2000

Journal of Actuarial Practice, Volume 8, Nos. 1 and 2 (2000)

Colin Ramsay , Editor
University of Nebraska - Lincoln, cramsay@unl.edu

Journal of Actuarial Practice Volume 8, Nos. 1 and 2, 2000

ARTICLES

Journal of Actuarial Practice

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. One of the goals of the Journal of Actuarial Practice is to improve communication between the practicing and academic actuarial communities. In addition, the Journal provides a forum for the presentation and discussion of ideas, issues (controversial or otherwise), and methods of interest to actuaries.

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:

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

REVIEW PROCESS

A paper submitted to the Journal first is screened by the editor for suitability. If it i deemed suitable, copies are sent to several independent referees. The name of th author(s) of the paper under consideration is usually anonymous to the referees, and th identities of referees are never revealed to the author(s).

The paper is reviewed for content and clarity of exposition. Papers do not have to contai original ideas to be acceptable. On the basis of the referee reports, the editor makes on of the following decisions: (1) accept subject to minor revisions, (2) accept subject to majo revisions, or (3) reject.

The editor communicates the recommendation to the author(s) along with copies of th referees' reports. The entire process is expected to take two to four months.

See inside back cover for instructions to authors.

Absalom Press, Inc.
P.O. Box 22098, Lincoln, NE 68542-2098, USA.

Journal of Actuarial Practice ISSN 1064-6647

Realistic Pension Funding: A Stochastic Approach

Shih-Chieh Chang*

Abstract†

The process funding pension plans is viewed as a dynamic control process. Two performance measures are introduced to evaluate the effectiveness of plan contributions: the cost-induced performance measure (CIPM) and the ratio-induced performance measure (RIPM). A dynamic programming approach is used to determining the optimal contributions with the objective of minimizing the performance measure. The methodology developed is applied to a sample of members of Taiwan's Public Employees Pension Plan (Tai-PERS). We show that RIPM produces more stable results than those using CIPM.

Key words and phrases: *contributions, control theory, dynamic programming, performance measure*

1 Introduction

Following Haberman and Sung (1994), the process of pension plan funding is viewed as a stochastic control system where the plan's financial obligations are affected by random events (such as pension plan turnover, investment returns, deaths, retirements, etc.). The methods of control theory can thus be used to assist plan administrators in choosing optimal contributions. Contributions may optimized with

*Shih-Chieh Chang, Ph.D., is a professor of risk management and insurance in the College of Commerce at National Chengchi University in Taiwan. Dr. Chang received his bachelor's degree in mathematics from National Taiwan University and his Ph.D. in statistics from the University of Wisconsin-Madison. His research interests are in stochastic models for actuarial science and the Taiwanese financial market data.

Dr. Chang's address is: Department of Risk Management and Insurance, College of Commerce, National Chengchi University, Taipei, Taiwan, R.O.C. Internet address: *bchang@nccu.edu.tw*

†The author wishes to thank the three anonymous referees and (especially) the editor for their detailed comments and numerous suggestions on an earlier draft of this paper.

respect to a performance measure or with respect to maximizing the probability of plan solvency during the period of budget planning.

Several authors have studied pension funding using control theory or other similar methods including Bowers et al., (1982), O'Brien (1986, 1987), Bacinello (1988), Dufresne (1988, 1989), Haberman (1992, 1993, 1994), Daykin et al., (1994), Haberman and Sung (1994), Haberman and Wong (1997), Schäl (1998), and Chang(1999). Survey articles on applications of control theory to insurance in general include Martin-Löf (1994) and Runggaldier (1998).

The objective of this paper is to use control theory techniques to determine the optimal employer contributions. The optimality criterion used is based on the risks associated with the stability and security of the pension plan. Two types of risks associated with pension plan stability and security (as proposed by Haberman and Sung (1994)) are used: (i) the contribution rate risk, and (ii) the solvency risk. An objective function associated with these two risks is constructed and dynamic programming methods are then used to derive the optimal contributions that minimize the solvency risk subject to specific constraints. Details of this dynamic optimization can be found in Chang (1999). The methodology is applied to the Taiwan Public Employees Retirement System (Tai-PERS).

The paper is organized as follows: Section 2 describes two performance measures and the dynamic optimization scheme. An expression is given for the minimum contribution. Section 3 contains a mathematical description of the benefit structure of Tai-PERS. Section 4 presents an application of the proposed methodology to the Tai-PERS data.

2 The Model

2.1 The Basic Notation

Let T denote the plan administrator's planning time horizon, i.e., the administrator is interested in achieving certain results related to plan stability and solvency by time T. For $t = 0, 1, 2, \ldots, T-1$, the following actuarial notations are used throughout this paper:

$$
\begin{aligned}
i_t &= \text{The actual rate of return on pension fund assets in } [t, t+1);\\
v_t &= \text{The discount factor from time } t \text{ to time 0, with } v_0 = 1;\\
C_t &= \text{Contributions paid at time } t \text{ for the plan year } [t, t+1);\\
B_t &= \text{Total benefits paid during the plan year } [t, t+1);\\
\mathrm{NC}_t &= \text{The normal cost at time } t;\\
F_t &= \text{Total pension plan assets at time } t\text{, excluding } C_t;\\
\mathrm{AL}_t &= \text{Total plan accrued liability at time } t+;\\
\mathfrak{I}_t &= \text{All plan information in } [0, t);\\
\beta_t &= \text{The risk-weighted ratio at time } t\text{; and}\\
\eta &= \text{The target funding ratio.}
\end{aligned}
$$

The parameter β_t is needed to balance the tradeoff between the solvency risk and the stability risk (discussed later). To generalize the model, η is used as the target funding ratio with the understanding that the funding ratio may not necessarily be equal to the accrued liability, i.e., η is not necessarily equal to one.

We assume that for the period of the planning horizon, the plan valuations occur at times $0, 1, \ldots, T-1$ using the *entry-age-normal cost method*,[1] and that $\mathfrak{I}_t$ contains all of the information gathered from all previous plan valuations. This information includes:

- v_k, for $k = 0, 1, 2, \ldots, t$;
- i_k, for $k = 0, 1, 2, \ldots, t-1$;
- C_k, for $k = 0, 1, 2, \ldots, t-1$;
- NC_k, for $k = 0, 1, 2, \ldots, t$;
- AL_k, for $k = 0, 1, 2, \ldots, t$;
- F_k, for $k = 0, 1, 2, \ldots, t$;
- β_k, for $k = 0, 1, 2, \ldots, t$;
- B_k, for $k = 0, 1, 2, \ldots, t$; and
- The information pertaining to plan demographics including employment and salary information for each employee that entered or left the plan during the period $[0, t)$.

[1] For details on pension cost methods see, for example, Shapiro (1985) or Anderson (1990).

Thus the only information unknown at time t is the rate of return for the coming year, i_t, and the actual contribution, C_t.

2.2 The Optimization Equation and Its Solution

We will define the performance measures used to construct the optimization equation.

2.2.1 Performance Measures

A performance measure is a nonnegative function that is used to determine the set of parameters of the model that best fit the data over the entire planning horizon. Two performance measures are defined: the cost-induced performance measure (CIPM) and the ratio-induced performance measure (RIPM). The CIPM was proposed by Haberman and Sung (1994) as a discounted quadratic deviation risk measure to obtain the optimal contribution. The RIPM was proposed by Chang (1999). It employs relative performance ratios to measure the discounted quadratic deviation in achieving the optimal funding status.

The CIPM is defined as Γ, a function of the contributions:

$$\begin{aligned}\Gamma &\equiv \Gamma\left(C_0, \ldots, C_{T-1}\right) \\ &= E\left[\sum_{t=0}^{T-1}\left[v_t\left(C_t - \mathrm{NC}_t\right)^2 + v_{t+1}\beta_{t+1}\left(F_{t+1} - \eta \mathrm{AL}_{t+1}\right)^2\right]\right]\end{aligned} \tag{1}$$

and RIPM is defined as J, a function of the contributions:

$$\begin{aligned}J &\equiv J\left(C_0, \ldots, C_{T-1}\right) \\ &= E\left[\sum_{t=0}^{T-1}\left[v_t\left(1 - \frac{C_t}{\mathrm{NC}_t}\right)^2 + v_{t+1}\beta_{t+1}\left(1 - \frac{F_{t+1}}{\eta \mathrm{AL}_{t+1}}\right)^2\right]\right].\end{aligned} \tag{2}$$

Each performance measure can be split into two risk measure components: the contribution rate risk and solvency risk as shown in Table 1.

2.2.2 The Optimization Equation

The optimization problem at time $t \leq T-1$ is to determine the sequence of contribution rates $C_0, C_1, \ldots, C_{T-1}$ that minimize the chosen

Table 1
Risk and Performance Measures

Risk Measures	Performance Measures: CIPM	RIPM
Contribution Rate Risk:	$v_t\,(C_t - \mathrm{NC}_t)^2$	$v_t\left(1 - \frac{C_t}{\mathrm{NC}_t}\right)^2$
Solvency Risk:	$v_{t+1}\,(F_{t+1} - \eta \mathrm{AL}_{t+1})^2$	$v_{t+1}\left(1 - \frac{F_{t+1}}{\eta \mathrm{AL}_{t+1}}\right)^2$

performance measure. In the case of the RIPM, J must be minimized, i.e., the optimization equation is:

$$\min_{C_0,\dots,C_{T-1}} J = \min_{C_0,\dots,C_{T-1}} E\left[\sum_{s=0}^{T-1}\left[v_s\left(1-\frac{C_s}{\mathrm{NC}_s}\right)^2 + v_{s+1}\beta_{s+1}\left(1-\frac{F_{s+1}}{\eta\mathrm{AL}_{s+1}}\right)^2\right]|\mathfrak{I}_t\right].$$

If we set

$$J_{k,T} = E\left[\sum_{s=k}^{T-1}\left[v_s\left(1-\frac{C_s}{\mathrm{NC}_s}\right)^2 + v_{s+1}\beta_{s+1}\left(1-\frac{F_{s+1}}{\eta\mathrm{AL}_{s+1}}\right)^2\right]|\mathfrak{I}_t\right]$$

for $k = t, t+1, \dots, T-1$, and

$$J_{0,t} = E\left[\sum_{s=0}^{t-1}\left[v_s\left(1-\frac{C_s}{\mathrm{NC}_s}\right)^2 + v_{s+1}\beta_{s+1}\left(1-\frac{F_{s+1}}{\eta\mathrm{AL}_{s+1}}\right)^2\right]|\mathfrak{I}_t\right],$$

then the optimization equation at t can be written as

$$\min_{C_0,\dots,C_{T-1}} J = \min_{C_0,\dots,C_{T-1}} J_{0,t} + J_{t,T}.$$

At time t, however, $C_0, C_1, \dots, C_{t-1}$ are already known so $J_{0,t}$ is known.

$$= J_{0,t} + \min_{C_t,\dots,C_{T-1}} J_{t,T}.$$

Hence only $C_t, C_{t+1}, \dots, C_{T-1}$ and $J_{t,T}$ need to be determined. In other words, we need only to minimize $J_{t,T}$ subject to certain funding constraints. The full optimization problem thus is to

$$\min_{C_t,\ldots,C_{T-1}} J_{t,T} = E\left[\sum_{s=t}^{T-1}\left[v_s\left(1-\frac{C_s}{\mathrm{NC}_s}\right)^2 + v_{s+1}\beta_{s+1}\left(1-\frac{F_{s+1}}{\eta \mathrm{AL}_{s+1}}\right)^2\right] | \mathfrak{I}_t\right] \tag{3}$$

subject to the funding constraint

$$F_{t+1} = (F_t + C_t - B_t)(1 + i_t). \tag{4}$$

Equation (4) assumes B_t is paid in a lump sum at the start of the year. As we will see, this is clearly not the case under Tai-PERS because its benefits are paid twice per year or as a lump sum.

2.2.3 Some Assumptions

For $k = t, t+1, \ldots, T-1$, let $\Omega_k(F_t)$ be a function of $F(t)$, where

$$\Omega_k(F_t) = \min_{C_k,\ldots,C_{T-1}} J_{k,T}. \tag{5}$$

As pension plans usually evaluate their financial status annually, it is reasonable to assume that F_t is a Markov process. As $\Omega_k(F_t)$ involves the term $E\left[F_{s+1}^2 | \mathfrak{I}_t\right]$, then we can write

$$\Omega_k(F_t) = a_1(k) F_t^2 + a_2(k) F_t + a_3(k) \tag{6}$$

where, for $t = 0, 1, \ldots, T$, the terms $a_1(t)$, $a_2(t)$, and $a_3(t)$ are sequences of constants. Following Haberman and Sung (1994), we use the boundary conditions $a_1(T) = a_2(T) = a_3(T) = 0$.

For mathematical simplicity, the sequence of annual rates of return, $\{i_t\}$, is assumed to consist of independent and identically distributed normal variables with constant mean θ and variance σ^2, i.e.,

$$i_t \sim N\left(\theta, \sigma^2\right). \tag{7}$$

The first two moments of i_t are

$$E[1 + i_t] = 1 + \theta \quad \text{and} \quad E\left[(1 + i_t)^2\right] = \sigma^2 + (1 + \theta)^2.$$

The moments of $F_{t+1} \mid F_t$ are

$$E\left[F_{t+1} \mid F_t\right] = (1+\theta)\left(F_t + C_t - B_t\right)$$

and

$$E\left[F_{t+1}^2 \mid F_t\right] = \left(\sigma^2 + (1+\theta)^2\right)\left(F_t + C_t - B_t\right)^2.$$

The sequence of valuation discount factors $\{v_t\}$, however, is assumed to be independent of the sequence of i_ts.

2.2.4 The Solution to the Optimization Equation

The solution follows that proposed by Chang (1999).

$$\begin{aligned}\Omega_t\left(F(t)\right) &= \min_{C_t,\ldots,C_{T-1}} \left\{ v_t \left(1 - \frac{C_t}{\mathrm{NC}_t}\right)^2 \right. \\ &\quad \left. + v_{t+1}\beta_{t+1} E\left[\left(1 - \frac{F_{t+1}}{\eta \mathrm{AL}_{t+1}}\right)^2 | \mathfrak{I}_t\right] + J_{t+1,T} \right\} \\ &= \min_{C_t} \left\{ v_t \left(1 - \frac{C_t}{\mathrm{NC}_t}\right)^2 \right. \\ &\quad \left. + v_{t+1}\beta_{t+1} E\left[\left(1 - \frac{F_{t+1}}{\eta \mathrm{AL}_{t+1}}\right)^2 | \mathfrak{I}_t\right] + \Omega_{t+1}\left(F_t\right) \right\} \\ &= \min_{C_t} R\left(C_t\right), \end{aligned} \tag{8}$$

where

$$
\begin{aligned}
R\left(C_t\right) = {} & v_t\left(1-\frac{C_t}{\mathrm{NC}_t}\right)^2+v_{t+1}\beta_{t+1}E\left[\left(1-\frac{F_{t+1}}{\eta \mathrm{AL}_{t+1}}\right)^2 \mid \mathfrak{I}_t\right] \\
& +E\left[\left(a_1(t+1)F_{t+1}^2+a_2(t+1)F_{t+1}+a_3(t+1)\right)\mid \mathfrak{I}_t\right] \\
= {} & v_t\left(1-\frac{C_t}{\mathrm{NC}_t}\right)^2 \\
& +v_{t+1}\beta_{t+1}\left[1-2\frac{(1+\theta)\left(F_t+C_t-B_t\right)}{\eta \mathrm{AL}_{t+1}}\right. \\
& \left.+\frac{\left(\sigma^2+(1+\theta)^2\right)\left(F_t+C_t-B_t\right)^2}{\left(\eta \mathrm{AL}_{t+1}\right)^2}\right] \\
& +a_1(t+1)\left(\sigma^2+(1+\theta)^2\right)\left(F_t+C_t-B_t\right)^2 \\
& +a_2(t+1)(1+\theta)\left(F_t+C_t-B_t\right)+a_3(t+1). \qquad (9)
\end{aligned}
$$

The first derivative of $R\left(C_t\right)$ is

$$
\begin{aligned}
\frac{dR\left(C_t\right)}{dC_t} = {} & \frac{-2v_t}{\mathrm{NC}_t}\left(1-\frac{C_t}{\mathrm{NC}_t}\right) \\
& +v_{t+1}\beta_{t+1}\left[\frac{-2(1+\theta)}{\eta \mathrm{AL}_{t+1}}+2\left(\sigma^2+(1+\theta)^2\right)\frac{\left(F_t+C_t-B_t\right)}{\left(\eta \mathrm{AL}_{t+1}\right)^2}\right] \\
& +2a_1(t+1)\left(\sigma^2+(1+\theta)^2\right)\left(F_t+C_t-B_t\right) \\
& +a_2(t+1)(1+\theta). \qquad (10)
\end{aligned}
$$

Setting equation (10) to zero and solving for the optimal C_t^* yields

$$
C_t^* = \frac{D_t^{(\mathrm{R})}-H_t^{(\mathrm{R})}\times F_t}{G_t^{(\mathrm{R})}} \qquad (11)
$$

where

$$D_t^{(\mathrm{R})} = \frac{2v_t}{\mathrm{NC}_t^2} + \frac{2v_{t+1}\beta_{t+1}(1+\theta)}{\eta \mathrm{AL}_{t+1}} + \frac{2v_{t+1}\beta_{t+1}\left(\sigma^2+(1+\theta)^2\right)}{\eta^2 \mathrm{AL}_{t+1}^2} B_t$$
$$+ 2a_1(t+1)\left(\sigma^2+(1+\theta)^2\right)B_t - a_2(t+1)(1+\theta) \quad (12)$$
$$H_t^{(\mathrm{R})} = \frac{2v_{t+1}\beta_{t+1}(1+\theta)}{\eta \mathrm{AL}_{t+1}} + 2a_2(t+1)\left(\sigma^2+(1+\theta)^2\right) \quad (13)$$
$$G_t^{(\mathrm{R})} = \frac{2v_t}{\mathrm{NC}_t^2} + \frac{2v_{t+1}\beta_{t+1}\left(\sigma^2+(1+\theta)^2\right)}{\eta^2 \mathrm{AL}_{t+1}^2}$$
$$+ 2a_1(t+1)\left(\sigma^2+(1+\theta)^2\right). \quad (14)$$

As

$$\Omega_t(F_t) \equiv R(C_t^*),$$

it follows that

$$a_1(t)F_t^2 + a_2(t)F_t + a_3(t)$$
$$\equiv v_t\left(1-\frac{C_t^*}{\mathrm{NC}_t}\right)^2 + v_{t+1}\beta_{t+1}E\left[\left(1-\frac{F_{t+1}}{\eta \mathrm{AL}_t}\right)^2 |\mathfrak{I}_t\right]$$
$$+ a_1(t+1)E\left[F_{t+1}^2|\mathfrak{I}_t\right] + a_2(t+1)E\left[F_{t+1}|\mathfrak{I}_t\right] + a_3(t+1)$$

which yields

$$a_1(t) = \frac{v_t (H_t^{(R)})^2}{(G_t^{(R)})^2 \mathrm{NC}_t^2} + v_{t+1}\beta_{t+1} \frac{\sigma^2 + (1+\theta)^2}{\eta \mathrm{AL}_t^2} \left(1 - \frac{H_t^{(R)}}{G_t^{(R)}}\right)^2$$
$$+ a_1(t+1)\left(\sigma^2 + (1+\theta)^2\right)\left(1 - \frac{H_t^{(R)}}{G_t^{(R)}}\right)^2 \tag{15}$$

$$a_2(t) = \frac{-2H_t^{(R)}}{G_t^{(R)}\mathrm{NC}_t}\left(\frac{D_t^{(R)}}{G_t^{(R)}\mathrm{NC}_t} - 1\right) - 2v_{t+1}\beta_{t+1}\left(\frac{(1+\theta)}{\eta \mathrm{AL}_t}\right)\left(1 - \frac{H_t^{(R)}}{G_t^{(R)}}\right)$$
$$+ 2v_{t+1}\beta_{t+1}\frac{\sigma^2 + (1+\theta)^2}{\eta^2 \mathrm{AL}_{t+1}^2}\left(1 - \frac{H_t^{(R)}}{G_t^{(R)}}\right)\left(\frac{D_t^{(R)}}{G_t^{(R)}} - B_t\right)$$
$$+ 2a_1(t+1)\,\sigma^2 + (1+\theta)^2\left(1 - \frac{H_t^{(R)}}{G_t^{(R)}}\right)\left(\frac{D}{G_t^{(R)}} - B_t\right)$$
$$+ a_2(t+1)(1+\theta)\left(1 - \frac{H_t^{(R)}}{G_t^{(R)}}\right). \tag{16}$$

To ensure that C_t^* indeed yields a minimum for $R(C_t)$, we must have $\frac{d^2R(C_t)}{dC_t^2} > 0$. But

$$\frac{d^2R(C_t)}{dC_t^2} = \frac{2v_t}{\mathrm{NC}_t^2} + v_{t+1}\beta_{t+1}\frac{1 + 2\theta + \theta^2 + \sigma^2}{(\eta \mathrm{AL}_{t+1})^2}$$
$$+ a_1(t+1)\left((1+\theta)^2 + \sigma^2\right)$$
$$> 0,$$

which implies that

$$a_1(t+1) > -\left(\frac{v_t}{\left((1+\theta)^2 + \sigma^2\right)\mathrm{NC}_t^2} + \frac{v_{t+1}\beta_{t+1}}{(\eta \mathrm{AL}_{t+1})^2}\right). \tag{17}$$

So C_t^* is the unique optimum contribution for year t.

Finally, the optimal contributions for a planning period of T years under CIPM are determined from the following equations:

$$C_t^* = \frac{D_t^{(C)} - H_t^{(C)} \times F_t}{G_t^{(C)}} \tag{18}$$

where

$$D_t^{(C)} = \left(2v^t NC_t + 2\left(v^{t+1}\beta + a_1(t+1)\right)\left((1+\theta)^2 + \sigma^2\right)\right) B_t + \left(2v^{t+1}\beta AL_{t+1} - a_2(t+1)\right)(1+\theta) \tag{19}$$

$$H_t^{(C)} = 2v^t NC_t + 2\left(v^{t+1}\beta + a_1(t+1)\right)\left((1+\theta)^2 + \sigma^2\right); \tag{20}$$

$$G_t^{(C)} = 2v^t + 2\left(v^{t+1}\beta + a_1(t+1)\right)\left((1+\theta)^2 + \sigma^2\right). \tag{21}$$

The $\{a_1(t)\}$ and $\{a_2(t)\}$ sequences used in equations (19), (20), and (21) are different from that used in the RIPM method. They are defined recursively as:

$$a_1(t) = v^t \frac{(v^{t+1}\beta + a_1(t+1))((1+\theta)^2 + \sigma^2)}{v^t + (v^{t+1}\beta + a_1(t+1))((1+\theta)^2 + \sigma^2)} \tag{22}$$

and

$$a_2(t) = 2(NC_t - B_t)a_1(t) + \frac{v^t\left(a_2(t+1) - 2v^{t+1}\beta AL_{t+1}\right)(1+\theta)}{v^t + (v^{t+1}\beta + a_1(t+1))((1+\theta)^2 + \sigma^2)}. \tag{23}$$

3 Taiwan's Public Employees Retirement System

The Taiwan Public Employees Retirement System (Tai-PERS) is a large defined benefit public retirement system that provides retirement and ancillary benefits to public employees, i.e., employees of the national government or any local government in Taiwan. Retirement benefits are calculated according to length of service and final salary at retirement. The present funding policy requires both employer and employee contributions to a public trust fund for a maximum 35 years. Each participant in Tai-PERS contributes 2.8% of covered monthly salary while the employer (national or local government) contributes 5.2% of the participant's covered monthly salary .

According to current Tai-PERS regulations, it is "mandatory" for employees to retire at the age of 65. If the employee is still healthy and willing to continue to work beyond age 65, however, Tai-PERS authorizes the employee to work for a maximum of 5 years more under the condition that officials approve the employee's application for late retirement. Thus the oldest retirement age is 70.

Retirees have three options at retirement: (i) a lump-sum retirement payment, (ii) a monthly pension with the cost of living adjustment that is decided by the government annually, or (iii) a combination of a lump-sum payment and a monthly pension. Participants who have worked for more than 5 years and less than 15 years can have a lump-sum single retirement payment.

3.1 More Notation

Some of the actuarial notation used to describe Tai-PERS are:

y is the normal retirement age ($y = 65$ years by law);

t is the current time;

x is the current age;

m is the minimum entry age into the plan;

$S_{x,j,t}$ is the base monthly salary of the participant j at age x at time t;

s_x is the salary scale function at age x;

${}_hp_x^{(T)}$ is the probability that an active participant age x is still active after h years;

$q_x^{(r)}$ is the probability that an active participant age x retires within a year;

$q_x^{(d)}$ is the probability that an active participant age x dies within a year;

$q_x^{(l)}$ is the probability that an active participant age x is laid off and disabled within a year;

$q_x^{(w)}$ is the probability that an active participant age x withdraws within a year;

$v = \frac{1}{1+j}$ is the discount factor under the assumed interest rate j;

$h(z)$ is the cost of living adjustment (COLA) function for a retiree z years after retirement, $z \geq 0$;

$A_{x,t}$ is the set of active participants in plan age x at time t;

$Y_{x,j,t}$ is the number of years of service (employment) up to age x for participant $j \in A_{x,t}$;

$[z]$ denotes the largest integer less than or equal to z;

$M^{sin}_{x,j,t}$ is the factor (based on the number of years of service) used to calculate a single lump-sum payment for participant $j \in A_{x,t}$, where

$$M^{sin}_{x,j,t} = \begin{cases} \min(53, 1.5[Y_{x,j,t}]) & \text{if } Y_{x,j,t} = [Y_{x,j,t}] \\ \min(53, 1.5[Y_{x,j,t}]) & \text{if } Y_{x,j,t} < [Y_{x,j,t}] + 0.5 \\ \min(53, 1.5[Y_{x,j,t}] + 1.5) & \text{if } Y_{x,j,t} \geq [Y_{x,j,t}] + 0.5; \end{cases}$$

$M^{mon}_{x,j,t}$ is the factor (based on the number of years of service) used to calculate a monthly pension for participant $j \in A_{x,t}$. For participants with at least 15 years of service ($[Y_{x,j,t}] \geq 15$),

$$M^{mon}_{x,j,t} = \begin{cases} \min(35, [Y_{x,j,t}]) & \text{if } Y_{x,j,t} = [Y_{x,j,t}] \\ \min(35, [Y_{x,j,t}] + 0.5) & \text{if } 0.5 > Y_{x,j,t} - [Y_{x,j,t}] > 0 \\ \min(35, [Y_{x,j,t}] + 1) & \text{if } 1 > Y_{x,j,t} - [Y_{x,j,t}] \geq 0.5; \end{cases}$$

${}^{\text{cola}}a^{(2)}_x$ is the cost of living adjustments annuity at age x, i.e.,

$${}^{\text{cola}}a^{(2)}_x = \sum_{k=1}^{\infty} \frac{1}{2} v^{\frac{k}{2}} \, {}_{\frac{k}{2}}p_x \prod_{w=1}^{k} \left(1 + h(\frac{w}{2})\right)^{\frac{1}{2}};$$

${}^{\text{cola}}s^{(2)}_{\overline{n}|}$ is the n-year cost of living adjustments annuity-certain, i.e.,

$${}^{\text{cola}}s^{(2)}_{\overline{n}|} = \sum_{k=1}^{2n} \frac{1}{2} v^{\frac{k}{2}} \prod_{w=1}^{k} \left(1 + h(\frac{w}{2})\right)^{\frac{1}{2}};$$

PVFLS^j_t is the actuarial present value of the lump-sum (single) payment to participant $j \in A_{x,t}$; and

PVFMO^j_t is the actuarial present value of the monthly pension to participant $j \in A_{x,t}$.

It is important to note at this point that a public employee's salary structure under Tai-PERS differs from that of public employees in the United States. In Taiwan a public employee's salary is divided into two components. One component is the base salary (i.e., $S_{x,j,t}$), which varies according to the employee's years of service and job rank. The other is the merit pay or bonus that varies according the employee's work performance. These two components are roughly 50% of each employee's salary. For example, an employee may have monthly salary of \$5,000 (with base salary of \$2,500 and an additional bonus of \$2,500, while another employee may have month salary of \$5,000 but with base salary of \$2,700 and an additional bonus of \$2,300. In this study, the base salary (i.e., roughly half of participant's earned salary) is used to calculate the retirement benefits. Hence $S_{x,j,t}$ is assumed to be half of participant's actual monthly salary.

Without losing accuracy and to avoid messy and complex presentation, some minor modifications in formulating the various benefit payments are used.

3.2 Normal Retirement Benefits

According to the current Taiwanese public employees retirement and compensation law, it is mandatory for the employees to retire at the age of 65 unless permission is granted otherwise. A public employee can voluntarily retire at the age of 60 after 5 years of service or at any age after 25 years of service. Employees who have worked for more than 15 years are eligible to choose any one of the following five benefit payment options:

- Option 1: A lump-sum single payment;
- Option 2: A monthly benefit;
- Option 3: One half of the lump-sum payment and one half of the monthly pension;
- Option 4: One third of the lump-sum payment and two thirds of the monthly pension; or
- Option 5: One fourth of the lump-sum payment and three fourths of the monthly pension.

Participants who have worked more than 5 years but less than 15 years, however, only have the lump-sum payment option. There is no retirement benefit for the participants who have worked less than 5 years.

The so-called monthly pension under Tai-PERS is paid to retirees twice per calendar year on specific dates, namely January 16 and July 16. The payment on January 16 is for the period between January and June and the payment on July 16 is for the period between July and December each year. Regulations specify that the last employer (at the time of retirement) provides the first payment and Tai-PERS provides the rest of the benefits every six months. So, for example, consider a person born on May 8, 1934, hired on May 8, 1959, and who retired on May 8, 1999 at age 65 with a pension of \$3,000 per month. In this case, she will receive $\$3,000(1+\frac{24}{31}) = \$5,323$ from her last employer on May 8, 1999 for payment between May 8, 1999 and June 30, 1999. Then she will receive \$18,000 from Tai-PERS every six months starting on July 16, 1999 until she dies.

3.2.1 Lump-Sum Benefits

For participants who choose the lump-sum payment option, it is assumed that the payment is made in the middle of the retiring age. The lump-sum retirement benefit is 1.5 times the monthly final salary for each year of service, with a maximum benefit of 53 times the monthly salary. To encourage early retirements, if the participant chooses to retire at age 55, he or she receives an extra lump-sum benefit of five more months credited to his or her years of service.

Based on the information at time t, the actuarial present value of the lump-sum payment of participant $j \in A_{x,t}$ is

$$\mathrm{PVFLS}_t^j = \sum_{k=0}^{y-x-1} {}^{\sin}B^r_{x+k,j,t}\, {}_kp_x^{(T)}\, q_{x+k}^{(r)}\, v^{k+\frac{1}{2}}, \tag{24}$$

where

$$^{\sin}B^r_{x+k,j,t} = 2S_{x,j,t}\left(\frac{s_{x+k}}{s_x} M^{sin}_{x,j,t} + 5I\{x+k=55\} \times \frac{s_{55}}{s_x}\right)$$

and $I\{A\}$ is the indicator function for an event (set) A, i.e.,

$$I\{A\} = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{otherwise.} \end{cases}$$

3.2.2 Monthly Benefits

For participants who choose the monthly or the mixed payment option, the monthly pension is calculated according to the specific percent of the participant's salary. This base monthly benefit for a participant currently age x who retires at age $x+k$, ${}^{mon}B^{r}_{x+k,j,t}$, is 2% per year for each year of service, with a maximum benefit of 70% of final salary, i.e.,

$$ {}^{mon}B^{r}_{x+k,j,t} = 2S_{x,j,t}\frac{s_{x+k}}{s_x}\, 2\% \times M^{mon}_{x,j,t} \times 12. $$

The base monthly benefit is increased semi-annually with cost of living adjustments.

In addition to their monthly retirement benefit, retirees who choose the monthly benefit are entitled to two extra death benefits: (i) a payment of six months of final salary, and (ii) an extra payment as an inducement to encourage retirees to choose the monthly benefit option. Specifically, consider participant j who is currently age x who retires at age $x+k$ then dies at age $x+k+u$:

(i) The six months of final salary payment is

$$ 2S_{x,j,t}\frac{s_{x+k}}{s_x} \times 6. $$

(ii) The extra payment as an inducement is determined as follows: the quantities

$$ 2S_{x,j,t}\frac{s_{x+k}}{s_x}\, M^{sin}_{x,j,t}\,(1+i)^u $$

(as if the retiree had chosen the lump-sum option) and

$$ 2S_{x,j,t}\frac{s_{x+k}}{s_x} \times 2\% \times M^{mon}_{x+k,j,t} \times 12 \times {}^{\text{cola}}s^{(2)}_{\overline{u}|} $$

(the accumulated value of the monthly pension that the retiree has already received) are compared. A benefit is paid when the accumulated value of the monthly pension he received is less than the payment based on lump-sum option. Thus the inducement benefit can be written as

$$ \max\{0, 2S_{x,j,t}\,\frac{s_{x+k}}{s_x}\,(M^{sin}_{x,j,t}\,(1+i)^u - 0.24M^{mon}_{x+k,j,t}\,{}^{\text{cola}}s^{(2)}_{\overline{u}|})\}. $$

The size of the total extra death benefit for a person who retires at age $x + k$ is

$$\mathrm{XDB}_{x+k,u,j,t} = 2S_{x,j,t}\frac{s_{x+k}}{s_x} \times \left[6 + \max\left\{0, M^{sin}_{x,j,t}(1+i)^u - 0.24M^{mon}_{x+k,j,t}\ {}^{\mathrm{cola}}s^{(2)}_{\overline{u}|}\right\}\right]$$

and the actuarial present value of the monthly retirement benefit plus the extra death benefit is

$$\mathrm{PVFMO}^j_t = \sum_{k=0}^{y-x-1} {}_kp^{(\tau)}_x\, q^{(r)}_{x+k}\left[{}^{mon}B^r_{x+k,j,t}\, v^{k+\frac{1}{2}}\ {}^{\mathrm{cola}}a^{(2)}_{k+1/2} + \sum_{u=0}^{\infty} v^{\frac{u}{2}}\ {}_{\frac{u+1}{2}|\frac{1}{2}}q_{x+k}\,\mathrm{XDB}_{x+k,u,j,t}\right] \tag{25}$$

where ${}_{\frac{u+1}{2}|\frac{1}{2}}q_{x+k}$ is based on post-retirement mortality. Because we only have the retiree's age at retirement (recorded as an integer), 1/2 year is added to the recorded retirement age in order to adjust for the fractional part.

3.3 Benefits Other than Normal Retirement Benefits

3.3.1 Withdrawal Refund Benefits

In Tai-PERS both employers and employees make monthly contributions to the Tai-PERS trust fund. If C^j_z is the annual overall contribution rate (expressed as a percentage of salary) at age z for participant j, then 65% of C^j_z comes from the employer and 35% of C^j_z comes from the employee.

An employee who withdraws from service can receive his or her accumulated contributions. In addition, if the employee withdraws at exact age 35 or 45, the employee can also receive the employer's matching contribution. Based on the information available at time t, let $\mathrm{AC}^j_{x,z,t}$ denote the projected accumulated contributions at age $z \geq x$ for participant j who is currently age x. It follows that

$$\mathrm{AC}^j_{x,z,t} = \sum_{k=m}^{z-1} 24 \times 0.35 \times C^j_z S_{x,j,t}\frac{s_k}{s_x} a^{(12)}_{\overline{1}|}(1+i)^{z-k}.$$

The actuarial present value of withdrawal refund benefit is:

$$\mathrm{PVFWB}^{j}_{x,t} = \sum_{z=x}^{y-1} \left[1 + \frac{0.65}{0.35} \left(I\{z = 35\} + I\{z = 45\} \right) \right] \times \mathrm{AC}^{j}_{x,z,t} \, {}_{z-x}p^{(\tau)}_{x} \, q^{(w)}_{z} \, v^{(z-x)+\frac{1}{2}}. \tag{26}$$

3.3.2 Death Benefits for Active Participants

The size and form of the death benefit for active participants depends on the participant's years of service at the time of death. According to the regulation, a single payment is paid to his spouse if $Y_{x,j,t}$ is less than 15 years, while a monthly pension plus an extra benefit is paid when $Y_{x,j,t}$ is equal or more than 15 years.

For participants who worked for less 15 years the single payment death benefit is:

$$2S_{x,j,t} \frac{s_{x+k}}{s_x} M^{sin}_{x,j,t}.$$

For participants who worked for more than 15 years, however, the death benefit is paid in two forms:

(i) A semi-annual annuity paying an annual benefit of five months salary for 10 years to the surviving spouse, i.e.,

$$10S_{x,j,t} \frac{s_{x+k}}{s_x} {}^{\mathrm{cola}}\ddot{a}^{(2)}_{x+k:\overline{10}|}$$

where ${}^{\mathrm{cola}}\ddot{a}^{(2)}_{x+k:\overline{10}|}$ is the actuarial present value of the semi-annual pension paid to the spouse. To simplify the calculations, we assume that the spouse and the participant have the same age.

(ii) A lump-sum payment of 15 months salary for 15 years of service with an increase of half a month's salary for each year of service beyond 15 years to a maximum 25 months, i.e.,

$$2S_{x,j,t} \frac{s_{x+k}}{s_x} (15 + \min\{10, 0.5\,[Y_{x,j,t} - 14.5]\}.$$

Thus the actuarial present value of the total death payments is:

$$\text{PVFRB}_{x,t}^{j} = \sum_{k=0}^{y-x-1} \text{DB}_{x,x+k,t}^{j}\ {}_{k}p_{x}^{(\tau)}\ q_{x+k}^{(d)}\ v^{k+\frac{1}{2}}, \tag{27}$$

where

$$\begin{aligned}\text{DB}_{x,x+k,t}^{j} = 2S_{x,j,t}\frac{S_{x+k}}{S_x}\Big[&M_{x,j,t}^{sin} I\{Y_{x,j,t} < 15\} \\ &+ 5^{\text{cola}}\ddot{a}_{x+k:\overline{10|}}^{(2)} I\{Y_{x,j,t} \geq 15\} \\ &+ (15 + \min\left\{10, 0.5\,[Y_{x,j,t} - 14.5]\}I\{Y_{x,j,t} \geq 15\right\}\Big].\end{aligned}$$

3.3.3 Termination and Disability Benefits

Tai-PERS provides a comprehensive compensation plan for its members. According to the current Taiwan public employees retirement and compensation law, employees are terminated once they receive notice from their employer. Employees are then eligible for termination benefits, which are the same as the lump-sum benefit for retirement, i.e., 1.5 times the monthly final salary for each year of service, with a maximum benefit of 53 times the monthly salary and it excludes the extra benefit at age 55.

When an employee is incapable of fulfilling his or her responsibility or performing similar other jobs, retirement is mandatory under Taiwanese law. Disability benefits are the same as termination benefits.

$$\text{PVFDISB}_{x,t}^{j} = \sum_{k=0}^{y-x-1} \text{DIS}_{x+k,t}^{j}\,{}_{k}p_{x}^{(\tau)} q_{x+k}^{(d)} v^{k+\frac{1}{2}}, \tag{28}$$

where

$$\text{DIS}_{x+k,t}^{j} = 2S_{x,j,t}\frac{S_{x+k}}{S_x} M_{x,j,t}^{sin}.$$

4 Application of the Methodology to Tai-PERS

4.1 Data Description

There were about 271,215 active participants in the Tai-PERS system in 1996. A sample of 3,823 participants was chosen used to evaluate the performance of the minimization scheme proposed in Section 2.

The average age of the sample of employees was 42.99 years and the average number of years of service was 15.6 years. Generally speaking, the employees in the sample were relatively older than the Tai-PERS population as a whole. Accordingly, the contribution rates obtained from the sample vary from that for Tai-PERS. The distribution of the sample is shown in Table 2. The distribution of new entrants into the sample is given in Table 3.

A service table (Table 4) is constructed based on the experience data collected from July 1, 1995 to June 30, 1996. Though the earliest entry age into Tai-PERS is around 18 (after graduation from high school), making the youngest possible retirement age around 23, we set the youngest entry age to be 20. Tables 2, 3, and 4 are used to project the evolution of our sample of employees.

