniques

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Undeveloped Losses</th>
<th>Straight Average</th>
<th>1 High &amp; 1 Low</th>
<th>2 High &amp; 2 Low</th>
<th>Modified 1 High &amp; 1 Low</th>
<th>Modified 2 High &amp; 2 Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>$5,327,859</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>1972</td>
<td>$4,995,827</td>
<td>1.0123</td>
<td>1.0123</td>
<td>1.0123</td>
<td>1.0123</td>
<td>1.0123</td>
</tr>
<tr>
<td>1973</td>
<td>$5,175,219</td>
<td>1.0502</td>
<td>1.0502</td>
<td>1.0502</td>
<td>1.0502</td>
<td>1.0502</td>
</tr>
<tr>
<td>1974</td>
<td>$4,166,594</td>
<td>1.0934</td>
<td>1.0918</td>
<td>1.0918</td>
<td>1.0919</td>
<td>1.0919</td>
</tr>
<tr>
<td>1975</td>
<td>$3,662,977</td>
<td>1.1751</td>
<td>1.1626</td>
<td>1.1626</td>
<td>1.1632</td>
<td>1.1632</td>
</tr>
<tr>
<td>1976</td>
<td>$3,367,532</td>
<td>1.3897</td>
<td>1.3699</td>
<td>1.3529</td>
<td>1.3711</td>
<td>1.3541</td>
</tr>
<tr>
<td>1977</td>
<td>$2,686,208</td>
<td>1.8660</td>
<td>1.8414</td>
<td>1.8023</td>
<td>1.8460</td>
<td>1.8079</td>
</tr>
<tr>
<td>Total:</td>
<td>$31,199,705</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ultimate Losses:

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Straight Average</th>
<th>1 High &amp; 1 Low</th>
<th>2 High &amp; 2 Low</th>
<th>Modified 1 High &amp; 1 Low</th>
<th>Modified 2 High &amp; 2 Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>$5,327,859</td>
<td>$5,327,859</td>
<td>$5,327,859</td>
<td>$5,327,859</td>
<td>$5,327,859</td>
</tr>
<tr>
<td>1972</td>
<td>$5,057,365</td>
<td>$5,057,365</td>
<td>$5,057,365</td>
<td>$5,057,365</td>
<td>$5,057,365</td>
</tr>
<tr>
<td>1973</td>
<td>$5,434,955</td>
<td>$5,434,955</td>
<td>$5,434,955</td>
<td>$5,434,955</td>
<td>$5,434,955</td>
</tr>
<tr>
<td>1974</td>
<td>$4,555,688</td>
<td>$4,549,008</td>
<td>$4,549,008</td>
<td>$4,549,326</td>
<td>$4,549,326</td>
</tr>
<tr>
<td>1975</td>
<td>$4,304,274</td>
<td>$4,258,744</td>
<td>$4,258,744</td>
<td>$4,260,887</td>
<td>$4,260,887</td>
</tr>
<tr>
<td>1976</td>
<td>$4,680,012</td>
<td>$4,613,276</td>
<td>$4,555,846</td>
<td>$4,617,162</td>
<td>$4,560,077</td>
</tr>
<tr>
<td>1977</td>
<td>$5,012,583</td>
<td>$4,946,406</td>
<td>$4,841,310</td>
<td>$4,958,844</td>
<td>$4,856,293</td>
</tr>
<tr>
<td>1978</td>
<td>$4,812,972</td>
<td>$4,734,976</td>
<td>$4,695,357</td>
<td>$4,766,110</td>
<td>$4,735,286</td>
</tr>
<tr>
<td>1979</td>
<td>$5,606,883</td>
<td>$5,411,560</td>
<td>$5,393,340</td>
<td>$5,500,387</td>
<td>$5,511,474</td>
</tr>
<tr>
<td>Total:</td>
<td>$44,792,590</td>
<td>$44,334,149</td>
<td>$44,113,782</td>
<td>$44,472,894</td>
<td>$44,293,520</td>
</tr>
<tr>
<td>Accident Year</td>
<td>Straight Average</td>
<td>1 High &amp; 1 Low</td>
<td>2 High &amp; 2 Low</td>
<td>Modified 1 High &amp; 1 Low</td>
<td>Modified 2 High &amp; 2 Low</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>1971</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1972</td>
<td>$61,538</td>
<td>$61,538</td>
<td>$61,538</td>
<td>$61,538</td>
<td>$61,538</td>
</tr>
<tr>
<td>1973</td>
<td>$259,736</td>
<td>$259,736</td>
<td>$259,736</td>
<td>$259,736</td>
<td>$259,736</td>
</tr>
<tr>
<td>1974</td>
<td>$389,094</td>
<td>$382,414</td>
<td>$382,414</td>
<td>$382,732</td>
<td>$382,732</td>
</tr>
<tr>
<td>1975</td>
<td>$641,297</td>
<td>$595,767</td>
<td>$595,767</td>
<td>$597,910</td>
<td>$597,910</td>
</tr>
<tr>
<td>1976</td>
<td>$1,312,480</td>
<td>$1,245,744</td>
<td>$1,188,314</td>
<td>$1,249,630</td>
<td>$1,192,545</td>
</tr>
<tr>
<td>1977</td>
<td>$2,326,375</td>
<td>$2,260,198</td>
<td>$2,155,102</td>
<td>$2,272,636</td>
<td>$2,170,085</td>
</tr>
<tr>
<td>1978</td>
<td>$3,441,028</td>
<td>$3,363,032</td>
<td>$3,323,413</td>
<td>$3,394,166</td>
<td>$3,363,342</td>
</tr>
<tr>
<td>1979</td>
<td>$5,161,338</td>
<td>$4,966,015</td>
<td>$4,947,795</td>
<td>$5,054,842</td>
<td>$5,065,929</td>
</tr>
<tr>
<td>Total:</td>
<td>$13,592,885</td>
<td>$13,134,444</td>
<td>$12,914,077</td>
<td>$13,273,189</td>
<td>$13,093,815</td>
</tr>
</tbody>
</table>
This is the indicated bias when the sample size is large with $p = 12.5\%$, $\mu_1 = 1.2636$, and $\sigma_1^2 = 0.2155$. On the other hand, a further simulation test based on 5,000 replicates indicates that the bias is -0.80 percent when middle six of eight are used for the average. This simulated bias is smaller than the theoretical bias when sample size is large.