4.2 Assumptions

The basic actuarial assumptions are:

Service Table: The Tai-PERS service table is based on plan experience in 1995-1996 (see Table 3);

Post-Retirement Mortality: Follows the 1989 Taiwan Standard Ordinary life table (1989 TSO) annuity table;

Actuarial Cost Method: Individual entry age normal (EAN) cost method;

Salary Scale: $s_x = (1.035)^x$, i.e., salaries increase by 3.5% annually;

Inflation Rate: $h(z) = (1.03)^z$, i.e., a 3% annual inflation rate;

Valuation Interest Rate: 6%, i.e., $v_t = (1.06)^{-t}$;

Target Fund Ratio: $\eta = 75\%$ annually;

Risk Measurement Weight: $\beta_t = 60\%$ annually;

Fund Return Parameters: $\theta = 8\%$ and $\sigma^2 = 0.0004$;

Initial Fund: $F_0 = 373,211,585$ NT;

Table 2
Distribution of Sample of Tai-PERS Employees

Age x	No. of Employees	Average Years of Service: Total	Average Years of Service: In Plan	Base Monthly Salary (NT$)
21	1	2.39	1.39	9630.00
22	6	2.40	1.31	9740.83
23	9	3.42	1.33	13792.67
24	27	2.69	1.30	14557.78
25	32	3.25	1.33	14580.00
26	60	3.82	1.39	15038.92
27	75	3.80	1.39	17748.53
28	88	4.59	1.41	16865.91
29	93	5.21	1.44	15914.41
30	85	5.68	1.48	18528.71
31	112	6.44	1.50	17409.33
32	119	6.78	1.49	18942.90
33	157	7.40	1.47	18531.15
34	152	8.67	1.49	18933.52
35	173	10.08	1.50	19567.11
36	162	9.57	1.49	19321.60
37	184	11.19	1.50	20812.99
38	179	11.89	1.50	21111.17
39	169	12.73	1.50	22736.51
40	146	13.31	1.51	24158.42
41	169	14.17	1.51	24670.59
42	137	14.34	1.51	24502.26
43	120	15.27	1.51	25039.63
44	133	16.42	1.51	26975.98
45	124	16.74	1.51	25803.67

Table 2 (contd.)
Age Distribution of Tai-PERS Sample

Age x	No. of Employees	Average Years of Service Total	Average Years of Service In Plan	Base Monthly Salary (in NT$)
46	105	18.32	1.51	28092.10
47	116	18.79	1.51	26684.96
48	86	19.49	1.51	28444.19
49	84	21.56	1.50	28695.89
50	65	22.67	1.51	27389.92
51	42	23.48	1.51	31958.93
52	45	22.49	1.51	30450.67
53	51	23.03	1.51	29586.47
54	60	25.55	1.51	29098.58
55	50	26.71	1.51	31298.30
56	47	28.83	1.51	33984.26
57	33	28.60	1.51	31777.27
58	30	29.76	1.51	33272.83
59	31	32.72	1.51	31489.19
60	32	33.51	1.51	31162.97
61	29	34.35	1.51	38002.76
62	37	35.09	1.51	35186.08
63	34	34.86	1.51	33188.38
64	31	37.90	1.51	36694.52
65	50	36.74	1.51	40514.10
66	29	37.36	1.51	38078.62
67	4	41.78	1.51	47010.00
68	10	39.21	1.51	46951.00
69	3	39.86	1.51	47010.00
70	6	42.95	1.51	45535.00

Table 3
The Recruitment Distribution of New Entrants

Age	20-24	25-29	30-34	35-39	40-44
Percentage	20%	60%	18%	1%	1%

Table 4
Service Table of Tai-PERS

x	$l_x^{(\tau)}$	$q_x^{(d)}$	$q_x^{(r)}$	$q_x^{(w)}$	$q_x^{(l)}$
20	100,000	0.06%	0.00%	0.03%	0.00%
21	99,905	0.12%	0.00%	1.18%	0.00%
22	98,608	0.13%	0.00%	1.32%	0.01%
23	97,166	0.10%	0.00%	0.54%	0.01%
24	96,530	0.08%	0.00%	0.35%	0.01%
25	96,106	0.05%	0.00%	0.45%	0.01%
26	95,612	0.05%	0.01%	0.69%	0.01%
27	94,888	0.08%	0.00%	0.78%	0.01%
28	94,057	0.07%	0.01%	0.67%	0.01%
29	93,347	0.06%	0.01%	0.57%	0.03%
30	92,730	0.07%	0.00%	0.57%	0.02%
31	92,109	0.06%	0.00%	0.56%	0.02%
32	91,524	0.07%	0.00%	0.52%	0.04%
33	90,942	0.07%	0.00%	0.47%	0.06%
34	90,394	0.06%	0.00%	0.47%	0.06%
35	89,858	0.08%	0.01%	0.50%	0.11%
36	89,235	0.08%	0.00%	0.39%	0.08%
37	88,738	0.09%	0.00%	0.38%	0.11%
38	88,221	0.08%	0.01%	0.34%	0.16%
39	87,706	0.12%	0.01%	0.29%	0.17%
40	87,187	0.07%	0.01%	0.26%	0.18%
41	86,728	0.12%	0.00%	0.24%	0.23%
42	86,215	0.12%	0.03%	0.20%	0.26%
43	85,693	0.11%	0.02%	0.20%	0.21%
44	85,235	0.15%	0.08%	0.16%	0.25%
45	84,684	0.14%	0.18%	0.17%	0.23%

Table 4 (contd.)
Service Table of Tai-PERS

x	$l_x^{(\tau)}$	$q_x^{(d)}$	$q_x^{(r)}$	$q_x^{(w)}$	$q_x^{(l)}$
46	84,071	0.20%	0.22%	0.16%	0.27%
47	83,351	0.25%	0.30%	0.06%	0.22%
48	82,661	0.22%	0.32%	0.07%	0.21%
49	81,984	0.27%	0.26%	0.08%	0.22%
50	81,310	0.25%	2.30%	0.04%	0.09%
51	79,128	0.25%	1.83%	0.06%	0.15%
52	77,325	0.21%	1.62%	0.09%	0.13%
53	75,736	0.28%	1.91%	0.14%	0.18%
54	73,835	0.43%	1.34%	0.04%	0.12%
55	72,410	0.38%	9.92%	0.04%	0.08%
56	64,863	0.45%	1.90%	0.04%	0.17%
57	63,210	0.21%	2.84%	0.10%	0.07%
58	61,182	0.43%	3.42%	0.12%	0.12%
59	58,679	0.50%	9.37%	0.05%	0.10%
60	52,805	0.45%	8.90%	0.09%	0.09%
61	47,776	0.47%	5.12%	0.11%	0.07%
62	45,015	0.55%	6.07%	0.09%	0.04%
63	41,982	0.56%	6.53%	0.04%	0.12%
64	38,938	0.63%	7.58%	0.07%	0.19%
65	35,643	0.24%	98.82%	0.45%	0.00%
66	176	0.00%	29.85%	1.01%	0.00%
67	122	0.00%	20.32%	1.22%	0.00%
68	95	0.00%	17.86%	0.00%	0.00%
69	78	0.00%	57.93%	0.00%	0.00%
70	33	1.52%	70.97%	0.00%	0.00%

Initial Plan Size: 3,823 employees; and

Benefits: For simplicity, we assume that all benefits are paid at the beginning of each year. To accomplish this, some adjustments are done in simulating the payment streams.

An additional adjustment is needed to determine the retirement benefits for the entire plan. This is done as follows: Let ϕ_i denote the probability that a retiree chooses option i, $i = 1, 2, 3, 4, 5$, with $\sum \phi_i = 1$. The actuarial present value of retirement benefits for all active employees in Tai-PERS at time t is:

$$\text{PVFRB}_t = \sum_{x=m}^{y-1} \sum_{j \in A_{x,t}} \text{PVFRB}_{x,t}^j \tag{29}$$

where $\text{PVFRB}_{x,t}^j$ denotes the actuarial present value of retirement benefits for employee j who is age x at time t in Tai-PERS.

$\text{PVFRB}_{x,t}^j$ can be written as a weighted sum of the the five options listed at the start of Section 2, i.e.,

$$\begin{aligned} \text{PVFRB}_{x,t}^j = {} & \text{PVFLS}_t^j \times \phi_1 + \text{PVFMO}_t^j \times \phi_2 \\ & + \left[\frac{1}{2}\text{PVFLS}_t^j + \frac{1}{2}\text{PVFMO}_t^j\right] \times \phi_3 \\ & + \left[\frac{1}{3}\text{PVFLS}_t^j + \frac{2}{3}\text{PVFMO}_t^j\right] \times \phi_4 \\ & + \left[\frac{1}{4}\text{PVFLS}_t^j + \frac{3}{4}\text{PVFMO}_t^j\right] \times \phi_5. \end{aligned} \tag{30}$$

In reality, the ϕ_is will depend on an individual's age at retirement, number of years of service, financial circumstances, health status, etc. So, in general, it will be difficult to estimate each individual's ϕ_i. Thus the ϕ_is actually used in our simulations are aggregate probabilities for the entire plan and are estimated using the Tai-PERS experience in 1995–1997.

It must be pointed out that the work force and benefit payments used in the objective function are simulated under the assumption of an open group with a constant size. That is, we assume that the overall number of employees in the sample is held at a constant level, as is required by Taiwanese government policy in order to reduce the financial burden of Tai-PERS.

4.3 The Results and Analyses

Using the data and plan assumptions in Sections 4.1 and 4.2 and three planning periods (T = 10, 15, and 20 years), the benefits (B_t), plan experience, and the interest earned on investments (i_t) are simulated for 20 years. They are used to estimate actuarial accrued liabilities, normal costs, and fund assets. The optimal contributions for each planning period under RIPM are determined from equations (11) through (14). The optimal contributions for each planning period under CIPM are determined from equations (18) through (21). The results are shown in Tables 5 through 10.

Figures 1 through 3 show the optimal contribution ratios (C_t/NC_t) and funding ratios (F_t/AL_t) for the various planning periods. These figures show that these ratios vary over time, suggesting that they might be influenced by the demographic assumptions from our simulations. There is a consistent pattern of contribution ratios gradually decreasing from about 140% in 1997 to 100% at the end of the planning period under RIPM. There is a much wider variation under CIPM, from 160% in 1997 to roughly 60% at the end of the planning period. The optimal funding ratios tend to move in the opposite direction. They are gradually increasing by years from 70% in 1998 to about 97% at the end of the planning period under CIPM, while from 70% in 1998 to 100% at the end of the planning period under RIPM. The contribution ratios under CIPM seem to fluctuate more than under RIPM.

Figures 4 and 5 show the optimal contributions under CIPM and RIPM, respectively, for the various planning periods. These figures show that the size of the optimal contributions in any year decreases as we extend the length of the planning period.

The ratios of optimal contributions ($C_t^{\mathrm{RIPM}}/C_t^{\mathrm{CIPM}}$) are also plotted in Figure 6 for comparison. These ratios behave similarly during the first seven years or so. They then fluctuate significantly. Thus, we can expect different optimal contributions using RIPM and CIPM after the first seven to ten years.

Table 5

Optimum Contributions Using Haberman and Sung (1994) with T = 10

t	F_t	B_t	i_t	NC_t	AL_t	C_t^*
0	373,211,585	106,636,560	0.081	264,658,176	585,530,240	413,806,784
1	735,466,112	48,948,364	0.076	254,203,072	851,652,800	315,874,944
2	1,078,517,154	89,903,272	0.072	250,750,880	1,152,353,024	290,852,704
3	1,371,757,669	95,409,192	0.080	247,360,208	1,461,198,464	275,160,192
4	1,675,629,429	67,699,656	0.086	243,759,120	1,775,253,248	247,427,728
5	2,014,240,058	76,361,992	0.085	240,339,776	2,101,083,136	230,731,024
6	2,352,101,536	65,474,592	0.090	238,243,632	2,454,753,024	217,235,248
7	2,728,191,745	102,758,760	0.091	235,766,864	2,813,157,888	214,558,960
8	3,098,192,936	134,871,760	0.091	233,146,752	3,183,145,472	213,573,856
9	3,467,139,484	193,478,128	0.089	231,159,440	3,578,989,056	169,976,656
10	3,750,704,679	161,202,912		227,970,288	3,953,183,488	

Table 6
Optimum Contributions Using Chang (1999) with $T = 10$

t	F_t	B_t	i_t	NC_t	AL_t	C_t^*
0	373,211,585	106,636,560	0.081	264,658,176	585,530,240	337,189,323
1	652,645,634	48,948,364	0.076	254,203,072	851,652,800	300,062,452
2	972,393,725	89,903,272	0.072	250,750,880	1,152,353,024	283,749,928
3	1,250,364,179	95,409,192	0.080	247,360,208	1,461,198,464	271,027,799
4	1,540,061,470	67,699,656	0.086	243,759,120	1,775,253,248	259,663,643
5	1,880,346,539	76,361,992	0.085	240,339,776	2,101,083,136	250,553,324
6	2,228,378,413	65,474,592	0.090	238,243,632	2,454,753,024	244,624,247
7	2,623,226,718	102,758,760	0.091	235,766,864	2,813,157,888	239,496,195
8	3,010,889,344	134,871,760	0.091	233,146,752	3,183,145,472	235,155,913
9	3,395,413,562	193,478,128	0.089	231,159,440	3,578,989,056	232,006,863
10	3,740,144,404	161,202,912		227,970,288	3,953,183,488	

Table 7
Optimum Contributions Using Haberman and Sung (1994) with $T = 15$

t	F_t	B_t	i_t	NC_t	AL_t	C_t^*
0	373,211,585	106,636,560	0.076	264,658,176	585,530,240	413,807,424
1	731,995,223	48,948,364	0.101	254,203,072	851,652,800	318,112,256
2	1,102,432,124	89,903,272	0.067	250,750,880	1,152,353,024	275,453,824
3	1,374,074,884	95,409,192	0.072	247,360,208	1,461,198,464	273,680,352
4	1,663,767,839	67,699,656	0.090	243,759,120	1,775,253,248	255,104,096
5	2,018,193,822	76,361,992	0.106	240,339,776	2,101,083,136	228,293,248
6	2,400,078,024	65,474,592	0.100	238,243,632	2,454,753,024	186,703,904
7	2,772,486,847	102,758,760	0.081	235,766,864	2,813,157,888	187,585,408
8	3,088,821,372	134,871,760	0.094	233,146,752	3,183,145,472	228,123,600
9	3,479,795,126	193,478,128	0.080	231,159,440	3,578,989,056	216,944,064
10	3,783,522,098	161,202,912	0.055	227,970,288	3,953,183,488	222,646,480
11	4,057,552,838	209,269,152	0.081	225,249,024	4,343,038,464	288,066,880
12	4,472,912,082	186,510,336	0.080	221,934,336	4,712,074,752	247,495,104
13	4,896,608,792	222,970,272	0.080	219,645,904	5,146,928,640	221,517,792
14	5,286,769,028	225,303,168	0.076	216,970,176	5,565,712,384	139,811,344
15	5,599,154,131	225,283,040		214,121,360	5,987,062,784	

Table 8
Optimum Contributions Using Chang (1999) with $T = 15$

t	F_t	B_t	i_t	NC_t	AL_t	C_t^*
0	373,211,585	106,636,560	0.076	264,658,176	585,530,240	350,178,999
1	663,540,043	48,948,364	0.101	254,203,072	851,652,800	313,132,211
2	1,021,568,503	89,903,272	0.067	250,750,880	1,152,353,024	293,796,067
3	1,307,374,409	95,409,192	0.072	247,360,208	1,461,198,464	281,704,276
4	1,600,879,698	67,699,656	0.090	243,759,120	1,775,253,248	270,224,146
5	1,966,115,867	76,361,992	0.106	240,339,776	2,101,083,136	259,003,289
6	2,376,445,902	65,474,592	0.100	238,243,632	2,454,753,024	249,622,442
7	2,815,687,083	102,758,760	0.081	235,766,864	2,813,157,888	242,331,492
8	3,194,703,571	134,871,760	0.094	233,146,752	3,183,145,472	237,606,661
9	3,605,954,219	193,478,128	0.080	231,159,440	3,578,989,056	233,316,809
10	3,937,456,489	161,202,912	0.055	227,970,288	3,953,183,488	229,179,783
11	4,226,892,750	209,269,152	0.081	225,249,024	4,343,038,464	226,784,402
12	4,589,761,801	186,510,336	0.080	221,934,336	4,712,074,752	222,857,052
13	4,996,197,398	222,970,272	0.080	219,645,904	5,146,928,640	220,173,079
14	5,392,872,435	225,303,168	0.076	216,970,176	5,565,712,384	217,197,027
15	5,796,679,404	225,283,040		214,121,360	5,987,062,784	

Table 9
Optimum Contributions Using Haberman and Sung (1994) with $T = 20$

t	F_t	B_t	i_t	NC_t	AL_t	C_t^*
0	373,211,585	106,636,560	0.057	264,658,176	585,530,240	413,807,424
1	718,864,907	48,948,364	0.061	254,203,072	851,652,800	326,569,440
2	1,056,906,812	89,903,272	0.083	250,750,880	1,152,353,024	304,776,480
3	1,377,752,643	95,409,192	0.063	247,360,208	1,461,198,464	271,311,552
4	1,651,177,015	67,699,656	0.067	243,759,120	1,775,253,248	263,213,824
5	1,971,131,272	76,361,992	0.082	240,339,776	2,101,083,136	258,606,032
6	2,330,794,365	65,474,592	0.080	238,243,632	2,454,753,024	231,329,568
7	2,696,381,396	102,758,760	0.080	235,766,864	2,813,157,888	236,605,056
8	3,056,646,538	134,871,760	0.084	233,146,752	3,183,145,472	248,848,608
9	3,435,770,039	193,478,128	0.107	231,159,440	3,578,989,056	245,305,776
10	3,859,855,541	161,202,912	0.058	227,970,288	3,953,183,488	173,510,544
11	4,097,383,674	209,269,152	0.062	225,249,024	4,343,038,464	262,651,472
12	4,410,027,561	186,510,336	0.085	221,934,336	4,712,074,752	289,607,296
13	4,898,313,408	222,970,272	0.103	219,645,904	5,146,928,640	232,266,864
14	5,414,224,333	225,303,168	0.083	216,970,176	5,565,712,384	144,879,232
15	5,774,126,863	225,283,040	0.088	214,121,360	5,987,062,784	182,089,664
16	6,202,151,121	340,608,224	0.088	212,082,800	6,427,823,616	176,958,464
17	6,570,279,274	384,232,000	0.050	209,096,448	6,821,858,816	157,806,992
18	6,658,007,323	422,253,632	0.091	206,210,160	7,183,346,176	271,490,016
19	7,096,278,014	510,401,472	0.080	202,645,072	7,431,084,544	104,960,320
20	7,226,104,098	564,830,528		199,309,184	7,703,323,648	

Table 10

Optimum Contributions Using Chang (1999) with T = 20

t	F_t	B_t	i_t	NC_t	AL_t	C_t^*
0	373,211,585	106,636,560	0.057	264,658,176	585,530,240	364,317,179
1	666,575,492	48,948,364	0.061	254,203,072	851,652,800	328,944,297
2	1,003,965,740	89,903,272	0.083	250,750,880	1,152,353,024	313,678,298
3	1,330,043,764	95,409,192	0.063	247,360,208	1,461,198,464	297,754,027
4	1,628,575,732	67,699,656	0.067	243,759,120	1,775,253,248	285,976,833
5	1,971,303,895	76,361,992	0.082	240,339,776	2,101,083,136	275,267,053
6	2,349,014,952	65,474,592	0.080	238,243,632	2,454,753,024	265,528,632
7	2,752,994,621	102,758,760	0.080	235,766,864	2,813,157,888	257,003,484
8	3,139,819,127	134,871,760	0.084	233,146,752	3,183,145,472	249,404,354
9	3,526,500,242	193,478,128	0.107	231,159,440	3,578,989,056	243,021,150
10	3,957,741,588	161,202,912	0.058	227,970,288	3,953,183,488	234,354,560
11	4,265,346,112	209,269,152	0.062	225,249,024	4,343,038,464	231,038,516
12	4,554,893,586	186,510,336	0.085	221,934,336	4,712,074,752	226,938,696
13	4,987,526,269	222,970,272	0.103	219,645,904	5,146,928,640	222,891,400
14	5,502,303,373	225,303,168	0.083	216,970,176	5,565,712,384	218,218,211
15	5,948,870,582	225,283,040	0.088	214,121,360	5,987,062,784	214,753,310
16	6,427,793,436	340,608,224	0.088	212,082,800	6,427,823,616	212,199,100
17	6,854,136,765	384,232,000	0.050	209,096,448	6,821,858,816	208,924,539
18	7,009,570,609	422,253,632	0.091	206,210,160	7,183,346,176	206,373,028
19	7,408,653,370	510,401,472	0.080	202,645,072	7,431,084,544	202,667,418
20	7,668,993,166	564,830,528		199,309,184	7,703,323,648	

Figure 1
Optimal Contribution Ratios (C_t/NC_t)
And Funding Ratios (F_t/AL_t) for $T = 10$ Under CIPM and RIPM

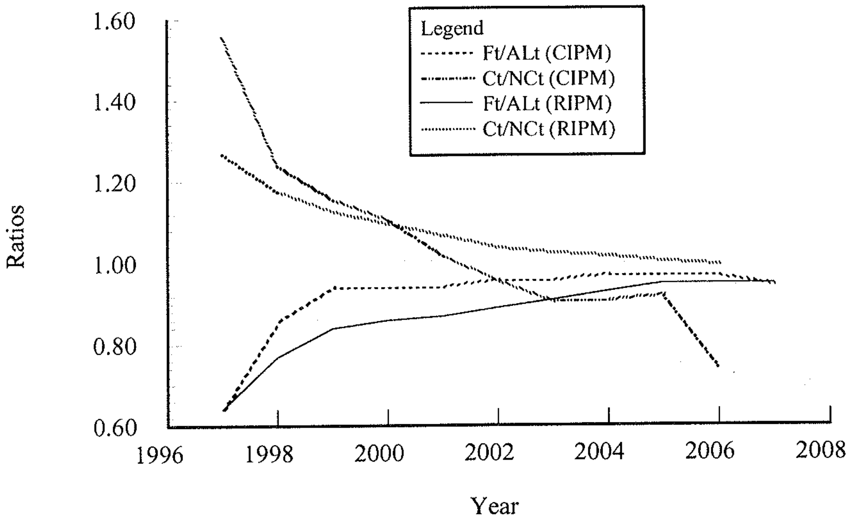


Figure 2
Optimal Contribution Ratios (C_t/NC_t)
And Funding Ratios (F_t/AL_t) for $T = 15$ Under CIPM and RIPM

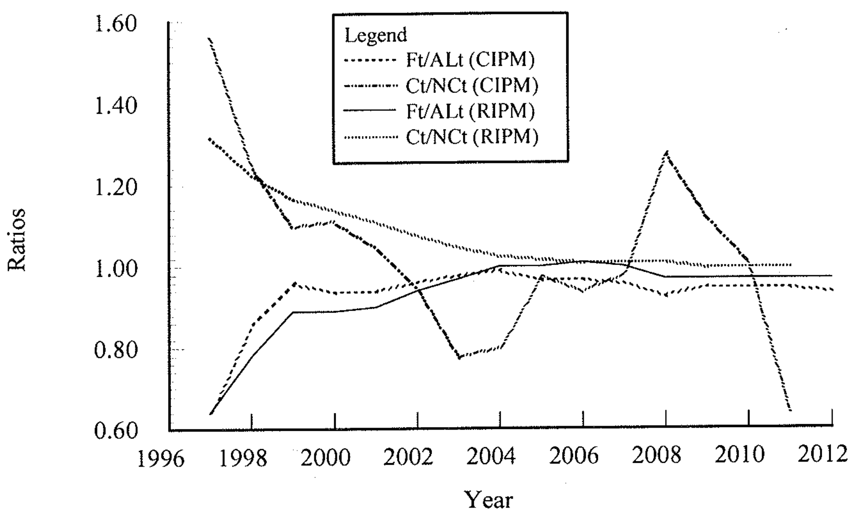


Figure 3
Optimal Contribution Ratios (C_t/NC_t)
And Funding Ratios (F_t/AL_t) for T = 20 Under CIPM and RIPM

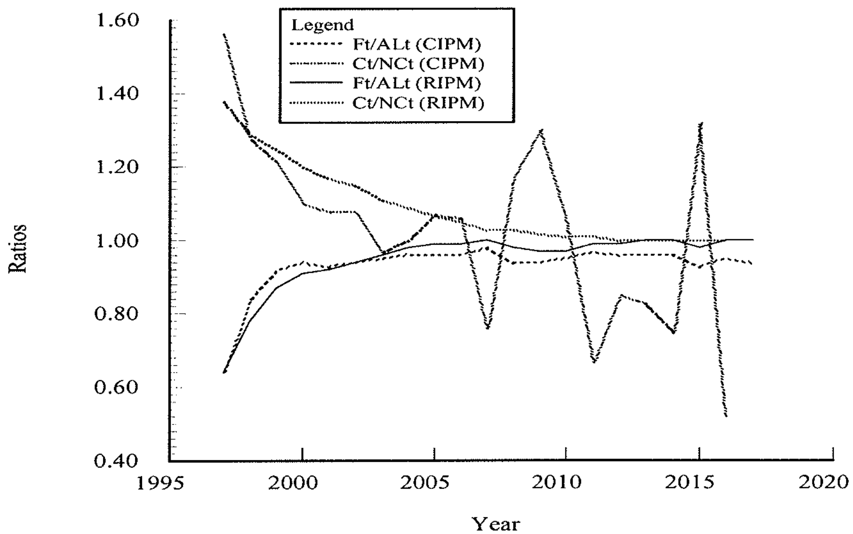


Figure 4
Optimal Contributions (C_t) for T = 10, 15 and 20 Under CIPM

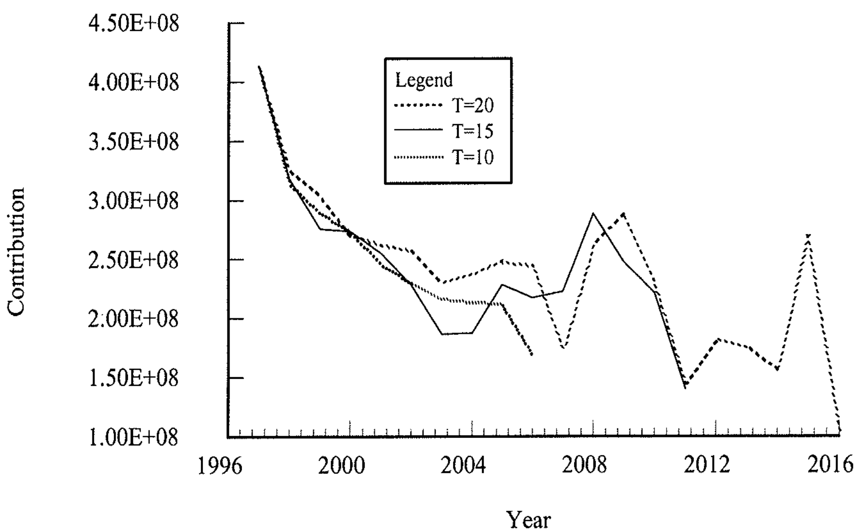


Figure 5
Optimal Contributions (C_t) for $T = 10$, 15 and 20 Under RIPM

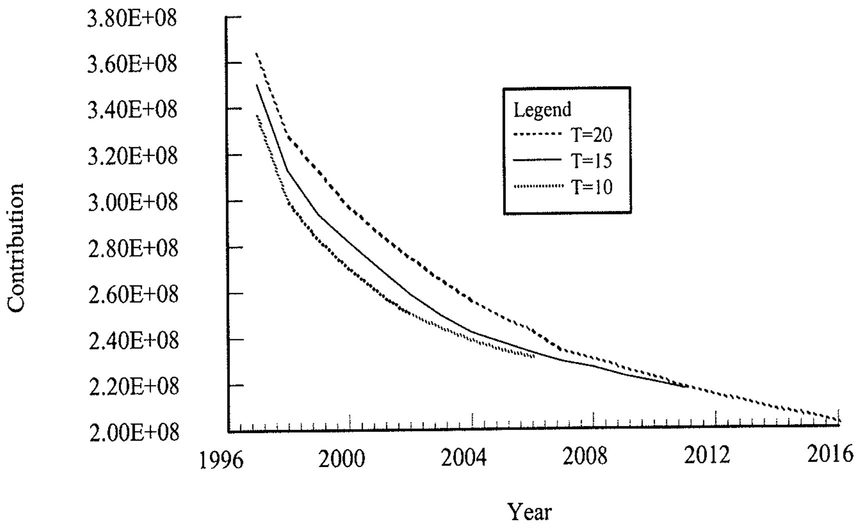
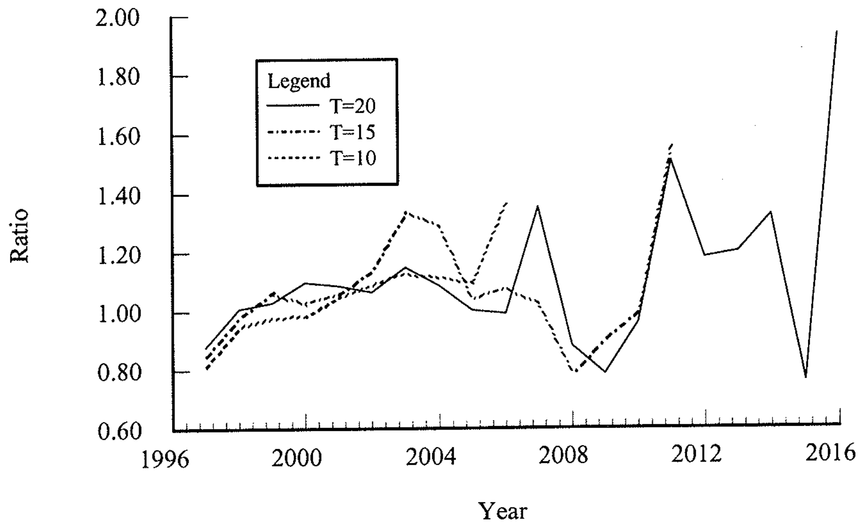


Figure 6
Ratio of Optimal Contributions ($C_t^{\text{RIPM}}/C_t^{\text{CIPM}}$) for $T = 10$, 15 and 20



Some empirical observations worth noting are the following:

1. If RIPM is chosen for its robustness with respect to the variation cross years, then a long-term stable decreasing trend would be found in contribution rates.

2. When CIPM is adopted, much more variable contribution rates may occur. Based on our analysis, a longer projection period results in much more unstable contribution rates.

3. There is no clear better performance measure for funding ratios in *RIPM* and CIPM within different time frames.

4. The effects on these ratios are diminished when we reduce the length of the planning horizon. It explains that the effects due to policy intervention in contribution ratios are significant in longer time frames under both performance measures.

5. RIPM is a relatively stable strategy for decision making that provides more consistent and smoother optimal solutions under policy intervention.

5 Closing Comments

As the percentage of the population past retirement age increases, pension-related topics have taken on a new significance, and much attention has been focused on the implementation of better retirement systems. We hope the approach presented here can be used in this effort.

A summary of the advantages of this approach is listed below:

1. With the ready availability of today's high speed computers, the plan administrator can forecast the plan's future cash flows;

2. The optimal contribution can be estimated under various scenarios based on specific plan investment and recruiting strategies;

3. The optimal funding and actuarial status of the plan can be estimated under specific performance measure implemented through a computerized system; and

4. Running an extensive set of scenarios will clarify the interaction between the plan liability and the investment performance.

As we mentioned previously, Figures 4 and 5 show that the size of the optimal contributions in any year decreases as we extend the length of the planning period. One area for further study is to determine if it is better to use a single planning period of length $2n$ years (20 years, say) or use two planning periods of n years (10 years) each.

References:

Anderson, A.W. *Pension Mathematics For Actuaries.* Winsted, Conn.: Actex Publications, 1990.

Bacinello, A.R. "A Stochastic Simulation Procedure for Pension Scheme." *Insurance: Mathematics and Economics* 7 (1988): 153-161.

Bowers, N.L., Gerber, H. U., Hickman, J. C., Jones, D. A. and Nesbitt, C. J. "Notes On the Dynamics of Pension Funding." *Insurance: Mathematics and Economics* 1 (1982): 261-270.

Chang, S.C. "Optimal Pension Funding Through Dynamic Simulations: the Case of Taiwan Public Employees Retirement System." *Insurance: Mathematics and Economics* 24 (1999): 187-199.

Daykin, C.D., Pentikainen, T. and Pesonen, M. *Practical Risk Theory for Actuaries*, Monographs on Statistics and Applied Probability 53, London, U.K.: Chapman and Hall, 1994.

Dufresne, D. "Moments of Pension Fund Contributions and Fund Levels when Rates of Return are Random." *Journal of the Institute of Actuaries* 115 (1988): 535-544.

Dufresne, D. "Stability of Pension Systems When Rates of Rreturn Are Random." *Insurance: Mathematics and Economics* 6 (1989): 129-134.

Haberman, S. "Pension Funding with Time Delays: A Stochastic Approach." *Insurance: Mathematics and Economics* 11 (1992): 179-89.

Haberman, S. "Pension Funding with Time Delays and Autoregressive Rates of Investment Return." *Insurance: Mathematics and Economics* 13 (1993): 45-56.

Haberman, S. "Autoregressive Rates of Return and the Variability of Pension Contributions and Fund Levels for a Defined Benefit Pension Scheme." *Insurance: Mathematics and Economics* 14 (1994): 219-240.

Haberman, S. and Sung, J. H. "Dynamic Approaches to Pension Funding." *Insurance: Mathematics and Economics* 15 (1994): 151-162.

Haberman, S. and Wong, L. Y. "Moving Average Rates of Return and the Variability of Pension Contributions and Fund Levels for A Defined Benefit Pension Scheme." *Insurance: Mathematics and Economics* 20 (1997): 115–135.

Martin-Löf, A. "Lectures on the Use of Control Theory in Insurance." *Scandinavian Actuarial Journal* 1 (1994): 1-25.

O'Brien, T. "A Stochastic-Dynamic Approach to Pension Funding." *Insurance: Mathematics and Economics* 5 (1986): 141–146.

O'Brien, T. "A Two-Parameter Family of Pension Contribution Functions and Stochastic Optimization." *Insurance: Mathematics and Economics* 6 (1987): 129–134.

Runggaldier, W.J. "Concept and Methods for Discrete and Continuous Time Control under Uncertainty." *Insurance: Mathematics and Economics* 22 (1998): 25–39.

Schäl, M. "On Piecewise Deterministic Markov Control Processes: Control of Jumps and of Risk Processes in Insurance." *Insurance: Mathematics and Economics* 22 (1998): 75–91.

Shapiro, A.F. "Contributions to the Evolution of Pension Cost Analysis." *Journal of Risk and Insurance* 52 (1985): 81–99.

Journal of Actuarial Practice Vol. 8, 2000

Risk Sources in a Life Annuity Portfolio: Decomposition and Measurement Tools

Mariarosaria Coppola,* Emilia Di Lorenzo,† and Marilena Sibillo‡

Abstract§

The paper considers a model for a homogeneous portfolio of whole life annuities immediate. The aim is to study two risk factors: the investment risk and the insurance risk. A stochastic model of the rate of return is used to study these risk factors. Measures of the insurance risk and the investment risk for the entire portfolio are suggested. The problem of the longevity risk is presented, and its consequences with different projections of the mortality tables are analyzed. The model is applied to some concrete cases, and several illustrations show the importance of the two components of the riskiness in terms of the number of policies in the portfolio. Understanding these risks will allow insurance companies to control, to some extent, the overall risk of their annuity portfolios.

Key words and phrases: *Ornstein-Uhlenbeck process, investment risk, insurance risk, longevity risk, moments of insurance functions*

*Mariarosaria Coppola, Ph.D., is a researcher in the area of stochastic analysis of financial and insurance risk at the University of Naples "Federico II".

Dr. Coppola's address is: Dipartimento di Matematica e Statistica, Facolta' di Economia, Universita' degli Studi di Napoli "Federico II", via Cintia, Complesso Monte S.Angelo 80126 Napoli, ITALY. Internet address: *coppola@dms.unina.it*

†Emilia Di Lorenzo, Ph.D., is an associate professor of financial mathematics at the University of Naples "Federico II". She has been grant-holder at the Italian National Research Council and researcher at the University of Salerno.

Dr. Di Lorenzo's address is: Dipartimento di Matematica e Statistica, Facolta' di Economia Universita' degli Studi di Napoli "Federico II", via Cintia, Complesso Monte S.Angelo 80126 Napoli, ITALY. Internet address: *diloremi@dmsna.dms.unina.it*

‡Marilena Sibillo, Ph.D., is an associate professor of financial mathematics at the University of Sassari.

Dr. Sibillo's address is: Istituto Economico e Aziendale, Facolta' di Economia, Universita' degli Studi di Sassari, Via Rolando 1 07100 Sassari, ITALY. Internet address: *sibillo@ssmain.uniss.it*

§Work performed as part of the research project "Interazione fra rischio finanziario e rischio assicurativo: modelli e strumenti di controllo" of the Programma di Ricerca di Interesse Nazionale 1998 "Modelli per la gestione di rischi finanziari, assicurativi e operativi."

1 Introduction

Most of the problems faced by an insurer managing a portfolio of life insurance policies are based on the investment risk (due to interest rates) and insurance risk (due to mortality) and on their interactions. Because of the nature of these risks, most of the research has been done on the present value of a single policy within a framework whereby both interest rates and mortality are random. Recently the focus has shifted to similar problems concerning an entire portfolio of policies. Among the contributions in this area are Norberg (1993), Parker (1993), (1994a), (1994b), (1996), and (1997), and Frees (1998).

Norberg (1993) gave the first two moments of the present value of stochastic payment streams and applied them to a portfolio of temporary insurance contracts. Parker (1993) studied moments of the present value of future cash flows modeling the force of interest by (i) a white noise, (ii) a Wiener process, and (iii) an Ornstein-Uhlenbeck process. Parker found moments of the present value of a portfolio of benefits relating to life policies (1994a) and endowment insurance policies (1994b) by modeling the force of interest using a Vasicek model; see Vasicek (1977). Parker (1996) proposed two methods to obtain the limiting distribution of the present value of a portfolio of benefits. Parker (1997) provided an interesting paper on the interaction between investment and insurance risks for a portfolio of life insurance policies with random curtate future lifetimes. Using the Vasicek model for the rate of return Parker considered the variance as a measure of the riskiness of a portfolio and divided it into insurance and investment risks. Frees (1998) showed the utility of the coefficient of determination for quantifying the relative importance of each source of uncertainty where there are more than two sources of risks.

The aim of the paper is to study the risk of an annuity portfolio by dividing this risk into two components: an investment risk and an insurance risk. We offer some ways of controlling these by means of the variability measures of the expected value of the life annuities portfolio with respect to each of these two components.

In dealing with a portfolio of life insurance policies, it is well-known that the effect of accidental deviations of mortality can be reduced by using pooling techniques. But as pointed out in Marocco and Pitacco (1998) and Olivieri (1998), however, in the case of a portfolio of life annuities, a phenomenon not controllable by pooling techniques is the *longevity risk*, which is the systematic deviations of the actual number of deaths from the expected number of deaths due to the improvements in future mortality. The longevity risk produces actuarial losses in the case of a life annuity portfolio, while in the case of life insurance contracts it produces actuarial gains. For these reasons it seems particularly useful to include suitable projections of mortality improvements in the case of a life annuity portfolio.

In Section 2 we propose the random variables in a portfolio of homogeneous whole life annuities immediate and we obtain the first two moments of the present value of the portfolio and of the average cost per policy. Section 3 presents a description of the stochastic process used to model the instantaneous rate of return, while in Section 4 we consider the two sources of risk and their measures for the entire portfolio; the longevity risk is introduced also. In Section 5, the model is applied and several illustrations concerning the importance of the two components of the riskiness, as they relate to the number of policies in portfolio, are presented.

2 Portfolio of Life Annuities

Let us consider a portfolio of c homogeneous whole life annuity-immediate policies. These policies are assumed to have been issued to c lives each age x and pay an annual benefit of one unit payable at the end of each year to each of the survivors. For $i = 1, 2, \dots, c$, let T_i be the random variable representing the curtate-future-lifetime of the ith life insured and let Z_i be the random variable representing the present value of the lifetime annuity benefits for the ith annuitant:

$$Z_i = \begin{cases} 0 & \text{if } T_i = 0; \\ \sum_{h=1}^{T_i} e^{-y(h)} & \text{if } T_i = 1, 2, \dots, \end{cases} \tag{1}$$

where:

$$y(t) = \int_0^t \delta_s \, ds, \quad t > 0,$$

with δ_s being the random instantaneous rate of return at time s that is used for discounting the payments.

Moreover we suppose (see, for example, Bowers et al., 1987, Chapters 3 and 8, and Parker 1994a) that the following assumptions hold:

(i) For $i = 1, 2, \dots, c$, the T_is are independent and identically distributed;

(ii) Given knowledge of $y(h)$ for $h = 1, 2, \ldots$, the Z_is are independent and identically distributed for $i = 1, 2, \ldots, c$; and

(iii) For $i = 1, 2, \ldots, c$, the T_is and δ_s are mutually independent.

The random Z_i variables are independent only when conditioning on the knowledge of the sequence of $y(h)$s for $h = 1, 2, \ldots$. In general they are not independent, as the same rates of return are used for discounting the payments.

For our valuations it is necessary to compute the first and the second moments of Z_i that are:

$$E[Z_i] = E[E[Z_i \mid T_i]] = \sum_{h=1}^{\infty} {}_h p_x E[e^{-y(h)}] \tag{2}$$

$$E[Z_i^2] = \sum_{h=1}^{\infty} {}_h p_x E[e^{-2y(h)}] + 2\sum_{h=2}^{\infty} {}_h p_x \sum_{r=1}^{h-1} E[e^{-y(r)} e^{-y(h)}]. \tag{3}$$

The proof of equation (3) is easily derived as follows:
Proof:

$$\begin{aligned}
E[Z_i^2] &= E[E[Z_i^2 \mid \{y(h)\}_{h=1}^{\infty}]] \\
&= \sum_{h=1}^{\infty} E[(\sum_{k=1}^{h} e^{-y(k)})^2]\, {}_{h|1}q_x \\
&= \sum_{h=1}^{\infty} E[(\sum_{k=1}^{h} e^{-y(k)})^2]({}_h p_x - {}_{h+1}p_x) \\
&= E[e^{-2y(1)}]p_x + \sum_{h=1}^{\infty} \left\{ E[(\sum_{k=1}^{h+1} e^{-y(k)})^2] - E[(\sum_{k=1}^{h} e^{-y(k)})^2]\, {}_{h+1}p_x \right\} \\
&= E[e^{-2y(1)} p_x + \sum_{h=2}^{\infty} {}_h p_x (\sum_{r=1}^{h-1} 2e^{-y(r)} e^{-y(h)} + e^{-2y(h)})]
\end{aligned}$$

and equation (3) holds. □

Let $Z(c)$ denote the total present value for the entire portfolio of c annuities, i.e.,

$$Z(c) = \sum_{i=1}^{c} Z_i . \tag{4}$$

The first two moments of $Z(c)$ are:

$$E[Z(c)] = c \sum_{h=1}^{\infty} {}_hp_x E[e^{-y(h)}] \tag{5}$$

$$\begin{aligned} E[Z(c)^2] &= E[\sum_{i=1}^{c} Z_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{c} Z_i Z_j] \\ &= \sum_{i=1}^{c} E[Z_i^2] + \sum_{\substack{i,j=1 \\ i \neq j}}^{c} E[Z_i Z_j]. \end{aligned} \tag{6}$$

Next we need an expression for $E[Z_i Z_j]$. But, by virtue of assumptions (i), (ii), and (iii) (Parker 1994a),

$$\begin{aligned} E[Z_i Z_j] &= E[E[Z_i Z_j \mid \{y(h)\}_{h=1}^{\infty}]] \\ &= E[E[Z_i \mid \{y(h)\}_{h=1}^{\infty}] E[Z_j \mid \{y(h)\}_{h=1}^{\infty}]] \\ &= E[E[Z_1 \mid \{y(h)\}_{h=1}^{\infty}] E[Z_2 \mid \{y(h)\}_{h=1}^{\infty}]] \\ &= E[Z_1 Z_2] \\ &= E[\sum_{h=1}^{T_1} e^{-y(h)} \sum_{k=1}^{T_2} e^{-y(k)}] \\ &= E[E[\sum_{h=1}^{T_1} e^{-y(h)} \sum_{k=1}^{T_2} e^{-y(k)} \mid \{y(r)\}_{r=1}^{\infty}]] \\ &= E[\sum_{h=1}^{\infty} {}_hp_x e^{-y(h)} \sum_{k=1}^{\infty} {}_kp_x e^{-y(k)}] \\ &= \sum_{h=1}^{\infty} \sum_{k=1}^{\infty} {}_hp_x \, {}_kp_x E[e^{-y(h)-y(k)}]. \end{aligned} \tag{7}$$

Therefore equation (6) can be written as:

$$E[Z(c)^2] = cE[Z_i^2] + \sum_{\substack{i,j=1 \\ i \neq j}}^{c} \sum_{h=1}^{\infty} \sum_{k=1}^{\infty} {}_h p_x {}_k p_x E[e^{-y(h)-y(k)}]$$
$$= cE[Z_i^2] + c(c-1) \sum_{h=1}^{\infty} \sum_{k=1}^{\infty} {}_h p_x {}_k p_x E[e^{-y(h)-y(k)}]. \tag{8}$$

Finally, from equations (5) and (8), we can obtain the variance of $Z(c)$.

For our analysis it will be useful to consider the average cost per policy, $Z(c)/c$, of the portfolio under consideration.

3 Stochastic Rate of Return

One of the problems facing insurance companies is the financial risk arising from fluctuations of their rate of return. To investigate this problem we follow Di Lorenzo, Sibillo, and Tessitore (1997) and model the instantaneous global rate of return ($Y(t)$) as a sum of two components: a deterministic component ($\delta(t)$) and a stochastic component ($X(t)$) that describes the deviations of the instantaneous global rate of return from its expected value, $\delta(t)$. This means that $Y(t)$ can be written as:

$$Y(t) = \delta(t) + X(t). \tag{9}$$

We suppose that $\delta(t)$ is determined by forecasts based on the existing investments. In addition, $\{X(t), 0 \leq t < +\infty\}$ is an Ornstein-Uhlenbeck process, with parameters $\beta > 0$ and $\sigma > 0$ and initial value $X(0) = 0$. $X(t)$ is characterized by the following stochastic differential equation:

$$dX(t) = -\beta X(t)dt + \sigma dW(t) \tag{10}$$

where $W(t)$ is a standard Wiener (Brownian motion) process.

It follows from equation (9) that the stochastic present value at time 0 of a payment of one monetary unit at time t is given by:

$$
\begin{aligned}
e^{-y(t)} &= e^{-\int_0^t Y(s)ds} \\
&= e^{-\int_0^t (\delta(s)+X(s))ds} \\
&= v(t)F(t)
\end{aligned} \tag{11}
$$

where

$$v(t) = e^{-\int_0^t \delta(s)ds} \tag{12}$$

and

$$F(t) = e^{-\int_0^t X(s)ds}. \tag{13}$$

Clearly $v(t)$ is the deterministic discounting factor and $F(t)$ is the stochastic discounting factor. $F(t)$ is log normally distributed with parameters $-E[\int_0^t X(s)ds]$, and $\mathrm{Var}[\int_0^t X(s)ds]$ and its rth moment about the origin is given by the formula

$$E[(F(t))^r] = \exp\{-rE[\int_0^t X(s)ds] + \frac{1}{2}r^2\mathrm{Var}[\int_0^t X(s)ds]\}. \tag{14}$$

Using the fact that $E[X(t)] = 0$ and letting:

$$\phi(t) = \mathrm{Var}[\int_0^t X(s)ds] \tag{15}$$

we obtain (Crow and Shimizu 1988):

$$E[F(t)] = e^{\frac{1}{2}\phi(t)} \tag{16}$$

and

$$\mathrm{Var}[F(t)] = e^{\phi(t)}[e^{\phi(t)} - 1]. \tag{17}$$

Finally, according to Di Lorenzo, Sibillo, and Tessitore (1997), the autocovariance function can be written as follows:

$$\mathrm{Cov}[F(h), F(k)] = e^{\frac{1}{2}(\phi(h)+\phi(k))}[e^{\Phi(h,k)} - 1] \tag{18}$$

where:

$$\Phi(h,k) = \mathrm{Cov}[\int_0^h X(s)ds, \int_0^k X(s)ds].$$

4 Measures of Sources of Uncertainty

As Frees (1998) points out, it is important to identify the factors affecting the total risk. To this end, we will consider mortality and stochastic interest as risk factors and make actuarial valuations using an instantaneous total rate of return (interest income plus capital gains and losses) represented by the stochastic process defined in equations (9) and (10). Moreover, we will take into account the mortality component, both relating to the riskiness caused by random mortality deviations, and to the riskiness caused by improvements in mortality trend.

After identifying the risk factors, we must study ways to manage them. The risk control tools are different depending on the risk components considered. For example,

- The risk due to random deviations of the numbers of deaths from their expected values can be controlled by means of pooling techniques and reinsurance;
- The investment risk can be controlled by various well-known financial risk management techniques such as immunization techniques and hedging strategies (Frees 1998); and
- The longevity risk (due to an improved mortality trend) can be controlled by using projected mortality tables that are constructed on the basis of forecasts of the future mortality trend (Marocco and Pitacco 1998 and Olivieri 1998).

In light of the above considerations, it is important to quantify the contribution of each risk factor to the total riskiness of the portfolio. It is for this purpose that we want to study the mortality and investment components of the life annuity portfolio considered in Section 2.

4.1 Insurance and Investment Risk Measures

For valuation purposes, it seems reasonable to adopt a simple measure of the two risk components affecting the portfolio. We adopt a well-known formula for the decomposition of the variance and apply it to the variance of the present value of the annuity portfolio.

First we observe that $\text{Var}[Z(c)]$, the variance of the present value of the portfolio considered in our study, can be decomposed in two ways as follows (Parker 1997):

$$\text{Var}[Z(c)] = E[\text{Var}[Z(c) \mid \{T_i\}_{i=1}^{c}]] + \text{Var}[E[Z(c) \mid \{T_i\}_{i=1}^{c}]] \qquad (19)$$

and

$$\mathrm{Var}[Z(c)] = E[\mathrm{Var}[Z(c) \mid \{y(k)\}_{k=1}^{\infty}]] + \mathrm{Var}[E[Z(c) \mid \{y(k)\}_{k=1}^{\infty}]]. \tag{20}$$

In equation (19), $\mathrm{Var}[E[Z(c) \mid \{T_i\}_{i=1}^{c}]]$ provides a measure of the variability of $Z(c)$ caused by cash flows connected to random events (mortality, survival), after averaging out the effect of the stochastic discounting factors. Thus, we have the following definition:

Definition 1. *The insurance risk measure is* $\mathrm{Var}[E[Z(c) \mid \{T_i\}_{i=1}^{c}]]$.

Analogously, $E[\mathrm{Var}[Z(c) \mid \{T_i\}_{i=1}^{c}]]$ is an average over cash flows connected to random events of the variability in $Z(c)$ due to the stochastic rate of return, and it can be considered as an investment risk measure. In equation (20), however, $\mathrm{Var}[E[Z(c) \mid \{y(k)\}]]$ is a measure of the variability of $Z(c)$ due to the effect of the stochastic discounting factors as the effect of random events connected with mortality and survival have been averaged out, so it is a measure of the investment risk. Thus, we have the following definition:

Definition 2. *The investment risk measure is* $\mathrm{Var}[E[Z(c) \mid \{y(k)\}_{k=1}^{\infty}]]$.

We choose equation (20) for our valuations, because, as Parker (1997) explains, it allows us to clearly relate the risk components to the number of policies. We get:

$$\begin{aligned}\mathrm{Var}[E[Z(c) \mid \{y(k)\}_{k=1}^{\infty}]] &= \mathrm{Var}[E[\sum_{i=1}^{c} Z_i \mid \{y(k)\}_{k=1}^{\infty}]] \\ &= \mathrm{Var}[c \sum_{h=1}^{\infty} {}_hp_x e^{-y(h)}] \\ &= c^2 \sum_{h=1}^{\infty} \sum_{k=1}^{\infty} {}_hp_x {}_kp_x \mathrm{Cov}[e^{-y(h)}, e^{-y(k)}] \end{aligned} \tag{21}$$

also given by:

$$\begin{aligned}\mathrm{Var}[E[Z(c) \mid \{y(k)\}_{k=1}^{\infty}]] &= c^2 \sum_{h=1}^{\infty} \sum_{k=1}^{\infty} {}_hp_x {}_kp_x E[e^{-y(h)-y(k)}] \\ &\quad - (c \sum_{h=1}^{\infty} {}_hp_x E[e^{-y(h)}])^2\end{aligned}$$

and

$$
\begin{aligned}
E[\mathrm{Var}[Z(c) \mid \{y(k)\}_{k=1}^{\infty}]] &= E[\mathrm{Var}[\sum_{i=1}^{c} Z_i \mid \{y(k)\}_{k=1}^{\infty}]] \\
&= E[c\mathrm{Var}[Z_i \mid \{y(k)\}_{k=1}^{\infty}]] \\
&= cE[E[Z_i^2 \mid \{y(k)\}_{k=1}^{\infty}]] \\
&\quad - cE[(E[Z_i \mid \{y(k)\}_{k=1}^{\infty}])^2]. \qquad (22)
\end{aligned}
$$

With regard to the average cost per policy, $Z(c)/c$, we get:

$$
\mathrm{Var}[E[\frac{Z(c)}{c} \mid \{y(k)\}_{k=1}^{\infty}]] = \sum_{h=1}^{\infty}\sum_{k=1}^{\infty} {}_hp_x {}_kp_x \mathrm{Cov}(e^{-y(h)}, e^{-y(k)}) \qquad (23)
$$

and

$$
\begin{aligned}
E[\mathrm{Var}[\frac{Z(c)}{c} \mid \{y(k)\}_{k=1}^{\infty}]] &= \frac{1}{c}(E[E[Z_i^2 \mid \{y(k)\}_{k=1}^{\infty}]] \\
&\quad - E[[E[Z_i \mid \{y(k)\}_{k=1}^{\infty}]]^2]). \qquad (24)
\end{aligned}
$$

4.2 The Longevity Risk

Together with the risk due to accidental deviations of death frequencies from their expected values, the improvements of mortality trends at adult ages have consequences on all life insurance contracts. As life annuities are contracts pertaining to survival benefits, the calculation of present values should be based on mortality tables with built-in mortality projections, because unexpected improvements in future mortality at the older ages could result in an underestimation of future costs and result in actuarial losses.

Definition 3. *The longevity risk is the systematic deviation of the actual number of deaths from their expected values across the older ages.*

By analyzing mortality trend in terms of survival functions, two aspects known as *rectangularization* and *expansion* emerge. *Rectangularization* refers to the higher concentration of deaths around the mode of the curve of deaths, lowering the risk for the insurer. *Expansion* refers to the random advancement of the mode of curve of deaths toward the ultimate life time (Olivieri and Pitacco 1999) and hence a higher risk for the insurer. Longevity risk is the result of rectangularization and expansion acting jointly (Marocco and Pitacco 1998). It can be mitigated by using projected mortality tables;, that is, tables constructed on the basis of a forecast of the future mortality trend (Pitacco 1998).

5 Numerical Illustrations

Let us consider a portfolio of c whole life annuities immediate as described in Section 2. We will quantify the insurance and investment risks on the basis of equations (21) to (24) and four different mortality tables.

Following Olivieri (1998), we assume that the basic distribution of future lifetimes can be represented by a Weibull distribution, i.e., the survival function from age 0 to age x, $s(x)$, is given by:

$$s(x) = e^{-(x/\alpha)^{\gamma}}, \quad x > 0,$$

where $\alpha > 0$ and $\gamma > 0$ are constant parameters. The projected survival function from age 0 to age x is also assumed to follow a Weibull distribution. The basic mortality table and the three projected tables with increasing survival probabilities are based on the parameters α and γ suggested by Olivieri (1998). These parameter values are given below.

Parameter Values

Survival Tables	α	γ
Basic	82.7	7.00
Pessimistic Projection	83.5	8.00
Realistic Projection	85.2	9.15
Optimistic Projection	87.0	10.45

The parameters β and σ of the force of interest process (equation (9)) used in our calculations are determined in a manner similar to Di Lorenzo, Sibillo, and Tessitore (1997). As the Ornstein-Uhlenbeck process, $X(t)$, (equation (9)) represents the deviations of the force of interest from its expected values, we use the differences between the actual observed rates and the corresponding forecasted rates. Then by means of the covariance equivalence principle (Pandit and Wu 1983 and Parker 1994), we can estimate β and σ from these differences.

Using data from Italian short-term (three months) bonds, regularly reported in *Statistical Bulletin*, we obtain $\delta = 0.09$, $\beta = 0.11$, and $\sigma = 0.005$.

Tables 1 and 2 show the mean, variance, investment risk component, and insurance risk component of the present value of a portfolio of c annuities issued at age 65. Table 1 is based on $c = 15$, while Table 2 is based on $c = 1000$.

Tables 3 and 4 show the mean, variance, investment risk component, and insurance risk component of the present value of the average cost per policy of a portfolio of c annuities issued at age 65. Table 3 is based on $c = 15$, while Table 4 is based on $c = 1000$.

Tables 5 and 6 show the mean, variance, investment risk component, and insurance risk component of the present value of a portfolio of c annuities issued at age 45. Table 5 is based on $c = 15$, while Table 6 is based on $c = 1000$.

Tables 7 and 8 show the mean, variance, investment risk component, and insurance risk component of the present value of the average cost per policy of a portfolio of c annuities issued at age 45. Table 7 is based on $c = 15$, while Table 8 is based on $c = 1000$.

Table 1
Present Value of Annuity Portfolio at Age 65 with c = 15

		Projections		
	Basic	Pessimistic	Realistic	Optimistic
$E[Z(c)]$	106.654	110.001	114.706	120.257
$\mathrm{Var}[Z(c)]$	199.384	196.662	196.012	197.376
$\mathrm{Var}[E[Z(c)\|\{y\}]]$	94.698	102.631	114.973	131.174
$E[\mathrm{Var}[Z(c)\|\{y\}]]$	104.686	94.031	81.039	66.202

Table 2
Present Value of Annuity Portfolio at Age 65 with c = 1000

		Projections		
	Basic	Pessimistic	Realistic	Optimistic
$E[Z(c)]$	7110.24	7333.41	7647.04	8017.12
$\mathrm{Var}[Z(c)]$	427861.00	462405	516394.00	587408.00
$\mathrm{Var}[E[Z(c)\|\{y\}]]$	420882.00	456136.00	510992.00	582995.00
$E[\mathrm{Var}[Z(c)\|\{y\}]]$	6979.00	6269.00	5402.00	4413.00

From Tables 1 and 2 we observe that the mean value of Z(c) increases with the projection; the global variance, for c = 15, decreases, except for the optimistic projection, while it always increases for c = 1000. Analyzing the two risk components we note that for both values of c the financial risk increases with the projection, while the insurance risk decreases.

Tables 3 and 4 show a similar behavior to Tables 1 and 2, respectively. The numerical results for the global variance are confirmed if we study it as function of the c:

Table 3
Present Value of Average Cost per Policy at Age 65 with $c = 15$

		Projections		
	Basic	Pessimistic	Realistic	Optimistic
$E[\frac{Z(c)}{c}]$	7.11024	7.33341	7.64704	8.01712
$\text{Var}[\frac{Z(c)}{c}]$	0.88614	0.87404	0.87116	0.87725
$\text{Var}[E[\frac{Z(c)}{c} \mid \{y\}]]$	0.42088	0.45613	0.51099	0.58299
$E[\text{Var}[\frac{Z(c)}{c} \mid \{y\}]]$	0.46526	0.41791	0.36017	0.29426

Table 4
Present Value of Average Cost per Policy at Age 65 with $c = 1000$

		Projections		
	Basic	Pessimistic	Realistic	Optimistic
$E[\frac{Z(c)}{c}]$	7.11024	7.33341	7.64704	8.01712
$\text{Var}[\frac{Z(c)}{c}]$	0.42786	0.46240	0.51639	0.58740
$\text{Var}[E[\frac{Z(c)}{c} \mid \{y\}]]$	0.42088	0.45613	0.51099	0.58299
$E[\text{Var}[\frac{Z(c)}{c} \mid \{y\}]]$	0.00698	0.00627	0.00540	0.00441

$$\begin{aligned}
\text{Var}[\frac{Z(c)}{c}]_{\text{pess}} &= -53.7790 + \frac{60.5039 + 54.2351(c-1)}{c} \\
&= 0.4561 + \frac{6.2688}{c} \\
\text{Var}[\frac{Z(c)}{c}]_{\text{real}} &= -58.4772 + \frac{64.3908 + 58.9882(c-1)}{c} \\
&= 0.5110 + \frac{5.4026}{c} \\
\text{Var}[\frac{Z(c)}{c}]_{\text{opt}} &= -64.2742 + \frac{69.2707 + 64.8572(c-1)}{c} \\
&= 0.5830 + \frac{4.4135}{c}.
\end{aligned}$$

So the variance related to the pessimistic projection is greater than the variance related to the realistic projection for $c < 16$; moreover, the variance related to the realistic projection is greater than the variance related to the optimistic projection for $c < 14$.

Table 5
Present Value of Annuity Portfolio at Age 45 with $c = 15$

		Projections		
	Basic	Pessimistic	Realistic	Optimistic
$E[Z(c)]$	143.506	145.974	148.913	151.390
$\mathrm{Var}[Z(c)]$	263.391	264.082	269.890	276.967
$\mathrm{Var}[E[Z(c)\mid\{y\}]]$	227.782	239.457	254.678	268.497
$E[\mathrm{Var}[Z(c)\mid\{y\}]]$	35.609	24.625	15.212	8.470

Table 6
Present Value of Annuity Portfolio at Age 45 with $c = 1000$

		Projections		
	Basic	Pessimistic	Realistic	Optimistic
$E[Z(c)]$	9567.07	9731.6	9927.55	10092.7
$\mathrm{Var}[Z(c)]$	1014670	1065890	1132920	1193880
$\mathrm{Var}[E[Z(c)\mid\{y\}]]$	1012360	1064250	1131900	1193320
$E[\mathrm{Var}[Z(c)\mid\{y\}]]$	2310	1640	1020	560

For all values of c, the financial risk increases and the insurance risk decreases when the projection increases. We observe that the decreasing behavior of the insurance risk is stronger when the number of policies is small. From a mathematical point of view, we can justify this behavior by means of equation (24) in which the dependence of $E[\mathrm{Var}[\frac{Z(c)}{c} \mid \{y(k)\}]]$ on c is evident.

For every fixed survival table, the global variance of $\frac{Z(c)}{c}$ decreases as c increases. In particular, the financial risk takes the same value (from equation (23) we see that $\mathrm{Var}[E[\frac{Z(c)}{c} \mid \{y(k)\}]]$ does not depend on c), while the insurance risk decreases to zero as c tends to infinity (see equation (24)).

We can repeat analogous considerations about Tables 5, 6, 7, and 8. Observe that for $x = 45$ the global variance always increases; in fact we have:

Table 7
Present Value of Average Cost per Policy at Age 65 with $c = 15$

		Projections		
	Basic	Pessimistic	Realistic	Optimistic
$E[\frac{Z(c)}{c}]$	9.56706	9.73160	9.92753	10.0926
$\mathrm{Var}[\frac{Z(c)}{c}]$	1.17062	1.17369	1.19951	1.23096
$\mathrm{Var}[E[\frac{Z(c)}{c} \mid \{y\}]]$	1.01236	1.06425	1.13190	1.19332
$E[\mathrm{Var}[\frac{Z(c)}{c} \mid \{y\}]]$	0.15826	0.10944	0.06761	0.03764

Table 8
Present Value of Average Cost per Policy at Age 65 with $c = 1000$

		Projections		
	Basic	Pessimistic	Realistic	Optimistic
$E[\frac{Z(c)}{c}]$	9.56706	9.73160	9.92753	10.0926
$\mathrm{Var}[\frac{Z(c)}{c}]$	1.01467	1.06589	1.13292	1.19388
$\mathrm{Var}[E[\frac{Z(c)}{c} \mid \{y\}]]$	1.01236	1.06425	1.13190	1.19332
$E[\mathrm{Var}[\frac{Z(c)}{c} \mid \{y\}]]$	0.00231	0.00164	0.00102	0.00056

$$
\begin{aligned}
\mathrm{Var}[\frac{Z(c)}{c}]_{\mathrm{pess}} &= -94.5385 + \frac{97.2269 + 95.599(c-1)}{c} \\
&= 1.0605 + \frac{1.6279}{c} \\
\mathrm{Var}[\frac{Z(c)}{c}]_{\mathrm{real}} &= -98.3670 + \frac{100.497 + 99.495(c-1)}{c} \\
&= 1.1280 + \frac{1.0020}{c} \\
\mathrm{Var}[\frac{Z(c)}{c}]_{\mathrm{opt}} &= -101.6260 + \frac{103.368 + 102.814(c-1)}{c} \\
&= 1.1880 + \frac{0.5460}{c}.
\end{aligned}
$$

The variance related to the pessimistic projection is greater than the variance related to the realistic projection for $c < 10$; moreover, the variance related to the realistic projection is greater than the variance related to the optimistic projection for $c < 8$.

6 Summary and Concluding Remarks

We have analyzed and quantified two risk sources for a portfolio of life annuities: the investment risk and the insurance risk. This analysis was done in a framework in which both mortality and rates of returns are random.

The global rate of return is modeled as the sum of two components: a deterministic one, which considers the existing investments of the company, and a stochastic one, representing the deviations of the real rate of return from its anticipated values. The stochastic component is an Ornstein-Uhlenbeck process with a mean reversion level of zero.

We also consider the longevity risk, the risk due to the improvements in mortality trend. The effects of the mortality improvements are investigated using different projected mortality tables.

On the basis of the numerical examples presented, we may conclude that the insurance risk decreases when the projection increases. On the other hand, the financial risk increases when the projection increases, because the company could be exposed for a longer period to a risk of systematic nature. Moreover, the mean value of the present value of the cash flows connected to the portfolio increases when the projection increases, because the insurer could bear bigger costs.

In conclusion, the numerical results presented in Section 6 show how the use of projected mortality tables allows the insurer to front the risk of greater costs and how the exposure to the financial risk and to the insurance risk varies, depending on the longevity of the lives insured.

One area for future research is the development of the model presented in the paper, focusing on the effect of the randomness of the projections in the valuations concerning the considered portfolio. Such research can lead to the determination of the systematic risk component due to the type of randomness depicted by the survival functions used for constructing mortality tables.

References.

Arnold, L. *Stochastic Differential Equations: Theory and Applications.* New York, N.Y.: John Wiley & Sons, 1974.

Beekman, J.A. and Fuelling, C.P. "Interest and Mortality Randomness in Some Annuities." *Insurance: Mathematics and Economics* 9 (1990): 185–196.

Beekman, J.A. and Fuelling, C.P. "Extra Randomness in Certain Annuity Models." *Insurance: Mathematics and Economics* 10 (1991): 275-287.

Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A., and Nesbitt, C.J. *Actuarial Mathematics.* Schaumburg, Ill.: Society of Actuaries, 1986.

Crow, E.L. and Shimizu, K. (Editors) *Lognormal Distribution, Theory and Applications.* New York, N.Y.: Marcel Dekker, Inc., 1988.

Di Lorenzo, E., Sibillo, M., and Tessitore, G. "A Stochastic Model for Financial Evaluations. Applications to Actuarial Contracts." *Proceedings of the Second International Conference on "Applied Stochastic Models and Data Analysis."* Held in Capri, Italy. (1997): 115-120.

Frees, E.W. "Stochastic Life Contingencies with Solvency Considerations." *Transactions of the Society of Actuaries* 42 (1990): 91-148.

Frees, E.W. "Relative Importance of Risk Sources in Insurance Systems." *North American Actuarial Journal* (1998): 34-52.

Gard, T.C. *Introduction to Stochastic Differential Equations. Pure and Applied Mathematics.* New York, N.Y.: Marcel Dekker, Inc., 1988.

Marocco, P. and Pitacco, E. (1998) "Longevity Risk and Life Annuity Reinsurance." *Transactions of the 26th International Congress of Actuaries.* Held in Birmingham, England. 6 (1998): 453-479.

Norberg, R. "A Solvency Study in Life Insurance." In *Proceedings of the 3rd AFIR International Colloquium* (1993): 821-830.

Norberg, R. "Stochastic Calculus in Actuarial Science." *Surveys in Applied and Industrial Mathematics,* 1997.

Olivieri, A. "Per una Quantificazione del Longevity Risk." *Atti del Convegno AMASES.* Held in Genova, Italy. (1998): 383-399.

Olivieri, A. and Pitacco, E. *Funding Sickness Benefits for the Elderly.* Dipartimento di Matematica Applicata alle Scienze Ec., Stat. e Att. "B. De Finetti," Università di Trieste, Trieste, Italy, 1999.

Pandit, S.M. and Wu, S.M. *Time Series and System Analysis with Applications.* New York, N.Y.: John Wiley and Sons, 1983.

Parker, G. "Distribution of the Present Value of Future Cash Flows." In *Proceedings of the 3rd AFIR International Colloquium* (1993): 831-840.

Parker, G. "Moments of the Present Value of a Portfolio of Policies." *Scandinavian Actuarial Journal* (1994a): 53-67.

Parker, G. "Stochastic Analysis of a Portfolio of Endowment Insurance Policies." *Scandinavian Actuarial Journal* (1994b): 119-130.

Parker, G. "A Portfolio of Endowment Policies and its Limiting Distribution." *Astin Bulletin* 26, no. 1, (1996): 25-33.

Parker G. "Stochastic Analysis of the Interaction Between Investment and Insurance Risks." *North American Actuarial Journal* (1997): 55-84.

Pitacco, E. *Tavole di mortalità proiettate e loro impieghi in ambito attuariale. Appunti introduttivi.* Quaderni del Dipartimento di Matematica Applicata alle Scienze Economiche, Statistiche e Attuariali "B. de Finetti" n. 1/1997, Università di Trieste, Trieste, Italy, 1997.

Vasicek, O. "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics* 5, no. 2, (1977): 177-188.

A Comparative Study of the Performance of Loss Reserving Methods through Simulation

Prakash Narayan* and Thomas Warthen†

Abstract‡

Actuaries are often asked to provide a range or confidence level for the loss reserve along with a point estimate. Traditional methods of loss reserving do not provide an estimate of the variance of the estimated reserve, and actuaries use various ad hoc methods to derive a range for the indicated reserve. We use a Monte Carlo simulation method to compare various loss reserve estimation methods, including traditional methods and regression-based methods of loss reserving.

Key words and phrases: *loss development factor, loss triangle, severity, reporting delays, regression, loss ratio*

---

*Prakash Narayan, Ph.D., A.C.A.S., M.A.A.A., is an actuary at RLI Insurance. He is a graduate of Banaras Hindu University, India and obtained his Ph.D. degree from Monash University, Australia. Dr. Narayan has published several papers in statistical journals. His current research interests include loss reserving methods, property catastrophe loss management, and dynamic financial models of insurance. Dr. Narayan worked for Towers Perrin Reinsurance and Alexander & Alexander as a consultant. His work experience includes loss reserving, development of catastrophe models, and setting retentions and designing reinsurance programs. He is also a member of the American Academy of Actuaries Task Force on Risk-Based Capital.

Dr. Narayan's address is: RLI Insurance, 9025 N Lindbergh Drive, Peoria IL 61615-1431, USA. Internet address: *pnayaran@rlicorp.com*

†Tom Warthen, F.C.A.S., M.A.A.A., is vice president of actuarial services and director of RLI Insurance Group. He earned his B.S. degree in mathematics from the Johns Hopkins University. Prior to joining RLI, Tom was an actuarial consultant at Towers Perrin/Tillinghast, as well as an actuary at Balis & Company, both located in Philadelpia, Pennsylvania.

Mr. Warthen's address is: RLI Insurance, 9025 N Lindbergh Drive, Peoria IL 61615-1431, USA. Internet address: *twarthen@rlicorp.com*

‡The authors would like to thank Janet E. Duncan, F.C.A.S., and an anonymous referee for many helpful suggestions that significantly improved the quality of the presentation.

1 Introduction

Loss reserving, or projecting losses to their ultimate value, is an important actuarial function. The loss development factor (LDF) method attempts to estimate the pattern with which losses for a given cohort of claims change over time. This method produces a point estimate of the required reserve and is the most commonly used actuarial technique for projecting losses to their ultimate value. Actuaries are often asked to provide a range or the variability associated with the point estimate of the loss reserve.

Mack (1994) developed a methodology to estimate the variability of the estimated loss reserves when the LDF method is used. His method may not be appropriate in many situations, however, as the selection of the development factors is often judgmental. Holmberg (1994) has also presented a model by which actuaries can estimate the variability of their loss reserve estimates. Regression modeling of the loss triangle, which can provide both a point estimate and the variability associated with the point estimate, is receiving increasing attention from actuaries. Regression methods provide an estimator of the variance more directly. These methods, however, are rarely used by actuaries because of the methods' complexity. It is desirable to thoroughly test a new methodology before it can be accepted as an appropriate technique and used in practice. Comparisons of forecasting methods based on historical data are not generally considered an objective method for testing forecasting methods. Such studies are likely to be biased by the preference of the investigator.

Alternatively, statistical simulation is a well-accepted technique for comparing various methods of estimation when the properties of the estimators cannot be studied analytically. Stanard (1985) used this technique to compare various traditional methods of loss reserving. We shall apply the same technique to compare the traditional methods with the regression method of loss reserving. Our study uses a variety of methods to simulate the loss triangles.

We have selected the LDF method as one to compare because it is the most commonly used traditional actuarial method. We have included the Bühlmann complementary loss ratio method (which Standard refers to as the additive model), because this method was the best of the tested methods per the Stanard (1985) study. We compare these loss reserve estimation methods and regression methods. The various loss reserve estimation regression models considered in this study differ in the number of the parameters used in modeling the loss triangle.

Our approach is to simulate random loss triangles with a variety of methods and estimate the corresponding loss reserves using the loss development method, Bühlmann complementary loss ratio method, and log-regression models. We assume that the ultimate losses (and hence the reserves) are known with certainty. We compute the deviations of the estimated reserves from the actual reserves derived by various methods. We expect this deviation to be small for a good reserving method. We use several criteria to compare the estimated deviations of actual versus estimated reserves under the various reserving methods.

In the second section the particular methods of simulating random loss triangles are described. We do not claim that these methods capture all the intricacies of the claims process. Our methods also do not generate loss triangles that incorporate the effects of structural changes in the loss process. We also require that incremental losses be positive in our generated triangles. In reality, this constraint may be violated in some actual loss triangles. We believe, however, that our methods generate loss data triangles that are stochastic and do not provide an apparent advantage to any particular method of loss reserve estimation. A particular method of reserve estimation may incorporate some underlying assumptions about the claims process and will obviously provide a better estimate of the loss reserve if those assumptions are valid. In practice it may not be possible to test the assumptions underlying a particular loss reserve estimation method. If a statistical test is applied, it can only detect a gross violation of the assumptions and cannot confirm that those assumptions are true.

Loss development factor methods have an extensive history of use in actuarial practice that preceded the investigation and documentation of the assumptions underlying these methods. Given the current and historical familiarity with loss development factor methods, the assumptions underlying these methods are in some sense secondary to the methods themselves. Given their widespread historical use and technical adequacy as loss development estimation methods, loss development factor methods would be used by actuaries even if no studies about the underlying assumptions were ever published. This is a major consideration, which leads us to use a variety of methods to simulate the random loss triangles.

We are comparing a traditional loss development factor loss reserve estimation method, the Bühlmann complementary loss ratio method, and three fixed regression loss estimation models to estimate the loss reserves. These methods and models are briefly described below. We also discuss the criteria used to compare the results of the simulations. One can definitely define comparison criteria other than those used here. The criteria used are comprehensive, and an estimator performing better in the criteria considered will likely be a good estimator with respect to other reasonable criteria. We have also provided a brief summary of the results of the simulations for the aggregate loss reserves in this section. Appendix A provides the individual accident year results of our computations. We end with several observations of the results and some conclusions based on this simulation study.

2 Simulating Random Loss Triangles

Modeling a claims process to generate the random elements of a loss triangle is complicated. There does not appear to be any study that derives a severity distribution for losses where the individual loss amount may change over time. Stanard (1985) and Pentikäinen and Rantala (1995) describe methods of simulating random loss triangles. Their methods are fundamentally different. The Stanard method is based on a loss severity distribution of individual claim amounts whereas Pentikäinen and Rantala use an aggregate stochastic claim process.

The various methods of loss triangle simulation used here do not satisfy the assumptions underlying the various methods of loss reserving compared. For example, for log-regression modeling, it is assumed that the incremental losses are independent. This assumption is violated by all the methods used for simulating the random loss triangles. Similarly the random loss triangle simulation methods do not satisfy the basic requirement of the LDF method that the future development is determined by the latest available data. One can infer that our study tests the robustness of the various methods of loss reserving against data sets that do not conform to the assumptions underlying the reserve estimation methods.

We have used four different techniques for simulating the loss triangles. The Pentikäinen and Rantala (1995) method is one of them. As we shall see later, the log-regression method of loss reserving requires that the incremental losses be positive. If this is not the case, some subjective judgments need to be made. One way to treat such incidences is to delete such observations from the data set. To be uniform and consistent, we have selected loss triangle simulation methods that will generate positive incremental losses. Stanard's method does not satisfy this requirement and is not used.

For all the methods in this study, 11 accident years are considered. It is further assumed that the losses completely mature at the 11th year of development, i.e., the first accident year is at the ultimate loss level and no further development is expected. Because we require complete knowledge of the ultimate losses for a proper comparison of the reserve estimation results of the different estimation methods, we generate a complete history for each accident year. In estimating the reserves, only the top half of the loss triangle is available to the actuary as data. The top half is used to estimate the lower half of the triangle, particularly the last (right) column, which represents the projection of ultimate losses. For $i, j = 1, 2, \ldots$, let $S_{i,j}$ denote the incremental losses for the accident year i at the end of the calendar year $i + j - 1$, and let $L_{i,j}$ denote the cumulative losses for the accident year i at the end of the calendar year $i + j - 1$, i.e.,

$$L_{i,j} = \sum_{k=1}^{j} S_{i,k}.$$

The ultimate loss for accident year i, L_i, is given by

$$L_i = \lim_{j \to \infty} L_{i,j}.$$

For simplicity the losses are assumed to be fully developed after 11 years, i.e., $L_i = L_{i,11}$. In addition, we consider only 11 accident years.

2.1 Random Reporting Factor

The steps of this random loss triangle generation method for accident year i ($i = 1, \ldots, 11$) are:

Step 1: Generate N_i, the number of losses for accident year i, as a Poisson random variable with mean 100.

Step 2: Generate N_i claim amount variables $\{C_{i,1}, C_{i,2}, \ldots, C_{i,N_i}\}$ where each $C_{i,k}$ is log-normally distributed with parameters $\mu = 7.3659$ and $\sigma = 1.517427$.[1] These parameters correspond to a loss severity mean of 5000 and a coefficient of variation of 3. The ultimate losses for accident year i is

$$L_i = L_{i,11} = 1.06^{(i-1)} \sum_{k=1}^{N_i} C_{i,k}.$$

[1]A random variable X is said to be log-normally distributed with parameters μ and σ if $\ln X$ is normally distributed with mean μ and variance σ^2.

Step 3: Generate ten random numbers $U_{i,j}$, for $j = 1, \ldots, 10$, that are uniform on $(0, 1)$.

Step 4: For $j = 1, \ldots, 10$ compute

$$T_{i,j} = \frac{1}{10} + \frac{1}{2}U_{i,j} + \frac{1}{2}\ln(j) \tag{1}$$

and

$$X_{i,j} = \sum_{k=1}^{j} T_{i,k}. \tag{2}$$

Step 5: The simulated cumulative loss for accident year 1 at lag (delay) j, $L_{1,j}$ is given by

$$L_{i,j} = L_{i,11}(1 - e^{-X_{i,j}}).$$

Note that the accident year losses are inflated by 6 percent per year.

Though this method may look like a development factor model, it does not strictly satisfy the assumptions of the loss development factor model. The ratio of the expected losses $E[L_{i,j+1}]/E[L_{i,j}]$ is a constant, not the conditional expectation. It also does not satisfy the assumption of independence of incremental losses underlying the log regression models of loss reserves.

2.2 Random Backward Development Factor

This method is similar to method 1 except that the factors are computed in reverse order. The steps of the method for accident year i are:

Step 1: Generate N_i, the number of losses for accident year i, as a Poisson random variable with mean 100.

Step 2: Generate N_i claim amount variables $\{C_{i,1}, C_{i,2}, \ldots, C_{i,N_i}\}$ where each $C_{i,k}$ is log-normally distributed with parameters $\mu = 7.3659$ and $\sigma = 1.517427$. The ultimate loss for accident year i is

$$L_i = L_{i,11} = 1.06^{(i-1)} \sum_{k=1}^{N_i} C_{i,k}.$$

Step 3: For $j = 1, 2, \ldots, 10$, generate the log-normal variates $Y_{i,11-j}$ with parameters $\mu_{i,j} = a_j$ and $\sigma_{i,j} = b_j$ where

$$a_j = \frac{(j + (j-1)^2)}{100}$$

and

$$b_j = \frac{(j + (j-1)^2)}{500}.$$

Note that $Y_{i,j}$ is a randomly generated development factor for the development period j to $j+1$.

Step 4: Losses reported at the end of year 10 for the accident year i, $L_{i,10}$, are $L_i / Y_{i,10}$. The reported losses at earlier valuation dates are computed by dividing by $Y_{i,j}$ successively, i.e.,

$$L_{i,j} = \frac{L_{i,j+1}}{Y_{i,j}}, \quad j = 10, 9, \ldots, 2, 1.$$

The a_j and b_j parameters are selected so that $\Pr[Y_{i,j} > 1] = 1 - \epsilon$ for very small ϵ.

2.3 Individual Losses with Changing Severity

This method is based on the ideas of Stanard (1985) and Bühlmann, Schnieper, and Straub (1980). As in Stanard, we assume an exponential delay in reporting and settlement with the added assumption that the severity distribution varies with delay. The claim amounts are assumed to follow a Pareto distribution with parameters λ and θ.[2]

[2]A random variable X is said to have a Pareto distribution with parameters λ and θ if

$$\Pr[X \le x] = 1 - (1 + \frac{x}{\lambda})^{-\theta} \quad x > 0.$$

As each individual claim develops, the percentile level of the individual loss is assumed to remain constant over time but the parameters λ and θ are assumed to change until the claim is settled. In other words, if the kth claim in accident year i is initially of size $C_{i,k}$, the percentile level of the claim is $U_{i,k}$ where

$$U_{i,k} = 1 - \left(1 + \frac{C_{i,k}}{\lambda}\right)^{-\theta}.$$

For $k = 1, 2, \ldots, N_i$, the kth claim in accident year i, $C_{i,k}$, is assumed to have three random characteristics measured from the beginning of the accident year: the date of occurrence, $X_{i,k,1}$, which is a uniform variate on $(0, 1)$; the reporting delay, $X_{i,k,2}$, which is exponentially distributed with mean 2; and the settlement delay, $X_{i,k,3}$, which is exponentially distributed with mean 2. As we require that the ultimate values of a claim be known within 11 calendar years after it occurred, we truncate both $X_{i,k,1} + X_{i,k,2}$ and $X_{i,k,1} + X_{i,k,2} + X_{i,k,3}$ at 11 if they exceed 11. This provides loss amounts for each claim for delays for $j = 1, 2, \ldots, 11$. Specifically, let $r_{i,k}$ and $R_{i,k}$ be nonnegative integers such that

$$r_{i,k} = \min\{\lfloor (X_{i,k,1} + X_{i,k,2}) \rfloor,\ 11\} \tag{3}$$

$$R_{i,k} = \min\{\lfloor (X_{i,k,1} + X_{i,k,2} + X_{i,k,2}) \rfloor,\ 11\} \tag{4}$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. It follows that the kth claim in accident year i is reported in calendar year $i + r_{i,k}$ and is settled in calendar year $i + R_{i,k}$. The estimated loss after delay j is $\hat{C}_{i,k,j}$, which is defined as:

$$\hat{C}_{i,k,j} = \begin{cases} 0 & \text{if } j = 1, 2, \ldots, r_{i,k}; \\ \lambda(j)\left(\dfrac{1}{(1 - U_{i,k})^{1/\theta(j)}} - 1\right) & \text{if } j = r_{i,k} + 1, \ldots, R_{i,k}; \\ \lambda(j)\left(\dfrac{1}{(1 - U_{i,k})^{1/\theta(R_{i,k})}} - 1\right) & \text{if } j = R_{i,k} + 1, \ldots, 11; \end{cases} \tag{5}$$

where $\lambda = \lambda_1$, $\theta = \theta_1$

$$\lambda(j) = 50(20 + j - 1)(1.06)^{j-1}$$

and

$$\theta(j) = (50 - (j - 1))/20.$$

Note that the middle expression for $\hat{C}_{i,k,j}$ in equation (5) can be written as

$$\left(\frac{1}{(1-U_{i,k})^{1/\theta(j)}}-1\right)=\left(\left(1+\frac{C_{i,k}}{\lambda}\right)^{\theta/\theta(j)}-1\right)$$

and that if a claim is settled at delay j then $\hat{C}_{i,j,k}$ remains constant at later valuations.

The estimated claim amount $\hat{C}_{i,k,j}$ increases over time (as j increases). In actual practice, however, the estimated claim amount may decrease from an earlier to a later valuation for some claims. The procedure used here will always increase the severity of the loss from one valuation to the next. This is done to force the incremental losses to be positive.

The steps of the method for accident year i are:

Step 1: Generate N_i, the number of losses for accident year i, as a Poisson random variable with mean 100.

Step 2: Generate N_i claim amount variables $\{C_{i,1}, C_{i,2}, \ldots, C_{i,N_i}\}$ where each $C_{i,k}$ is a Pareto distribution with parameters $\lambda = 1000$ and $\theta = 2.5$. The corresponding $\{U_{i,1}, U_{i,2}, \ldots, U_{i,N_i}\}$ are also determined.

Step 3: For the $k = 1, 2, \ldots N_i$, generate uniform $(0,1)$ variates $X_{i,k,1}$ for the occurrence date and exponential variates $X_{i,k,2}$ and $X_{i,k,3}$ with mean 2 and 5 respectively for the reporting delay and the settlement delay. The quantities $r_{i,k}$ and $R_{i,k}$ are calculated according to equations (3) and (4).

Step 4: Calculate the $\hat{C}_{i,k,j}$s for $j = 1, 2, \ldots 11$. Note that the ultimate loss for accident year i is

$$L_i = L_{i,11} = 1.06^{(i-1)} \sum_{k=1}^{N_i} \hat{C}_{i,k,11}. \tag{6}$$

2.4 Pentikäinen-Rantala Method

This method is based on the procedure described by Pentikäinen and Rantala (1995). Our implementation differs slightly from theirs. We shall describe the computational steps of this method briefly; the reader is encouraged to review the original Pentikäinen-Rantala paper for a complete explanation of their method. The computational steps of this method are:

Step 1: We assume a reporting pattern for a cohort of aggregate losses. This pattern is assumed not to change over time and includes pure IBNR. Specifically, let $X(j)$ denote the proportion of the losses in accident year i reported in calendar year $i+j-1$, $j = 1, 2, \ldots, 11$. The pattern used is $X(1) = 0.220$, $X(2) = 0.180$, $X(3) = 0.150$, $X(4) = 0.120$, $X(5) = 0.100$, $X(6) = 0.080$, $X(7) = 0.060$, $X(8) = 0.040$, $X(9) = 0.027$, $X(10) = 0.016$, $X(11) = 0.007$.

Step 2: Claims for the accident year i reported at delay j are given by

$$S_{i,j} = K \times X(j) \times XP(i) \times q(i,j) \times \text{INF}(i + j - 1) \qquad (7)$$

where K is constant parameter related to the total losses for accident year 1;

$$XP(i) = ((1.01)(1.06))^{i-1} \quad \text{Exposure and inflation growth;}$$
$$q(i,j) = 0.4 + 0.6q(i, j-1) + \epsilon_{i,j}$$

where $q(i,0) = 1$ and $\epsilon_{i,j} \sim N(0, 0.05)$

$$\text{INF}(t) = \prod_{k=1}^{t} (1 + \delta(k));$$
$$\delta(k+1) = \max(0.06 + 0.7(\delta(k) - 0.06) + \omega_k)$$

and $\delta(1) = 0.06$ and $\omega_k \sim N(0, 0.015)$.

This method is based on randomizing the aggregate losses of all the claims for an accident year. Claim reporting and inflation are modeled by autoregressive processes. We further restrict the inflation rate to a minimum of 3 percent. This method also has an exposure growth of 1 percent.

We note that in the simulation of random loss triangles by the methods of Sections 2.1, 2.2, and 2.3, individual claim severity is unlimited. In practice individual losses will have an upper limit in most cases. Occurrence of an individual large loss in the simulation process may cause an individual accident year loss to be out of line with other accident year losses in an individual loss triangle.

It is worth stating that the computations for the simulations were performed in Excel. We have, however, implemented our own module to generate the uniform random variate.

3 Methods of Loss Reserving

One can see that each of the four methods of generating loss triangles in Section 2 has several parameters. As these parameters are changed, the simulated triangles may exhibit significantly different development patterns. A particular method of loss reserving, considered best with a selected set of loss triangle generation parameters, need not be better for any other set of loss triangle generation parameters. The simulation conducted here emphasizes a variety of methods of loss triangle generation rather than the sensitivity of the loss triangle generation methods over a range of possible parameters.

Let us assume that there is no further claim development beyond year n or, equivalently, that $L_{i,n}$ is the ultimate loss value for the accident year i. (Recall in Section 2 that $n = 11$.) Further assume that all $S_{i,j}$ are positive and let

$$Z_{i,j} = \ln(S_{i,j}). \tag{8}$$

To simplify the later exposition of our estimation process, let us further assume that the accident year loss inflation rate is 6 percent and there is no exposure growth except for the Pentikäinen-Rantala method in which constant exposure growth of 1 percent is assumed. Our problem is to estimate $L_{i,j}$ for $i = 1, 2, \ldots, n$ and $j = n+2-i, n+3-i, \ldots, n$ given that $L_{i,j}$ is known for $i = 1, 2, \ldots, n$ and $j = 1, \ldots, n+1-i$.