4 Summary and Conclusions

For many insurance applications, high-low averages result in lower estimates than straight averages because insurance data exhibit a long-tailed property. This downward bias is shown in a simple reserving example in which high-low averages are used to select loss development factors. These averages, however, can be modified using equation (2).

The analysis and data given in this paper are far from complete. The levels of the downward bias that would most likely exist in practice have not been reviewed. Also, the level of the downward bias for a limited volume of data has not been fully explored. There needs to be more in-depth research in this area.

References


Disclosure and Confidentiality Requirements of Corporate Pension Plan Actuaries

Theodore Konshak*

Abstract

Corporate pension plan actuaries are subject to the standards of the Joint Board for the Enrollment of Actuaries. The Joint Board is empowered to establish such standards under the provisions of the Employee Retirement Income Security Act of 1974, a federal law. In consideration of these statutory standards, this article will discuss whether standards published by professional actuarial organizations have any applicability. The contrast between the disclosure requirements of federal law and the confidentiality standards of the Society of Actuaries will be highlighted.

Key words and phrases: ERISA, actuarial standards, fiduciaries, enrolled actuary, plan auditor

1 Introduction

Defined benefit pension plans promise to pay a monthly income to each participant for the remainder of the participant's lifetime or for the lifetimes of both the participant and his or her spouse. Money is deposited into a trust fund, is invested by the pension plan trustees, and is periodically withdrawn by the plan administrator to pay retirees their monthly benefits. In the United States, the minimum amount to be deposited into the trust fund is based on an annual actuarial valuation performed by an enrolled actuary under the terms of the 1974 Employee Retirement Income Security Act (ERISA).