Two traditional methods of loss reserving and three regression models are used. The two traditional methods are the loss development factor method and the Bühlmann complementary loss ratio method. The loss development factor method is the most commonly used actuarial technique. The Bühlmann loss ratio method was chosen for this analysis because this method outperforms other actuarial methods in the simulation study by Stanard (1985).

The three regression models we have selected for comparison are similar; the differences among them lie in the number of parameters fitted. These methods are described next.

Loss Development We compute:

$$f_{i,j} = \frac{L_{i,j+1}}{L_{i,j}}$$

$$f_j = \frac{1}{n-j} \sum_{i=1}^{n-j} f_{i,j}$$

$$u_k = \prod_{j=k}^{n-1} f_j$$

and the estimated $\bar{L}_{i,n}$ is given by

$$\bar{L}_{i,n} = L_{i,n-i+1} u_{n-i+1}.$$

Bühlmann Complementary Loss Ratio Method This method of loss reserving has not been commonly applied in North America and is suitable for application to paid loss data. It is based on the presumption that the proportion of losses paid at a particular delay remains constant over time. This proportion is estimated from the historical loss experience and is used to forecast the future. We compute:

$$\bar{M}_j = \frac{1}{n-j+1} \sum_{i=1}^{n-j+1} S_{i,j}(1+r)^{n-i} \quad \text{for } j = 2, 3, \ldots, n \text{ and}$$

$$S_{i,j} = \bar{M}_j (1+r)^{i-n} \quad \text{for } j = n+2-i, \ldots, n \text{ and } i = 2, 3, \ldots, n$$

where r is rate of inflation for losses and is assumed to be 6 percent in our simulation.

Regression Models Our discussion of the regression models considered in our analysis is brief. These models are discussed in greater detail by Zehnwirth (1994) and Verrall (1994) among others. We have used an unbiased estimator for the loss reserves as recommended by Verrall (1994) rather than Bayes or maximum likelihood estimates (MLE). In these models the incremental losses are assumed to follow some stochastic distribution. Usually some transformation is applied to the incremental losses before the model parameters are estimated. Although various transformations have been investigated, the logarithmic transformation is most commonly used. Let us describe the methodology briefly with the log transformation for completeness. Readers not familiar with the methodology are encouraged to review the papers by Verrall (1994) and Zehnwirth (1994).

Recall equation (8). We assume that

$$Z_{i,j} = \mu + \alpha_i + \beta_j + \epsilon_{i,j}$$

where μ, α_i, and β_j are the constant parameters of the model and the $\epsilon_{i,j}$ are error terms that are assumed independent identically distributed normal variates with mean 0 and variance σ^2 and are the error terms or the random noise. We make the usual assumption that $\alpha_1 = 0$ and $\beta_1 = 0$ to make the model of full rank. The parameters of the model are estimated by the least squares method. Under the assumption of the normality of the error terms, the estimates are also MLEs. We use the unbiased estimate for the forecasting and require that the errors are independent and normally distributed.

The three regression models investigated in this paper are:

Model 1: α_i and β_j are all different for $i, j = 2, 3, \ldots, n$.

Model 2: $\alpha_i = (i-1)\alpha$ for $i = 2, \ldots, 11$ and β_j are different for $j = 2, 3, \ldots, n$.

Model 3: $\alpha_i = (i-1)\alpha$ and $\beta_j = (j-1)\beta + \gamma \ln(j)$ for $i, j = 2, 3, \ldots, n$.

In the actual application of regression models, one will select the model that provides the best fit to the data based on the evaluation of the residuals and other statistics of the fitted models. Such an approach is not feasible in simulation. Zehnwirth (1994) emphasizes parsimony when applying the regression models for forecasting. The number of parameters used in the three regression models is 21, 12, and 4, respectively. The difference among the three models lies in the number of parameters used to fit the data. As defined, regression model 1 has too many parameters and model 3 too few to capture the essence of a random loss triangle. Model 2 and model 3 assume some underlying relationships among the model 1 parameters. In selecting these regression models, our purpose is not to compare these models with each other, but to see the effect of using fewer parameters in regression modeling.

The parameters are estimated by the least squares method and used to forecast ultimate losses. We refer the reader to Verrall (1994), who provides a complete description of the estimation method and an unbiased estimator of the lower triangle for model 1. The other regression models require revisions to the design matrix and modification of the appropriate equations from those described in Verrall (1994).

4 Comparison of Procedures

We have generated 5000 hypothetical loss triangles for each of the simulation methods described earlier. For each of the 5000 sets of hypothetical data, the reserves are estimated by the loss development method, Bühlmann complementary loss ratio method, and the regression loss reserve estimation methods. The deviations between the loss reserve estimates and the actual reserves are computed.

An important property of a good estimator is that it is unbiased. Stanard (1985) used this criterion for comparing various loss reserve estimators. If an estimator is unbiased, the average deviation of estimated versus actual reserves over many simulations will be negligible.

Between two unbiased estimators, statisticians prefer the estimator with the smaller variance. Between biased estimators, the estimator with the minimum mean square error is preferred. In our context, this means that the average squared deviations between the estimated and actual reserves should be small. This is an important criterion for a reserve estimation method in the insurance context.

The reserves are an important component of the insurer's financial reporting. A reserving method that provides estimates with small biases, but for which the individual simulation (data set) estimates vary a lot from the actual reserves, may not be an appropriate reserve estimation method. One will prefer the reserve estimates to be closer to the true value. We use root mean square error (RMSE) and the average absolute deviation of the estimated versus the actual reserve to test the closeness of the reserve estimators to the actual reserve values. We also compute the average percentage error. A reserve estimation method that generates a smaller percentage error in the estimate is better. Another criterion used to compare the various loss reserving methods is to compute the correlation between the actual reserves and the estimated reserves. One would expect a high correlation for a good reserving method.

We compare the reserve estimates for each of the loss reserve estimation methods, for each of the random loss triangle simulation methods. Our comments follow:

Random Reporting Factor: The Bühlmann complementary loss ratio method is the best loss reserve estimation method for the random reporting factor method of random loss triangle simulation. The regression models of loss reserve estimation perform better than the loss development factor method. The correlation for all the reserving methods is low and surprisingly is smallest for the Bühlmann complementary loss ratio method. Regression model 2 performs slightly better than regression model 1. The main difference between these models is that regression model 2 estimates accident year inflation and allows one parameter for that model component, whereas regression model 1 allows an inflation parameter for each accident year. Our results indicate that parsimony in the regression model is important and that over-parametrization may provide inferior results.

Random Backward Development Factor: Regression model 3 appears to be the best method for this loss simulation method based on aggregate combined accident years' forecast. The Bühlmann complementary loss ratio method is superior in the individual accident year forecasts. Regression model 3 does not capture the payout pattern correctly. The other regression models perform better than the loss development factor method. The Bühlmann method again shows poor correlation with the actual reserves, while the other methods show a reasonable correlation level. We conclude that the Bühlmann method and regression model 2 perform better for this method of random loss triangle simulation than the other tested methods.

Individual Losses with Changing Severity: The loss development method performs well for this loss simulation method. Regression model 2 appears to be better overall. Regression model 3 performs poorly, perhaps because of an insufficient number of model parameters.

Pentikäinen and Rantala Method: Regression models 1 and 2 outperform the other methods. The loss development factor method performs better than the Bühlmann complementary loss ratio method and regression model 3.

Table 1 summarizes our results for each of the four methods of random simulation of the hypothetical loss triangles. Tables 2 through 5 provide similar statistics for individual accident years. These tables show that one of the three regression models considered in this analysis is generally better than the LDF method.

Table 1
Summary of Results the Four Methods
Of Random Simulation of Hypothetical Loss Triangles

	Forecast Method				
	Loss Dev.	Bühlmann	Regression		
	Method	Loss Ratio	Model 1	Model 2	Model 3
Five Thousand Iterations Under Method 1					
Actual Total Reserve: Average = 1,108,298, Std. Dev. = 244,287					
Bias	151,681	5,222	36,486	31,240	51,367
RMSE	466,055	266,874	395,819	328,870	341,537
AAD	364,628	204,674	314,829	254,069	263,444
APE	16.84%	4.84%	6.22%	6.75%	8.69%
CORR	0.25	0.09	0.25	0.15	0.14
Five Thousand Iterations Under Method 2					
Actual Total Reserve: Average = 3,665,734, Std. Dev. = 485,206					
Bias	157,684	(8,088)	55,356	15,393	3,125
RMSE	512,092	639,187	481,727	542,257	519,705
AAD	391,022	485,769	373,282	420,438	403,056
APE	4.38%	1.23%	1.58%	0.79%	0.47%
CORR	0.70	0.11	0.70	0.57	0.58
Five Thousand Iterations Under Method 3					
Actual Total Reserve: Average = 1,634,559, Std. Dev. = 252,631					
Bias	30,566	(83,039)	(144,192)	(52,327)	(176,089)
RMSE	413,137	441,109	375,367	299,099	340,506
AAD	356,932	347,340	314,629	259,057	280,243
APE	1.39%	-4.36%	-9.49%	-3.31%	-9.52%
CORR	0.62	0.39	0.68	0.66	0.32
Five Thousand Iterations Under Method 4					
Actual Total Reserve: Average = 3,183,654, Std. Dev. = 330,776					
Bias	10,106	(21,441)	5,326	4,789	34,136
RMSE	186,688	186,916	183,351	195,148	201,012
AAD	147,536	147,830	145,029	153,675	157,283
APE	0.23%	-0.24%	0.07%	0.06%	0.98%
CORR	0.89	0.84	0.89	0.88	0.88

Notes: Loss Dev. = Loss Development; Std. Dev. = Standard Deviation; RMSE = Root Mean Square Error; AAD = Average Absolute Deviation; APE = Average Percentage Error; and CORR = Correlation between the Actual Reserves and the Estimated Reserves.

Table 2
Random Reporting Factor

	Forecast Method				
	Loss Dev. Method	Bühlmann Loss Ratio	Regression Model 1	Regression Model 2	Regression Model 3
AY	Bias				
2	0	0	(0)	0	(0)
3	1	1	(1)	(1)	3
4	3	3	(5)	(4)	7
5	(1)	3	(26)	(23)	(17)
6	21	12	(68)	(62)	(144)
7	60	(10)	(211)	(158)	(607)
8	371	(60)	(383)	(105)	(1,228)
9	1,365	(121)	(844)	1,135	665
10	9,755	1,141	985	7,587	14,377
11	140,106	4,253	37,041	22,871	38,312
Total	151,681	5,222	36,486	31,240	51,367
AY	RMSE				
2	7	8	4	8	5
3	25	31	17	29	25
4	101	122	72	116	107
5	388	478	291	451	425
6	1,371	1,772	1,066	1,662	1,583
7	4,478	6,041	3,722	5,660	5,390
8	13,713	18,441	12,389	17,657	16,815
9	37,972	51,449	37,716	51,583	50,046
10	110,938	117,820	118,183	125,616	126,393
11	441,193	218,265	368,764	258,292	265,187
Total	466,055	266,874	395,819	328,870	341,537

Notes: Loss Dev. = Loss Development; AY = Accident Year;
RMSE = Root Mean Square Error.

Table 2 (continued)
Random Reporting Factor

	Forecast Method				
	Loss Dev. Method	Bühlmann Loss Ratio	Regression Model 1	Regression Model 2	Regression Model 3
AY	Average Absolute Deviations				
2	5	6	3	6	4
3	19	23	12	21	18
4	74	90	51	85	79
5	288	356	209	332	312
6	1,018	1,295	772	1,198	1,119
7	3,373	4,430	2,750	4,150	3,876
8	10,440	13,784	9,320	13,206	12,359
9	29,290	37,826	28,973	38,281	37,003
10	86,296	88,748	91,752	96,232	97,421
11	346,382	166,051	296,479	197,807	203,655
Total	364,628	204,674	314,829	254,069	263,444
AY	Average Percentage Errors				
2	25.67%	36.00%	9.95%	35.38%	19.36%
3	22.20%	33.13%	9.27%	29.45%	35.19%
4	20.02%	29.68%	8.80%	25.10%	29.48%
5	16.10%	26.60%	7.65%	21.81%	21.16%
6	14.12%	23.31%	7.46%	19.31%	15.14%
7	11.63%	20.28%	6.92%	17.33%	11.96%
8	10.17%	17.67%	7.00%	16.12%	12.08%
9	8.67%	15.01%	6.57%	15.10%	14.19%
10	9.89%	13.16%	7.30%	14.80%	17.14%
11	30.23%	10.57%	12.90%	13.69%	16.36%
Total	16.84%	4.84%	6.22%	6.75%	8.69%

Notes: Loss Dev. = Loss Development; AY = Accident Year;
RMSE = Root Mean Square Error.

Table 3
Random Backward Development Factor

	Forecast Method				
	Loss Dev. Method	Bühlmann Loss Ratio	Regression Model 1	Regression Model 2	Regression Model 3
AY	Bias				
2	2	11	(22)	(22)	16,639
3	119	208	69	40	32,153
4	222	360	156	(54)	38,008
5	694	(470)	534	(1,165)	27,398
6	1,321	(904)	664	(1,643)	2,799
7	3,705	(643)	1,683	(678)	(26,084)
8	8,940	(3,630)	3,237	(2,019)	(47,440)
9	23,597	1,501	12,390	5,744	(37,721)
10	42,679	(1,854)	16,585	5,805	(15,255)
11	76,406	(2,667)	20,062	9,384	12,628
Total	157,684	(8,088)	55,356	15,393	3,125
AY	RMSE				
2	1,564	2,837	1,591	2,786	17,107
3	4,614	9,596	4,791	9,176	33,712
4	11,084	24,442	11,681	22,880	44,274
5	22,705	50,969	24,318	46,958	52,965
6	40,045	87,472	43,509	78,937	77,117
7	64,981	133,834	70,337	121,306	120,900
8	105,093	188,472	111,145	171,608	174,206
9	154,998	222,490	157,449	210,056	208,665
10	231,105	269,200	222,925	271,969	268,947
11	320,345	287,746	293,720	318,988	317,808
Total	512,092	639,187	481,727	542,257	519,705

Notes: Loss Dev. = Loss Development; AY = Accident Year;
RMSE = Root Mean Square Error.

Table 3 (continued)
Random Backward Development Factor

	Forecast Method				
	Loss Dev.	Bühlmann	Regression		
	Method	Loss Ratio	Model 1	Model 2	Model 3
AY	Average Absolute Deviations				
2	1,187	2,082	1,206	2,056	16,639
3	3,553	7,119	3,681	6,848	32,192
4	8,525	18,316	8,996	17,259	39,640
5	17,267	37,913	18,525	34,859	43,380
6	30,623	63,847	33,027	57,937	57,143
7	49,699	99,623	53,807	89,929	86,860
8	78,098	133,685	82,422	122,794	119,958
9	115,161	163,563	118,176	156,441	151,540
10	167,451	193,324	163,199	200,545	196,412
11	233,783	211,057	219,487	240,241	239,523
Total	391,022	485,769	373,282	420,438	403,056
AY	Average Percentage Errors				
2	4.73%	13.54%	4.23%	13.27%	360.75%
3	3.43%	12.41%	3.18%	11.90%	171.65%
4	2.61%	11.08%	2.51%	10.42%	77.29%
5	2.30%	9.75%	2.21%	8.90%	31.55%
6	2.01%	8.80%	1.74%	7.86%	9.62%
7	2.25%	9.06%	1.74%	8.10%	0.62%
8	2.65%	8.33%	1.55%	7.37%	-2.07%
9	4.37%	8.81%	2.64%	8.09%	0.93%
10	6.14%	8.52%	2.75%	8.36%	5.41%
11	9.19%	8.02%	2.65%	8.63%	9.01%
Total	4.38%	1.23%	1.58%	0.79%	0.47%

Notes: Loss Dev. = Loss Development; AY = Accident Year;
RMSE = Root Mean Square Error.

Table 4
Individual Losses with Changing Severity

	Forecast Method				
	Loss Dev.	Bühlmann	Regression		
	Method	Loss Ratio	Model 1	Model 2	Model 3
AY	Bias				
2	1,048	3,668	788	3,751	(9,394)
3	2,867	(5,109)	508	(5,560)	(19,900)
4	(25,022)	(28,645)	(27,849)	(28,329)	(43,090)
5	36,381	(18,287)	17,366	(17,199)	(31,036)
6	29,767	40,903	24,706	39,074	23,710
7	(64,522)	(77,462)	(72,000)	(80,096)	(93,727)
8	40,487	6,886	16,237	6,239	(2,950)
9	3,817	9,224	(7,890)	13,152	8,934
10	8,224	41,085	(9,994)	52,394	46,612
11	(2,480)	(55,304)	(86,063)	(35,753)	(55,246)
Total	30,566	(83,039)	(144,192)	(52,327)	(176,089)
AY	RMSE				
2	31,330	36,022	30,391	35,694	14,678
3	52,869	43,391	51,745	43,108	38,400
4	115,037	126,487	114,408	125,039	118,422
5	171,424	78,184	130,753	78,614	75,765
6	61,614	69,934	58,478	64,412	40,414
7	222,605	230,023	221,536	232,337	262,153
8	110,891	107,524	86,577	95,115	80,851
9	117,997	55,953	119,472	77,183	74,298
10	94,518	97,249	96,401	90,863	86,043
11	259,037	222,586	260,944	173,258	169,985
Total	413,137	441,109	375,367	299,099	340,506

Notes: Loss Dev. = Loss Development; AY = Accident Year;
RMSE = Root Mean Square Error.

Table 4 (continued)
Individual Losses with Changing Severity

	Forecast Method				
	Loss Dev. Method	Bühlmann Loss Ratio	Regression Model 1	Regression Model 2	Regression Model 3
AY	Average Absolute Deviations				
2	19,830	24,090	18,965	24,010	12,509
3	37,669	31,125	36,969	31,640	24,974
4	72,059	77,787	70,388	76,060	62,687
5	105,641	59,338	88,162	61,260	54,471
6	35,254	45,292	33,024	44,186	31,115
7	111,397	120,745	106,008	113,100	119,757
8	94,362	87,744	76,900	78,766	66,943
9	90,948	48,065	88,693	66,927	63,907
10	81,845	75,610	77,164	75,122	66,540
11	202,172	161,009	209,386	135,948	127,446
Total	356,932	347,340	314,629	259,057	280,243
AY	Average Percentage Errors				
2	47.87%	94.00%	44.07%	95.99%	-17.10%
3	49.86%	20.49%	42.22%	20.10%	-13.53%
4	57.18%	63.35%	48.78%	62.45%	16.07%
5	119.19%	12.26%	80.54%	16.42%	0.49%
6	56.20%	77.73%	48.35%	74.57%	48.74%
7	5.03%	-0.81%	0.06%	-3.61%	-7.82%
8	42.12%	29.37%	27.35%	25.73%	17.56%
9	15.32%	8.67%	9.34%	14.34%	13.17%
10	6.43%	21.75%	0.22%	25.76%	23.66%
11	20.04%	3.88%	-2.87%	5.32%	-1.39%
Total	1.39%	-4.36%	-9.49%	-3.31%	-9.52%

Notes: Loss Dev. = Loss Development; AY = Accident Year;
RMSE = Root Mean Square Error.

Table 5
Pentikäinen and Rantala Method

	Forecast Method				
	Loss Dev. Method	Bühlmann Loss Ratio	Regression Model 1	Regression Model 2	Regression Model 3
AY	Bias				
2	2	1	(2)	11	6,037
3	26	(44)	12	11	9,160
4	114	(52)	78	118	9,389
5	283	(217)	213	190	7,237
6	313	(680)	197	162	(755)
7	456	(1,090)	222	473	(9,710)
8	1,225	(2,113)	855	549	(14,271)
9	1,837	(3,756)	1,257	497	(7,450)
10	3,050	(5,083)	1,984	1,456	7,505
11	2,801	(8,407)	509	1,322	26,995
Total	10,106	(21,441)	5,326	4,789	34,136
AY	RMSE				
2	672	608	620	607	6,101
3	1,897	1,724	1,759	1,733	9,399
4	4,070	3,650	3,822	3,687	10,222
5	7,375	6,387	6,997	6,478	9,896
6	12,478	10,855	11,986	10,987	10,942
7	19,710	17,324	19,117	17,606	19,634
8	29,727	26,563	29,027	27,007	29,730
9	40,540	37,157	39,974	37,662	37,813
10	54,967	51,969	54,481	53,464	54,281
11	74,329	71,882	73,934	76,120	82,500
Total	186,688	186,916	183,351	195,148	201,012

Notes: Loss Dev. = Loss Development; AY = Accident Year;
RMSE = Root Mean Square Error.

Table 5 (continued)
Pentikäinen and Rantala Method

	Forecast Method				
	Loss Dev. Method	Bühlmann Loss Ratio	Regression Model 1	Regression Model 2	Regression Model 3
AY	Average Absolute Deviations				
2	536	482	493	480	6,037
3	1,499	1,363	1,388	1,368	9,160
4	3,229	2,912	3,029	2,924	9,401
5	5,839	5,070	5,547	5,105	8,119
6	9,887	8,663	9,495	8,759	8,728
7	15,628	13,742	15,153	13,986	15,864
8	23,414	21,143	22,868	21,350	23,951
9	31,748	29,300	31,319	29,800	30,088
10	43,420	41,011	43,107	41,990	42,469
11	58,767	56,865	58,504	60,077	64,294
Total	147,536	147,830	145,029	153,675	157,283
AY	Average Percentage Errors				
2	0.42%	0.51%	0.32%	0.56%	89.77%
3	0.38%	0.25%	0.29%	0.34%	40.30%
4	0.44%	0.35%	0.35%	0.48%	18.60%
5	0.46%	0.23%	0.37%	0.38%	7.81%
6	0.30%	0.06%	0.21%	0.25%	-0.35%
7	0.25%	0.08%	0.15%	0.30%	-3.72%
8	0.37%	0.00%	0.26%	0.25%	-3.71%
9	0.34%	-0.13%	0.23%	0.15%	-1.36%
10	0.40%	-0.09%	0.25%	0.22%	1.06%
11	0.26%	-0.22%	0.02%	0.16%	2.82%
Total	0.23%	-0.24%	0.07%	0.06%	0.98%

Notes: Loss Dev. = Loss Development; AY = Accident Year;
RMSE = Root Mean Square Error.

The Bühlmann method is slightly better in some cases, but we assumed that the inflation rate is known for the Bühlmann method. We are therefore using additional information for this method and are obtaining slightly better answers. Such information will ordinarily not be available in actual practice. Actual loss data will be tainted by both exposure changes and the inflationary loss cost changes that will vary over time. For most of the methods of random loss generation the effect of inflation has a minimal impact on the ultimate answer derived by traditional methods. Inflation affects the weighting given to individual accident years in the total reserve. For regression models, inflation will affect the forecast in a more complicated fashion.

Although no particular method can be identified as superior in every situation, the regression models generally performed well. It is worth noting that we have not performed a sensitivity analysis of the individual methods of simulating the loss triangles. By changing the inflation rate or the reporting pattern, for example, one may find that the performance of the individual methods of loss reserving will be different. We suspect, however, that the overall performance will be similar.

5 Closing Comments

Regression modeling provides an appropriate tool for estimating loss reserves. Regression methods do not provide the best answers in all situations, but are stable and have the added advantage of providing directly the variance or the confidence interval for the reserve estimate. The regression models studied are a priori fixed. In actual practice, the structure of the models will be determined from a much wider set of possible models based on an analysis of the data under review. Testing and selection of an appropriate loss reserving regression model should improve the ultimate loss reserve forecast in actual application.

Actuaries do not apply the more traditional LDF method blindly. The array of development factors is typically examined carefully before a selection of particular factors entering the reserve estimation is made. The appropriateness of the LDF method is determined for the given data set before the results of any such analysis are accepted. Professional judgment and the selection of an appropriate model are more important when regression loss reserve estimation methods are used. Therefore, an important step is missing for the regression methods as applied in this study. For the Bühlmann method, we assume knowledge of the inflation rate in addition to what is assumed known for other methods. In practice, inflation will not be known precisely, and the loss triangle will be distorted by exposure changes and inflation. This method may therefore not be as well-behaved in practice as in the simulation studies presented here.

The point estimation of the loss reserve has been the primary focus of this study, and we have not considered the variability of loss reserves around the point estimate. Verrall (1994) has outlined the procedures for computing the variance of the forecast including both the forecasting error and the parameter uncertainty.

The overall performance of the LDF method is satisfactory. The closeness of the answers of the various methods assures us that the actuarial methods of loss reserve estimation are generally well behaved. These results also tell us that regression modeling provides estimates similar to traditional actuarial methods, and one should not hesitate to use them. Given the advantage that regression methods also estimate the variability of the estimated reserve, it is expected that their use in the actuarial field will increase.

References

Bühlmann, H., Straub, E. and Schnieper, R. "Claims Reserves in Casualty Insurance Based on a Probabilistic Model." *Mitteilungen der Vereinidung Schweizerischer Versicherungsmathematiker* 1 (1980): 21–46.

Holmberg, R.D. "Correlation and the Measurement of Loss Reserve Variability." *Casualty Actuarial Society Forum* (Spring 1994): 247–277.

Mack, T. "Measuring the Variability of Chain Ladder Reserve Estimates." *Casualty Actuarial Society Forum* (Spring 1994): 101–182.

Pentikäinen, T., and Rantala, J. "A Simulation Procedure for Comparing Different Claims Reserving Methods." *Casualty Actuarial Society Forum* (Fall 1995): 128–156.

Stanard, J.N. "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques." *Proceedings of the Casualty Actuarial Society* 72 (1985): 124–148.

Verrall, R.J. "Statistical Methods for Chain Ladder Techniques." *Casualty Actuarial Society Forum* (Spring 1994): 393–446.

Zehnwirth, B. "Probabilistic Loss Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals and Risk Based Capital." *Casualty Actuarial Society Forum* (Spring 1994): 447–605.

Concentration in the Property and Liability Insurance Market by Line of Insurance

Edward Nissan* and Regina Caveny†

Abstract‡

This paper uses an National Association of Insurance Commissioners (NAIC) 1995 data set to examine the concentration of property and liability insurance by line of insurance in the U.S. The primary measure of concentration used is the Herfindahl index. The largest 100 affiliates are divided into three largest sets of 20, 30, and 50. We find that the homeowners line is the most concentrated line and commercial auto physical damage is the least concentrated line, with the top 20 affiliates commanding the largest between-set and within-set contributions.

Key words and phrases: *Herfindahl index, coefficient of variation, variance decomposition, between-set and within-set contributions, numbers equivalent*

*Edward Nissan is a professor in the Department of Economics and International Business at the University of Southern Mississippi. He holds a Ph.D. in statistics from Texas A & M University. His research areas include income inequality, regional economics, industry structure, and actuarial issues in insurance.

Dr. Nissan's address is: Department of Economics and International Business University of Southern Mississippi, PO Box 5072, Hattiesburg, MS 39406-5072, U.S.A. Internet address: *nissan@cba.usm.edu*

†Regina Caveny is an associate professor in the Department of Economics and International Business at the University of Southern Mississippi. She holds a Ph.D. in management science from the University of Alabama. Her work history includes work as a scientific programmer for a NASA project at the University of Kentucky and as an assistant statistician at the Statistical Consulting Service at Florida State University.

Dr. Caveny's address is: Department of Economics and International Business University of Southern Mississippi, PO Box 5072, Hattiesburg, MS 39406-5072, U.S.A. Internet address: *caveny@cba.usm.edu*

‡The authors are grateful for two referees and the editor for helpful comments and suggestions. The authors also wish to thank Claire Lenz of The National Association of Insurance Commissioners (NAIC) for her help and patience in handling the data on which this research is based.

1 Introduction

1.1 Background

Ng (1995) has observed that structure, conduct, and performance are key elements in comprehending the organization of any industry, a view shared by Adams and Brock (1995) in their survey of 11 major U.S. industries. But while the economics literature is rich in the study of the industrial organization of a majority of businesses in the U.S., Joskow (1973) has noted a lack of such research in the U.S. insurance industry. Joskow attempts to rectify by addressing organizational issues such as business concentration, economies of scale, and ease of entry.

Mayers and Smith (1988) address the issue of the alternative ownership characteristics of property and casualty insurance, indicating that Lloyds, common stock, mutual, and reciprocal are the main types of ownership. Cummins and Weiss (1991) wrote a seminal paper on the subject addressing recent problems of property and liability insurance such as pricing, rate regulation, anticompetitive practices, market concentration, and profitability. More recent contributions are provided by Chidambaran et al., (1997) and Bajtelsmit and Bouzouita (1998).

An important facet of an industry's structure is concentration. Concentration depicts the level of control of an industry by a few dominant firms. For instance, Chidambaran et al., (1997) find that among 18 lines of insurance, concentration levels in private passenger automobile insurance are the highest. Similarly, Bajtelsmit and Bouzouita (1998) examine the relationship between profitability and concentration in automobile insurance. The underlying focus of the research in these papers is the connection of concentration with excess profits through collusion to restrict supply and thus gain artificially high premiums.

Another way to look at concentration[1] is to take into account the relationship between concentration and corporate demand for insurance, an approach exemplified by Mayers and Smith (1982, 1990). They explain that though the demand for insurance by individuals demonstrates risk aversion, the incentives to purchase insurance by corporations, aside from risk aversion, are viewed as being part of financing policies, which include taxes and contracting costs. They also show that among the factors that affect the demand for insurance are business concentration, geographic concentration, and line of business.

[1]The authors are indebted to an anonymous referee for this viewpoint

1.2 Insurance Regulation

To protect against erosion and subversion of competition in the U.S., Congress passed a series of laws to protect consumers. Lereah (1985) explains that 1944 marked a reversal in regulation of insurance premiums, as a consequence of a Supreme Court decision, from primarily regulation by states to federal regulation as well when insurance is transacted across state borders. Subsequently, insurers became subject to antitrust regulation through federal laws such as the Sherman and Clayton Antitrust Acts. Through the McCarren-Ferguson Act of 1945, however, Congress allowed priority for regulation to the states unless effective state regulation was absent. For more on the regulation of insurance companies, see, for example, Black and Skipper (1994, Chapter 34), and Hamilton (1997).

The insurance industry in the last two decades has been accused of price-fixing and anticompetitive and monopoly practices, which may deserve attention by the federal government (Joskow and McLaughlin 1991, Chidambaran *et al.* 1997, Bajtelsmit and Bouzouita 1998). It is an article of faith among most economists that substantial seller concentration affects the social performance of an industry to the detriment of consumers (Caves 1967, Schutz 1995).[2] Brown (1914) made it clear that the tendency toward monopoly and the tendency toward concentration are closely interrelated. Weiss (1983) and Adams and Brock (1990) argue that bigness undermines performance, efficiency, innovation, and technical progress. Scherer and Ross (1990) provide a comprehensive picture for the proposition that monopoly prices reduce economic welfare due to distortions to efficiency that can be amplified through successive vertical stages of output. Wenders (1987) indicates also that part of the gains from monopoly power are dissipated into higher wages in addition to higher profits.

Costs under monopoly conditions may also be excessive because cost controls become lax and wasteful expenditures to strengthen and defend monopoly positions proliferate. Advertising, excessive product variety, and excessive government regulation costs compound further the decline in economic welfare due to concentration. Other dimensions of the impact of concentration on performance mentioned by Scherer and Ross include the possibility of redistribution of income

[2]Not all economists agree that control of a large share of a market by a few firms is necessarily bad. For instance, Brozen (1982) takes the view that bigness is the reward for efficiency. Concentration is indicative of a movement away from high-cost firms toward lower-cost, more efficient firms.

benefiting the shareholders and, to a certain extent, institutional investors.

1.3 Aim and Purpose

The property and liability insurance industry is composed roughly of 3,000 companies and about 40 major lines of property and liability insurance. Many of these companies deal with multiple lines of insurance. There is a great deal of affiliation among these companies, resulting in approximately 800 affiliations or groups. The individual companies in a group have considerable autonomy. Joint directorships and ownership of a group make it appropriate to consider an insurance group as one firm when investigating the questions of concentration. According to Huebner, Black, and Webb (1996), the affiliates had combined admitted assets[3] of $571.5 billion in 1993. Most of the affiliates operate in more than one state, and a substantial number operate nationwide.

A distinction is made between *industry concentration* where a few sellers dominate a specific market and *aggregate concentration* where a few conglomerate firms control large chunks of an economy. A conglomerate controls many lines of business. The lines may or may not be associated with high concentration. Nissan (1996) provides an assessment of aggregate concentration in the property and liability insurance for the years 1985, 1989, and 1993 for the largest 200 firms. The main finding of Nissan's work is a slight increase in aggregate concentration between 1985 and 1993 that is most pronounced among the top 20 firms.

While Nissan's concern is concentration on an aggregate basis, irrespective of the line of insurance, this paper takes Nissan's research a bit further: it provides an assessment of the degree of concentration by the major lines of property and liability insurance in the most recent year of data availability. The paper compares the degree of concentration among the major lines under consideration. Our focus is on concentration for 12 lines (as indicated in Tables 2 to 4) and we deal only with the largest 100 firms (affiliates) in each line (from Table 5 onward). Also, comparisons with previous studies on concentration are made. The analysis is conducted for all the affiliates in a line of insurance as well as by sets of hundreds.

[3]Admitted assets encompass all assets not prohibited by statute or regulation and include cash and bank deposits, real estate, mortgage loans, stocks, bonds, and other assets.

1.4 Database and Data Handling

The data used in this research were obtained from the National Association of Insurance Commissioners (NAIC). The association has existed some 125 years. The NAIC (1998) publication furnishes a brief synopsis of the extent and type of information gathered from some 5,500 of such insurance companies as life and health, property and liability, and the like. The data collected by NAIC account for 98 percent of all U.S. domiciled insurance companies.

This research benefits from the electronic company listing way of handling the data. The company listing for the property and liability insurance provides information on all member insurance companies including company name, company type (line), state of domicile, NAIC group code, and NAIC company code.

Because the data were provided in a disaggregated form covering some 40 different lines of insurance, it was necessary to undertake a massive effort to aggregate the 5,500 data points of the 40 lines in each state into a manageable number of groups of companies (affiliates or firms) which are then used as the database for this research. Throughout the rest of this paper the term *affiliates* or *firms* will denote a group of companies.

A note about economic data is in order. In economics one can only obtain a single realization (sample) of the economic process as we cannot stop the economy and restart the economic process to produce a new realization. Economists view the values obtained from that single realization as random values in the sense that if a new realization under the same conditions can be obtained, one would almost certainly not obtain identical numerical values. In fact, Darnell and Evans (1990, p. 25) state that

> It is our argument that any economic data may be conceived as being a random sample from a hypothetical population, and that this implies certain testable properties of the data which contribute to the design of the framework of data analysis.

Similar observations regarding economic data were made by others; see, for example, Griliches (1985) and McCloskey (1985).

2 Measurements

2.1 Measures of Concentration

There are two types of measures of concentration: static measures and dynamic measures (Tschoegl 1982). Static measures of concentration include the concentration ratio (CR_n) of market share of the n largest firms and weighted measures of the form

$$C_h = \sum_{i=1}^{N} h(z_i) z_i \tag{1}$$

where N is the number of firms in the market; $z_i \geq 0$ is the market share of the ith firm in the total amount, with $\sum z_i = 1$; and $h(z_i)$ is a nonnegative weight function; see Jacquemin and Kumps (1971). In this paper we take z_i as the share (proportion) of premiums written by firm i.

The weighted measures will differ depending on the selection of $h(\cdot)$. Two common weighted concentration measures are the Herfindahl index (where $h(z_i) = z_i$), and Theil's entropy (where $h(z_i) = -\ln(z_i)$). Another static measure suggested by Kwoka (1977) is the dominance index D given by

$$D = \sum_{i=1}^{N-1} (z_{(i)} - z_{(i+1)})^2$$

where $z_{(i)}$ is the market share of the ith largest firm, i.e., $z_{(i)} \geq z_{(i+1)}$. For this measure the emphasis is on the gaps between successive firms when they are ranked by size.

The concentration ratio is simple to construct, easy to understand, and hence widely used according to Hannah and Kay (1977). The choice of n is arbitrary. Typically, n is chosen as the largest four firms, eight firms, ten firms, 20 firms, or 50 firms. The Herfindahl is the most popular (Scherer and Ross 1990). The various static measures of concentration, however, are highly correlated, as indicated by Scherer and Ross (1990).

The dynamic measures, on the other hand, reflect the change in size of firms over a time period. An example of a dynamic measure, as pointed out by Tschoegl, is the index of market share instability developed by Hymer and Pashigian (1962). This index takes the form

$$I = \sum_{i=1}^{N-1} |z_{i,t} - z_{i,t-2}|$$

where $z_{i,t}$ is the ith firm's market share at time t. If a firm is not in the top 100 or 20 (depending on the market considered), at either time t or $t-2$, the market share for that time period is set to zero. A greater degree of change over the period (taken here as 2 years) induces a higher value for the index I, which implies a greater competitive turbulence and a greater amount of entry and exit of firms. Other dynamic measures are derived from standard stochastic growth models which take into account any first-order serial correlation in growth rates.

This paper uses two static concentration indexes to compare the concentration levels by line of insurance: the concentration ratio and the Herfindahl index, H

$$H = \sum_{i=1}^{N} z_i z_i = \sum_{i=1}^{N} z_i^2. \tag{2}$$

When one firm holds all shares, $H = 1$. When shares are held equally, $H = 1/N$. Thus $1/N \leq H \leq 1$. The Herfindahl index is used in merger guidelines by the Department of Justice Antitrust Division and the Federal Trade Commission in merger and monopolization cases. The Herfindahl index also has a wide appeal among economists, according to Clarke and Davies (1983), because of its origins in economic theory and history.

For a concentration index C_h, as expressed by equation (2), a concept known as *numbers equivalent*, M_h, provides an intuitive understanding of the extent of concentration. Specifically, M_h is the number of equally sized firms that will produce the given concentration index. For M_h firms, $z_i = 1/M_h$, for $i = 1, 2, \ldots, M_h$, so that

$$\begin{aligned} C_h &= \sum_{i=1}^{M_h} \frac{1}{M_h} h\left(\frac{1}{M_h}\right) \\ &= h\left(\frac{1}{M_h}\right). \end{aligned}$$

If h is invertible, then

$$M_h = \frac{1}{h^{-1}(C_h)}. \tag{3}$$

For the Herfindahl, $C_h = H$ and $h^{-1} = h$ so that

$$M_H = \frac{1}{H}. \tag{4}$$

The interpretation of M_H is that for the Herfindahl index for the N firms each with market share $1/N$, the index value will correspond to M_H equally sized firms.

2.2 Variance and Herfindahl Decompositions

Suppose the N firms in a line can be placed into G distinct sets for comparison purposes among the lines, and the sets are labeled $1, 2, \ldots, G$. Each line is composed of a number of affiliates or firms. Let A_g denote the set of indices in the gth set for a given line. The ith firm is in A_g if and only if $i \in A_g$. We assume that each firm appears in exactly one and only one set. The total variance and the square of the coefficient of variation of the z_is are S^2 and CV^2 respectively where

$$S^2 = \frac{1}{N}\sum_{i=1}^{N}[z_i - \bar{z}]^2 \tag{5}$$

and

$$\mathrm{CV}^2 = \frac{S^2}{\bar{z}^2} \tag{6}$$

with

$$\bar{z} = \frac{1}{N}\sum_{i}^{N} z_i.$$

Standard results of a statistical analysis of variance (ANOVA) shows that the variance can be decomposed into a *between sets sum of the squares* and a *within sets sum of the squares*.[4] The variance can be decomposed as

[4]For more on ANOVA see, for example, Scheffé (1959, Chapter 3) or Neter, Wasserman, and Craig (1990, Chapters 14).

$$S^2 = \sum_{g=1}^{G} \frac{N_g}{N}(\bar{z}_g - \bar{z})^2 \qquad \text{Between Sets}$$

$$+ \sum_{g=1}^{G} \frac{N_g}{N} \sum_{i \in A_g} \frac{1}{N_g}(z_i - \bar{z}_g)^2 \qquad \text{Within Sets.} \tag{7}$$

where N_g is the number of firms in A_g and $\bar{z}_g = \sum_{i \in A_g} z_i / N_g$.

Theil (1967) shows that CV^2 can be decomposed similarly into the *between-set squared coefficient of variation*, which measures between-set concentration, and the *within-set squared coefficients of variation*, which measures concentration between firms within a given set. Let

$$\bar{S}^2 = \sum_{g=1}^{G} \frac{N_g}{N}(\bar{z}_g - \bar{z})^2 \tag{8}$$

$$\overline{\mathrm{CV}}^2 = \frac{\bar{S}^2}{\bar{z}^2} \tag{9}$$

$$\bar{H} = \sum_{g=1}^{G} \bar{z}_g^2 \tag{10}$$

$$S_g^2 = \sum_{i \in A_g} \frac{1}{N_g}(z_i - \bar{z}_g)^2 \tag{11}$$

$$\mathrm{CV}_g{}^2 = \frac{S_g^2}{\bar{z}_g^2} \tag{12}$$

and

$$H_g = \sum_{i \in A_g} z_i^2. \tag{13}$$

It follows that

$$S^2 = \bar{S}^2 + \sum_{g=1}^{G} \frac{N_g}{N} S_g^2 \tag{14}$$

and

$$\bar{z}^2 \mathrm{CV}^2 = \bar{z}^2 \overline{\mathrm{CV}}^2 + \sum_{g=1}^{G} \frac{N_g}{N} \bar{z}_g^2 \mathrm{CV}_g{}^2, \tag{15}$$

which implies,

$$\begin{aligned} \mathrm{CV}^2 = {} & \overline{\mathrm{CV}}^2 \qquad \text{Between Sets Squared CV} \\ & + \sum_{g=1}^{G} \frac{N_g}{N} \frac{\bar{z}_g^2}{\bar{z}^2} {\mathrm{CV}_g}^2 \qquad \text{Within Sets Squared CV.} \end{aligned} \tag{16}$$

There is a relationship between H and CV. Clarke (1985) shows that

$$\mathrm{CV}^2 = N \times H - 1. \tag{17}$$

It immediately follows that

$$NH - 1 = G\bar{H} - 1 + \sum_{g=1}^{G} \frac{N_g}{N} \frac{\bar{z}_g^2}{\bar{z}^2} (N_g H_g - 1) \tag{18}$$

which implies,

$$H = \frac{G}{N}\bar{H} + \sum_{g=1}^{G} \frac{N_g}{N^2} \frac{\bar{z}_g^2}{\bar{z}^2} (N_g H_g - 1). \tag{19}$$

It is also useful to test the equality of variances and by implication the equality of the Herfindahl measures for the various pairs of lines of insurance by the test statistic:

$$F^* = \frac{S_k^2}{S_j^2} \tag{20}$$

where k is the number of firms in one line and j is the number of firms in another line, and S_k^2 and S_j^2 are the variances computed from equation (11). (As there are 12 lines, there will be 66 (= 12!/(2!10!) different pairs of tests.) The test statistic F^* is compared for significance with a tabular $F^*(\alpha, n_{k-1}, n_{j-1})$ for significance level α with n_{k-1} and n_{j-1} degrees of freedom, reflecting the number of firms in the two lines of insurance under consideration.

3 Empirical Results

Table 1 reports the distribution of the total premium written of approximately \$271 billion in 1995 among the major lines of property and liability insurance, listed in order of magnitude from the largest to the smallest. The private passenger auto liability line, with approximately \$60 billion of business accounting for 21.6 percent of the total, is the largest. The various lines of automobile insurance (private and commercial) together account for 44.2 percent of the total. Workers compensation with 12.3 percent and homeowners multiple peril with 9.3 percent are the next two largest lines.

Even though some of the lines contribute a small percentage to the total, such as fire (1.9 percent) and allied (1.3 percent), they command large dollar amounts. For the fire line, premium written is over \$5 billion. For the allied line, the premium written is almost \$4 billion.

Table 2 provides the distribution of premiums written for various years for the ten largest lines, which constitute almost 90 percent of the total. It shows that the percentages, with the exception of a slight increase in private passenger auto liability, remained virtually unchanged between 1981 and 1995.

Table 3 provides the extent to which property and liability insurance is dominated by the largest four, ten, and 50 firms for 1989 and 1995 as measured by the aggregate shares and the concentration ratio (CR_n). For private passenger auto liability damage for instance, the largest four, ten, and 50 firms in 1989 have shares of 43.2 percent, 56.6 percent, and 85.6 percent, respectively. For 1995, the corresponding percentages are 47.2, 62.2, and 88.4.

Another line of insurance that demonstrates large increases in concentration for the largest groups of companies is homeowners insurance where for the top four, ten, and 50 firms the respective concentrations for 1989 are 39.5 percent, 52.7 percent, and 82.1 percent. These concentrations increase in 1995 to 47.0 percent, 61.3 percent, and 87.7 percent. The other lines of insurance in Table 3 demonstrate, in general, similar increases in percentages between 1989 and 1995.

O'Neill (1996) provides data on aggregate concentration in non-manufacturing sectors in 1987: the largest 50 firms in banking held 48.3 percent of assets; in life insurance, the top 50 firms controlled 70 percent of assets and 48 percent of written premiums; for electric and gas utilities, the largest 40 firms controlled 64.5 percent of assets; and for retail trade and transportation, the top 50 firms controlled 23.0 percent and 74.0 percent of assets, respectively.

Table 1
Premiums Distributions of Property and Liability by Line of Insurance

Type of Insurance	Total ($1000s)	% of Total
Other Private Passenger Auto Liability	59,932,460	21.6
Private Passenger Auto Physical Damage	37,435,637	13.5
Workers Compensation	34,139,204	12.3
Homeowners Multiple Peril	25,846,863	9.3
Other Liability	22,512,579	8.1
Other Commercial Auto Liability	13,163,519	4.7
Commercial Multiple Peril (Non-Liability Portion)	11,176,538	4.0
Commercial Multiple Peril (Liability Portion)	9,777,047	3.5
Private Passenger Auto No-Fault (Personal Injury Protection)	7,491,338	2.7
Inland Marine	6,917,010	2.5
Medical Malpractice	6,164,639	2.2
Fire	5,350,244	1.9
Group Accident and Health	5,218,269	1.9
Commercial Auto Physical Damage	4,750,182	1.7
Allied Lines	3,734,984	1.3
Aggregate Write-Ins for Other Lines of Business	2,993,766	1.1
Surety	2,734,916	1.0
Product Liability	2,262,510	0.8
Ocean Marine	2,127,036	0.8
Mortgage Guaranty	2,076,288	0.7

Source: NAIC (1998) and calculations by the authors.

Table 1 (Contd.)

Premiums Distributions of Property and Liability by Line of Insurance

Type of Insurance	Total ($1000s)	% of Total
Aircraft (All Perils)	1,395,441	0.5
Farm Owners Multiple Peril	1,331,496	0.5
Earthquake	1,277,090	0.5
Other Accident Only	1,258,281	0.5
All Other A & H	1,105,694	0.4
Multiple Peril	956,723	0.3
Fidelity	954,182	0.3
Financial Guaranty	818,348	0.3
Boiler and Machinery	786,604	0.3
Guaranteed Renewable A & H	521,741	0.2
Credit	374,237	0.1
Nonrenewable for Stated Reasons Only	365,709	0.1
Commercial Auto No-Fault (Personal Injury Protection)	358,268	0.1
Credit A & H (Group and Individual)	312,622	0.1
Federal Employees Health Benefit Premium	250,048	0.1
Burglary and Theft	132,264	0.0
Collectively Renewable A & H	51,156	0.0
Glass	15,251	0.0
Noncancelable A & H	2,679	0.0
TOTAL	278,072,881	100.0

Source: NAIC (1998) and calculations by the authors.

Table 2
Premiums Distributions of Property and Liability By Major Line of Insurance (Percent of Total, in %)

Line of Insurance	1981	1985	1989	1992	1995
Private Passenger Auto Liability	19.8	19.3	21.1	20.5	21.6
Private Auto Physical Damage	14.1	14.5	14.2	13.2	13.5
Commercial Auto Liability	4.8	5.4	5.8	5.1	4.7
Commercial Auto Physical Damage	2.7	2.8	2.5	1.7	1.7
Homeowners	11.5	9.6	8.5	8.6	9.3
Fire and Allied	5.1	4.3	3.4	2.9	3.2
Commercial Multiple Peril	6.9	8.3	8.4	7.3	7.5
General Liability	6.1	7.9	8.8	7.9	8.1
Medical Malpractice	1.3	2.0	2.1	2.1	2.2
Workers Compensation	14.7	11.8	13.6	15.7	12.3
Other	13.0	14.1	11.6	15.0	15.9
Total Premiums ($ billions)	99.3	146.1	208.4	247.9	278.7

Sources: 1981, 1985, and 1989 entries (Cummins and Weiss 1991); 1992 and 1995 entries are calculated by the authors from NAIC (1998).

Table 3
Concentration Ratios for Premiums Written
By Property and Liability Insurance (Selected Years)

	Top 4 Firms		Top 10 Firms		Top 50 Firms	
Line of Insurance	1989	1995	1989	1995	1989	1995
Private Passenger Auto Liability	43.2	47.2	56.6	62.2	85.6	88.4
Private Auto Physical Damage	41.8	46.7	53.9	60.2	80.4	85.7
Commercial Auto Liability	19.6	20.6	38.1	37.9	80.9	81.4
Commercial Auto Physical Damage	19.6	18.9	35.0	35.3	78.0	81.3
Homeowners	39.5	47.0	52.7	61.3	82.1	87.7
Fire	[18.9]	30.0	[36.3]	45.3	[73.9]	83.0
Allied		22.7		43.3		82.9
Commercial Multiple Peril	[21.8]	24.2	[43.1]	45.1	[85.9]	87.7
Commercial Multiple Peril (Liability)		28.3		49.1		88.9
General Liability	32.6	40.0	51.9	59.8	84.5	88.6
Medical Malpractice	32.0	37.0	52.3	62.7	2.4	99.8
Workers Compensation	26.7	28.4	49.2	50.3	88.4	87.3

Sources: 1989 entries (Cummins and Weiss 1991); 1995 entries are calculated by the authors from NAIC (1998).

Table 4
Herfindahl Index of Concentration for Premiums Written by Major Property and Liability Insurance (All Affiliates)

Line of Insurance	1989	1995	CV^2	M_H
Private Passenger Auto Liability	0.0650	0.0836	7.36	11.96
Private Auto Physical Damage	0.0676	0.0850	7.50	11.76
Commercial Auto Liability	0.0214	0.0218	1.18	45.87
Commercial Auto Physical Damage	0.0313	0.0203	1.03	49.26
Homeowners	0.0573	0.0882	7.82	11.34
Fire	[0.0149]	0.0365	2.65	27.40
Allied		0.0268	1.68	37.31
Commercial Multiple Peril	[0.0263]	0.0293	1.93	34.13
Commercial Multiple Peril (Liability)		0.0359	2.59	27.86
General Liability	0.0450	0.0674	5.74	14.84
Medical Malpractice	0.0364	0.0574	4.74	17.42
Workers Compensation	0.0364	0.0331	2.31	30.21

Sources: 1989 entries (Cummins and Weiss, 1991).
1995 entries are calculated by the authors from NAIC (1998)

In contrast, the top 50 companies in property and liability insurance (Table 3) controlled much larger shares in premium written, ranging between 81.3 percent for commercial auto physical damage to 99.8 percent for medical malpractice in 1995. For 1989, the corresponding range is between 78 percent and 88.4 percent.

Table 4 provides the Herfindahl index of concentration for all affiliates for 1989 and 1995 and the numbers equivalent for 1995. The Herfindahl index is calculated using equation (1), CV^2 is calculated using equation (3), and the numbers equivalent is calculated using equation (2). $M_H = 49.26$ is the largest and $M_H = 11.76$ is the smallest among the equally sized firms. In contrast to the insurance industry, for instance, Adams and Brock (1995) report that the U.S. crude oil market had in 1991 a Herfindahl index of 0.02894, corresponding to $M_H = 36$.

The concentration index increased noticeably between 1989 and 1995 in almost all lines of property and liability insurance. The most noticeable increase, from 0.0650 in 1989 to 0.0850 in 1995, is for the private passenger auto physical damage line. The only slight decrease in concentration, from 0.0364 in 1989 to 0.0331 in 1995, is for the workers compensation line.

Table 5
Dollar Magnitudes of Premium Written
By Top 100 Affiliates and Proportions by Sets for 1995

	Line Codes	Total	Set 1	Set 2	Set 3
Private Passenger Auto Liability	1	54,312,023,302	0.756	0.149	0.095
Private Auto Physical Damage	2	33,635,781,692	0.754	0.150	0.095
Workers Compensation	3	26,326,667,644	0.674	0.226	0.100
Homeowners	4	23,412,901,987	0.765	0.151	0.085
General Liability	5	20,476,334,144	0.774	0.148	0.078
Commercial Auto Liability	6	11,848,100,515	0.582	0.281	0.137
Commercial Multiple Peril	7	10,115,152,209	0.694	0.218	0.088
Commercial Multiple Peril (Liability)	8	8,949,287,219	0.705	0.212	0.083
Fire	9	4,402,251,361	0.663	0.223	0.114
Commercial Auto Physical Damage	10	4,233,965,153	0.569	0.293	0.138
Medical Malpractice	11	3,560,398,869	0.848	0.149	0.002
Allied	12	3,338,366,360	0.665	0.208	0.127

Notes: Set 1 (top 20 affiliates), Set 2 (next 30 affiliates), and Set 3 (bottom 50 affiliates).

Sources: NAIC (1998) and calculations by the authors.

4 Comparing the Top 100 Affiliates

4.1 Defining the Sets

Because the top 100 affiliates in each line command the largest share of premium written and to make the comparison for concentration among the lines meaningful, the analysis for 1995 will concentrate on the premiums written by the top 100 affiliates for each line. Also, in order to take advantage of the disaggregation procedure of equation (16) whereby the total CV^2 can be split into a between-set component and within-set component, the top 100 affiliates are grouped into the top 20 affiliates (Set 1), the next 30 affiliates (Set 2), and the remaining 50 affiliates (Set 3).

Table 5 provides a dollar amount summary of the 100 affiliates of the largest 12 lines arranged from the largest magnitude to the smallest. Coding facilitates comparisons and analysis. The table reveals that the top 20 affiliates of each line control a substantial proportion of premiums written, ranging from 0.569 for commercial auto physical damage to 0.848 for medical malpractice.

4.2 Decomposition Results

The results of the computations from equation (16) for the top 100 affiliates, whereby the square of the coefficient of variation is decomposed into between-set and within-set components, are shown in Tables 6 and 7.[5] Tables 6 and 7 arrange the lines of the top 100 affiliates according to their S^2 and by implication by CV^2 magnitude. Thus, the homeowners line (code 4) at $CV^2 = 8.60$ is the most concentrated line. The least concentrated line at $CV^2 = 1.28$ is recorded by the commercial auto physical damage line (code 10).

The total CV^2 from equation (16) is decomposed into the between-set and within-set components then expressed as totals in Table 6 and as proportions in Table 7. Thus, for the homeowners line, the total $CV^2 = 8.60$ is split into between-set concentration of a total amount 2.00 and within-set concentration of a total amount 6.60, translated into respective proportions in the amounts 0.233 and 0.767.

[5]Note that we are not using analysis of variance to test between and within sets of firms in a given line. This disaggregation was made only to produce Tables 6 and 7. We are only testing for equality of total variances of two lines at a time; each line is composed of 100 firms.

Table 6
Decomposition of Total CV^2 of the Top 100 Affiliates

			Between Sets				Within Sets			
Line	Code	Total	Set 1	Set 2	Set 3	Total	Set1	Set 2	Set 3	Total
Homeowners	4	8.60	1.59	0.07	0.34	2.00	6.58	0.01	0.00	6.60
Private Passenger APD	2	8.44	1.54	0.07	0.33	1.94	6.50	0.00	0.00	6.50
Private Passenger AL	1	7.36	1.54	0.08	0.33	1.95	5.40	0.01	0.00	5.41
General Liability	5	6.30	1.65	0.08	0.36	2.09	4.21	0.01	0.00	4.21
Fire	9	3.15	1.07	0.02	0.30	1.39	1.74	0.02	0.01	1.76
Medical Malpractice	11	2.84	1.01	0.20	0.25	1.46	1.37	0.02	0.00	1.38
Commercial MPL	8	2.82	1.27	0.03	0.35	1.65	1.14	0.03	0.00	1.17
Workers Compensation	3	2.51	1.12	0.02	0.32	1.46	1.03	0.01	0.01	1.05
Commercial MP	7	2.18	1.22	0.02	0.34	1.58	0.57	0.03	0.00	0.60
Allied	12	1.98	1.08	0.02	0.28	1.38	0.57	0.02	0.01	0.60
Commercial AL	6	1.45	0.73	0.00	0.26	0.99	0.42	0.02	0.01	0.45
Commercial APD	10	1.28	0.68	0.00	0.26	0.94	0.31	0.02	0.01	0.34

Notes: APD = Auto Physical Damage, AL = Auto Liability, MPL = Multiple Peril (Liability), MP = Multiple Peril.

Table 7
Proportional Decomposition of Total CV^2 of the Top 100 Affiliates

			Between Sets				Within Sets			
Line	Code	Total	Set 1	Set 2	Set 3	Total	Set1	Set 2	Set 3	Total
Homeowners	4	1.00	0.185	0.008	0.040	0.233	0.765	0.001	0.000	0.767
Private Passenger APD	2	1.00	0.182	0.009	0.039	0.230	0.770	0.000	0.000	0.770
Private Passenger AL	1	1.00	0.209	0.011	0.045	0.265	0.734	0.001	0.000	0.735
General Liability	5	1.00	0.262	0.013	0.057	0.332	0.668	0.002	0.000	0.668
Fire	9	1.00	0.340	0.006	0.095	0.441	0.552	0.006	0.003	0.559
Medical Malpractice	11	1.00	0.356	0.070	0.088	0.514	0.482	0.007	0.000	0.486
Commercial MPL	8	1.00	0.450	0.011	0.124	0.585	0.404	0.011	0.000	0.415
Workers Compensation	3	1.00	0.446	0.008	0.128	0.582	0.410	0.004	0.004	0.418
Commercial MP	7	1.00	0.560	0.009	0.156	0.725	0.261	0.014	0.000	0.275
Allied	12	1.00	0.545	0.010	0.142	0.697	0.288	0.010	0.005	0.303
Commercial AL	6	1.00	0.504	0.000	0.179	0.683	0.290	0.014	0.007	0.317
Commercial APD	10	1.00	0.531	0.000	0.203	0.734	0.242	0.016	0.008	0.266

Notes: APD = Auto Physical Damage, AL = Auto Liability, MPL = Multiple Peril (Liability), MP = Multiple Peril.

Table 7 shows general increases in the between-set proportions and simultaneous decreases in the within-set proportions as one moves from the largest $CV^2 = 8.60$ to the smallest $CV^2 = 1.28$ with corresponding between-set proportions increasing from 0.233 to 0.734 and corresponding within-set proportions declining from 0.767 to 0.266. This suggests that among the major lines of the insurance industry, the effects of within-set concentration are highest among the largest 20 affiliates as observed under Set 1 in Tables 6 and 7. In either case, as observed in Tables 6 and 7, the biggest contributors to between-sets and within-sets components of CV^2 are the largest 20 affiliates (Set 1).

Inspection of equation (16) indicates that in computing CV^2 and its decomposition the between-set concentration is due to the presence of heterogeneous sets that have mean premiums written that differ from the overall 100 affiliates' mean. The contribution of each set to the variance is weighted by its respective share of the sample. For Set 1 the weight is 20/100; for Set 2 the weight is 30/100; and for Set 3 the weight is 50/100.

The within-set concentration, on the other hand, is due to the aggregate contribution to CV^2 within each of the three groups, again weighted by their respective shares. Thus, for the between-set concentration case, the mean premiums written by the largest 20 affiliates differ considerably from the overall mean. For the within-set concentration, the premium written by the largest affiliates among the 20 affiliates differ significantly from the mean premiums written of their own group.

Using the relationship between the squared coefficient of variation CV^2 and the Herfindahl index provided in equation (17), the test statistic for testing the equality of two variances, F^*, is helpful. The procedure is to align the CV^2 values from Table 6 in ascending order and to use equation (20) to test for statistical significance of the F^* ratio. The results are compared with the tabular F^* for the five percent significance level, which for large samples equals approximately 1.35.

The first two rows of Table 8 show the arrangement in ascending order and the corresponding line codes. The entries in columns 3 through 14 represent the F^* ratios for all possible comparisons of two specific lines. An asterisk ($*$) in the matrix indicates there is no significance between the two lines. For example, the $*$ in column 3 indicates that there is not a significant difference in the concentration of the homeowners and the private passenger auto physical damage lines. All other comparisons with the homeowners line (rest of column 3) show a significant difference. The complete matrix indicates that in the majority of cases there are noticeable differences in concentration among the different lines.

5 Closing Comments

A large body of literature pertaining to aggregate concentration describes, depending on the economic philosophy of the researchers, the adverse or beneficial consequences of increasing aggregate concentration. Some argue that firms that practice in markets characterized by high levels of concentration will likely exhibit tacit or collusive behavior, providing these firms with monopoly profits. Others argue that the greater efficiency of the leading firms leads to greater concentration and, thus, a positive relation between concentration and profitability is a result of efficiency rather than monopolization. In the majority of cases, however, government agencies concerned with promoting competition among firms in an industry treat increases in concentration with disfavor. Government fears that concentration will result in lessening of competition and enhancement of corporate and political power, both of which can be detrimental to the interests of consumers.

While the property and liability insurance industry is relatively exempt from federal laws through help from the McCarren-Ferguson Act of 1945, the act is blamed as being the cause of a variety of problems faced by that same industry (Joskow and McLaughlin 1991). There are calls for the repeal of the act making the industry subject to antitrust rules because of the perceived excessive control of the industry by a few large firms. Therefore, research into the level of concentration by line of insurance for property and liability is useful in pointing out those lines with the largest concentration.

This paper extends the previous literature in two ways. First, the paper updates the previous findings. Second, it provides finer statistical tools and testing procedures to compute the level of concentration by the line of insurance and to identify the lines that are most concentrated.

In this regard, the paper shows that there was an increase in concentration (sometimes sizable) between 1989 and 1995 for almost all the lines of property and liability insurance. A plausible explanation is the extensive acquisition of firms by other firms, a phenomenon that has prevailed throughout U.S. industries in recent years. The second important result of this research is an ordering of the various lines of property and liability insurance by their degree of concentration as shown in the matrix of Table 8. One notices here that the most concentrated is the homeowners line, followed closely by private passenger auto physical damage.

Table 8
***F*-ratios for Testing Hypothesis of Equality of Herfindahl Index for Premium Written by Top 100 Affiliates of Property and Liability Insurance**

		Code											
		4	2	1	5	9	11	8	3	7	12	6	10
Line	Code	8.60	8.44	7.36	6.30	3.15	2.84	2.82	2.51	2.18	1.98	1.45	1.28
HOME	4	1.000											
PPAPD	2	1.019*	1.000										
PPAL	1	1.168	1.147*	1.000									
GENLB	5	1.365	1.340*	1.168	1.000								
Fire	9	2.730	2.679	2.337	2.000	1.000							
MEDMP	11	3.028	2.972	2.592	2.218	1.109*	1.000						
CMLPL	8	3.050	2.993	2.610	2.234	1.117*	1.007*	1.000					
WORKC	3	3.426	3.363	2.932	2.510	1.255*	1.131*	1.124*	1.000				
CMLP	7	3.945	3.872	3.376	2.890	1.445	1.303*	1.294*	1.151*	1.000			
Allied	12	4.343	4.263	3.717	3.182	1.591	1.434	1.424	1.268*	1.101*	1.000		
CMAL	6	5.931	5.821	5.076	4.345	2.172	1.959	1.945	1.731	1.503	1.366	1.000	
CMAPD	10	6.719	6.594	5.750	4.922	2.461	2.219	2.203	1.961	1.703	1.547	1.133*	1.000

Source: NAIC (1998) and calculations by equation (5). *Notes*: HOME = Homeowners, PPAPD = Private Passenger Auto Physical Damage, PPAL = Private Passenger Auto Liability, GENLB = General Liability, MEDMP = Medical Malpractice, CMLPL = Commercial Multiple Peril (Liability), WORKC = Workers Compensation, CMLP = Commercial Multiple Peril, CMAL = Commercial Auto Liability, CMAPD = Commercial Auto Physical Damage. A superscripted * indicates no statistical significance in concentration between the corresponding pairs of insurance lines at 5% level.

The above two conclusions present a picture of the property and liability insurance industry as being highly concentrated. Concentration alone does not preclude competitiveness, however, as observed by Joskow and McLaughlin (1991). They claim that most major lines of property and liability insurance are provided by a large number of firms; that firms can add any line they choose as part of their business, implying no monopolization exists by any of the lines; that entry and exit of firms are relatively easy; and that many commercial customers may opt to purchase insurance from other established firms. In addition, self-insurance and the availability of industry insurance groups are other alternatives. These factors may combine to exert pressure on the industry to be competitive.

References

Adams, W. and Brock J. "Efficiency, Corporate Power, and the Bigness Complex." *Journal of Economic Education* 21(1990): 30-50.

Adams, W. and Brock J. *The Structure of American Industry.* Englewood Cliffs, New Jersey: Prentice Hall, 1995.

Bajtelsmit, V.L. and Bouzouita, R. "Market Structure and Performance in Private Passenger Automobile Insurance." *The Journal of Risk and Insurance* 65 (1998): 503-514.

Black, J., Jr. and Skipper, H.D., Jr. *Life Insurance* Twelfth Edition. Englewood Cliffs, N.J.: Prentice Hall, 1994.

Brown, W.J. *The Prevention and Control of Monopolies.* London, United Kingdom: John Murray, 1914.

Brozen, Y. *Concentration, Mergers, and Public Policy.* New York, New York: MacMillan Publishing Company, 1982.

Caves, R. *American Industry: Structure, Conduct, and Performance.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1967.

Chidambaran, N.K., Pugel, T.A. and Saunders, A. "An Investigation of the Performance of the U.S. Property-Liability Insurance Industry." *The Journal of Risk and Insurance* 64 (1997): 371-381.

Clarke, R. *Industrial Economics.* New York, New York: Basil Blackwell, 1985.

Clarke, R. and Davies, S.W. "Aggregate Concentration, Market Concentration and Diversification." *The Economic Journal* 93 (1983): 183-192.

Cummins, J.D. and Weiss, M.A. "The Structure, Conduct, and Regulation of the Property-Liability Insurance Industry." In R.W. Kopeke and R.E. Randall (eds.) *The Financial Condition and Regulation of Insurance Companies.* Boston, Mass.: Federal Reserve Bank of Boston, 1991: 117-154.

Darnell, A.C. and Evans, J. L. *The Limits of Econometrics.* Hants., England: Edward Elgar, 1990.

Griliches, Z. "The Use and Abuse of Econometrics." *American Economic Review* 75 (1985): 196-200.

Hamilton, K.L. *The Changing Nature of Insurance Regulation.* Malvern, Pa.: Insurance Institute of America, 1997.

Hannah, L. and Kay, J.A. *Concentration in Modern Industry: Theory, Measurement and the U.K. Experience.* London, United Kingdom: The MacMillan Press, 1977.

Huebner, S.S., Black, K. and Webb, B.L. *Property and Liability Insurance.* Upper Saddle, New Jersey: Prentice Hall, 1996.

Hymer, S. and Pashigian, P. "Turnover of Firms as a Measure of Market Behavior." *Review of Economics and Statistics* 44 (1962): 82-87

Jacquemin, A.P. and Kumps, A. "Changes in the Size Structure of the Largest European Firms: An Entropy Measure." *The Journal of Industrial Economics* 20 (1971): 59-70.

Joskow, P.L. "Cartels, Competition and Regulation in the Property-Liability Insurance Industry." *The Bell Journal of Economics and Management Science* 4 (1973): 375-427.

Joskow, P.L. and McLaughlin, L. "McCarran-Ferguson Act Reform: More Competition or More Regulation?" *Journal of Risk and Uncertainty* 4 (1991): 373-401.

Kwoka, J.F., Jr. "Large Firm Dominance and Price-Cost Margins in Manufacturing Industries." *Southern Economic Journal* 44 (1977): 183-189.

Lereah, D.A. *Insurance Markets: Information Problems and Regulation.* New York, New York: Praeger Publishers, 1985.

Mayers, D. and Smith, C.W., Jr. "On the Corporate Demand for Insurance: Evidence from the Reinsurance Market." *Journal of Business* 63 (1990): 19-40.

Mayers, D. and Smith, C.W., Jr. "On the Corporate Demand for Insurance." *Journal of Business* 55 (1982): 281-296.

Mayers, D. and Smith, C.W., Jr. "Ownership Structure Across Lines of Property-Casualty Insurance." *Journal of Law and Economics* 31 (1988): 351-378.

McCloskey, D. N.. "The Loss Function Has Been Mislaid: The Rhetoric of Significance Tests." *American Economic Review* 75 (1985): 201-205.

NAIC. *NAIC Database Products 1998.* Kansas City, NAIC Database Sales, 1998.

Neter, J., Wasserman, W. and Craig, A.T. *Applied Linear Statistical Models.* 3rd Edition. Boston, Mass.: Irwin, 1990.

Ng, L.F.Y. "Changing Industrial Structure and Competitive Patterns of Manufacturing and Non-Manufacturing in a Small Open Economy: An Entropy Measurement." *Managerial and Decision Economics* 16 (1995): 547-563.

Nissan, E., "Concentration in American Property-Casualty Companies." *Journal of Actuarial Practice* 4 (1996): 131-142.

O'Neill, P.B. "The Trend of Aggregate Concentration in the United States: Problems of Scope and Measurement." *American Journal of Economics and Sociology* 55 (1996): 197-211.

Scheffé, H. *The Analysis of Variance.* New York, N.Y.: John Wiley & Sons, 1959.

Scherer, F.M. and Ross, D. *Industrial Market Structure and Economic Performance.* Boston, Massachusetts: Houghton Mifflin Company, 1990.

Schutz, E. "Market Power." *Journal of Economic Issues* 4 (1995): 1147-1170.

Tschoegl, A.E. "Concentration Among International Banks." *Journal of Banking and Finance* 4 (1982): 567-578.

Theil, H. *Economics and Information Theory.* Amsterdam, Holland: North Holland Publishing Company, 1967.

Wenders, J.T. "On Perfect Rent Dissipation." *The American Economic Review* 77 (1987): 456-459.

Weiss, L.W. "The Extent and Effects of Aggregate Concentration." *The Journal of Law and Economics* 26 (1983): 429-455.

Journal of Actuarial Practice Vol. 8, 2000

Safe-Side Requirements in Life Insurance: A Corporate Perspective

Annamaria Olivieri* and Ermanno Pitacco†

Abstract‡

Safe-side requirements concern the assumptions used to calculate premiums in relation to a set of more realistic assumptions. Roughly, safe-side requirements express the capability of premiums to generate positive margins. In a strictly actuarial framework, safe-side requirements are given in terms of some notion of expected profit, calling for assumptions that let such profit be nonnegative. An expected profit of zero, however, is not a realistic aim for the insurer.

We investigate the notion of conservative assumptions by adopting a unconventional approach. Our focus is the management of the financial resources coming both from premiums and from shareholders' capital. This leads to a general structure that includes as particular cases the results obtainable in a strictly actuarial environment.

Key words and phrases: *technical basis, expected profit, portfolio fund, shareholders' capital, opportunity cost of capital, discounted cash flow*

---

*Annamaria Olivieri, Ph.D., is professor of financial mathematics at the University of Parma. She received her Ph.D. in mathematics for the analysis of financial markets from the University of Brescia, Italy, and was an assistant professor at the University of Trieste, Italy. She is the author of several papers in actuarial science.

Dr. Olivieri's address is: Dipartimento di Economia, Sezione di Matematica "E. Levi," University of Parma, ITALY. Internet address: *annamaria.olivieri@econ.univ.trieste.it*

†Ermanno Pitacco is a full professor of actuarial mathematics at the University of Trieste. He is a member of the ASTIN Committee, contributing editor of *Insurance: Mathematics & Economics*, member of the Education Committee of the "Group Consultatif des Associations d'Actuaries des Pays des Communautes Europeennes." He has authored several textbooks and papers in actuarial science.

Prof. Pitacco's address is: Dipartimento di Matematica Applicata "Bruno de Finetti," University of Trieste, ITALY. Internet address: *ermanno.pitacco@econ.univ.trieste.it*

‡This research was partially funded by the Italian MURST, within the project "Modelli attuariali analitici e numerici per l'innovazione nelle assicurazioni di persone." A preliminary version was presented to the XVI Euro, held in Brussels in July 1998.

The authors are grateful to the anonymous referees for several useful remarks that have helped in the preparation of this paper.

1 Introduction

Though the concept of safe-side requirements has always been an integral part of actuarial science, it is only recently that it has become a core topic in Europe. This is especially true in Italy, as a consequence of the European Third Directive concerning life insurance regulation (Directive EEC, No. 96 of 1992).[1] Because tariffs are no longer subject to approval, other tools must be used to monitor the stability of insurance companies.

Expressions such as "prudent," "prudent valuation," and so on are widely used in current legislation[2] in order to define freedom in choosing premium ratings. These terms are often used vaguely or dubiously.

In this paper we have tried to outline a general structure that can help in understanding what a safe-side requirement is and to what it refers. We have adopted two approaches: (i) a purely actuarial approach, referring to some classical results of actuarial mathematics such as the notion of expected profit and the contribution formula of Homans (1863) (see also Haberman and Sibbett 1995, pages 287-297) as well as to some ideas developed in the framework of multistate models (see Hoem (1988) and Olivieri (1999)); and (ii) a corporate approach introducing the notion of shareholders' capital, which is not usually considered in traditional actuarial mathematics. The corporate approach has proved to be more general than the purely actuarial one. The results obtained in the actuarial framework are special cases of the corporate approach. These approaches lead to a unitary formal structure.

Our basic set of assumptions are:

(a) Only life insurance policies issued on a single life are considered. Policies involving more than one life or health insurance policies are disregarded;

(b) As we are considering only net premiums, expenses and expense loadings are also disregarded;

(c) Benefits and premiums are specified at policy issue and remain unchanged throughout the insured period (hence, financial adjustments of benefits and/or premiums are not permitted);

[1]Directive EEC, No. 96 of 1992 has been published on the Gazette Officielle des Communautes Europeennes, loi n. 360 du 9/12/1992.

[2]The legislation is the Italian D.Lgs. 174/95 (published on the Gazzetta Ufficiale della Repubblica Italiana, n. 56, 18/5/1995), which introduces European legislation in Italy.

(d) Premiums are assumed to be paid at the start of each year and death benefits are paid at the end of the year of death (thus we adopt a time-discrete approach); and

(e) Though the underlying processes may vary, we consider expected values only (some remarks on the possibility of considering higher moments or, in general, probability distributions are made in Section 6).

The model is simple, but it makes the interpretation of the results easier. Moreover, comparisons with the traditional actuarial model are immediate.

Section 2 introduces the notion of prudence; this notion is referred to as a "first order basis" in comparison to a given set of more realistic assumptions that are referred to as a "second order basis." A formal structure is introduced under which various safe-side requirements can be classified. In Section 3 various definitions of safe-side requirements are given in a strictly actuarial framework. Section 4 adopts a corporate approach; some comparisons between the two approaches are then made. Section 5 discusses some numerical examples.

2 Technical Bases and Prudence

2.1 Some Preliminary Aspects

One of the objectives of actuarial valuations is to assess the adequacy of premiums and to forecast future payments by the insurer and the insured. This objective requires the choice of a convenient set of basic assumptions on which to base the forecast. In life insurance, such basic assumptions include demographic assumptions (e.g., mortality, morbidity, and lapse rates) and financial assumptions (e.g., interest and inflation rates). Throughout this paper, the set of assumptions used to derive net premiums will be called the *technical basis*.

When life insurance premiums are calculated according to the equivalence principle,[3] the expected profit for the insurer is zero if a realistic technical basis is used. Hence, it is necessary to:

(i) Use a realistic technical basis and adopt premium principles other than the equivalence principle in order to include an explicit safety loading into premiums; or to

[3]The equivalence principle states that the actuarial present value of premiums is equal to the actuarial present value of benefits.

(ii) Adopt the equivalence principle and use a conservative technical basis in order to include an implicit safety loading into premiums.

In both cases, the safety loading leads to a positive expected profit. Among European life insurance companies, choice (ii) is commonly made.

The *first order technical basis* (briefly TB_1) is the set of conservative assumptions (i.e., assumptions favorable to the insurer) used in choice (ii). This set is relatively easy to define. It includes the valuation rate of interest (usually constant) as well as a mortality table.

The *second order technical basis* (briefly TB_2) is the set of realistic assumptions. The concept of second order basis is more complex. TB_2 must give a realistic description of the scenario facing the insurance company and the insured person. Thus, it should include assumptions about policyholder behavior, investment performance, and company behavior and assumptions about macro- and micro-economic forces.

It is likely that in the early days of actuarial practice the two concepts emerged simultaneously. Given the contractual relevance of TB_1, TB_2 has usually been expressed as a simple shift of TB_1. Hence, TB_1 and TB_2 were usually assumed to have the same structure; for example, for insurance policies TB_1 includes a lower interest rate and higher mortality than that included in TB_2, while for annuities TB_1 includes a lower interest rate and lower mortality than in TB_2.

In these days of easy access to high speed computers, it is important to adopt a more flexible structure for TB_2. For example,the financial assumption may include a (deterministic) term structure of interest rates or a convenient stochastic model; as to the demographical aspect, a projected table can represent the future trend of mortality. (A stochastic model can express the uncertainty of the projection.)

For the sake of simplicity and for obtaining results that can be compared to the traditional model, we will adopt a conventional structure for TB_2. On the other hand, a deterministic term structure of interest rates as well as a (deterministic) projected mortality table would not add significance to the considerations discussed below.

2.2 Formal Aspects

As we have stated in Section 2.1, safe-side requirements will be referred to as TB_1 in relation to a given realistic basis TB_2. We need a yardstick to assess whether TB_1 is on the safe side with respect to TB_2. From a formal viewpoint, this yardstick is represented by a vector-valued mapping of the two technical bases. Given the dynamic nature of life insurance contracts and the length of these contracts, the vector could quantify safe-side requirements imposed on any single year. In other cases, the elements of the vector will represent the different components of the contract that can generate safety loadings and thus contribute to the expected profit. We will consider only deterministic mappings obtained as expected values of random variables.

If Γ denotes the particular insurance policy, we will deal with functions Φ that map (TB_1, TB_2, Γ) to quantities (typically expected profits) that are used to assess prudence. These quantities are specified in the proposed safe-side requirements. In particular, Γ allows us to specify benefits and to determine premiums and reserves. If Φ is a vector-valued function, its elements will be denoted by $\phi_1, \phi_2, \ldots,$.

As we have stated in Section 2.1, we adopt a constant rate of interest i and a given set of mortality rates q_y, $y = 0, 1, \ldots$ for TB_1 and i^* and q_y^*, $y = 0, 1, \ldots$ for TB_2. Thus we have

$$\begin{aligned} TB_1 &= (i, \{q_y\}) \\ TB_2 &= (i^*, \{q_y^*\}). \end{aligned}$$

As usual, we put $p_y = 1 - q_y$ and $p_y^* = 1 - q_y^*$. Traditional actuarial notation is used whenever possible.

We will focus on insurance policies with the following characteristics:

- x is the issue age, and ω is the limiting age of the mortality table;
- Term n years, where $n = 1, 2, \ldots, \omega - x$;
- The death benefit, paid at the end of the tth policy year of death, is C_t, where $t = 1, 2, \ldots, n$;
- Sum S is paid in case of survival to age $x + n$ (when n is finite); and
- The premiums are paid at the start of each policy year (if the insured is then alive). The premium paid at time k (i.e., age $x + k$) is P_k, $k = 0, 1, \ldots, n - 1$. The cases with single premiums and premiums payable for at most m years (with $m \leq n$) are included; annuities are also included, by letting $P_k < 0$ (thus paid by the insurer).

Such a general insurance structure includes many practical policies such as term, whole life, endowment, pure endowment insurances, and immediate and deferred annuities.

We assume that premiums are calculated according to the equivalence principle (obviously, with TB_1). The net premium reserve at time t, V_t, is defined as

$$\begin{aligned} V_t = & \sum_{h=0}^{n-t-1} C_{t+h+1}\, v^{h+1}\, {}_hp_{x+t}\, q_{x+h+t} + S\, v^{n-t}\, {}_{n-t}p_{x+t} \\ & - \sum_{h=0}^{n-t-1} P_{t+h}\, v^h\, {}_hp_{x+t} \end{aligned} \tag{1}$$

where $v = 1/(1+i)$. The boundary conditions are $V_0 = 0$ and $V_n = S$ (when n is finite). V_t satisfies the recurrence equation

$$(V_t + P_t)(1+i) = (C_{t+1} - V_{t+1})\, q_{x+t} + V_{t+1}. \tag{2}$$

3 Safe-Side Requirements in an Actuarial Framework

3.1 Profits and Second Order Reserves

From the recurrence equation (2), we get an expression for u^*_{t+1}, the annual profit at the end of the $t+1$st policy year, which is obtained evaluating the assets and liabilities using the realistic basis TB_2. We have

$$u^*_{t+1} = (V_t + P_t)(1+i^*) - (C_{t+1} - V_{t+1})\, q^*_{x+t} - V_{t+1} \tag{3}$$

from which we obtain the contribution formula of Homans (1863)

$$u^*_{t+1} = (V_t + P_{t+1})\,(i^* - i) + (C_{t+1} - V_{t+1})\,(q_{x+t} - q^*_{x+t}) \tag{4}$$

where the financial and demographic components of profits are:

$$\begin{aligned} {}_fu^*_{t+1} &= (V_t + P_{t+1})\,(i^* - i) && \text{Financial Component} \quad (5) \\ {}_du^*_{t+1} &= (C_{t+1} - V_{t+1})\,(q_{x+t} - q^*_{x+t}) && \text{Demographic Component.} \quad (6) \end{aligned}$$

The total future expected profit at time 0 is u^* where

$$u^* = \sum_{h=0}^{n-1} u_{h+1}^* \, {}_h p_x^* \, (1+i^*)^{-(h+1)}. \tag{7}$$

After substituting equation (3) into equation (7), and some algebraic manipulations, we obtain the following expression for total profit:

$$\begin{aligned} u^* &= \sum_{h=0}^{n-1} P_h \, {}_h p_x^* \, (1+i^*)^{-h} \\ &\quad - \sum_{h=0}^{n-1} C_{h+1} \, {}_h p_x^* \, q_{x+h}^* \, (1+i^*)^{-(h+1)} - S \, {}_n p_x^* \, (1+i^*)^{-n}. \end{aligned} \tag{8}$$

As before, total profit can be split into the financial and demographic components:

$$_f u^* = \sum_{h=0}^{n-1} {}_f u_{h+1}^* \, {}_h p_x^* \, (1+i^*)^{-(h+1)} \tag{9}$$

$$_d u^* = \sum_{h=0}^{n-1} {}_d u_{h+1}^* \, {}_h p_x^* \, (1+i^*)^{-(h+1)}. \tag{10}$$

More generally, we can define the expected future profit after time t, i.e., in the interval $[t, n]$, as

$$u^*(t,n) = \sum_{h=0}^{n-t-1} u_{t+h+1}^* \, {}_h p_{x+t}^* \, (1+i^*)^{-(h+1)}.$$

Specifically, $u^* = u^*(0,n)$. Also this expected profit can be split into its financial and demographic components; moreover, a result similar to equation (8) holds for $u^*(t,n)$.

Finally, we define V_t^* as the second order reserve calculated using the realistic assumptions of TB_2

$$\begin{aligned} V_t^* = &\sum_{h=0}^{n-t-1} C_{t+h+1} \, {}_h p_{x+t}^* q_{x+t+h}^* \, (1+i^*)^{-(h+1)} + S \, {}_{n-t} p_{x+t}^* \, (1+i^*)^{-(n-t)} \\ &- \sum_{h=0}^{n-t-1} P_{t+h} \, {}_h p_{x+t}^* \, (1+i^*)^{-h}. \end{aligned} \tag{11}$$

3.2 Safe-Side Requirement (SSR) Definitions

Let us now turn to various safe-side requirements, which are referred to as TB_1 in relation to TB_2. We will then analyze the links between the various definitions; we will verify that some safe-side requirements (SSR) imply others.

Definition 1 (Naive SSR):

We say that TB_1 is on the safe side (with respect to TB_2) if and only if

- 1. Benefits are payable in case of survival only,
  2. $i \leq i^*$, and
  3. $q_{x+h-1} \leq q^*_{x+h-1}$ for $h = 1, 2, \ldots, n$;

  or
- 1. Benefits are payable in case of death only,
  2. $i \leq i^*$, and
  3. $q_{x+h-1} \geq q^*_{x+h-1}$ for $h = 1, 2, \ldots, n$.

In this case prudence is directly measured on the single elements of the two technical bases. The corresponding mapping Φ is a vector with $n + 1$ elements, given by

$$\phi_h = q_{x+h-1} - q^*_{x+h-1} \quad \text{for } h = 1, 2, \ldots, n$$
$$\phi_{n+1} = i^* - i.$$

The safe-side requirement can be easily stated in terms of the mapping Φ. Such a definition can be applied only to a restricted number of cases; for example, it cannot be used for policies that contain both survival and death benefits.

Definition 2 (Financial/Demographic Annual Profits SSR):

Let us consider the mapping $\Phi = [\phi_1, \phi_2, \ldots, \phi_{2n}]$ where

$$\phi_{2t-1} = {}_f u^*_t \quad \text{for } t = 1, 2, \ldots, n \tag{12}$$
$$\phi_{2t} = {}_d u^*_t \quad \text{for } t = 1, 2, \ldots, n. \tag{13}$$

We say that TB_1 is on the safe side if and only if $\phi_h \geq 0$ for $h = 1, 2, \ldots, 2n$.

In this case, as in the following ones, the measurement of prudence relies on expected present values, which consist in equations (12) and (13) of the components of expected annual profit. From the definition of ${}_fu_t^*$ and ${}_du_t^*$ (see equations (5) and (6)) conditions on the elements of TB_1 can be derived. The condition concerning the demographical assumption is based on the sum at risk $(C_{t+1} - V_{t+1})$; hence, it is more widely applicable than simply requiring that $q_{x+h-1} \leq q^*_{x+h-1}$ or $q_{x+h-1} \geq q^*_{x+h-1}$ for $h = 1, 2, \dots, n$.

Definition 3 (Annual Profits SSR):

Let $\Phi = [\phi_1, \phi_2, \dots, \phi_n]$, where

$$\phi_t = u_t^*, \quad t = 1, 2, \dots, n.$$

We say that TB_1 is on the safe side if and only if $\phi_t \geq 0$ for $t = 1, 2, \dots, n$.

Definition 4 (Total Profit SSR):

Let

$$\Phi = u^*.$$

We say that TB_1 is on the safe side if and only if $\Phi \geq 0$.