In addition to providing actuarial services under ERISA, actuarial service providers also can earn income by providing an employer with

*Theodore Konshak is an enrolled actuary with Negotiated Pension Plans, Ltd. Mr. Konshak's address is: Negotiated Pension Plans, Ltd., 1107 Wilson Avenue, Green Bay, WI 54303-4206, USA.
a wide variety of services including, but not limited to, recordkeeping services for defined contribution pension plans and consulting services on the legal requirements of the pension plan. This paper reviews the required disclosures of the enrolled actuary under ERISA and the confidentiality provisions of the Professional Code of Conduct of the Society of Actuaries to determine if the statutory standards of ERISA are being subverted, inadvertently or otherwise, by these professional standards.

2 Enrolled Actuaries and Plan Auditors

Corporate pension plans are classified into two general types based on whether the provisions of the plan define the deposits or the withdrawals:

- In a defined benefit plan, the plan provisions define a series of monthly withdrawals payable during the lifetime of the participant. The minimum deposit under federal law is determined by an actuarial valuation performed by an enrolled actuary.

- In a defined contribution plan, the plan provisions define the deposits that can and will be made. Withdrawals are these deposits and the investment earnings on these deposits. Defined contribution plans do not require the services of an enrolled actuary because the deposit is specified by the terms of the plan.

A pension plan auditor examines the financial statements of the pension plan. In addition to determining whether the money is actually there, the pension plan auditor reviews the statement of where the money has gone. Both defined benefit and defined contribution pension plans with at least 100 participants require the services of an auditor.

Under ERISA administrators of pension plans must engage both an enrolled actuary and pension plan auditor on the behalf of all plan participants. For enrolled actuaries, this is the actual language of Section 103(a)(4)(A) of ERISA.

Engagement on the behalf of all plan participants is a legal requirement to ensure the impartiality of actuarial determinations and the integrity of pension plan audits. The accuracy of audited pension records should not reflect the effort an employer is willing to exert. The minimum deposit required under federal law should not reflect the amount the employer currently is willing to contribute. Impartiality of actuarial determinations and integrity of pension plan audits are duties and obligations of the enrolled actuary and pension plan auditor.
The administrators of pension plans who hire the enrolled actuary and pension plan auditors are fiduciaries under Section 21(A) of ERISA. This means plan administrators must discharge their duties solely in the interest of plan participants (Section 404(a)(1) of ERISA). Fiduciaries are the responsible parties and may be held legally liable for any misconduct.

Enrolled actuaries are accredited and regulated by the Joint Board for the Enrollment of Actuaries, a federal board consisting of three members appointed by the Secretary of the Treasury and two members appointed by the Secretary of Labor. The Joint Board is administered by an executive director appointed by the Secretary of the Treasury.

Section 3042 of ERISA describes the enrollment of actuaries by the Joint Board:

The Joint Board shall, by regulations, establish reasonable standards and qualifications for persons performing actuarial services with respect to plans to which this Act applies and, upon application by any individual, shall enroll such individual if the Joint Board finds that such individual satisfies such standards and qualifications ... The Joint Board may, after notice and an opportunity for a hearing, suspend or terminate the enrollment of an individual under this section if the Joint Board finds that such individual—(1) has failed to discharge his duties under this Act, or (2) does not satisfy the requirements for enrollment as in effect at the time of his enrollment.

The Standards of Performance for Enrolled Actuaries were published by the Joint Board under Title 20, Chapter VIII, Part 901, Subpart C of the Code of Federal Regulations.

The executive director of the Joint Board maintains a roster of all persons whose enrollment to perform actuarial services under ERISA has been suspended or terminated. This roster contains the names of only two suspended enrolled actuaries.

To date the Joint Board has devoted most of its limited resources to the accreditation of enrolled actuaries through examinations and continuing education. Regulation of enrolled actuaries through its standards of performance has not been a priority. These statutory standards of performance may be expanded and be more strictly enforced if the composition of the Joint Board is changed to include representatives of the Pension Benefit Guaranty Corporation (PBGC).1

---

1 The Pension Benefit Guaranty Corporation (PBGC) is the federal agency that insures the payment of benefits from failed pension plans. As in other insolvency insurance
A representative of the PBGC presently can participate in discussions of the Joint Board, but cannot vote on any issues. As the minutes of the March 15, 1995 meeting of the Joint Board indicate, the PBGC has attempted in the past to obtain voting representation on the Joint Board. Unreasonable actuarial determinations of the variable rate premium paid by corporate sponsors to the PBGC could intensify their efforts to obtain such representation.

3 Actuarial Codes of Conduct

3.1 Professional

Actuarial organizations publish codes of professional conduct to govern the relationship between consulting actuary and client/employer. The professional codes of conduct of virtually every actuarial organization in the United States are identical to the Code of the Society of Actuaries. Uniformity is necessary because the Actuarial Board for Counseling and Discipline (ABCD) enforces these actuarial codes of conduct for the professional actuarial organizations sponsoring its activities.

The following is from Precept 10 and Annotation 10-1, respectively, of the Code of Professional Conduct of the Society of Actuaries (1996)²:

An actuary shall not disclose to another party any confidential information obtained through professional services performed for a principal (i.e., client or employer) unless authorized to do so by the client or employer or required to do so by law. “Confidential information” refers to information not in the public domain of which the actuary becomes aware in conjunction with the rendering of professional services to a principal. It may include ... information which the actuary has reason to believe that the principal would not wish to be divulged.

3.2 Criminal

Under Title I, Section 103(d) of ERISA the results of the annual actuarial valuation are to be disclosed on an actuarial statement prepared by pools sponsored by the federal government (e.g., the FDIC for banks and the FSLIC for savings and loan associations), the PBGC is financed by premiums paid by the corporate sponsors of the defined benefit pension plans covered under the program.

the enrolled actuary (i.e., the Schedule B prepared by the enrolled actuary and attached to the annual Form 5500 filed by the plan administrator). Section 103(d) of ERISA also requires the disclosure of information including, but not limited to, a statement of actuarial assumptions and methods used to determine costs and justifications for any change in those actuarial assumptions or methods. Section 103(d)(13) of ERISA requires disclosure of "such other information as may be necessary to fully and fairly disclose the actuarial position of the plan."

Section 1027 of the Criminal Codes states:

> Whoever, in any document required by Title I of the Employee Retirement Income Security Act of 1974 (as amended from time to time) to be published, ... knowingly conceals, covers up, or fails to disclose any fact the disclosure of which is required by such title or is necessary to verify, explain, clarify, or check the accuracy and completeness any report required by such title to be published ... shall be fined not more than $10,000, or imprisoned not more than five years, or both.

ERISA created an enrolled actuary engaged on behalf of all plan participants, imposed duties and obligations under ERISA on that enrolled actuary, and subjected him or her to the standards of the Joint Board for the Enrollment of Actuaries. Professional actuarial organizations also have standards governing the relationship between consulting actuary and client/employer. Confidentiality and full disclosure are examples of when the duties of the enrolled actuary are diametrically opposed to the standards of professional actuarial organizations.

### 3.3 Contrast to Pension Plan Auditor

Under the terms of Section 103(a)(3)(A) of ERISA the pension plan auditor examines the books and records of the pension plan under generally accepted auditing standards. In contrast, actuarial services are not performed under ERISA according to generally accepted actuarial standards. That term does not appear and would not have to appear in ERISA because the Joint Board for the Enrollment of Actuaries establishes the standards for enrolled actuaries.

The statement preceding the enrolled actuary's signature on Schedule B (Form 5500) makes no reference to generally accepted actuarial standards:

> To the best of my knowledge, the information supplied in this schedule and on the accompanying statements, if any, is
complete and accurate and, in my opinion, each assumption used in combination represents my best estimate of anticipated experience under the plan. Furthermore, in the case of a plan other than a multiemployer plan, each assumption used is (a) reasonable (taking into account the experience of the plan and reasonable expectations) or (b) would, in the aggregate, result in a total contribution equivalent to that which would be determined if each such assumption were reasonable. In the case of a multiemployer plan, the assumptions used, in the aggregate, are reasonable (taking into account the experience of the plan and reasonable expectations).

Because actuarial determinations under ERISA are calculated according to standards established by the Joint Board for the Enrollment of Actuaries and not according to the standards of the actuarial profession, enrolled actuaries have no legal basis for citing any actuarial standard under the Code of Professional Conduct of the Society of Actuaries or any other actuarial organization. Auditing standards have legal recognition and standing under ERISA. The standards of professional actuarial organizations do not enjoy such recognition or standing.