Definition 5 (Financial/Demographic Total Profit SSR):

Let $\Phi = [\phi_1, \phi_2]$, where

$$\phi_1 = {}_fu^*$$
$$\phi_2 = {}_du^*.$$

We say that TB_1 is on the safe side if and only if $\phi_1 \geq 0$, $\phi_2 \geq 0$.

Definition 6 (Residual Profits SSR):

Let $\Phi = [\phi_1, \phi_2, \dots, \phi_n]$, where

$$\phi_t = u^*(t-1, n) \quad t = 1, 2, \dots, n.$$

We say that TB_1 is on the safe side if and only if $\phi_t \geq 0$ for $t = 1, 2, \dots, n$.

Definition 7 (Reserves SSR):

Let $\Phi = [\phi_1, \phi_2, \dots, \phi_n]$ where

$$\phi_t = V_{t-1} - V^*_{t-1} \quad t = 1, 2, \dots, n$$

We say that TB_1 is on the safe side if and only if $\phi_t \geq 0$ for $t = 1, 2, \ldots, n$.

This requirement means that in each year what is set aside to meet future net liabilities (V_t) must be at least equal to the realistic value of future net liabilities themselves (V_t^*).

The quantity $V_{t-1} - V_{t-1}^*$ is called the "Loewy increment." (See Hoem (1988) and Loewy (1917).) From equations (1) and (11), after some manipulations we get

$$V_{t-1} - V_{t-1}^* = u^*(t, n) \tag{14}$$

which becomes, in particular, $V_0 - V_0^* = u^*(0, n)$, i.e.,

$$V_0^* = -u^*. \tag{15}$$

Relations among the previous definitions can be easily found. Each definition involves a different degree of strictness. For example, a TB_1 which complies with Definition 2 also complies with Definition 3, the latter being less strict than the former. Let us adopt the following notation:

- (Di) denotes Definition i, e.g., $(D3)$ refers to Definition 3; and
- $(Di) \Rightarrow (Dj)$ means that a first order technical basis TB_1 that is on the safe side according to (Di) is also on the safe side according to (Dj).

It can be easily verified that

$$(D1) \Rightarrow (D2) \Rightarrow (D3) \Rightarrow (D6) \Leftrightarrow (D7) \Rightarrow (D4)$$

and

$$(D2) \Rightarrow (D5) \Rightarrow (D4).$$

Note that a severe requirement that involves a higher premium could produce high profits per policy. High premiums may reduce the demand of the insurance policy, however, and profits for the entire portfolio.

The discussion in this section allows to verify the link between the first two safe-side requirements common in actuarial practice (i.e., $(D1)$ and $(D2)$, which are severe) and the seventh safe-side requirement (i.e., $(D7)$, which has been recently proposed in actuarial literature, see Hoem (1988)). Moreover, the definitions above described give a general picture that allows to understand the meaning of the expression "prudent actuarial valuation."

Definitions $(D1)$ through $(D7)$, as well as equations (14) and (15), show that the notion of prudence is linked to some notion of implied expected profit. Having used only expected values, it is impossible to consider explicitly measures of demographic or financial riskiness. As a result, the (positive) lower bound for profit can be chosen arbitrarily. It would be interesting, considering definition $(D4)$, to require conditions such as $\Phi \geq u'$ where u' represents a minimum value for expected profit (or safety loading in terms of single premium) to be fixed in relation to the riskiness of the insurance contract.

We will briefly comment on an explicit consideration of risk in Section 6. In Section 4 we will introduce a structure that, although based on expected values only, is more general than the one just described. This structure allows us to single out positive lower bounds for profit.

4 Safe-Side Requirements and the Cost of Capital: Toward a Corporate Approach

4.1 Portfolio Fund, Discounted Cash Flow (DCF)

The safe-side requirements analyzed in Section 3 are minimal requirements in a corporate perspective. It is not enough for the policy to generate profits, but such profits must be enough to pay some minimum return to the invested capital. Because shareholders' capital is allocated at a portfolio level, in this section we refer not to a policy only, but to a portfolio of life insurance.

For simplicity, however, we consider a cohort of homogeneous policies, issued at the same time, identical in terms of insurance policy, age at entry, term, benefits, and premiums. Under such assumptions, S will be the total amount paid in case of survival of all policies at maturity. The amount actually paid at time n is a random variable that depends on the random number of survivors. At time 0, its expected value is given by $S \times {}_np_x^*$ (according to TB_2); similarly, $V_t \times {}_tp_x^*$ represents the (expected) portfolio reserve at time t. Similar relations hold for the other quantities.

For the sake of brevity, we adopt the following notation

$$\hat{V}_t = V_t \times {}_tp_x^*$$
$$\hat{u}_{t+1}^* = u_{t+1}^* \times {}_tp_x^*.$$

For $t = 0, 1, \ldots, n$, let Z_t denote the expected portfolio fund accumulated at time t (according to the information available at time 0) and let K_t ($t = 0, 1, \ldots, n$) be the shareholders' capital flow withdrawn from ($K_t > 0$) or paid to ($K_t < 0$) the portfolio fund. The sign of K_t is determined from the point of view of shareholders. The behavior of the portfolio fund can then be described by

$$Z_{t+1} = Z_t\,(1 + i^*) + P_t\,(1 + i^*)\,{}_tp_x^* - C_{t+1}\,{}_tp_x^* q_{x+t}^* - K_{t+1} \tag{16}$$

for $t = 0, 1, \ldots, n-1$, where we assume $Z_0 = -K_0$ with $K_0 \leq 0$. As K_t is assumed to be deterministic, no mortality factor is needed.

The analysis of cash flows is also considered in recent actuarial models, for example in profit testing techniques (see Goford (1985)). In that framework, however, shareholders' flows are not included—the main aim is the assessment of technical profit.

We will consider the sequence $\{K_t\}$ as given flows. In concrete terms, flows depend both on corporate strategies and insurance regulation. In any case, we assume

$$K_n = Z_{n-1}\,(1 + i^*) + P_n\,(1 + i^*)\,{}_np_x^* - C_n\,{}_{n-1}p_x^* q_{x+n-1}^* - S\,{}_np_x^* \tag{17}$$

so that $Z_n = 0$. Moreover, we suppose that flows K_t are chosen so that $Z_t \geq \hat{V}_t$ for $t = 1, 2, \ldots, n-1$. We define "free portfolio fund" as the (nonnegative) quantity $Z_t - \hat{V}_t$, which consists of the financial resources in excess of the expected reserve. $Z_t - \hat{V}_t$ represents the shareholders' capital globally linked to the portfolio at time t.

Let $G(\rho)$ denote the discounted cash flow (DCF) at time 0 for shareholders, calculated with a rate ρ, i.e.,

$$G(\rho) = \sum_{t=0}^{n} K_t\,(1 + \rho)^{-t}. \tag{18}$$

The rate ρ represents the yield required from shareholders on the capital invested in the portfolio (i.e., the opportunity cost of capital). We assume that $\rho \geq i^*$. The DCF can be split into a sequence of periodic contributions. (This notion has been proposed, for a financial operation in general, by Peccati (1989).)

Let $g_{t+1}(\rho)$ be the contribution at time $t+1$ to the DCF $G(\rho)$, evaluated at time 0. The splitting of DCF is based on the notion of outstanding capital at each time t, i.e., the capital invested at that time, which, as seen above, is given by the free portfolio fund $Z_t - \hat{V}_t$. The annual contribution to DCF can then be defined by amending the annual shareholders' flow K_{t+1} with the variation in the free portfolio fund. Hence, for $t = 0, 1, \dots, n-1$, we have

$$g_{t+1}(\rho) = \frac{-(Z_t - \hat{V}_t)}{(1+\rho)^t} + \frac{(K_{t+1} + Z_{t+1} - \hat{V}_{t+1})}{(1+\rho)^{t+1}}. \tag{19}$$

The structure of equation (19) is coherent with that of annual profit as defined in conventional life insurance mathematics. In the latter quantity, however, only debt capital (i.e., the reserve) is taken into consideration (see equation (3)). It can be easily verified that

$$G(\rho) = \sum_{t=0}^{n-1} g_{t+1}(\rho). \tag{20}$$

Subtracting equation (3) (previously multiplied by ${}_tp_x^*$) from (16) we obtain

$$Z_{t+1} - \hat{V}_{t+1} = (Z_t - \hat{V}_t)(1+i^*) - K_{t+1} + \hat{u}^*_{t+1}. \tag{21}$$

Solving for $Z_t - \hat{V}_t$, with initial condition $Z_0 - \hat{V}_0 = -K_0$, we obtain

$$Z_t - \hat{V}_t = -\sum_{h=0}^{t} K_h (1+i^*)^{t-h} + \sum_{h=1}^{t} \hat{u}^*_h (1+i^*)^{t-h}. \tag{22}$$

Substituting equation (21) into equation (19) we get

$$g_{t+1}(\rho) = [\hat{u}^*_{t+1} - (Z_t - \hat{V}_t)(\rho - i^*)](1+\rho)^{-(t+1)} \tag{23}$$

for $t = 0, 1, \dots, n-1$. Equation (23) shows that the periodical contribution to DCF is equal to the annual profit amended by the loss incurred by investing the free portfolio fund at rate i^* instead of the required ρ.

From equation (18), noting that K_n is equal to the accumulated value of the insurance profit plus the accumulated value of shareholders' capital flows (as can be checked by substituting equation (22) into equation (17)), we obtain the following expression for DCF:

$$\begin{aligned} G(\rho) = (1+\rho)^{-n} \Bigg[& u^* \times (1+i^*)^n \\ & + \sum_{t=0}^{n} K_t \times \left((1+\rho)^{n-t} - (1+i^*)^{n-t}\right) \Bigg]. \end{aligned} \tag{24}$$

Expression (24) can be easily interpreted in terms of the loss originating from the difference between ρ and i^*. Obviously, equation (24) could be obtained also by discounting back to time 0 the annual losses (i.e., by substituting (23) and (22) into (20)).

Equation (24) allows us to separate two components of DCF:

- The *technical component,*

  $$(1+\rho)^{-n} \times u^* \times (1+i^*)^n,$$

  which stems from the technical management of the insurance portfolio, and

- The *capital component,*

  $$(1+\rho)^{-n} \sum_{t=0}^{n} K_t \times \left((1+\rho)^{n-t} - (1+i^*)^{n-t}\right),$$

  which stems from the management of shareholders' capital flows.

4.2 More Safe-Side Requirements

The notions of DCF, splitting the DCF into annual contributions and then splitting the annual contributions themselves, suggest other safe-side requirements. We will consider mappings of the form $\Phi = \Phi(TB_1, TB_2, \Gamma, \rho)$.

Definition 8 (DCF Annual Contributions SSR):

Let $\Phi = [\phi_1(\rho), \phi_2(\rho), \ldots, \phi_n(\rho)]$ where

$$\phi_t(\rho) = g_t(\rho) \quad t = 1, 2, \ldots, n,$$

ρ is the opportunity cost of capital, and $g(\rho)$ is defined in equation (23). For a given ρ, we say that TB_1 is on the safe side if and only if $\phi_t(\rho) \geq 0$ for $t = 1, 2, \ldots, n$, i.e., if and only if

$$\hat{u}_t^* \geq (Z_{t-1} - \hat{V}_{t-1})(\rho - i^*) \quad t = 1, 2, \ldots, n. \tag{25}$$

Definition 8 (D8) leads to some interesting observations, especially when compared with (D3). Considering the nonnegativity of the free portfolio fund, the lower bound for $\hat{u}_t^*$ depends on the difference between the rates ρ and i^* (which can reasonably be assumed to be nonnegative). Note in particular that:

(i) If we require a yield of ρ on the free portfolio fund that is higher than i^*, equation (25) expresses a more severe condition than (D3), as a positive lower bound for the expected annual profit is imposed (unless the free portfolio fund is equal to zero). In addition, the entity of the lower bound depends on the value of the free portfolio fund; hence, it depends on the strategy concerning the choice of $\{K_t\}$. (We will comment on some particular cases later.);

(ii) If $\rho = i^*$, we have (D3);

(iii) The (overall) riskiness of the insurance business can be introduced into the safe-side requirement by properly choosing the rate ρ ($\rho > i^*$), which may include a risk premium.

Definition 9 (DCF SSR):

Let $\Phi = G(\rho)$, where ρ is the opportunity cost of capital, and $G(\rho)$ is defined in equation (18). For a given ρ, we say that TB_1 is on the safe side if and only if $G(\rho) \geq 0$. According to equation (24), $G(\rho) \geq 0$ is achieved when

$$u^* \geq (1 + i^*)^{-n} \sum_{t=0}^{n} -K_t \left((1 + \rho)^{n-t} - (1 + i^*)^{n-t}\right). \tag{26}$$

For certain choices of the sequence $\{K_t\}$, the capital component of DCF can be positive, in which case equation (26) gives a negative bound for total profit.[4] Equation (26) must then be modified as follows

$$u^* \geq \max\left\{0, (1+i^*)^{-n} \sum_{t=0}^{n} -K_t \left((1+\rho)^{n-t} - (1+i^*)^{n-t}\right)\right\}. \quad (27)$$

According to equation (27), we say that a TB_1 is on the safe side if and only if $G(\rho) \geq 0$ and $u^* \geq 0$.

Equation (27) can now be compared to (D4). The lower bound that is now required for total profit u^* depends both on the spread $\rho - i^*$ as well as on the sequence of shareholders' capital flows $\{K_t\}$. If $\rho = i^*$ and/or the capital component of DCF is positive, we find $u^* \geq 0$, i.e., the requirement of (D4).

It is interesting to examine some cases related to particular choices of $\{K_t\}$.

(i) Consider an initial investment followed by one withdrawal only at maturity, i.e., $K_0 < 0$, $K_1 = K_2 = \cdots = K_{n-1} = 0$. From equation (27) we find

$$u^* (1+i^*)^n \geq -K_0 \left[(1+\rho)^n - (1+i^*)^n\right]. \quad (28)$$

Equation (28) points out the need for the total profit to compensate the shareholders for missed returns occurring when the initial investment of shareholders' equity incurrs a return of i^* instead of the required ρ. In terms of annual profits, equations (25) and (22) lead to

$$\hat{u}^*_{t+1} \geq [-K_0 (1+i^*)^t + \sum_{h=1}^{t} \hat{u}^*_h (1+i^*)^{t-h}] (\rho - i^*), \quad (29)$$

which shows that the expected annual profit must cover the missed yield, equal to $\rho - i^*$, obtained on shareholders' equity accumulated at the beginning of the year. As shown in equation (29),

[4]This is unacceptable from the point of view of looking for premiums to meet the benefits. Selling some insurance policies at a loss, however, might be profitable for an entire portfolio in the long run if other policies generate cash flows that can be invested profitably elsewhere. Throughout the rest of this paper we will disregard this opportunity, as it is not allowed by insurance regulations in many countries.

shareholders' capital (which is equal to the free portfolio fund) at time t comes from the accumulated value of the initial investment and previous (undistributed) annual profits.

(ii) Consider the case where $K_0 < 0$ and $K_h = \hat{u}_h^*$ for $h = 1, 2, \ldots, n-1$. In general terms, equations (25) and (22) lead to

$$\begin{aligned} \hat{u}_{t+1}^* \geq [-K_0\,(1+i^*)^t - \sum_{h=1}^{t} K_h\,(1+i^*)^{t-h} \\ + \sum_{h=1}^{t} \hat{u}_h^*\,(1+i^*)^{t-h}]\,(\rho - i^*). \end{aligned} \tag{30}$$

When $K_h > 0$ for $h = 1, 2, \ldots, t$, equation (30) shows the decrease in the lower bound due to shareholders' capital withdrawals. When $K_h = \hat{u}_h^*$, $h = 1, 2, \ldots, t$, equation (30) becomes

$$\hat{u}_{t+1}^* \geq [-K_0\,(1+i^*)^t]\,(\rho - i^*). \tag{31}$$

The loss is incurred only on the accumulated value of the initial investment. Multiplying both sides of equation (31) by $(1+i^*)^{-(t+1)}$ and summing with respect to t, we obtain

$$u^* \geq -n K_0\,(\rho - i^*)\,(1+i^*)^{-1}. \tag{32}$$

The lower bound for total profit provided by equation (32) is more severe than that coming from equation (27), as equation (32) is obtained discounting the annual lower bounds with a higher factor. It can be easily verified that the lower bound implied by equation (28) is greater than that implied by equation (32) (and, therefore, than that implied by equation (27)). When $K_h = \hat{u}_h^*$, $h = 1, 2, \ldots, n-1$, it is possible to reinvest annual profits at the rate ρ, hence weakening the bound on total profit.

(iii) As shown by the above mentioned examples, the notion of prudence depends on the strategy of shareholders' equity (as well as on the yield ρ). It is interesting to consider an objective strategy. To this aim, let us define the sequence $\{M_t\}$, $t = 0, 1, \ldots, n-1$, where M_t is the minimum solvency margin that must be assigned (according to insurance regulation) to the insurance portfolio at time t. Suppose the shareholders set their capital flows such that

$$Z_t - \hat{V}_t = M_t \quad \text{for} \quad t = 0, 1, \ldots, n-1.$$

The consequent bounds on profit can be interpreted on one hand as those implied by the opinion on the riskiness of the insurance business expressed by current legislation (through $\{M_t\}$), and on the other hand by shareholders (through ρ).

5 Numerical Examples

We consider two types of policies: (i) a 15-year endowment insurance with face value of 1,000 monetary units that is issued to an Italian male age 50; and (ii) a 15-year deferred whole life annuity with annual benefits of 100 monetary units that is issued to an Italian male age 50. In both policies premiums are level and paid for 15 years, the second order rate of interest is $i^* = 0.06$, and the opportunity cost of capital is $\rho = 0.08$.

5.1 Endowment Insurance

For the endowment insurance, the second order level of mortality is derived from the Italian Table SIM1992 (which is referred to the Italian male population, observed in 1992). Denoting by $q_x^{SIM1992}$ the rate of mortality calculated according to Table SIM1992, we assume $q_x^* = 0.70\, q_x^{SIM1992}$.

In Table 1 the traditional approach is adopted; thus prudence is analyzed only according to profits. Three different technical bases are used as examples.

- In the first example (Columns (2) to (4) in Table 1), the TB_1 ($i = 0.03$, $q = 1.2\, q^*$) complies with all safe-side requirements.
- In the second example (Columns (5) to (7) in Table 1), the TB_1 ($i = 0.03$, $q = 0.8\, q^*$) is used. This example does not satisfy (D2). The financial profit in each year, however, is enough to allow positive annual profits. Thus, (D3) and (D4) as well as (D6) and (D7), are satisfied.

- In the third example (Columns (8) to (10) in Table 1), a higher technical rate of interest has been chosen. Total profit is dramatically reduced as compared to the first example.

In Tables 2 to 4 the corporate approach is implemented. In each table, a different strategy of shareholders' capital flows has been adopted. In Table 2, we have an initial investment ($K_0 < 0$) and a final withdrawal ($K_1 = K_2 = \cdots = K_{n-1} = 0$, $K_n > 0$). The amount of the initial investment has been chosen according to a reasonable solvency fund to be assigned to the portfolio. Both of the technical bases, ($i = 0.03, q = 1.2\,q^*$) and ($i = 0.03, q = 0.8\,q^*$), satisfy (D8) and (D9) of prudence.

In Tables 3 and 4 withdrawals of shareholders' capital are permitted also at times $t = 1, 2, \ldots, n-1$. The requirements on annual and total profits are relaxed; note, however, that in Table 4 the TB_1 ($i = 0.03, q = 0.8\,q^*$) cannot be adopted as the free portfolio fund becomes negative.

5.2 Deferred Annuity

For the deferred annuity, mortality rates are taken from a projected table, which is obtained from Table SIM1992 using an exponential projection model; it reflects the future expected (decreasing) trend of mortality. We point out that the limiting age of the mortality table is $\omega = 109$.

In Table 5 the traditional approach is adopted. It is difficult to cover financial losses with mortality profits (and mortality losses with financial profits) throughout the whole insurance period. In Table 6 the strategy $K_0 < 0$, $K_1 = K_2 = \cdots = K_{n-1} = 0$ is examined (also in this case, K_0 has been chosen according to a reasonable solvency fund to be assigned to the portfolio). Because of the length of the insurance contract, such strategy is unsatisfactory when (D8) is assumed. In Table 7 withdrawals of shareholders' capital at time $t = 1, 2, \ldots$ are considered. (Their amount has been chosen according to the behavior of annual profits and to the interests required on the initial investment.) Because of the length of the contract, when a sequence $\{K_t\}$ is given, there is not much freedom in the choice of TB_1.

Similar results can be obtained when expenses and other loadings are considered.

Table 1
Endowment Insurance: Traditional Approach

	$i = 3\%, q_x = 1.2\, q_x^*$, and $P = 55.572$			$i = 3\%, q_x = 0.8\, q_x^*$, and $P = 54.441$			$i = 4\%, q_x = 1.2\, q_x^*$, and $P = 51.479$		
t	$\hat{u}_t^*$	${}_f\hat{u}_t^*$	${}_d\hat{u}_t^*$	$\hat{u}_t^*$	${}_f\hat{u}_t^*$	${}_d\hat{u}_t^*$	$\hat{u}_t^*$	${}_f\hat{u}_t^*$	${}_d\hat{u}_t^*$
1	2.274	1.667	0.607	1.026	1.633	-0.607	1.639	1.030	0.609
2	3.883	3.265	0.618	2.614	3.232	-0.618	2.644	2.021	0.623
3	5.532	4.896	0.636	4.230	4.866	-0.636	3.686	3.042	0.644
4	7.224	6.559	0.664	5.869	6.533	-0.664	4.769	4.094	0.675
5	8.932	8.250	0.682	7.549	8.230	-0.681	5.868	5.173	0.696
6	10.668	9.969	0.699	9.258	9.956	-0.698	6.997	6.280	0.716
7	12.415	11.713	0.702	11.008	11.708	-0.700	8.136	7.414	0.722
8	14.173	13.481	0.692	12.796	13.485	-0.690	9.291	8.576	0.715
9	15.942	15.273	0.669	14.619	15.286	-0.667	10.459	9.764	0.695
10	17.714	17.085	0.629	16.480	17.106	-0.626	11.635	10.979	0.656
11	19.482	18.916	0.565	18.382	18.943	-0.561	12.811	12.219	0.592
12	21.240	20.765	0.475	20.324	20.794	-0.471	13.984	13.484	0.500
13	22.982	22.629	0.354	22.306	22.656	-0.350	15.149	14.775	0.374
14	24.704	24.507	0.197	24.331	24.526	-0.195	16.299	16.090	0.209
15	26.398	26.398	0.000	26.398	26.398	0.000	17.430	17.430	0.000
	u^*	120.174		u^*	108.885		u^*	79.299	

Table 2
Endowment Insurance: Corporate Approach

		$i = 3\%$ and $q = 1.2\,q^*$			$i = 3\%$ and $q = 0.8\,q^*$		
t	K_t	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$
0	-20.000	20.000	20.000	-	20.000	20.000	-
1	0.000	76.900	23.474	0.400	75.702	22.226	0.400
2	0.000	136.765	28.765	0.469	134.301	26.174	0.445
3	0.000	199.673	36.023	0.575	195.870	31.975	0.523
4	0.000	265.663	45.408	0.720	260.446	39.762	0.640
5	0.000	334.883	57.065	0.908	328.173	49.697	0.795
6	0.000	407.397	71.157	1.141	399.109	61.936	0.994
7	0.000	483.357	87.841	1.423	473.403	76.660	1.239
8	0.000	562.884	107.284	1.757	551.172	94.055	1.533
9	0.000	646.081	129.663	2.146	632.512	114.318	1.881
10	0.000	733.057	155.158	2.593	717.529	137.657	2.286
11	0.000	823.953	183.949	3.103	806.359	164.299	2.753
12	0.000	918.908	216.225	3.679	899.133	194.480	3.286
13	0.000	1018.070	252.181	4.325	995.997	228.455	3.890
14	0.000	1121.587	292.016	5.044	1097.089	266.494	4.569
15		0.000	0.000	5.840	0.000	0.000	5.330
		K_n	335.935		K_n	308.881	
		$G(\rho)$	85.901		$G(\rho)$	77.372	
		$G^T(\rho)$	90.791		$G^T(\rho)$	82.262	
		$G^C(\rho)$	-4.890		$G^C(\rho)$	-4.890	
		$L(u^*)$	6.473		$L(u^*)$	6.473	

Notes: $L(\hat{u}_t^*)$ denotes the lower bound for $\hat{u}_t^*$; $L(u^*)$ denotes the lower bound for u^*; $G^T(\rho)$ denotes the technical component of DCF; and $G^C(\rho)$ denotes the capital component of DCF.

Table 3
Endowment Insurance: Corporate Approach

		$i = 3\%$ and $q = 1.2q^*$			$i = 3\%$ and $q = 0.8q^*$		
t	K_t	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$
0	-20.000	20.000	20.000	-	20.000	20.000	-
1	1.333	75.567	22.141	0.400	74.369	20.893	0.400
2	1.333	134.019	26.018	0.443	131.554	23.428	0.418
3	1.333	195.428	31.778	0.520	191.625	27.730	0.469
4	1.333	259.830	39.575	0.636	254.614	33.930	0.555
5	1.333	327.367	49.548	0.792	320.657	42.181	0.679
6	1.333	398.097	61.856	0.991	389.809	52.636	0.844
7	1.333	472.165	76.649	1.237	462.211	65.468	1.053
8	1.333	549.688	94.088	1.533	537.975	80.859	1.309
9	1.333	630.759	114.342	1.882	617.191	98.996	1.617
10	1.333	715.482	137.583	2.287	699.955	120.083	1.980
11	1.333	803.991	163.986	2.752	786.396	144.336	2.402
12	1.333	896.414	193.732	3.280	876.640	171.987	2.887
13	1.333	992.894	227.005	3.875	970.820	203.279	3.440
14	1.333	1093.567	263.996	4.540	1069.069	238.473	4.066
15		0.000	0.000	5.280	0.000	0.000	4.769
		K_n	306.234		K_n	279.180	
		$G(\rho)$	87.530		$G(\rho)$	79.001	
		$G^T(\rho)$	90.791		$G^T(\rho)$	82.262	
		$G^C(\rho)$	-3.261		$G^C(\rho)$	-3.261	
		$L(u^*)$	4.316		$L(u^*)$	4.316	

Notes: $L(\hat{u}_t^*)$ denotes the lower bound for $\hat{u}_t^*$; $L(u^*)$ denotes the lower bound for u^*; $G^T(\rho)$ denotes the technical component of DCF; and $G^C(\rho)$ denotes the capital component of DCF.

Table 4
Endowment Insurance: Corporate Approach

		$i = 3\%$ and $q = 1.2\,q^*$			$i = 3\%$ and $q = 0.8\,q^*$		
t	K_t	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$
0	-20.000	20.000	20.000	-	20.000	20.000	-
1	2.000	74.900	21.474	0.400	73.702	20.226	0.400
2	4.000	130.645	22.645	0.429	128.181	20.054	0.405
3	6.000	187.186	23.536	0.453	183.383	19.488	0.401
4	8.000	244.427	24.172	0.471	239.210	18.526	0.390
5	10.000	302.373	24.554	0.483	295.662	17.186	0.371
6	12.000	360.936	24.695	0.491	352.648	15.475	0.344
7	14.000	420.108	24.592	0.494	410.154	13.411	0.309
8	16.000	479.840	24.240	0.492	468.128	11.011	0.268
9	18.000	540.054	23.637	0.485	526.486	8.291	0.220
10	20.000	600.669	22.769	0.473	585.141	5.269	0.166
11	22.000	661.621	21.617	0.455	644.027	1.967	0.105
12	24.000	722.836	20.154	0.432	703.062	-1.591	0.039
13	26.000	784.235	18.346	0.403	762.161	-5.380	-0.032
14	28.000	845.721	16.151	0.367	821.223	-9.372	-0.108
15		0.000	0.000	0.323	0.000	0.000	-0.187
		K_n	43.518		K_n	16.464	
		$G(\rho)$	97.152		$G(\rho)$	88.623	
		$G^T(\rho)$	90.791		$G^T(\rho)$	82.262	
		$G^C(\rho)$	6.361		$G^C(\rho)$	6.361	
		$L(u^*)$	0.000		$L(u^*)$	0.000	

Notes: $L(\hat{u}_t^*)$ denotes the lower bound for $\hat{u}_t^*$; $L(u^*)$ denotes the lower bound for u^*; $G^T(\rho)$ denotes the technical component of DCF; and $G^C(\rho)$ denotes the capital component of DCF.

Table 5

Deferred Annuity: Traditional Approach

	$i = 3\%$, $q = 0.8\,q^*$, and $P = 73.689$			$i = 3\%$, $q = q = 1.1\,q^*$ and $P = 64.979$			$i = 4\%$, $q = 0.8\,q^*$ and $P = 64.979$		
t	$\hat{u}_t^*$	${}_f\hat{u}_t^*$	${}_d\hat{u}_t^*$	$\hat{u}_t^*$	${}_f\hat{u}_t^*$	${}_d\hat{u}_t^*$	$\hat{u}_t^*$	${}_f\hat{u}_t^*$	${}_d\hat{u}_t^*$
1	2.275	2.211	0.064	1.921	1.949	-0.028	1.286	1.232	0.054
2	4.614	4.476	0.138	3.889	3.950	-0.061	2.625	2.508	0.117
3	7.025	6.798	0.226	5.903	6.003	-0.100	4.021	3.828	0.193
4	9.515	9.176	0.338	7.959	8.109	-0.150	5.483	5.193	0.290
5	12.080	11.611	0.469	10.062	10.270	-0.208	7.008	6.604	0.404
6	14.732	14.102	0.630	12.206	12.486	-0.280	8.608	8.062	0.546
7	17.464	16.649	0.816	14.396	14.758	-0.362	10.278	9.568	0.710
8	20.284	19.251	1.033	16.628	17.088	-0.459	12.026	11.123	0.904
9	23.197	21.909	1.288	18.901	19.476	-0.574	13.860	12.727	1.133
10	26.207	24.620	1.587	21.214	21.923	-0.709	15.784	14.380	1.404
11	29.312	27.385	1.928	23.567	24.430	-0.862	17.797	16.083	1.715
12	32.517	30.201	2.316	25.960	26.999	-1.039	19.908	17.836	2.072
13	35.823	33.067	2.756	28.390	29.629	-1.239	22.118	19.639	2.479
14	39.236	35.982	3.255	30.856	32.324	-1.468	24.438	21.493	2.945
15	42.768	38.943	3.825	33.352	35.083	-1.731	26.879	23.397	3.481
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 5 (continued)
Deferred Annuity: Traditional Approach

	$i = 3\%, q = 0.8\,q^*$, and $P = 73.689$			$i = 3\%, q = q = 1.1\,q^*$ and $P = 64.979$			$i = 4\%, q = 0.8\,q^*$ and $P = 64.979$		
t	$\hat{u}_t^*$	${}_f\hat{u}_t^*$	${}_d\hat{u}_t^*$	$\hat{u}_t^*$	${}_f\hat{u}_t^*$	${}_d\hat{u}_t^*$	$\hat{u}_t^*$	${}_f\hat{u}_t^*$	${}_d\hat{u}_t^*$
45	2.946	1.099	1.847	0.076	0.811	-0.735	2.513	0.709	1.804
46	2.317	0.807	1.510	-0.011	0.588	-0.599	1.999	0.521	1.478
47	1.777	0.577	1.201	-0.060	0.414	-0.474	1.550	0.373	1.177
48	1.332	0.400	0.932	-0.084	0.283	-0.367	1.174	0.259	0.915
49	0.967	0.268	0.699	-0.088	0.186	-0.274	0.861	0.174	0.688
50	0.680	0.173	0.508	-0.081	0.118	-0.199	0.612	0.112	0.500
51	0.458	0.107	0.352	-0.066	0.071	-0.137	0.416	0.069	0.347
52	0.296	0.063	0.233	-0.050	0.041	-0.091	0.271	0.041	0.230
53	0.184	0.035	0.149	-0.036	0.022	-0.058	0.170	0.023	0.147
54	0.104	0.018	0.085	-0.022	0.011	-0.033	0.097	0.012	0.085
55	0.059	0.009	0.050	-0.015	0.005	-0.020	0.056	0.006	0.050
56	0.027	0.004	0.023	-0.007	0.002	-0.010	0.026	0.002	0.023
57	0.012	0.001	0.011	-0.004	0.001	-0.005	0.012	0.001	0.011
58	0.003	0.001	0.002	-0.001	0.000	-0.001	0.003	0.000	0.002
59	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	u^*	326.378		u^*	239.886		u^*	206.533	

Table 6

Deferred Annuity: Corporate Approach

		$i = 3\%$ and $q = 1.2\,q^*$			$i = 3\%$ and $q = q = 0.8\,q^*$		
t	K_t	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$
0	-60.000	60.000	60.000	-	60.000	60.000	-
1	0.000	141.711	65.875	1.200	128.917	64.886	1.200
2	0.000	227.997	74.441	1.317	201.695	71.404	1.298
3	0.000	319.112	85.932	1.489	278.550	79.709	1.428
4	0.000	415.321	100.603	1.719	359.703	89.974	1.594
5	0.000	516.890	118.719	2.012	445.380	102.381	1.799
6	0.000	624.105	140.574	2.374	535.823	117.131	2.048
7	0.000	737.259	166.473	2.811	631.280	134.438	2.343
8	0.000	856.665	196.745	3.329	732.015	154.530	2.689
9	0.000	982.652	231.747	3.935	838.307	177.662	3.091
10	0.000	1115.564	271.859	4.635	950.445	204.106	3.553
11	0.000	1255.761	317.482	5.437	1068.736	234.149	4.082
12	0.000	1403.626	369.048	6.350	1193.501	268.106	4.683
13	0.000	1559.559	427.014	7.381	1325.080	306.310	5.362
14	0.000	1723.986	491.871	8.540	1463.834	349.127	6.126
15	0.000	1897.356	564.151	9.837	1610.140	396.953	6.983
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 6 (continued)
Deferred Annuity: Corporate Approach

		$i = 3\%$ and $q = 1.2\,q^*$			$i = 3\%$ and $q = q = 0.8\,q^*$		
t	K_t	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$
45	0.000	5347.312	5311.426	100.160	3697.691	3662.637	69.059
46	0.000	5658.625	5632.428	106.229	3910.027	3884.394	73.253
47	0.000	5990.748	5972.151	112.649	4137.235	4119.008	77.688
48	0.000	6344.604	6331.812	119.443	4379.880	4367.322	82.380
49	0.000	6721.183	6712.688	126.636	4638.575	4630.223	87.346
50	0.000	7121.551	7116.129	134.254	4913.987	4908.648	92.604
51	0.000	7546.867	7543.556	142.323	5206.849	5203.584	98.173
52	0.000	7998.388	7996.465	150.871	5517.969	5516.070	104.072
53	0.000	8477.487	8476.437	159.929	5848.243	5847.204	110.321
54	0.000	8985.666	8985.126	169.529	6198.667	6198.133	116.944
55	0.000	9524.543	9524.293	179.703	6570.324	6570.076	123.963
56	0.000	10095.543	10095.778	190.486	6964.413	6964.307	131.402
57	0.000	10701.576	10701.537	201.916	7382.215	7382.177	139.286
58	0.000	11343.644	11343.632	214.031	7825.123	7825.110	147.644
59		0.000	0.000	226.873	0.000	0.000	156.502
		K_n	12024.263		K_n	8294.630	
		$G(\rho)$	68.250		$G(\rho)$	28.470	
		$G^T(\rho)$	108.334		$G^T(\rho)$	68.554	
		$G^C(\rho)$	-40.084		$G^C(\rho)$	-40.084	
		$L(u^*)$	120.762		$L(u^*)$	120.762	

Table 7
Deferred Annuity: Corporate Approach

		$i = 3\%$ and $q = 1.2\,q^*$			$i = 3\%$ and $q = q = 0.8\,q^*$		
t	K_t	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$
0	-60.000	60.000	60.000	-	60.000	60.000	-
1	6.000	135.711	59.875	1.200	122.917	58.886	1.200
2	8.000	213.637	60.081	1.197	187.335	57.044	1.178
3	11.000	292.891	59.711	1.202	252.328	53.488	1.141
4	13.000	374.526	59.808	1.194	318.908	49.179	1.070
5	16.000	457.648	59.476	1.196	386.138	43.138	0.984
6	18.000	543.308	59.777	1.190	455.026	36.334	0.863
7	21.000	630.614	59.828	1.196	524.635	27.793	0.727
8	24.000	719.621	59.701	1.197	594.972	17.487	0.556
9	27.000	810.386	59.480	1.194	666.040	5.396	0.350
10	30.000	902.961	59.256	1.190	737.843	-8.497	0.108
11	33.000	997.403	59.124	1.185	810.377	-24.209	-0.170
12	36.000	1093.766	59.188	1.182	883.641	-41.754	-0.484
13	39.000	1192.108	59.562	1.184	957.629	-61.141	-0.835
14	43.000	1291.488	59.373	1.191	1031.335	-83.372	-1.223
15	46.000	1392.907	59.703	1.187	1105.692	-107.495	-1.667
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 7 (Continued)
Deferred Annuity: Corporate Approach

		$i = 3\%$ and $q = 1.2\,q^*$			$i = 3\%$ and $q = q = 0.8\,q^*$		
t	K_t	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$	Z_t	$Z_t - \hat{V}_t$	$L(\hat{u}_t^*)$
45	6.000	94.002	58.116	1.154	-1555.619	-1590.673	-29.947
46	6.000	84.116	57.920	1.162	-1664.481	-1690.115	-31.813
47	5.000	76.769	58.172	1.158	-1776.744	-1794.971	-33.802
48	5.000	70.786	57.994	1.163	-1893.938	-1906.496	-35.899
49	4.000	66.936	58.441	1.160	-2015.672	-2024.024	-38.130
50	4.000	64.049	58.628	1.169	-2143.515	-2148.854	-40.480
51	4.000	61.915	58.604	1.173	-2278.103	-2281.368	-42.977
52	4.000	60.339	58.416	1.172	-2420.080	-2421.979	-45.627
53	4.000	59.155	58.105	1.168	-2570.089	-2571.128	-48.440
54	4.000	58.234	57.695	1.162	-2728.765	-2729.299	-51.423
55	4.000	57.465	57.215	1.154	-2896.753	-2897.001	-54.586
56	4.000	56.783	56.675	1.144	-3074.689	-3074.796	-57.940
57	4.000	56.127	56.088	1.134	-3263.233	-3263.272	-61.496
58	4.000	55.469	55.457	1.122	-3463.053	-3463.065	-65.265
59		0.000	0.000	1.109	0.000	0.000	-69.261
		K_n	58.797		K_n	-3670.836	
		$G(\rho)$	230.785		$G(\rho)$	191.005	
		$G^T(\rho)$	108.334		$G^T(\rho)$	68.554	
		$G^C(\rho)$	122.451		$G^C(\rho)$	122.451	
		$L(u^*)$	0.000		$L(u^*)$	0.000	

6 Some Final Remarks

This paper first considers a purely actuarial approach to prudence and then shows how it is possible to introduce risk measures in a structure based on expected values only. Risk is introduced through the cost of capital and the amount of shareholders' capital (i.e., the free portfolio fund) linked to the insurance portfolio.

Safe-side requirements can also be formulated in a stochastic framework by considering the distribution of the random profits. Below is a possible definition:

Definition 10 (A Possible Stochastic Safe-Side Requirement):

Let R denote the present value at time 0 of the future random profits, and let

$$\Pr[R \le r] = F_R(r \mid TB_1, TB_2).$$

We can say that a TB_1 is on the safe side if and only if

$$F_R(0 \mid TB_1, TB_2) \le p$$

where p is a given bound.

Similar (and more strict) definitions can be given considering annual random profits or their components. Several analytical results can be used in this approach; among the more recent contributions, we mention Hesselager and Norberg (1996), which deals with multistate models.

An individual approach (i.e., based on a single contract) has the disadvantage that in order to limit the safety loading, low levels of p must be chosen. A collective approach based on the entire portfolio may help in quoting competitive premiums; however, forecasts on the future size and composition of the portfolio are required, thus leading to a further element of uncertainty in the choice of the first order basis.

Finally, we must emphasize that an approach based on expected values only has the advantage that the results, relative to a whole portfolio, are linear in respect of those relative to single cohorts. On the other hand, a stochastic approach to prudence may lead to a more comprehensive classification of the notion of safe-side requirements.

References

Goford, J. "The Control Cycle: Financial Control of a Life Assurance Company." *Journal of the Institute of Actuaries Students' Society* 28 (1985): 99-114.

Haberman S. and Sibbett T.A. (eds.) *History of Actuarial Science.* London, England: Pickering & Chatto, 1995.

Hesselager, O. and Norberg, R. "On Probability Distributions of Present Values in Life Insurance." *Insurance: Mathematics & Economics* 18 (1996): 35-42.

Hoem, J.M. "The Versatility of the Markov Chain as a Tool in the Mathematics of Life Insurance." *Transactions of the 23rd International Congress of Actuaries* R (1988): 171-202.

Homans S. "On the Equitable Distribution of Surplus." *Journal of the Institute of Actuaries* 11 (1863): 121-129

Loewy, A. "Zur Theorie und Anwendung der Intensitäten in der Versicherungsmathematik." *Sitzungsberichte der Heidelberger Akademic der Wissenschaften* 8A (1917).

Peccati, L. "Multiperiod Analysis of a Levered Portfolio." *Rivista di Matematica per le Scienze Economiche e Sociali* 12, no. 1 (1989): 157-166.

Olivieri, A. "Safe-Side Requirements in the Framework of Multiple State Models for the Insurances of the Person." *Applied Stochastic Models in Business and Industry* 15 (1999): 393-408.

Olivieri, A. and Pitacco E. "Sulla Nozione di Prudenzialità in Assicurazione Vita: Profili Attuariali e Finanziari." *Atti del XXI Convegno AMASES,* Roma (1997): 515-530.

Actuarial Analysis of Retirement Income Replacement Ratios

Robert Keng Heong Lian,* Emiliano A. Valdez,† and Chan Kee Low‡

Abstract§

A measure of level of post-retirement standard of living is the replacement ratio, i.e., percentage of final salary received as annual retirement income derived from savings. The replacement ratio depends on many factors including salary, salary increases, investment returns, and post-retirement mortality. Elementary life contingencies techniques are used to develop a replacement ratio formula and analyze its sensitivity to these factors.

Key words and phrases: *retirement planning, savings, interest rates, life annuity, social security*

*Robert Keng Heong Lian, M.Sc., A.S.A., is an associate professor of actuarial science at Nanyang Technological University. He received his M.Sc. in actuarial science from Northeastern University. Mr. Lian was the general manager/principal officer of British American Life Assurance before joining the university.

Mr. Lian's address is: Division of Actuarial Science and Insurance, Nanyang Business School, Nanyang Technological University, Singapore 639798. Internet address: *akhlian@ntu.edu.sg*

†Emiliano A. Valdez, Ph.D., F.S.A. is an assistant professor of actuarial science at Nanyang Technological University. He holds a Ph.D. in Business from the University of Wisconsin-Madison. In 1998, Dr. Valdez was a co-winner of the Society of Actuaries' Halmstad Prize and Edward A. Lew Award.

Dr. Valdez's address is: Division of Actuarial Science and Insurance, Nanyang Business School, Nanyang Technological University, SINGAPORE 639798. Internet address: *aemiliano@ntu.edu.sg*

‡Chan Kee Low, Ph.D., is an associate professor at Nanyang Technological University. His research interests include applications of statistical techniques to problems in accounting, economics, insurance, and finance.

Dr. Low's address is: Division of Actuarial Science and Insurance, Nanyang Business School, Nanyang Technological University, SINGAPORE 639798. Internet address: *acklow@ntu.edu.sg*

§This paper received a "Best Paper Award" at the 4th Asia-Pacific Risk and Insurance Association Conference in Perth, Australia in July 2000. The authors thank Prof. John H. Pollard of Macquarie University, Australia and the referees for their valuable comments and suggestions, and thank the editor for his assistance in developing Section 4.1.