3.4 Consequences of Suspension

One of the two suspended enrolled actuaries on the Joint Board roster is currently a trustee of a major actuarial foundation. The other continues to be a consulting pension actuary with a major employee benefits consulting firm and was the discussion leader at a workshop on public employee retirement systems at an annual enrolled actuaries meeting.³

At least in these two cases employment and status within the actuarial profession were not adversely impacted by suspension as an enrolled actuary. Legally the statutory standards of the Joint Board take precedence. In practice, does the lack of detrimental consequences for suspension as an enrolled actuary suggest minimal respect for the statutory standards of the Joint Board?

³Under the separation of powers doctrine, pension plans sponsored by state and local governments (i.e., public employee retirement plans) are exempted from the requirements of ERISA, a federal law, and their actuaries therefore do not have to be enrolled by the Joint Board.
Table 1

<table>
<thead>
<tr>
<th>Type of Change</th>
<th>Number of Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
</tr>
<tr>
<td>Investment Return</td>
<td>8</td>
</tr>
<tr>
<td>Retirement Ages</td>
<td>4</td>
</tr>
<tr>
<td>Salary Increase</td>
<td>7</td>
</tr>
<tr>
<td>Terminations of Employment</td>
<td>7</td>
</tr>
<tr>
<td>Disabilities</td>
<td>2</td>
</tr>
<tr>
<td>Deaths</td>
<td>7</td>
</tr>
<tr>
<td>Payment of Plan Expenses</td>
<td>3</td>
</tr>
<tr>
<td>Funding Method Changes</td>
<td>14</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>57</td>
</tr>
</tbody>
</table>

4 Changes in Actuarial Assumptions & Techniques

4.1 The Study

Section 103(d) of ERISA and the instructions for completing Schedule B require the enrolled actuary to attach a statement of actuarial assumptions and techniques, a statement of the changes in actuarial assumptions and techniques, and a statement justifying those changes in actuarial assumptions and techniques. In a study of a sample of 20 pension plans that changed actuarial consulting firms, Konshak (1995) tabulated the number of actual, reported, and justified changes. These results are reported in Table 1. The asset values of the pension plans sampled range from $207 million to $16 million.

The actual changes in presented in Table 1 are obtained by comparing the statement of actuarial assumptions and techniques attached to the 1993 Schedule B to the similar statement attached in the 1992 Schedule B. Reported changes and justifications are tabulated from the 1993 attachments to Schedule B.

According to the analysis of 1993 and 1992 Schedule B attachments there were 57 changes in actuarial assumptions and techniques required to be reported and justified. Thirty-six of these changes were reported by the enrolled actuary on the current Schedule B attachment. Only 13 of the reported changes (12 changes in actuarial assumptions and one change in actuarial techniques) were justified by the enrolled actuary.

The practice of failing to disclose and justify changes in actuarial assumptions and techniques is either incompetence (enrolled actuaries certify to their familiarity with those portions of ERISA relating directly or indirectly to the responsibilities of an enrolled actuary on the Form 5434-A application for their triennial renewal of enrollment) or a knowledgeable concealment (corporate pension plan actuaries choose not to be governed by these provisions of ERISA). In any event, these individuals have failed to discharge their duties under the terms of ERISA.

4.2 An Example of a Justification

Although most of the justifications for changes in actuarial assumptions and techniques are not immediately verifiable, one justification from the Konshak study will be analyzed for reasonableness.

The assumed rate of future employment terminations was changed and justified by the enrolled actuary solely on the basis of prior plan experience: “The ultimate withdrawal rates were increased by a factor of 10 in order to better reflect actual plan experience.” Enrolled actuaries who change actuarial assumptions and techniques in this study, however, normally would not have a personal and intimate knowledge of the prior plan experience. Pension plans in this example changed actuarial service providers.

The prior experience of the pension plan can be established by studying individual participant records if available from the prior enrolled actuary or corporate sponsor. Corporate sponsors, however, generally are unwilling to incur the substantial expense of performing such a study. The corporate sponsor in this case changed actuarial service providers to reduce its actuarial expenses and provided that as the reason for the change in enrolled actuary on the Schedule C (Form 5500) filed with the federal government: “Change due to receipt of more competitive bid from another actuary firm.”

The change in the assumed rate of future terminations of employment was the only change in actuarial assumptions or techniques made by the new enrolled actuary, and this assumption change significantly decreased the minimum deposit payable under federal law. The expla-
nation for the change in enrolled actuary implies an experience study would not have been performed due to cost considerations and suggests another reason for the change in withdrawal rates: cutting costs for the corporate sponsor.

5 Resolution of Conflicting Values

Pension actuaries believe they can balance their roles as consultants to the employer (subject to the standards of the Code of Professional Conduct) with their duties and obligations as enrolled actuaries (subject to the requirements of ERISA and the standards of the Joint Board). Technically, the previously quoted section of the Society of Actuaries' Code of Professional Conduct relating to confidentiality does not conflict with the statutory requirements of ERISA: "An actuary shall not disclose to another party any confidential information obtained through professional services . . . unless . . . required to do so by law." A written disclaimer in the Code, however, cannot erase a sentiment to maintain confidentiality through superficial rather than full disclosure.

The justification quoted in the previous section for the change in withdrawal rates could have been subjected to a peer review. If an experience study had been performed, a peer reviewer could have avoided the aforementioned challenge to the reasonableness of this justification by suggesting its inclusion: "Based on the results of an experience study, the ultimate withdrawal rates were increased by a factor of 10 in order to better reflect actual plan experience." None of the assumption changes made by the other enrolled actuaries included in my study were justified in terms of an experience study.

If an experience study had not been performed, a peer reviewer could make two criticisms. First, the rate of increase, "by a factor of 10," will draw the attention of the reader and should be eliminated. Second, even though the statement preceding the enrolled actuary's signature on the Schedule B necessitates consideration of both prior and expected experience by requiring reasonable assumptions "taking into account the experience of the plan and reasonable expectations," this assumption change should have been justified solely in terms of the expected future experience of the plan. Drawing attention to the prior plan experience makes a regulatory request for data and a challenge based on that prior plan experience more likely. If the justification were changed according to these recommendations, it would be consistent with the justifications provided by the other enrolled actuaries reviewed by the
Konshak study: "The withdrawal rates were increased to better reflect the expected future experience of the plan."

Of the 12 actuarial assumption changes justified by the enrolled actuaries included in the Konshak study, nine were justified solely on the basis of the expected future experience of the plan. Two of the assumption changes were justified both in terms of the prior and the expected future experience of the plan (consistent with the enrolled actuary's statement on the Schedule B). Only the justification being analyzed was justified solely in terms of prior plan experience. None of the enrolled actuaries supplied evidence or any specific information to support their assertions.

By failing to introduce the uncertainty involved in predicting the future experience of the plan, the justification analyzed in Section 4.2 above lacks the vagueness of the other justifications reviewed by the Konshak study and therefore could be subjected to a challenge on its reasonableness. By maximizing vagueness and minimizing specifics, the other justifications reflect a sentiment to conceal under the confidentiality provisions of the Code of Professional Conduct rather than a sentiment to disclose under the statutory requirements of ERISA.

6 Conclusions

The codes of conduct of professional actuarial organizations and their standards of actuarial practice are irrelevant for the enrolled actuary performing actuarial services under ERISA. Actuarial codes and standards can only confuse the issue and provide opportunities to subvert, inadvertently or otherwise, the intent of statutory standards. Giving any credibility to the confidentiality provisions of any professional actuarial code would be irresponsible and contrary to the disclosure requirement of federal law.

Enforcement from federal agencies is a reasonable and expected result for those pension actuaries who believe their professional codes of silence are above the law. This enforcement to date has been passive and lacking, but the Joint Board may actively search for enrolled actuaries failing to discharge their duties under ERISA when the PBGC has more influence with the Joint Board.
Discussion of Theodore Konshak's "Disclosure and Confidentiality Requirements of Corporate Pension Plan Actuaries"

Richard Daskais*

Mr. Konshak's paper expresses concerns that enrolled actuaries are not properly discharging their duties to pension plan participants in choosing actuarial assumptions. The paper cites an apparent conflict between the disclosure requirements of ERISA and the confidentiality requirements of the actuarial professional organizations' codes of conduct.