1 Introduction

Planning for retirement is increasingly important because individuals are living longer, demanding a better quality of life, and are facing a higher cost of living. Individuals must continually assess how much to save during their working years and how much retirement income will be sufficient. These types of decisions are generally influenced by several factors including the education, level of income, age at entry into the work force, expected retirement age, investment returns, survivorship trends, inflation, wage increases, and expense pattern.

In retirement planning, the individual first must consider the various sources of income he or she will have at retirement. In several countries, the primary sources of income are generally derived from social security benefits, employer-provided retirement benefits, and personal savings. As illustrated in Rejda (1988, Chapter 4, pages 64-65), retirement benefits from social security or other public pension schemes and personal savings and investments generally form a large proportion of post-retirement income in the U.S.

Once the various sources of retirement income have been identified, the individual must determine whether he or she will have sufficient income for retirement. As Cordell (1999) observes, post-retirement financial adequacy is often measured by the level of standard of living enjoyed by the individual during the period just prior to retirement. This level of standard of living is best quantified in terms of the annual salary of the individual just prior to retirement.

In this paper, we will use the replacement ratio, RR_x, to express the amount of annual retirement income as a fraction of the individual's final annual salary just prior to retirement. The amount of annual retirement income can be determined by converting savings into a life annuity and aggregating all annuity payments. In effect, we have

$$\mathrm{RR}_x = \frac{\text{Annual Retirement Income}}{\text{Final Annual Salary}}. \tag{1}$$

So, for example, consider a retiree age 65 with a final salary of \$100,000. The retiree has savings of \$50,000 in a bank account, \$575,000 in a defined contribution plan,[1] receives \$7,200 per year from a social insurance scheme, and receives \$12,000 per year from a defined benefit

[1] In a defined contribution plan, the employer and employee make specific contributions into a fund for the employee. The accumulated contributions at retirement are used to purchase an annuity pension benefits.

plan.[2] If it costs \$11.541 to purchase an annuity due paying \$1 per year for life, then the annual retirement income is:

$$\begin{aligned}\text{Annual Retirement Income} &= \$7,200 + \$12,000 + \frac{\$50,000 + \$575,000}{11.541} \\ &= \$73,354.75,\end{aligned}$$

yielding a replacement ratio of 73.355%.

There are several reasons to purchase a fixed life annuity at retirement. Alternatives to purchasing a life annuity, however, do exist. For example, the retiree can purchase an annuity certain for a period that is equal to the retiree's life expectancy at retirement. An obvious problem with this choice is that the retiree may outlive the annuity and be left with no source of income beyond that age. On the other hand, it can be argued that a larger amount of income can be derived with an annuity-certain if the retired individual does not live beyond the life expectancy at retirement. With a life annuity, the risk resulting from survivorship beyond the life expectancy at age of retirement is transferred to the company issuing the life annuity. Furthermore, any investment risk also is borne by the issuing company. With a fixed life annuity, the interest rate generally is guaranteed from issue by the company. From the perspective of the retired individual, there is no uncertainty that may arise from interest rate changes that may affect his or her annuity income. See Borch (1984) for more on these issues.

Inflation is another consideration. In the annuity market, it is possible to purchase an annuity contract that provides some floor of protection against inflation. In some countries, the law may require that income arising from disbursements of retirement benefits come in the form of a life annuity (McGill 1984, Chapter 6, pages 125–129).

We will now develop replacement ratio formulas and explore the various factors that can affect the replacement ratio at retirement. The results are general and intuitively appealing. They are not country-specific as we take into account that the various sources of income that are common to most countries (although the proportions derived from the various sources may differ significantly). We do not recommend the level of replacement ratio that is adequate for an individual's retirement. Our hope is to provide a tool that can be used to assist in retirement planning.

[2] In a defined benefit plan the employer promises to pay an annual pension benefit to the employee for life starting at retirement. The size of the benefit is determined by a specified mathematical formula.

The paper is organized as follows. In Section 2 we provide a simple model for calculating replacement ratios where the sole source of retirement income is derived from personal savings. Section 3 contains an analysis of the various partial derivatives of the replacement ratio. Section 4 gives several modifications of the simple model. In section 5, the case of Singapore is considered. The paper concludes in Section 6.

The following standard actuarial notation are used; see, for example, Kellison (1991). For $j \geq 0$ and $n = 1, 2, \ldots,$

$$\ddot{s}_{\overline{n}|j} = \sum_{t=0}^{n-1} (1+j)^{n-t};$$

$$s_{\overline{n}|j} = \sum_{t=1}^{n} (1+j)^{n-t};$$

$$\ddot{a}_{\overline{n}|j} = \sum_{t=0}^{n-1} (1+j)^{-t};$$

$$a_{\overline{n}|j} = \sum_{t=1}^{n} (1+j)^{-t};$$

$$(Ia)_{\overline{n}|j} = \sum_{t=1}^{n} t(1+j)^{-t};$$

$$(D\ddot{a})_{\overline{n}|j} = \sum_{t=0}^{n-1} (n-t)(1+j)^{-t};$$

$${}^{k}\ddot{a}_{r} = \sum_{t=0}^{\infty} (1+k)^{-t} {}_{t}p_{r};$$

where ${}^{k}\ddot{a}_{r}$ is the present value of a life annuity due paying 1 annually beginning age r and discounted at the effective rate of interest k, and ${}_{t}p_{r}$ is the probability that a person age r survives to age $r+t$, and

$${}^{k}(Ia)_{r} = \sum_{t=1}^{\infty} t(1+k)^{-t} {}_{t}p_{r}.$$

2 Calculating Replacement Ratios: A Simple Model

Consider an individual currently age x whose sole source of retirement income is derived from personal savings and saves a percentage s of annual salary for retirement. We assume that the person will retire at the normal retirement age r. Let $\mathrm{AS}(x)$ be this individual's actual annual salary at age x and assume that salaries increase at the annual rate of w. Salaries are assumed to be paid at the start of the year. So the expected annual salary from age $x+t$ to $x+t+1$ for a person currently age x is paid at age $x+t$ and is given by

$$\mathrm{ES}_x(x+t) = \mathrm{AS}(x)(1+w)^t . \tag{2}$$

The individual will save the amount of $s\,\mathrm{ES}(x+t)$ at age $x+t$. If savings accumulate at the annual effective interest rate of i, then the portion of the retirement income derived from savings at age $x+t$ will be

$$Z(x+t,r) = s\,\mathrm{AS}(x)(1+w)^t(1+i)^{r-x-t} . \tag{3}$$

Let FS_x denote the accumulated future savings resulting from the annual savings made from age x to retirement, then

$$\begin{aligned}\mathrm{FS}_x &= \sum_{t=0}^{r-x-1} Z(x+t,r)\\ &= s\,\mathrm{AS}(x)\,G(i,w),\end{aligned} \tag{4}$$

where

$$G(i,w) = (1+i)^{r-x}\sum_{t=0}^{r-x-1}\left(\frac{1+w}{1+i}\right)^t .$$

It is straightforward to show that:

$$G(i,w) = \begin{cases} (1+w)^{r-x}\,\ddot{s}_{\overline{r-x}|\,j_1} & \text{if } i>w\\ (1+i)^{r-x}(r-x) & \text{if } i=w\\ \frac{(1+i)^{r-x}}{(1+j_2)}\,\ddot{s}_{\overline{r-x}|\,j_2} & \text{if } i<w \end{cases} \tag{5}$$

where

$$\begin{aligned} j_1 &= \left(\frac{1+i}{1+w}\right) - 1,\\ j_2 &= \left(\frac{1+w}{1+i}\right) - 1. \end{aligned}$$

Suppose that at age x, the individual has saved accumulated past of PS_x. At retirement, the individual can expect to have saved a total of

$$\mathrm{TS}_x(r) = \mathrm{PS}_x(1+i)^{r-x} + \mathrm{FS}_x. \tag{6}$$

The amount TS_x will then be used at age r to purchase a retirement life annuity with level annual payments of $B_x(r)$ given by

$$B_x(r) = \frac{\mathrm{TS}_x(r)}{{}^k\ddot{a}_r} \tag{7}$$

where ${}^k\ddot{a}_r$ is the actuarial present value of a life annuity due.

The replacement ratio at normal retirement age r for a person currently age x, $\mathrm{RR}_x(r)$, is given by

$$\mathrm{RR}_x(r) = \frac{B_x(r)}{\mathrm{ES}_x(r)}. \tag{8}$$

Note that

$$\begin{aligned}
\mathrm{RR}_x(r) &= \frac{\mathrm{PS}_x(1+i)^{r-x} + \mathrm{FS}_x}{\mathrm{AS}(x)(1+w)^{r-x}\,{}^k\ddot{a}_r} \\
&= \frac{\mathrm{PS}_x[(1+i)/(1+w)]^{r-x}}{\mathrm{AS}(x)\,{}^k\ddot{a}_r} + \frac{sG(i,w)(1+w)^{-(r-x)}}{{}^k\ddot{a}_r} \\
&= \mathrm{ps}_{x,r} + \mathrm{fs}_{x,r}
\end{aligned} \tag{9}$$

where

$$\mathrm{ps}_{x,r} = \left[\frac{1+i}{1+w}\right]^{r-x} \frac{c_x}{{}^k\ddot{a}_r} \tag{10}$$

and

$$\mathrm{fs}_{x,r} = \frac{s\,G(i,w)}{(1+w)^{r-x}\,{}^k\ddot{a}_r}. \tag{11}$$

The term $\mathrm{ps}_{x,r}$ can be intuitively interpreted as the portion of replacement ratio contributed by savings already made at age x accumulated to retirement, and the term $\mathrm{fs}_{x,r}$ can be intuitively interpreted as that portion contributed by savings from salary beginning from age x until retirement. Note that

$$c_x = \frac{\text{PS}_x}{\text{AS}\,(x)}$$

from equation (10) represents the proportion of past savings expressed as a percentage of salary at age x.

Some interesting results can be derived from $y\,(r)$

$$y\,(r) = \frac{\text{FS}_x}{\text{PS}_x\,(1+i)^{r-x}},$$

which is the ratio of future expected savings to past savings. If $i > w$, then

$$y\,(r) = \frac{s\ \text{AS}\,(x)}{\text{PS}_x}\left(\frac{1+w}{1+i}\right)^{r-x}\ddot{s}_{\overline{r-x}|\,j_1} = \frac{s}{c_x}\ddot{a}_{\overline{r-x}|\,j_1}$$

and therefore

$$\frac{\partial y\,(r)}{\partial r} = \left(\frac{s}{c_x}\right)\left(\frac{\delta_1}{d_1}\right)\left(\frac{1}{1+j_1}\right)^{r-x} \leq \frac{s}{c_x},$$

which gives an upper bound on the change y for increases in the normal retirement age. Here, $d_1 = j_1/(1+j_1)$ and $\delta_1 = \ln(1+j_1)$ are the usual discount and force of interest corresponding to interest rate j_1. If $i = w$, it is straightforward to show that

$$y\,(r) = \frac{s}{c_x}\,(r-x)$$

so that we have

$$\frac{\partial y\,(r)}{\partial r} = \frac{s}{c_x}.$$

If $i < w$, then

$$y\,(r) = \frac{s\ \text{AS}\,(x)}{\text{PS}_x}\left(\frac{\ddot{s}_{\overline{r-x}|\,j_2}}{1+j_2}\right) = \frac{s}{c_x}s_{\overline{r-x}|\,j_2}$$

and therefore

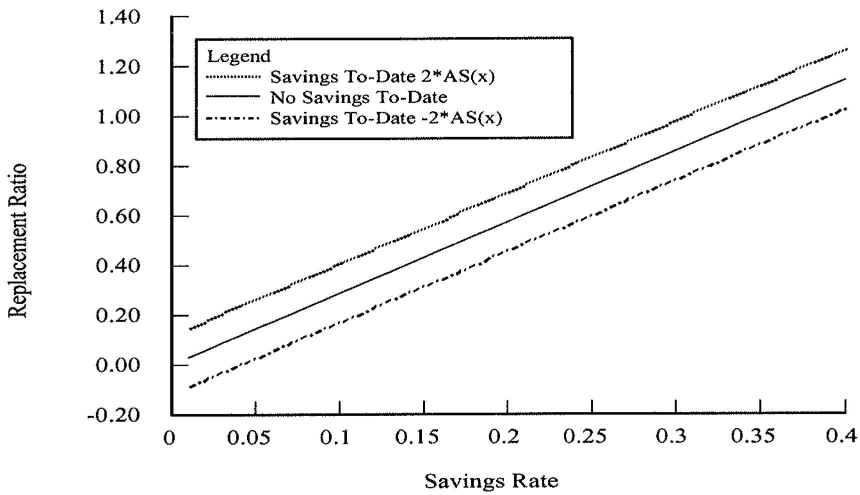
$$\frac{\partial y(r)}{\partial r} = \frac{s}{c_x}\frac{\delta_2}{j_2}(1+j_2)^{r-x} \geq \frac{s}{c_x},$$

which gives a lower bound on the change of the ratio for increases in the normal retirement age. To interpret these results, when the increase in the wages is greater than the interest earned on savings, then salary increases will raise FS_x relative to PS_x. For the other cases, a similar interpretation will hold.

For purposes of illustration, Figure 1 displays the relationship between the savings rate and the replacement ratio. In Figure 1, we assume the following: $x = 25$, $r = 65$, $i = 3$ percent, $w = 4$ percent, and a life annuity factor of $\ddot{a}_{65} = 11.541$ based on the United Kingdom's A1967-70 Mortality Table (ultimate). Savings to-date vary between -200 percent (a negative savings represents a net debt) and +200 percent of current annual salary. Notice the relationship between savings rate and replacement ratio is linear, with a greater intercept for larger savings-to-date. The linear relationship is positively sloping, which demonstrates the fact that increasing savings rate tends to increase the income replacement ratio at retirement. If the assumptions hold, a 25 year old individual with zero savings to-date must save about 35 percent of his or her annual salary to achieve a 100 percent replacement ratio at retirement at age 65.

Figure 2 displays the relationship of the age at retirement and the replacement ratio assuming $x = 25$, $i = 3$ percent, $w = 4$ percent, $s = 20$ percent, and life annuity factors based on the A1967-70 Mortality Table (ultimate). Savings to-date vary between -200 percent and +200 percent of current annual salary. Unlike Figure 1, Figure 2 displays a nonlinear relationship between retirement age and replacement ratio. It demonstrates that deferring the retirement, a choice that many can make today because of increasing life expectancy, will help increase the replacement ratio at retirement. According to the figure, a 25 year old individual with zero savings who intends to save 20 percent of salary annually can achieve a 100 percent replacement ratio when retirement is made at a little below age 75. Furthermore, for the same individual, if retirement were chosen at age 65, the replacement ratio will be below 60 percent.

Figure 1
Replacement Ratio $RR_x(65)$ and Savings Rate s



3 Sensitivity of the Replacement Ratio

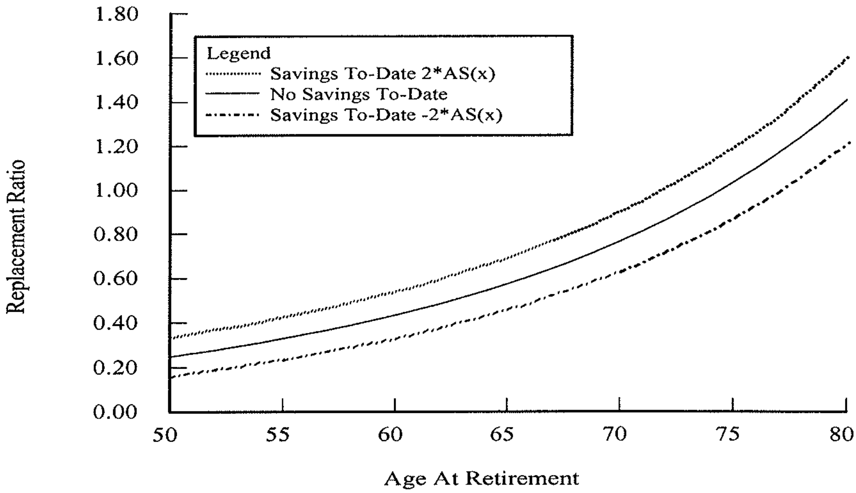
We now investigate the sensitivity of the RR_x to r, i, w, and x, respectively. To illustrate, the replacement ratio depends on the age at retirement r. Intuitively, for an individual who decides to retire at a later age, his or her replacement ratio is expected to increase because he or she will have a longer employment period which allows him or her to save more wealth for retirement. Furthermore, a later retirement age allows for a smaller life annuity factor ${}^k\ddot{a}_r$ (which decreases as r increases) and, as a result, a larger replacement ratio because the annuity is spread over a shorter expected retirement period.

3.1 The Effect of r on RR_x

We now show that, under certain conditions $\partial\, RR_x(r)/\partial r \geq 0$. Three different cases are considered.

Case 1: $i > w$. If the interest rate is larger than the rate of wage increase, the replacement ratio in equation (9) becomes

Figure 2
Replacement Ratio $\mathrm{RR}_x(r)$ and Age at Retirement r



$$\mathrm{RR}_x(r) = \frac{c_x (1+j_1)^{r-x} + s\, \ddot{s}_{\overline{r-x}|\, j_1}}{{}^k\ddot{a}_r} = \frac{f}{g} \tag{12}$$

where

$$f = c_x (1+j_1)^{r-x} + s\, \ddot{s}_{\overline{r-x}|\, j_1}$$

and

$$g = {}^k\ddot{a}_r.$$

Note that

$$\frac{\partial g}{\partial r} = g' = \sum_{t=0}^{\infty} v^t\, {}_tp_r\, (\mu_r - \mu_{r+t}) \tag{13}$$

where μ_y is the force of mortality at age $y \geq r$; see Jordan (1967). Now, consider the first derivative of equation (12):

$$\frac{\partial}{\partial r}\mathrm{RR}_x(r) = \frac{\partial}{\partial r}\left(\frac{f}{g}\right) = \left(\frac{f'g - fg'}{g^2}\right), \tag{14}$$

where

$$f' = \frac{\partial f}{\partial r} = (1+j_1)^{r-x}\,\delta_1\left(c_x + \frac{s}{d_1}\right) = \delta_1\left(f + \frac{s}{d_1}\right).$$

It is apparent that $f > 0$ and $g > 0$; similarly, $f' > 0$. If the post-retirement force of mortality is a nondecreasing function of age, i.e., $\mu_r \leq \mu_{r+t}$ for any $t > 0$, then $g' \leq 0$. Thus, from equation (14), $\partial\,\mathrm{RR}_x(r)/\partial r > 0$.

Case 2: $i = w$. If the interest rate and the rate of wage increase are the same, the replacement ratio in equation (9) becomes

$$\mathrm{RR}_x(r) = \frac{c_x + s(r-x)}{{}^k\ddot{a}_r} = \frac{f}{g}, \tag{15}$$

where

$$f = c_x + (r-x)s \quad \text{and } g = {}^k\ddot{a}_r.$$

Note that $f' = s > 0$. Under a nondecreasing force of mortality requirement, it is straightforward to see that $\partial\,\mathrm{RR}_x(r)/\partial r >$ 0.

Case 3: $i < w$. If the interest rate is smaller than the rate of wage increase, the replacement ratio in equation (9) becomes

$$\mathrm{RR}_x(r) = \frac{c_x(1+j_1)^{r-x} + s(1+j_1)^{r-x}\,s_{\overline{r-x}|j_2}}{{}^k\ddot{a}_r} = \frac{f}{g} \tag{16}$$

where

$$f = c_x(1+j_1)^{r-x} + s(1+j_1)^{r-x}\,s_{\overline{r-x}|j_2} \quad \text{and } g = {}^k\ddot{a}_r.$$

It follows that

$$f' = \frac{\delta_2}{(1+j_2)^{r-x}} \left(\frac{s - c_x j_2}{j_2} \right). \tag{17}$$

Under the condition $s > c_x j_2$, we have $f' > 0$. The reverse is true when $s \leq c_x j_2$. Consider the first derivative of equation (16). Both f and g are positive terms. Under the condition $s > c_x j_2$ we also have $f' > 0$. Furthermore, if the force of mortality is increasing with age, we have $g' \leq 0$. Thus, $\partial \mathrm{RR}_x(r) / \partial r > 0$.

We are now ready to state our first result.

Proposition 1. *If the post-retirement force of mortality μ_y is a nondecreasing function of $y \geq r$, then $\partial RR_x(r) / \partial r > 0$ if (i) $i \geq w$, or (ii) $i < w$ and $s > c_x j_2$.*

According to Proposition 1, there will be possible situations where $\mathrm{RR}_x(r)$ is a decreasing function of r. This does not mean, however, that the annuity payment is also a decreasing function of r. In other words, the retiree is not necessarily getting a smaller annuity payment by delaying retirement.

3.2 The Effect of i on $\mathrm{RR}_x(r)$

From equation (9), we have

$$\frac{\partial \mathrm{RR}_x(r)}{\partial i} = \frac{\partial \mathrm{ps}_{x,r}}{\partial i} + \frac{\partial \mathrm{fs}_{x,r}}{\partial i} \tag{18}$$

where

$$\begin{aligned} \frac{\partial \mathrm{ps}_{x,r}}{\partial i} &= \frac{c_x (r-x)}{{}^k\ddot{a}_r} \frac{1}{(1+w)} \left(\frac{1+i}{1+w} \right)^{r-x-1} \\ &= \frac{c_x (r-x)}{{}^k\ddot{a}_r} \frac{(1+j_1)^{r-x-1}}{(1+w)} \end{aligned} \tag{19}$$

and

$$\frac{\partial \mathrm{fs}_{x,r}}{\partial i} = \frac{s}{(1+w)^{r-x}\, {}^k\ddot{a}_r} \frac{\partial G(i,w)}{\partial i}. \tag{20}$$

We can easily see that

$$\frac{\partial \ddot{s}_{\overline{n}|j}}{\partial j} = \sum_{t=1}^{n-1} (n-t)(1+j)^{n-t-1} = (1+j)^{n-1}\left[n\ddot{a}_{\overline{n}|j} - (Ia)_{\overline{n}|j}\right]. \quad (21)$$

Equation (21) is obviously nonnegative because $n\ddot{a}_{\overline{n}|j} > (Ia)_{\overline{n}|j}$. We now consider the three cases: $i > w$, $i = w$, and $i < w$.

Case 1: $i > w$. If the interest rate is larger than the rate of wage increase, from equations (5) and (21), we have

$$\begin{aligned}\frac{\partial G(i,w)}{\partial i} &= (1+w)^{r-x}\frac{\partial \ddot{s}_{\overline{r-x}|j_1}}{\partial j_1}\frac{\partial j_1}{\partial i} \\ &= (1+w)^{r-x-1}\frac{\partial \ddot{s}_{\overline{r-x}|j_1}}{\partial j_1} \\ &= (1+i)^{r-x-1}\left[(r-x)\ddot{a}_{\overline{r-x}|j_1} - (Ia)_{\overline{r-x}|j_1}\right]\end{aligned} \quad (22)$$

which is obviously nonnegative.

Case 2: $i = w$. If the interest rate and the rate of wage increase are the same, from equation (5), we have

$$\frac{\partial G(i,w)}{\partial i} = (r-x)^2(1+i)^{r-x-1} \quad (23)$$

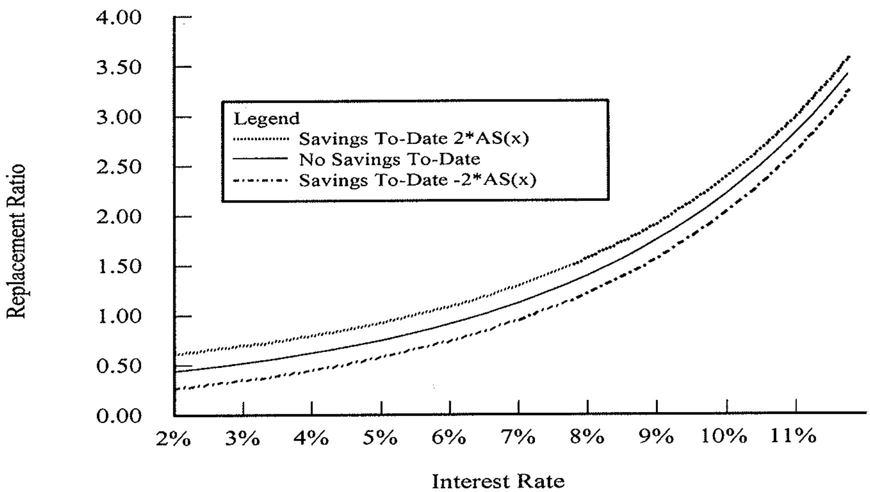
which is also obviously nonnegative.

Case 3: $i < w$. If the interest rate is smaller than the rate of wage increase from equations (5) and (21), we have

$$\begin{aligned}\frac{\partial G(i,w)}{\partial i} &= (1+w)^{r-x}\frac{\partial a_{\overline{r-x}|j_2}}{\partial j_2}\frac{\partial j_2}{\partial i} \\ &= (1+w)^{r-x}\frac{-1}{(1+j_2)}(Ia)_{\overline{r-x}|j_2}\left(\frac{-(1+w)}{(1+i)^2}\right) \\ &= (1+w)^{r-x}\frac{1}{1+i}(Ia)_{\overline{r-x}|j_2}\end{aligned} \quad (24)$$

which is also nonnegative.

Figure 3
Replacement Ratio $\mathrm{RR}_x(65)$ and Interest Rate i



In summary, combining equations (22) through (24) and the result from the three cases examined above, we have both

$$\frac{\partial\, \mathrm{ps}_{x,r}}{\partial i} \geq 0 \quad \text{and} \quad \frac{\partial\, \mathrm{fs}_{x,r}}{\partial i} \geq 0$$

which implies that

$$\frac{\partial \mathrm{RR}_x\,(r)}{\partial i} \geq 0$$

under all circumstances.

Figure 3 shows the relationship of the interest rate on savings and the replacement ratio. Figure 3 is based on the following information $x = 25$, $r = 65$, $w = 5$ percent, AS(25) = 20,000, $s = 20$ percent, and a life annuity factor of $\ddot{a}_x = 11.541$ based on the A1967-70 Mortality Table (ultimate). Savings to-date vary between −200 percent and +200 percent of current annual salary. Notice that there is a sharper increase for larger interest rates. This means that one way to increase income replacement ratio is to increase the interest earnings, assuming all other factors are fixed.

Due to the scaling of the graph, there appears to be little difference in the replacement ratio for differing amounts of savings, but such is not actually the case. For example at $i = 4$ percent, with savings of twice the current salary, the replacement ratio will be about 76 percent. With savings of minus twice the current salary (the individual is in debt), the replacement ratio is 42 percent and is almost halved.

3.3 The Effect of w on $\mathrm{RR}_x(r)$

From equation (9), we have

$$\frac{\partial \,\mathrm{RR}_x(r)}{\partial w} = \frac{\partial \,\mathrm{ps}_{x,r}}{\partial w} + \frac{\partial \,\mathrm{fs}_{x,r}}{\partial w}. \tag{25}$$

It is clear that

$$\frac{\partial \,\mathrm{ps}_{x,r}}{\partial w} < 0.$$

Turning to the second term, we have

$$\frac{\partial \,\mathrm{fs}_{x,r}}{\partial w} = \left(\frac{s}{k\ddot{a}_r}\right) \frac{\partial}{\partial w}\left[\frac{G(i,w)}{(1+w)^{r-x}}\right].$$

Consider the three cases: $i > w$, $i = w$, and $i < w$. It turns out that in all three cases,

$$\frac{\partial \,\mathrm{fs}_{x,r}}{\partial w} \leq 0 \quad \text{which implies that} \quad \frac{\partial \,\mathrm{RR}_x(r)}{\partial w} < 0.$$

Case 1: $i > w$. If the interest rate is larger than the rate of wage increase, from equations (5) and (21), we have

$$\frac{\partial \,\mathrm{fs}_{x,r}}{\partial w} = -\frac{s(1+j_1)^{r-x}}{(1+w)k\ddot{a}_r}\left((r-x)\ddot{a}_{\overline{r-x}|j_1} - (Ia)_{\overline{r-x}|j_1}\right). \tag{26}$$

Equation (26) is always negative.

Case 2: $i = w$. If the interest rate and the rate of wage increase are the same, from equation (5), we have

$$\frac{\partial\, \mathrm{fs}_{x,r}}{\partial w} = 0. \tag{27}$$

Case 3: $i < w$. If the interest rate is smaller than the rate of wage increase from equations (5) and (21), we have

$$\frac{\partial\, \mathrm{fs}_{x,r}}{\partial w} = -\frac{s\,(Ia)_{\overline{r-x}|\,j_2}}{(1+w)^k \ddot{a}_r}, \tag{28}$$

which is also always negative. In this case, when the investment earnings rate is lower than the wage increase rate, then a further increase in wages will decrease the replacement ratio. This is true because the replacement ratio is expressed as a percentage of the retiree's final salary just prior to retirement. In this case, accumulation from savings is not increasing fast enough to keep pace with increases in wages, which may be a reflection of inflation.

To summarize, we have $\partial\, \mathrm{RR}_x(r)/\partial w \leq 0$.

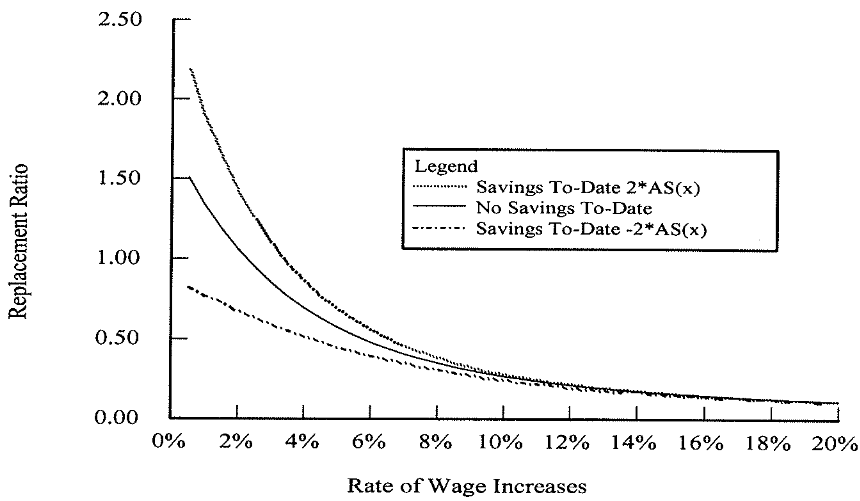
Figure 4 displays the relationship of the replacement ratio and the rate of increase in wages. We observe a declining replacement ratio under all circumstances. The decline in the replacement ratio caused by increasing rates of wages, however, should not be viewed negatively. When wages are particularly high, there is usually a tendency for individuals to increase their savings rate. In this analysis, we have assumed that the savings rate is constant regardless of the level of income. In reality, this may not be the case. We therefore caution the reader to carefully interpret these results.

The opposite effect that i and w have on the replacement ratio can be expected because of their impact on the numerator and denominator in the replacement ratio formula. The net impact of i and w on the replacement ratio depends on the change in the value of the composite rate of j_1 when $i > w$ or j_2 when $i < w$.

3.4 The Effect of x on $\mathrm{RR}_x(r)$

The current age can be viewed as the starting age for which the individual consciously saves for retirement. It is expected that delaying savings for retirement will lower the replacement ratio. We consider the following three cases.

Figure 4
Replacement Ratio $\mathrm{RR}_x(65)$ and Wage Increase w



Case 1: $i > w$. If interest rate is larger than the rate of wage increase, the replacement ratio in equation (9) becomes that in equation (12). We express it as f/g with

$$f = c_x (1 + j_1)^{r-x} + s\ddot{s}_{\overline{r-x}|j_1} \quad \text{and} \quad g = {}^k\ddot{a}_r.$$

Therefore, we have

$$\frac{\partial}{\partial x}\mathrm{RR}_x(r) = \frac{1}{g}\frac{\partial f}{\partial x} = -\frac{\delta_1}{g}(1 + j_1)^{r-x}\left(c_x + \frac{s}{d_1}\right). \quad (29)$$

Using this result, we have $\partial \mathrm{RR}_x(r)/\partial x < 0$. When investments are earning at a larger rate than the rate of increase of wages, a delay in savings for retirement will lower the replacement ratio. This is intuitively true—a delay in savings means forfeiting the opportunity to earn more through investing. A delay also means a lower amount of savings because of the shorter period to save.

Case 2: $i = w$. If the interest rate and the rate of wage increase are the same, the replacement ratio in equation (9) becomes equation (15) which can be expressed as f/g with

$$f = c_x + (r - x)s \quad \text{and} \quad g = {}^k\ddot{a}_r.$$

Consider the first derivative of equation (15):

$$\frac{\partial}{\partial x} \mathrm{RR}_x(r) = -\frac{s}{g} < 0. \tag{30}$$

While in both cases 1 and 2, the replacement ratio decreases with increasing age x, the rate of decrease is greater when $i > w$ than when $i = w$. This follows directly from

$$(1 + j_1)^{r-x} \delta_1 \left(c_x + \frac{s}{d_1}\right) > s.$$

Case 3: $i < w$. If interest rate is smaller than the rate of wage increase, the replacement ratio in equation (9) becomes (16) which can be expressed as f/g with

$$f = c_x (1 + j_1)^{r-x} + s(1 + j_1)^{r-x} s_{\overline{r-x}|j_2} \quad \text{and} \quad g = {}^k\ddot{a}_r.$$

Thus, we have

$$\frac{\partial}{\partial x} \mathrm{RR}_x(r) = \delta_2 (1 + j_2)^{x-r} \left(c_x - \frac{s}{j_2}\right). \tag{31}$$

It is also obvious that $\partial \mathrm{RR}_x(r)/\partial x < 0$ under the condition that $s > c_x j_2$.

We now state this as our next proposition.

Proposition 2. *$\partial RR_x(r)/\partial x < 0$ if (i) $i \geq w$; or (ii) $i < w$ and $s > c_x j_2$.*

The conditions in Propositions 1 and 2 are identical, but the signs of their respective partial derivatives are opposite. This makes sense because x and r would have the opposite effect on the savings accumulation period prior to retirement. Nevertheless, it should be noted that the impact of r on RR_x is much greater than the impact x has on RR_x. In addition to the impact on the savings accumulation period prior to retirement, r also has an effect on the cost in providing the retirement benefit.

4 Extensions to the Replacement Ratio

We now offer some suggestions on ways of extending the concept of replacement ratios introduced in Section 2. These extensions include considering (i) the impact of inflation during retirement, (ii) payments other than annual payments, and (iii) the incorporation of other possible sources of retirement income.

4.1 The Inflation-Adjusted Replacement Ratio, $\text{IARR}_x\,(r+z)$

Recall the central idea behind the replacement ratio: replace a fraction of the salary at retirement age r with a level annuity income. Suppose, however, after age r, the retiree could have continued working and could have expected annual salary increases of $100w'\%$. It seems reasonable to calculate the replacement ratio at age $r+z$, that is $z \geq 0$ years after retirement age r, based on the projected retirement benefit at age $r+z$ and the hypothetical post-retirement salary at age $r+z$.

Another factor to consider is the impact of inflation on both benefits and salaries. To ensure that there is no deterioration in the standard of living after the retirement age, it is important to have a benefit that increases annually to compensate for the corrosive effects of inflation and hence provide a measure of financial stability throughout the retirement period.

In defining the inflation-adjusted replacement ratio at age $r+z$ there are two factors to consider: (i) the hypothetical projected wage increases (that would have occurred had the retiree kept on working), and (ii) the expected inflation during the post retirement years. Specifically, we use the ratio of projected benefits to projected salary where both quantities are expressed in constant dollars to define the new replacement ratio.

Let $B_x\,(r+z)$ denote the annual retirement annuity income to be received at age $r+z$. Given a projected total savings of $\text{TS}_x\,(r)$, the benefits satisfy

$$\text{TS}_x\,(r) = \sum_{t=0}^{\infty} B_x\,(r+t)\,v^t\,{}_tp_r \tag{32}$$

where the right side of the equation gives the expected present value of the annuity benefits to be received during retirement.

Assuming a constant inflation rate of ζ per annum throughout the retirement period, the benefit at age $r+z$ expressed in inflation-adjusted (i.e., constant) dollars is $B_x\,(r+z)\,(1+\zeta)^{-z}$, while the projected hypothetical salary expressed in inflation-adjusted dollars is $\text{ES}_x\,(r)\,(1+w^{'})^{z}(1+\zeta)^{-z}$.

The inflation-adjusted replacement ratio at age $r+z$, $\text{IARR}_x\,(r+z)$, is thus defined as:

$$\begin{aligned}\text{IARR}_x\,(r+z) &= \frac{B_x\,(r+z)\,(1+\zeta)^{-z}}{\text{ES}_x\,(r)\,(1+w')^{z}(1+\zeta)^{-z}}\\ &= \frac{B_x\,(r+z)}{\text{ES}_x\,(r)\,(1+w')^{z}},\end{aligned} \tag{33}$$

which appears to be independent of inflation, ζ, even though it implicitly does depend on inflation.

To reduce the effects of inflation on the retiree's standard of living during retirement, the retiree may purchase an annuity with a benefit that increases annually at a constant rate $b^{'}$.[3] Thus

$$B_x\,(r+z) = B_x\,(r)\left(1+b^{'}\right)^{z}, \tag{34}$$

where

$$\begin{aligned}\text{TS}_x\,(r) &= \sum_{t=0}^{\infty} B_x\,(r)\,(1+b')^{t}\,v^{t}\,{}_tp_r\\ &= B_x\,(r)\sum_{t=0}^{\infty}\left(\frac{1+b'}{1+k}\right)^{t}{}_tp_r = B_x\,(r)\;{}^{k'}\ddot{a}_r\end{aligned} \tag{35}$$

and

$$1+k^{'} = \frac{1+k}{1+b^{'}}.$$

Thus, we can solve for

$$B_x(r) = \frac{\text{TS}_x\,(r)}{{}^{k'}\ddot{a}_r}.$$

It follows that

[3]Another option is to purchase a variable annuity that is linked to inflation. See, for example Black and Skipper (1994, pp. 159–161).

$$\begin{aligned}\mathrm{IARR}_x(r+z) &= \frac{B_x(r)}{\mathrm{ES}_x(r)}\left(\frac{1+b'}{1+w'}\right)^z \\ &= \mathrm{RR}_x(r)\left(\frac{1+b'}{1+w'}\right)^z \qquad (36)\\ &= \frac{\mathrm{TS}_x(r)}{\mathrm{ES}_x(r)\,{}^{k'}\ddot{a}_r}\left(\frac{1+b'}{1+w'}\right)^z. \qquad (37)\end{aligned}$$

To consider the effects of inflation on IARR_x, we split interest into two components: inflation and real interest. In addition, we split wage increases and benefit increases into two components: increases due to inflation and real increases. Thus,

$$k' = \zeta + k^* \qquad (38)$$

$$b' = \zeta + b^* \qquad (39)$$

and

$$w' = \zeta + w^*. \qquad (40)$$

where k^* is the real rate of interest, b^* the real rate of benefit increases and w^* is the real rate of wage increases. (Note that $-\infty < k^*, b^*, w^* < \infty$.) The IARR_x is now given by

$$\mathrm{IARR}_x(r+z) = \frac{\mathrm{TS}_x(r)}{\mathrm{ES}_x(r)\,{}^{k'}\ddot{a}_r}\left(\frac{1+\zeta+b^*}{1+\zeta+w^*}\right)^z \qquad (41)$$

with

$$1+k' = \frac{1+\zeta+k^*}{1+\zeta+b^*}.$$

Consider the rate of change of the replacement ratio with respect to ζ, and applying the chain rule of differentiation, we get

$$\begin{aligned}\frac{\partial}{\partial\zeta}\mathrm{IARR}_x(r+z) = -\frac{\mathrm{IARR}_x(r+z)}{(1+\zeta+b^*)}\Bigg[& z\frac{(b^*-w^*)}{(1+\zeta+w^*)} \\ &+\frac{(k^*-b^*)\,{}^{k'}(Ia)_r}{(1+\zeta+k^*)\,{}^{k'}\ddot{a}_r}\Bigg] \qquad (42)\end{aligned}$$

where ${}^{k'}(Ia)_r$ is the expected present value of an increasing life annuity immediate issued to a person age r. We have used the result

$$\frac{\partial k'}{\partial \zeta} = \frac{-(k^* - b^*)}{(1 + \zeta + b^*)^2}$$

$$\frac{\partial}{\partial \zeta}\left(\frac{1 + \zeta + b^*}{1 + \zeta + w^*}\right)^z = -z\frac{(b^* - w^*)}{(1 + \zeta + w^*)^2}\left(\frac{1 + \zeta + b^*}{1 + \zeta + w^*}\right)^{z-1}$$

and from Bowers, et al., (1997, Chapter 5), we have

$$\frac{\partial}{\partial k'}\left({}^{k'}\ddot{a}_r\right) = -\frac{1}{(1 + k')}{}^{k'}(Ia)_r .$$

The sign of the partial derivative is not clear as it depends on the sign of $(b^* - w^*)$ and $(k^* - b^*)$. For example, for U.S. hourly paid workers, the average annual percentage change in hourly earnings in constant dollars (w^*) ranged from -4.5 percent in 1980 to 1.5 percent in 1997.[4] On the other hand, as k^* is set by the insurer, it would reflect the insurer's conservative estimate of its expected real rate of return on its investments. In 1990, U.S. insurers earned on average 8.89 percent on its assets, which decreased to 7.17 percent in 1996. During the same period the annual rate of inflation dropped from 5.4 percent to 2.3 percent, so their real rate of return ranged between 3 and 6 percent.[5] The retiree sets b^*.

4.2 Adjustments for Monthly Factors

The definition of the replacement ratio in equation (8) assumes that

(1) The savings arising from salary are made at once at the beginning of each year, and

(2) The life annuity payable at retirement commences immediately and is paid also once a year.

In reality, employees are paid more frequently than once a year and the retired individuals want to receive income more often than once a

[4]Source: *Statistical Abstract of the Unites States 1998.* U.S. Department of Commerce, page 434, Table No. 693.

[5]Source: *Statistical Abstract of the Unites States 1998.* U.S. Department of Commerce, page 489, Table No. 772 (for inflation) and page 536, Table No. 854 (for insurers' returns).

year. It is easy to extend the replacement ratio definition to reflect such frequent payments. For the numerator, we adjust FS_x by multiplying it by $a_{\overline{1}|}^{(m)}$ assuming that salaries are paid m times during the year and the first such salary in the year commences at the end of the first $1/m$th of the year. Similarly, for the denominator, the life annuity factor ${}^k\ddot{a}_r$ will be replaced by ${}^k\ddot{a}_r^{(m)}$, where the annuity income is payable m times during the year and the first such income commences at the start of the first $1/m$th of the year.

Thus, from equations (8) and (9), the replacement ratio definition becomes

$$\mathrm{RR}_x^{(m)}(r) = \frac{\mathrm{PS}_x(1+i)^{r-x} + \mathrm{FS}_x a_{\overline{1}|}^{(m)}}{\mathrm{AS}(x)(1+w)^{r-x}\,{}^k\ddot{a}_r^{(m)}}. \tag{43}$$

In the case of monthly contributions and benefits, $m = 12$, and the m during the saving stage (numerator of equation (43)) is not necessarily the same as the m during the payout stage (denominator of equation (43)). The sensitivity analysis in Section 3 can therefore be repeated using equation (43) for the replacement ratio.