I do not believe the examples cited in the paper show that actuaries rely on the codes' confidentiality requirements to avoid disclosure required under ERISA. I believe the short justifications of assumption changes offered in the examples are consistent with the nature of the reporting of assumptions required on Schedule B and its attachments. At worst, the lack of justification (or a justification that is too short) is simply benign negligence on the part of the enrolled actuary; unfortunately, this negligence does not present a favorable view of the actuarial profession.

The paper implies that when plan sponsors change enrolled actuaries, the reason is often to reduce their contributions to pension plans. While this may sometimes be the case, I believe the more important abuses by enrolled actuaries in their choices of assumptions occur when the actuary knows his or her client's wishes and reflects these wishes in his or her choice of assumptions; no change in enrolled actuary is required. The most common abuse, in my opinion, is to choose conservative assumptions that are far from the actuary's best estimate in order to produce large deductible contributions. Conservative assumptions are used for small plans (often simply tax shelters for the principal participant or participants) and for larger plans whose costs are

*Richard Daskais, F.S.A., F.C.A., has been a consulting actuary specializing in pensions since 1957, except for a period from 1985 to 1989 when he was a vice president of Goldman, Sachs & Co. He was a member of the firm Daskais and Walls, Inc. in Chicago from 1966 to 1984.

Mr. Daskais' address is: 1174 Shellburn Lane, Ventura CA 93001-4055, USA.
passed to customers (e.g., plans of regulated utilities and government contractors).

My opinion of larger plans is based on my experience since the early 1970s in two areas: doing projections to analyze asset allocation policy where it was clear that the middle of the range of expected investment returns was much greater than the enrolled actuary's assumed investment return and consulting for government agencies on the level of pension cost reimbursement to contractors. In none of these activities was I or a colleague at my firm the enrolled actuary for the pension plan. Because of confidentiality responsibilities to clients (which may dismay the author), I cannot cite specific examples.

The paper refers to the "lack of detrimental consequences" for actuaries who have been disciplined by the Joint Board; i.e., there was apparently no discipline by actuaries' professional organizations. An important reason for the failure of the professional organizations to discipline is that there are significant disincentives for an individual to institute disciplinary proceedings against an actuary who has violated the organizations' codes of professional conduct. At best, there may be a major investment of time, and at worst the individual may find himself or herself a defendant in a defamation suit. Unfortunately, I have no easy solution to this problem; I am confident that the organizations' relevant committees have considered it.

Brian A. Jones*

Mr. Konshak's paper provides a useful summary of data from a 1995 study of the 5500 Schedule B of 20 pension plans. I do not find the rest of the paper equally useful, and I do not believe the author supports his conclusions. Specifically:¹

The codes of conduct of professional organizations and their standards of actuarial practice are irrelevant for the enrolled actuary performing actuarial services under ERISA. Actuarial codes and standards can only confuse the issue and provide opportunities to subvert, inadvertently or otherwise, the intent of statutory standards.

Not so. The Code of Professional Responsibility (CPR) binds all members of sponsoring organizations, and violation can lead to discipline.

---

*Brian A. Jones is an enrolled actuary and a member of a number of actuarial organizations and of the New York and D.C. Bars.

Mr. Jones's address is: 10 Clinton Street, Brooklyn NY 11201, USA.

¹Unless otherwise stated, all quotations are taken from Mr. Konshak's paper.
To the extent that the CPR would conflict with any applicable law or regulation, it specifically states that "the requirements of law or regulation shall take precedence." There cannot, therefore, be any real conflict between the two. If an observer purports to find such subversion in the text of the CPR, he or she simply has not read or understood it. If an actuary were to cite the CPR to justify such subversion, he or she would be perverting it.

The assertion that the "standards of professional actuarial organizations do not apply to enrolled actuaries" is simply incorrect (unless, of course, it is confined to enrolled actuaries who are not members of the sponsoring organizations). Repeated assertions in the body of the paper do not make it any less incorrect. Any broad claim such as this requires only one hypothetical to refute it: would the author seriously maintain that if an actuary found that a participant for whom he or she expected to compute a joint and survivor annuity was not, in fact, married to the apparent spouse, he or she could trumpet that fact to the world on the ground that ERISA and associated regulations do not discuss confidentiality and the CPR cannot apply? Any such actuary could and would be disciplined under the CPR. I doubt that a court would hold that the Joint Board's standards pre-empted the entire CPR.

Giving any credibility to the confidentiality provisions of any professional actuarial code would be irresponsible and contrary to the disclosure requirement of federal law.

Federal law takes precedence under the CPR. The issue of the superiority of federal law in the event of conflict is never reached. Giving credibility to confidentiality provisions in a way that conflicts with applicable law would be contrary to the CPR as well as federal law.

Enforcement from federal agencies is a reasonable and expected result for those pension actuaries who believe that their professional codes of silence are above the law.

The paper offers no support for the notion that actuaries hold such a belief. All that is demonstrated is a lack of written justification for assumption changes on some Schedule Bs. The work may be sloppy, but I doubt whether it is motivated by the confidentiality provisions of the CPR. There is nothing in the body of the paper to suggest any "sentiment to maintain confidentiality through superficial rather than full disclosure." (Indeed, it is not clear whether the alleged sentiment is being attributed to a limited number of actuaries or to the drafters of the CPR.)
Again, in discussion of two suspensions, the author asks whether "the lack of detrimental consequences for suspension as an enrolled actuary suggest[s] minimal respect for the statutory standards of the Joint Board?" Is it not likely that the relevant professional bodies simply decided that there was no need to impose any additional penalties on these two people beyond the Joint Board's action?

This enforcement to date has been passive and lacking but the Joint Board may actively search for enrolled actuaries failing to discharge their duties under ERISA when the PBGC has more influence with the Joint Board.

The author is entitled to his opinion, and some may agree. It seems unlikely that if PBGC does increase its influence with the Joint Board, however, it will put much emphasis on issues of client confidentiality.

Authors' Reply to Discussion

An example involving actuarial services performed under ERISA may clarify the difference in opinion. Under the Standards of Performance for Enrolled Actuaries published by the Joint Board at 901.20(h):

An enrolled actuary shall provide written notification of the nonfiling of any actuarial document he/she has signed upon discovery of the non-filing. Such notification shall be made to the office of the Internal Revenue Service, the Department of Labor, or the Pension Benefit Guaranty Corporation where such document should have been filed.

The corporate sponsor proudly notifies the enrolled actuary of the alteration and subsequent filing of an actuarial document the enrolled actuary signed and certified. The Code of Professional Conduct of the Society of Actuaries has materiality provisions under its Annotation 14-1. The Joint Board's Standards of Performance for Enrolled Actuaries have no materiality provisions.

If you fail to notify based on immateriality, you have used the sentiment of the Code of Professional Conduct to subvert the standards of the Joint Board. Let those governmental agencies decide if it is immaterial. Because enrolled actuaries are engaged on behalf of all plan participants, you are traveling down the wrong path if your decision is based on the potential response of that corporate sponsor ("they're going to call me a whistle-blower and ... "). Because the enrolled actuary
is an individual person under ERISA Section 3042, it is not a decision to be made by the actuarial consulting firm on its own behalf. And last, there is nothing to decide under 901.20(h).

Immateriality is a defense. Confidentiality is more of an underlying sentiment used to eliminate or diminish the need for a defense. Is the failure to justify actuarial assumption changes due to benign neglect? Sloppiness? These excuses would be more readily accepted if enrolled actuaries did not benefit from those mistakes.

When the enrolled actuary is changed, the plan administrator must provide an explanation for the change on the Schedule C attached to the Form 5500. The prior enrolled actuary must also be given a “Notice to Terminated Enrolled Actuary” containing the explanation for the change as disclosed on the Schedule C. The notice instructs the prior enrolled actuary to supply his or her comments on this explanation directly to the Office of Enforcement of the Pension and Welfare Benefits Administration, U.S. Department of Labor. Will the prior enrolled actuary be influenced by the confidentiality and professional courtesy provisions of the Code?

The joint and survivor annuity example presented by one of the discussants, Brian A. Jones, is not an actuarial service performed under ERISA. Hopefully Mr. Jones (or one of his non-actuarial associates) would inform the plan administrator of these facts. The plan administrator would need this information to properly discharge his or her duties under ERISA.

Is the success of saying “I didn’t hear that!” contingent upon a sense of confidentiality?