4.3 Accounting for Other Sources of Retirement Income

In the previous development of replacement ratios, it has been assumed that the sole source of retirement income is personal savings. This is unrealistic. In several countries, primarily in well-developed ones, other sources of retirement income include social security benefits and benefits from employer-sponsored pension plans. For either the public or private pension plan, employers generally make a contribution and "a significant part of labor's compensation consists of pension benefits" (Hsiao 1984).

Social security programs of several nations share common characteristics. Coverage is usually compulsory, i.e., every working individual is required to make a contribution, together with a portion of the employer's share, toward providing pension benefits. Although benefits are generally related to earnings, some programs pay benefits so that there is a certain level of standard of living attained by participants. For our purposes, we will assume that pension benefits derived from social security are fixed and pre-determined at retirement. We shall denote the annual social security benefit payable at age $r+z$ by SS_{r+z}.

With respect to employer-sponsored pension plans, there are two main types of plans: defined benefit plans and defined contribution plans. To simplify introducing these other sources of income into the replacement ratio formula, we shall denote the defined benefit payable at age $r + z$ as BB_{r+z}.

For the defined contribution pension plan, the contribution of the total employer and employee contribution at age $x + t$ is BC_{x+t} where

$$\mathrm{BC}_{x+t} = c\,\mathrm{AS}\,(x + t)$$

and c is the rate of total contributions as a percentage of salary. This amount will be assumed to earn interest at the annual rate of $i^{'}$. Thus, the amount of benefit derived from this defined contribution plan is expected to accumulate to

$$\begin{aligned}\sum_{t=0}^{r-x-1} BC_{x+t}\left(1+i^{'}\right)^{r-(x+t)} &= \sum_{t=0}^{r-x-1} c\,\mathrm{AS}\,(x+t)\left(1+i^{'}\right)^{r-x-t} \\ &= c\,\mathrm{AS}\,(x)\,G\left(i^{'},w\right) \qquad (44)\end{aligned}$$

at retirement age r.

In summary, the equivalent amount of annual retirement benefit at r can then be expressed as:

$$B_x(r) = \frac{\mathrm{TS}_x(r)}{{}^k\ddot{a}_r} + (SS_r + \mathrm{BB}_r) + \frac{c\,\mathrm{AS}\,(x)\,G\left(i^{'},w\right)}{{}^k\ddot{a}_r}.$$

5 The Case of Singapore

In Singapore, there are two unique elements that must be incorporated in the calculation of replacement ratios. First, employees and employers are required to make periodic contributes to the Singapore's Central Provident Fund (CPF), which is Singapore's major social security program to provide income at retirement. These contributions then accumulate with interest until the employee's retirement. The accumulated contributions are then used to purchase a retirement annuity. It is therefore similar to a defined-contribution plan. The CPF provides a major source of income during retirement.

The contributions are directly tied to earnings and the contribution rates are reviewed periodically and are generally linked to the country's economic performance. In good economic times, the contribution rates are higher and vice versa. For our purposes, CPF contribution rates can be viewed as forced savings, thus helping the individual boost his or her replacement ratio. For an in-depth discussion of the CPF program in Singapore, see Chen and Wong (1998).

The second unique element that must be considered when planning for retirement in Singapore is home ownership. Singapore is believed to have one of the highest percentage of home ownership in the world. According to the *1998 Singapore Yearbook of Statistics*, more than 90 percent of Singaporeans own a home in Singapore. Most homes in Singapore are generally referred to as HDB flats; the term HDB refers to the Housing Development Board, a government agency that develops and manages housing for Singaporeans. Because housing is generally considered an essential commodity in Singapore, the premium paid for owning one will decrease the amount that can be saved from personal income for retirement purposes. This is of major concern because the price of a house in Singapore is expensive relative to income. For example, a four-room HDB flat approximately costs $140,000, a five-room HDB flat can be double that amount, and landed properties can range in prices exceeding $1 million.[6]

Suppose that at age x, an individual purchases a house and borrows an amount of HL_x for a period of $r - x$ years at the housing loan rate of h. Assuming level annual amortization repayments of P_x, then

$$P_x = \frac{\mathrm{HL}_x}{\ddot{a}_{\overline{r-x}|h}}. \tag{45}$$

For simplicity, equation (45) assumes the first loan repayment is made at the time of the origination of the loan. Each year's salary therefore reduces the disposable income by an amount of P_x. Furthermore, assume that the rate of contribution to the CPF will be constant each year at c and that investments on CPF will earn an effective rate of i'. In Singapore, CPF members are allowed to repay their housing loans out of their accounts. Thus, accumulated future savings, FS'_x, become

$$\begin{aligned} \mathrm{FS}'_x = & \sum_{t=0}^{r-x-1} s\,\mathrm{ES}(x+t)\,(1+i)^{r-x-t} \\ & + \sum_{t=0}^{r-x-1} [c\,\mathrm{ES}(x+t) - P_x]\left(1+i'\right)^{r-x-t}. \end{aligned} \tag{46}$$

[6]As of October 2000, the average annual income in Singapore was approximately $36,000 (Singapore dollars), with US$1 = $1.75 (Singapore dollars).

Equation (46) can be simplified to:

$$
\begin{aligned}
\mathrm{FS}'_x &= s\,\mathrm{AS}(x) \sum_{t=0}^{r-x-1} (1+i)^{r-x} \left(\frac{1+w}{1+i}\right)^t \\
&\quad + c\,\mathrm{AS}(x) \sum_{t=0}^{r-x-1} \left(1+i'\right)^{r-x} \left(\frac{1+w}{1+i'}\right)^t \\
&\quad - P_x \sum_{t=0}^{r-x-1} \left(1+i'\right)^{r-x-t} \\
&= s\,\mathrm{AS}(x)\,G(i,w) + c\,\mathrm{AS}(x)\,G\left(i',w\right) - P_x \ddot{s}_{\overline{r-x}|\,i'} \\
&= \mathrm{AS}(x)\left[sG(i,w) + cG\left(i',w\right)\right] \\
&\quad - \mathrm{HL}_x \left(1+i'\right)^{r-x} \frac{\ddot{a}_{\overline{r-x}|\,i'}}{\ddot{a}_{\overline{r-x}|\,h}}.
\end{aligned} \tag{47}
$$

Again, similar to Section 2, we suppose that at age x the individual has saved a total of PS_x. At retirement, the individual can expect to have saved a total of

$$
\mathrm{TS}_x(r) = \mathrm{PS}_x (1+i)^{r-x} + \mathrm{FS}'_x. \tag{48}
$$

This amount will then be used to purchase a life annuity at retirement for which the level annual payments will be as given in equation (7). Therefore, the replacement ratio becomes

$$
\begin{aligned}
\mathrm{RR}_x(r) &= \frac{\mathrm{PS}_x (1+i)^{r-x}}{\mathrm{AS}(x)(1+w)^{r-x}\, k\ddot{a}_r} \\
&\quad + \frac{\mathrm{AS}(x)\left[sG(i,w) + c\,G\left(i',w\right)\right]}{\mathrm{AS}(x)(1+w)^{r-x}\, k\ddot{a}_r} \\
&\quad - \frac{\mathrm{HL}_x \left(1+i'\right)^{r-x} \ddot{a}_{\overline{r-x}|\,i'}}{\mathrm{AS}(x)(1+w)^{r-x}\, k\ddot{a}_r \ddot{a}_{\overline{r-x}|\,h}}.
\end{aligned} \tag{49}
$$

Using symbols earlier defined, we can express equation (49) as

Table 1
Effect of c on Replacement Ratios
In Singapore, with $h_x = 10$

	Retirement Age					
c	50	55	60	65	70	75
10%	-35.1%	-34.9%	-34.0%	-31.8%	-27.0%	- 17.5%
20%	-19.2%	-12.8%	-3.4%	10.7%	32.5%	67.0%
30%	-3.3%	9.4%	27.2%	53.2%	92.1%	151.4%
40%	12.6%	31.5%	57.9%	95.7%	151.7%	235.8%
50%	28.4%	53.6%	88.5%	138.3%	211.2%	320.3%

Table 2
Effect of h_x on Replacement Ratios
In Singapore, with $c = 30\%$

	Retirement Age					
h_x	50	55	60	65	70	75
0	74.8%	103.5%	143.1%	198.9%	279.6%	398.4%
5	35.7%	56.5%	84.2%	126.1%	185.9%	274.9%
10	-3.3%	9.4%	27.2%	53.2%	92.1%	151.4%
15	-42.4%	-37.7%	-30.7%	-19.6%	-1.7%	27.9%
20	-81.4%	-84.8%	-88.6%	-92.5%	-95.5%	-95.6%

$$\mathrm{RR}_x(r) = \frac{(1+j_1)^{r-x}}{{}^{k}\ddot{a}_r}\left[c_x + \frac{sG(i,w) + cG\left(i',w\right)}{(1+j_1)^{r-x}(1+w)^{r-x}} - \frac{\left(1+j_1'\right)^{r-x}\ddot{a}_{\overline{r-x}|i'}\,h_x}{(1+j_1)^{r-x}\,\ddot{a}_{\overline{r-x}|h}}\right] \tag{50}$$

where

$$h_x = \frac{HL_x}{\mathrm{AS}(x)}$$

is the housing loan expressed as a multiple of salary at age x.

For purposes of illustration, consider the case where we have a Singapore individual who is currently age 25 and is receiving an annual salary of \$20,000. Assume he or she has saved \$20,000 and that his or her salary is expected to increase at the rate of 3 percent per annum. Apart from contributions to the CPF, he or she expects to save annually 10 percent for retirement. Interest rates will be assumed at $i = 5$ percent, $i' = 4$ percent, and $h = 5$ percent.

Tables 1 and 2 display the result of varying the CPF contribution rates and the ratio of housing loans to salary, respectively, on the value of the replacement ratio at different retirement ages. Table 1 shows the effect of varying the contribution rates to the CPF, and Table 2 shows the effect of varying the housing loan ratio. To further illustrate the results, at a CPF contribution rate of say 30 percent, this same individual can retire at age 65 with a replacement ratio of 53.2 percent. Similarly, if he or she borrows money for housing at the amount five times current salary and CPF contribution rate stays at 30 percent, he or she can retire at age 65 with replacement ratio of a whopping 126.1 percent. The results in Tables 1 and 2 generally demonstrate that to improve replacement ratios, one can either:

- Decide to delay retirement;
- Increase contribution rates to the CPF scheme; or
- Control the budget for housing.

6 Concluding Remarks and Future Research

This paper proposes a financial measure, the replacement ratio, that can be used for individuals in planning for retirement. The replacement ratio, which is expressed as the proportion of the retirement income to that of the final salary at retirement, is an intuitively appealing financial construct. The paper does not recommend a suitable level of the replacement ratio: it is left to each person to determine the level that is appropriate.

The size of a person's retirement income depends on several factors such as wage increases, savings rate, interest rates, and inflation both prior to and during retirement. The replacement ratio developed in this paper readily allows us to examine the effect of changes of these factors to the replacement ratio. In Section 3, we examined the sensitivity of the replacement ratio with respect to changes in the retirement age r, the investment earnings rate i, the rate of wage increases w during employment periods, and the current age x (the age at which the individual consciously begins to set aside funds specifically for retirement purposes).

We found that under certain conditions, delaying the retirement age and increasing the return on investments have the effect of increasing the replacement ratio. On the other hand, delaying the age at which savings for retirement commences has the effect of reducing the replacement ratio. For wage increases, the replacement ratio will either increase or decrease depending upon the relationship of the rate of i and w. We caution the reader when interpreting the results when wage increases are concerned. Large increases in wages can cause the replacement ratio to be smaller—this should not be interpreted to mean that a lower replacement ratio leads to a deterioration of retirement income. When the base salary is large and yet savings have not been accumulating, the replacement ratio may be relatively low but the amount of retirement income can still be large.

Our work here is only a beginning to better understanding the various factors that can affect income at retirement. The discussion has been simplified to facilitate understanding. For example, we have assumed constant interest rates and constant rate of wage increases. Furthermore, we have assumed that interest rates, savings rates, and the rates of wage increases are all independent. In reality, this is not true. One possible future research is to examine empirical data that may support evidence of dependence and that may account for the time series nature of these variables. Some recent papers by Knox (1993) and Booth and Yakoubov (2000) that suggest the use of a stochastic approach may be helpful in these instances. It will be interesting to explore the link of these stochastic approaches in further developing the replacement ratio model recommended in this paper. Another possible future research area is the impact of mortality improvement, a phenomenon that is observed worldwide.

References

Black, J., Jr. and Skipper, H.D., Jr. *Life Insurance* Twelfth Edition. Englewood Cliffs, N.J.: Prentice Hall, 1994.

Booth, P. and Yakoubov, Y. "Investment Policy for Defined-Contribution Pension Schemes Members Close to Retirement: An Analysis of the Lifestyle Concept." *North American Actuarial Journal* 4 (2000): 1–19.

Borch, K. "Pension Plans and the Theory of Consumption and Saving." *22nd International Congress of Actuaries* 2 (1984): 111–116.

Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A., and Nesbitt, C.J. *Actuarial Mathematics*, 2nd edition. Schaumburg, Ill.: Society of Actuaries, 1997.

Chen, R. and Wong Kie Ann. "The Adequacy of the CPF Account for Retirement Benefits in Singapore." *Singapore International Insurance and Actuarial Journal* 2 (1998): 121–138.

Cordell, D.M. (ed.). *Fundamentals of Financial Planning.* Bryn Mawr, Pa.: The American College, 1992.

Gerber, H.U. *Life Insurance Mathematics*, 3rd edition. New York, N.Y.: Springer-Verlag, 1997.

Hsiao, W.C. "Impact of Inflation on Pension Benefits." *22nd International Congress of Actuaries* 2 (1984): 467–477.

Jordan, C.W. *Life Contingencies*, 2nd edition. Chicago, Ill.: Society of Actuaries, 1967.

Kellison, S.G. *The Theory of Interest*, 3rd edition. New York, N.Y.: Irwin/McGraw-Hill, 1991.

Knox, D.M. "A Critique of Defined Contribution Plans Using a Simulation Approach." *Journal of Actuarial Practice* 1 (1993): 49–66.

McGill, D.M. and Grubbs, D.S., Jr. *Fundamentals of Private Pensions.* Homewood, Ill.: Richard D. Irwin, Inc., 1984.

Rejda, G.E. *Social Insurance and Economic Security.* Englewood Cliffs, N.J.: Prentice Hall, 1988.

Life Contingencies with Stochastic Discounting Using Moving Average Models

Steven Haberman,* Russell Gerrard,† and Dimitrios Velmachos‡

Abstract§

This paper offers simplified procedures for calculating moments of functions in life contingencies when the random force of interest is modeled using an unconditional moving average process of order q, MA(q). It extends the MA(1) model that has been used for stochastic discounting. Using the more general MA(q) model allows actuaries to better capture the autocorrelation between successive interest rates in a time series.

Key words and phrases: *stochastic interest, present value, time series, normal distribution, net single premium*

*Steven Haberman, Ph.D., D.Sc., F.I.A., F.S.S., is a professor of actuarial science and Dean of the School of Mathematics at City University. He received his first degree in mathematics from Cambridge University. He joined the Prudential Assurance Company as an actuarial trainee and then the City University as a lecturer. He qualified as an F.I.A. in 1975. He has published widely in actuarial and related fields. His current research interests include mortality, morbidity, premium rating, and pension funding.

Dr. Haberman's address is: School of Mathematics, Actuarial Science and Statistics, City University, London EC1V 0HB, UNITED KINGDOM. Internet address: *S.Haberman@city.ac.uk*

†Russell Gerrard, Ph.D., F.S.S., is a senior lecturer in statistics at City University. He received his first degree in mathematics and Ph.D. in statistics from Cambridge University. Prior to joining City University, Dr. Gerrard worked at universities in Moscow, Zurich and Sussex. His research interests include stochastic modeling and applied probability.

Dr. Gerrard's address is: Department of Actuarial Science and Statistics, City University, London EC1V 0HB, UNITED KINGDOM. Internet address: *R.J.Gerrard@city.ac.uk*

‡Dimitrios Velmachos, M.Sc., is mathematician and former graduate student at City University.

Mr. Velmachos's address is: 83A Christchurch Rd., London SW2 3DH, UNITED KINGDOM. Internet address: *dvelmachos@hotmail.com*

§The authors would like to acknowledge the anonymous referees' helpful comments made on an earlier draft.

1 Introduction

1.1 Background

The theory of life contingencies comes from deterministic beginnings. Random fluctuations in risk factors such as mortality, morbidity, interest, and expenses historically have been ignored. Instead, actuaries traditionally have attempted to allow for random fluctuations by using conservative assumptions for each factor. For example, in the calculation of the present values of the liabilities for a policy, actuaries can assume that mortality follows a known mortality table or that the variability due to mortality (i.e., variability in future lifetimes) can be ignored because of the presence of a large number of identical liabilities in respect to different lives. Similarly, the interest rate may be assumed to be constant, or an implicit allowance may be made by adopting a conservative estimate of future interest rates (Bowers et al., 1997, Chapter 16, or Gerber 1995, Chapter 5).

A next step in the development of life contingencies was the *semi-stochastic approach*, which was to consider the time until decrement (death, disability, and so on) as a random variable in the calculation of the actuarial present value of actuarial functions, while the interest rate was assumed to be constant.

Actuaries were criticized for failing to account for the variability in interest rates in their financial calculations. Only since 1970 has there been interest in actuarial models that consider both the time until decrement and the investment rate of return as random variables.

Pollard (1971) and Boyle (1976) consider interest rate fluctuations by treating the force of interest as a random variable. Boyle (1976) examines the case in which the force of interest in any year is a normally distributed random variable that is independent of the force of interest in any other year. This simple assumption is explored further in Section 2.

Pollard (1971), on the other hand, models the force of interest using a stationary autoregressive process of order two. Panjer and Bellhouse (1980) and Bellhouse and Panjer (1981) develop a general theory for unconditional and conditional autoregressive models of order one and two of the force of interest.

Giaccotto (1986) has developed an algorithm for evaluating present value functions when interest rates are assumed to follow an ARIMA process. Also Wilkie (1976), Waters (1978), Westcott (1981), de Jong (1984), Dhaene (1989), and Frees (1990) consider stochastic interest models in the calculation of the standard actuarial functions of life insurance mathematics.

There remains a fundamental question: Is the stochastic nature used for the calculation of interest rates correct? Many actuaries remain skeptical about stochastic interest rate models because they believe that the results provided by such models are due to the peculiarities of the specific model rather than to any underlying reality.[1]

In this paper we do not consider the consequences of the choice of an incorrect interest model.

1.2 Objectives

In this paper, we concentrate on certain time series models of the force of interest called moving average processes.

Let i_k represent the random effective rate of interest from time k to time $k+1$, for $k = \ldots, -1, 0, 1, \ldots$. Without loss of generality, this period can be considered to be one year. The force of interest in the kth period, Δ_k is defined as

$$\Delta_k = \ln(1 + i_k).$$

The sequence of Δ_ks is assumed to be a moving average process of order q (MA(q), $q = 0, 1, 2, \ldots$) in the sense of Box and Jenkins (1976), i.e., for $k = \ldots, -1, 0, 1, \ldots$ the Δ_ks are defined as

$$\Delta_k = \delta + a_0\varepsilon_k + a_1\varepsilon_{k-1} + a_2\varepsilon_{k-2} + \ldots a_q\varepsilon_{k-q}. \tag{1}$$

Here δ is the mean about which the Δ_ks fluctuate, and the ε_ks are independent and identically distributed random variables with mean zero and variance σ^2. The coefficients $a_0, a_1, \ldots, a_q$ are usually constrained so that the roots of the characteristic polynomial equation in x

$$a_0 - \sum_{i=1}^{q} a_i x^i = 0$$

[1] The question of the significance of model sensitivity is the subject of current research. For example, Wright (1997) and Chadburn and Wright (2000) investigate the sensitivity of outcomes in pension funding models and life insurance asset-liability models, respectively, to the choice of stochastic asset model.

lie outside the unit circle. These constraints are required so that the model is invertible, i.e., that

$$\Delta_k - \delta = \sum_{j=0}^{\infty} b_j(\Delta_{k-j} - \delta) \tag{2}$$

for some sequence of constants b_j, $j = 0, 1, \ldots$. The invertibility requirement ensures that the right side of equation (2) is convergent (Box and Jenkins, 1976).

Two MA(q) models are considered:

- $q = 0$, in which case the Δ_ks are independent, identically distributed random variables; and
- $q = 1, 2, \ldots$, in which case the Δ_ks are dependent random variables.

These cases are considered in Sections 2 and 3, respectively.

For each MA(q) model, we derive the moments of the k-period discount factor (measured from time 0), v_k, where

$$v_k = \prod_{s=1}^{k} \frac{1}{(1 + i_k)} = e^{-\sum_{s=1}^{k} \Delta_s} \tag{3}$$

for $k = 1, 2, \ldots$ with $v_0 = 1$. In addition the moments of certain insurance and annuity functions are derived.

Throughout this paper we use the notation

$$M(\tau) = E[e^{\varepsilon\tau}],$$

which is the moment-generating function of ε. We assume $M(\tau)$ exists for some $\tau > 0$.

Finally, we give an expression for the coefficients of skewness and kurtosis for a random variable X with mean δ_X and variance σ_X^2:

$$\text{Coefficient of Skewness of } X = \frac{E[(X - \delta_X)^3]}{\sigma_X^3}$$

$$\text{Coefficient of Kurtosis of } X = \frac{E[(X - \delta_X)^4]}{\sigma_X^4}.$$

The coefficient of skewness measures the lack of symmetry in a probability distribution. The coefficient of kurtosis measures the extent to which the peak of a unimodal probability distribution departs from the shape of a normal distribution by being flatter or more pointed.

2 The Independent Interest Rate Model MA(0)

Let us consider the case where the ε_ks are assumed to be a sequence of i.i.d. normal random variables with mean 0 and variance σ^2. [2] This assumption also provides a benchmark assumption against which more complex models can be compared.

2.1 Some Basic Results

As ε_ks are i.i.d. normal random variables with mean 0 and variance σ^2, their moment generating function is

$$M(\tau) = e^{\frac{1}{2}\tau^2\sigma^2}. \tag{4}$$

The moment-generating function of Δ_k is thus

$$M_\Delta(\tau) = e^{\tau\delta}M(\tau). \tag{5}$$

The moments of v_k can easily be found as:

$$\begin{aligned} E[v_k^n] &= E[e^{-n\sum_{j=1}^{k}\Delta_j}] \\ &= (M_\Delta(-n))^k \\ &= e^{-nk\delta}\,(M(-n))^k \\ E[v_k^n] &= e^{-kd_n} \end{aligned} \tag{6}$$

where

$$d_n = n\delta - \frac{1}{2}n^2\sigma^2 \qquad \text{for } n = 1, 2, \ldots. \tag{7}$$

Similarly, for $r, s, m, n = 1, 2, \ldots$ and $r \geq s + 1$, we find that:

[2]The i.i.d. assumption means that $\ln(v_k)$ is a random walk, which is a desirable feature from a financial economics viewpoint. The random walk is a special case of a martingale, and its structure does not permit riskless arbitrage (Baxter and Rennie (1996)). No riskless arbitrage is a desirable quality in modern finance theory.

$$
\begin{aligned}
E[v_r^m v_s^n] &= E[e^{-m\sum_{j=1}^{r}\Delta_j + -n\sum_{k=1}^{s}\Delta_k}] \\
&= E[e^{\sum_{j=1}^{s} -(m+n)\Delta_j + \sum_{j=s+1}^{r} -m\Delta_j}] \\
&= (M_\Delta(-(m+n)))^s \, (M_\Delta(-m))^{(r-s)} \\
&= e^{-(m+n)s\delta} \, (M(-(m+n)))^s \, e^{-m(r-s)\delta} \, (M(-m))^{(r-s)} \\
&= e^{-(mr+ns)\delta + \frac{1}{2}(s(m+n)^2 + (r-s)m^2)\sigma^2} \\
&= e^{smn\sigma^2} e^{-(rd_m + sd_n)}. \qquad (8)
\end{aligned}
$$

Next, we let $\tilde{a}_{\overline{n}|}$ denote the stochastic equivalent for the traditional immediate annuity-certain $a_{\overline{n}|}$, i.e.,

$$\tilde{a}_{\overline{n}|} = \sum_{k=1}^{n} v_k.$$

It follows that

$$E[\tilde{a}_{\overline{n}|}] = \sum_{k=1}^{n} \theta^k \varphi^k \qquad (9)$$

where

$$\theta = e^{-\delta} \quad and \quad \varphi = e^{\frac{1}{2}\sigma^2}.$$

Waters (1978) provides simple expressions for some of the moments of $\tilde{a}_{\overline{n}|}$ that are useful for calculating the high order moments of certain actuarial functions. They will be used later in some numerical examples. The following results are obtained by Waters (1978):

$$E[\tilde{a}^2_{\overline{n}|}] = \sum_{k=1}^{n} \theta^{2k}\varphi^{4k} + 2\sum_{k=2}^{n}\sum_{j=1}^{k-1} \theta^{(k+j)}\varphi^{(k+3j)} \qquad (10)$$

$$E[\tilde{a}^3_{\overline{n}|}] = \sum_{k=1}^{n} \theta^{3k}\varphi^{9k} + 3\sum_{k=2}^{n}\sum_{j=1}^{k-1}[\theta^{(2k+j)}\varphi^{(4k+5j)} + \theta^{(k+2j)}\varphi^{(k+8j)}] + 6\sum_{k=3}^{n}\sum_{j=2}^{k-1}\sum_{i=1}^{j-1}\theta^{(i+j+k)}\varphi^{(5i+3j+k)} \tag{11}$$

$$\begin{aligned} E[\tilde{a}^4_{\overline{n}|}] = & \sum_{k=1}^{n} \theta^{4k}\varphi^{16k} \\ & + 4\sum_{k=2}^{n}\sum_{j=1}^{k-1}(\theta^{(k+3j)}\varphi^{(k+15j)} + \theta^{(3k+j)}\varphi^{(9k+7j)}) \\ & + 6\sum_{k=2}^{n}\sum_{j=1}^{k-1}\theta^{(2k+2j)}\varphi^{(4k+12j)} \\ & + 12\sum_{k=3}^{n}\sum_{j=2}^{k-1}\sum_{i=1}^{j-1}[\theta^{(i+j+2k)}\varphi^{(7i+5j+4k)} \\ & + \theta^{(i+2j+k)}\varphi^{(7i+8j+k)} + \theta^{(2i+j+k)}\varphi^{(12i+3j+k)}] \\ & + 24\sum_{k=4}^{n}\sum_{j=3}^{k-1}\sum_{i=2}^{j-1}\sum_{h=1}^{i-1}\theta^{(h+i+j+k)}\varphi^{(7h+5i+3j+k)} \end{aligned} \tag{12}$$

and

$$E[\tilde{a}_{\overline{n}|}v_m] = \begin{cases} \sum_{k=1}^{n}\theta^{(m+k)}\varphi^{(k+3m)} & \text{if } m \geq n \\ \sum_{k=1}^{m}\theta^{(m+k)}\varphi^{(m+3k)} & \\ \quad + \sum_{k=1}^{n}\theta^{(m+k)}\varphi^{(k+3m)} & \text{if } m < n. \end{cases} \tag{13}$$

2.2 Moments of Life Insurance Functions

This section is based on the approach of Frees (1990). We assume that there is only one decrement (mortality) and that mortality is independent of the sequence Δ_k. Let $T \equiv T(x)$ denote the random future time until death of a person age x and $K \equiv K(x)$ denote the largest integer less than or equal to T, i.e., $K \leq T < K+1$. The standard actuarial notation is used where possible.

Two types of general contracts are considered: insurance and annuity contracts. For $k = 0, 1, \ldots,$ a general insurance contract is considered with death benefit $b_{k+1} \geq 0$ is taken to be payable at the end of the policy year of death given that death occurs during the $k+1$ policy year, i.e., in year $(k, k+1)$. The random present value of this insurance benefit is

$$Z = v_{K+1} b_{K+1}. \tag{14}$$

For $k = 0, 1, \ldots,$ the general annuity contract pays c_k $(-\infty < c_k < \infty)$ at the start of the policy year, i.e., at time k, and payments continue for life. The random present value of the annuity benefits is then:

$$Y = \sum_{k=0}^{K} v_k c_k. \tag{15}$$

The nth moment of Z is easily derived as follows:

$$\begin{aligned} E[Z^n] &= E[E[v_{K+1}^n b_{K+1}^n | K]] \\ &= E[e^{-(K+1)d_n} b_{K+1}^n] \\ &= \sum_{k=0}^{\infty} e^{-(k+1)d_n} b_{k+1}^n {}_{k|1}q_x. \end{aligned} \tag{16}$$

This equation, however, is the net single premium (or actuarial present value) for a general whole life policy evaluated using a constant interest i_n^* given by

$$i_n^* = e^{d_n} - 1. \tag{17}$$

This interpretation only makes sense, of course, if $d_n \geq 0$.

For the case of general annuity contracts, the moments are more difficult to determine. The first two moments are:

$$
\begin{aligned}
E[Y] &= E[E[\sum_{s=0}^{K} v_s c_s | K]] \\
&= E[\sum_{s=0}^{K} c_s e^{-sd_1}] \\
&= \sum_{k=0}^{\infty} \sum_{s=0}^{k} c_s e^{-sd_1}\, {}_{k|1}q_x \\
&= \sum_{s=0}^{\infty} c_s e^{-sd_1}\, {}_s p_x
\end{aligned} \tag{18}
$$

and

$$
\begin{aligned}
E[Y^2] &= E[E[(\sum_{s=0}^{K} v_s c_s)^2 | K]] \\
&= E[E[\sum_{s=0}^{K} v_s^2 c_s^2 + 2 \sum_{r=1}^{K} \sum_{s=0}^{r-1} v_s c_s v_r c_r | K]] \\
&= E[\sum_{s=0}^{K} e^{-d_2 s} c_s^2] + 2E[\sum_{r=1}^{K} \sum_{s=0}^{r-1} e^{-d_1 s} e^{-(d_2 - d_1) r} c_r c_s] \\
&= \sum_{k=0}^{\infty} {}_{k|1}q_x \sum_{s=0}^{k} e^{-d_2 s} c_s^2 \\
&\quad + 2 \sum_{k=1}^{\infty} {}_{k|1}q_x \sum_{r=1}^{K} \sum_{s=0}^{r-1} e^{-d_1 s} e^{-(d_2 - d_1) r} c_r c_s .
\end{aligned} \tag{19}
$$

In addition,

$$
\begin{aligned}
E[ZY] &= E[E[b_{K+1} \sum_{s=0}^{K} v_s c_s | K]] \\
&= E[b_{K+1} \sum_{s=0}^{K} e^{-s(d_2 - d_1)} e^{-(K+1) d_1} c_s] \\
&= \sum_{k=0}^{\infty} {}_{k|1}q_x b_{k+1} \sum_{s=0}^{k} e^{-s(d_2 - d_1)} e^{-(k+1) d_1} c_s .
\end{aligned} \tag{20}
$$

For basic insurance and annuity products, the expressions for $E[Z]$ and $E[Y]$ can be simplified. For example, let Z be the present value of the benefit for a standard whole life insurance policy with death benefit of 1 issued to a life age x.

$$E[Z] = \sum_{k=0}^{\infty} {}_{k|1}q_x e^{-(k+1)d_1} = A_x^*,$$

which is the net single premium for a traditional whole life policy evaluated using a constant interest $i^* = e^{d_1} - 1$. This interpretation only makes sense, of course, if $d_1 \geq 0$. A similar result applies to term insurance and endowment insurance policies with face value 1, i.e., use a constant interest i^* to determine their net single premiums. This new force of interest d_1 (defined by equation (7)) consists of the mean force of interest δ less an allowance for the inherent volatility $(\sigma^2/2)$.

Let Y be the present value of a standard annuity due paying 1 per year for life issued to a life age x. Then,

$$E[Y] = \sum_{s=0}^{\infty} e^{-sd_1} {}_sp_x = \ddot{a}_x^*$$

which is the actuarial present value of a traditional whole life annuity evaluated using a constant interest i^*, i.e., constant force of interest d_1. A similar result applies to temporary annuities, i.e., use a constant interest i^*.

2.3 Examples

We will now derive the net single premium (mean), variance, skewness, and kurtosis of five basic insurance products: endowment, term (temporary) insurance, whole life, temporary annuity, and whole life annuity. The calculations are based on the assumption that all lives are subject to the mortality experience of the British A 1967-70 (Ultimate) mortality table, and that $\delta = 0.07$.

Let

$$Z_1 = \begin{cases} v_{K+1} & K = 0, 1, \ldots, n-1 \\ v_n & K = n, n+1, \ldots; \end{cases} \qquad n\text{-Year Endowment}$$

$$Z_2 = \begin{cases} v_{K+1} & K = 0, 1, \ldots, n-1 \\ 0 & K = n, n+1, \ldots; \end{cases} \qquad n\text{-Year Term Insurance}$$

$$Z_3 = v_{K+1}; \qquad \text{Whole Life}$$

$$Y_1 = \begin{cases} \ddot{a}_{\overline{K+1}|} & K = 0, 1, \ldots, n-1 \\ \ddot{a}_{\overline{n}|} & K = n, n+1, \ldots; \end{cases} \qquad n\text{-Year Temporary Annuity}$$

and

$$Y_2 = \ddot{a}_{\overline{K+1}|} \qquad \text{Whole Life Annuity.}$$

Equation (16) is used to determine the moments of Z_1, Z_2, and Z_3. The results of equations (9) to (12) combined with the methodology underlying equations (18) and (19) are used to determine the moments of Y_1 and Y_2. Note that

$$E[Z_1^m] = E[Z_2^m] + {}_m p_x E[v_n^m].$$

Tables 1 to 5 show the mean, standard deviation, skewness, and kurtosis for several values of σ, x, and n (as appropriate), and δ for the respective cases of a standard endowment (Z_1), term (Z_2), and whole life (Z_3) insurance, and a temporary (Y_1) and whole life (Y_2) annuity. Notice that in each example, the standard deviation of the random variable increases as σ increases, but the amount of increase is not significant.

The results in Tables 1 to 5 show that the greater part of the differences between these means is due to fluctuations in the age at death as opposed to fluctuations in the interest rates. This situation changes if we consider a large number of independent lives — then the fluctuations in interest rates would become more important; see, for example, Marceau and Gaillardetz (1999).

Table 1
Endowment Insurance for 20, 30, and 40 Years
Using the I.I.D. Model with δ = 7%

	$\sigma = 0.00$			$\sigma = 0.03$		
Age	20 yrs	30 yrs	40 yrs	20 yrs	30 yrs	40 yrs
			Net Single Premium			
25	0.2507	0.1310	0.0759	0.2530	0.1326	0.0771
30	0.2522	0.1354	0.0846	0.2544	0.1371	0.0858
35	0.2560	0.1444	0.0999	0.2582	0.1461	0.1012
40	0.2632	0.1599	0.1240	0.2654	0.1617	0.1255
45	0.2757	0.1848	0.1591	0.2779	0.1866	0.1607
			Standard Deviation			
25	0.0401	0.0550	0.0652	0.0524	0.0590	0.0669
30	0.0438	0.0625	0.0753	0.0553	0.0663	0.0771
35	0.0547	0.0787	0.0936	0.0644	0.0821	0.0955
40	0.0721	0.1017	0.1173	0.0798	0.1047	0.1194
45	0.0946	0.1289	0.1433	0.1008	0.1317	0.1455
			Coefficient of Skewness			
25	12.17	9.48	7.74	5.61	7.82	7.33
30	10.28	7.41	5.82	5.26	6.37	5.58
35	7.77	5.46	4.32	4.89	4.96	4.19
40	5.70	4.03	3.26	4.29	3.80	3.20
45	4.17	2.98	2.50	3.51	2.87	2.46
			Coefficient of Kurtosis			
25	165.99	106.91	76.52	59.54	82.93	71.18
30	122.68	68.83	46.31	50.67	56.49	43.72
35	71.70	38.79	26.67	39.35	34.28	25.64
40	39.42	21.94	16.08	27.59	20.42	15.65
45	21.87	12.83	10.21	17.85	12.31	10.01

Table 1 (contd.)
Endowment Insurance for 20, 30, and 40 Years
Using the I.I.D. Model with δ = 7%

Age	σ = 0.05			σ = 0.07		
	20 yrs	30 yrs	40 yrs	20 yrs	30 yrs	40 yrs
			Net Single Premium			
25	0.2570	0.1357	0.0793	0.2631	0.1404	0.0826
30	0.2584	0.1402	0.0881	0.2646	0.1449	0.0916
35	0.2622	0.1492	0.1036	0.2683	0.1540	0.1073
40	0.2694	0.1649	0.1281	0.2756	0.1698	0.1321
45	0.2819	0.1899	0.1636	0.2880	0.1949	0.1680
			Standard Deviation			
25	0.0700	0.0663	0.0702	0.0925	0.0774	0.0756
30	0.0724	0.0732	0.0806	0.0943	0.0839	0.0863
35	0.0797	0.0882	0.0991	0.1002	0.0980	0.1051
40	0.0927	0.1102	0.1233	0.1111	0.1191	0.1296
45	0.1115	0.1369	0.1498	0.1275	0.1453	0.1566
			Coefficient of Skewness			
25	2.70	5.82	6.67	1.74	4.14	5.80
30	2.69	5.03	5.19	1.77	3.82	4.67
35	2.86	4.24	3.99	1.92	3.49	3.73
40	2.94	3.44	3.09	2.08	3.03	2.95
45	2.73	2.69	2.40	2.10	2.48	2.32
			Coefficient of Kurtosis			
25	21.16	55.42	62.40	10.08	33.57	51.08
30	19.86	41.05	39.45	9.94	27.42	33.92
35	19.10	27.88	23.94	10.41	21.24	21.72
40	16.92	18.06	14.95	10.50	15.31	14.06
45	13.17	11.46	9.69	9.43	10.42	9.30

Table 2
Term (Temporary) Insurance for 20, 30, and 40 Years
Using the I.I.D. Model with δ = 7%

	$\sigma = 0.00$			$\sigma = 0.03$		
Age	20 yrs	30 yrs	40 yrs	20 yrs	30 yrs	40 yrs
			Net Single Premium			
25	0.0088	0.0158	0.0257	0.0088	0.0159	0.0259
30	0.0131	0.0253	0.0405	0.0132	0.0255	0.0410
35	0.0225	0.0425	0.0645	0.0226	0.0428	0.0651
40	0.0395	0.0705	0.0991	0.0397	0.0711	0.1001
45	0.0675	0.1125	0.1448	0.0679	0.1133	0.1461
			Standard Deviation			
25	0.0688	0.0763	0.0794	0.0693	0.0771	0.0804
30	0.0800	0.0901	0.0922	0.0806	0.0912	0.0936
35	0.1022	0.1129	0.1114	0.1030	0.1143	0.1131
40	0.1332	0.1414	0.1334	0.1344	0.1432	0.1354
45	0.1705	0.1711	0.1549	0.1720	0.1732	0.1572
			Coefficient of Skewness			
25	8.95	6.76	5.72	8.98	6.75	5.68
30	6.97	4.92	4.16	7.00	4.92	4.13
35	5.10	3.56	3.15	5.12	3.56	3.12
40	3.68	2.60	2.50	3.70	2.61	2.48
45	2.63	1.90	2.06	2.65	1.91	2.04
			Coefficient of Kurtosis			
25	91.00	58.89	47.46	91.45	58.58	46.68
30	56.77	33.44	27.65	57.14	33.29	27.18
35	31.21	18.68	17.00	31.47	18.62	16.74
40	17.07	11.08	11.54	17.23	11.05	11.37
45	9.52	7.05	8.39	9.61	7.02	8.27

Table 2 (contd.)
Term (Temporary) Insurance for 20, 30, and 40 Years
Using the I.I.D. Model with δ = 7%

	$\sigma = 0.05$			$\sigma = 0.07$		
Age	20 yrs	30 yrs	40 yrs	20 yrs	30 yrs	40 yrs
	Net Single Premium					
25	0.0089	0.0161	0.0264	0.0090	0.0165	0.0272
30	0.0133	0.0258	0.0418	0.0135	0.0264	0.0430
35	0.0228	0.0434	0.0664	0.0231	0.0444	0.0683
40	0.0400	0.0721	0.1019	0.0406	0.0736	0.1047
45	0.0685	0.1148	0.1485	0.0694	0.1171	0.1522
	Standard Deviation					
25	0.0700	0.0663	0.0702	0.0715	0.0807	0.0853
30	0.0724	0.0732	0.0806	0.0836	0.0962	0.1001
35	0.0797	0.0882	0.0991	0.1071	0.1209	0.1213
40	0.0927	0.1102	0.1233	0.1397	0.1514	0.1451
45	0.1115	0.1369	0.1498	0.1787	0.1831	0.1682
	Coefficient of Skewness					
25	9.03	6.73	5.60	9.12	6.74	5.51
30	7.05	4.93	4.09	7.15	4.96	4.04
35	5.18	3.58	3.09	5.26	3.61	3.07
40	3.75	2.62	2.45	3.82	2.65	2.43
45	2.68	1.91	2.01	2.74	1.94	1.98
	Coefficient of Kurtosis					
25	92.41	58.15	45.39	94.24	57.83	43.75
30	57.93	33.12	26.44	59.46	33.14	25.58
35	32.01	18.59	16.33	33.04	18.71	15.88
40	17.55	11.03	11.10	18.16	11.11	10.81
45	9.80	7.00	8.08	10.15	7.03	7.87

Table 3
Whole Life Insurance
Using the I.I.D. Model with $\delta = 7\%$

Age	$\sigma = 0.0$	$\sigma = 0.03$	$\sigma = 0.05$	$\sigma = 0.07$
		Net Single Premium		
25	0.0469	0.0477	0.0491	0.0513
30	0.0630	0.0639	0.0656	0.0683
35	0.0856	0.0867	0.0888	0.0920
40	0.1161	0.1174	0.1198	0.1236
45	0.1556	0.1572	0.1600	0.1643
		Standard Deviation		
25	0.0739	0.0750	0.0772	0.0807
30	0.0837	0.0852	0.0879	0.0924
35	0.1006	0.1023	0.1056	0.1108
40	0.1221	0.1241	0.1278	0.1339
45	0.1459	0.1481	0.1523	0.1590
		Coefficient of Skewness		
25	6.24	6.13	5.94	5.66
30	4.81	4.73	4.58	4.39
35	3.77	3.71	3.61	3.49
40	3.01	2.97	2.90	2.81
45	2.40	2.37	2.32	2.25
		Coefficient of Kurtosis		
25	55.44	53.89	51.18	47.35
30	34.98	33.98	32.31	30.07
35	22.08	21.54	20.64	19.49
40	14.50	14.20	13.71	13.10
45	9.77	9.59	9.32	9.00

Table 4
Temporary Annuity for 20, 30, and 40 Years
Using the I.I.D. Model with δ = 7%

	$\sigma = 0.0$			$\sigma = 0.03$		
Age	20 yrs	30 yrs	40 yrs	20 yrs	30 yrs	40 yrs
	Net Single Premium					
25	10.324	11.969	12.718	10.362	12.026	12.786
30	10.299	11.898	12.583	10.337	11.954	12.650
35	10.237	11.756	12.348	10.275	11.811	12.412
40	10.119	11.512	11.980	10.156	11.565	12.039
45	9.917	11.127	11.450	9.953	11.176	11.504
	Standard Deviation					
25	0.6188	0.8311	0.9759	0.9411	1.2325	1.4091
30	0.6796	0.9482	1.1273	0.9812	1.3111	1.5105
35	0.8505	1.1929	1.3982	1.1051	1.4929	1.7119
40	1.1188	1.5366	1.7473	1.3206	1.7731	1.9941
45	1.4637	1.9407	2.1276	1.6203	2.1244	2.3189
	Coefficient of Skewness					
25	-11.84	-9.27	-7.60	-3.21	-2.66	-2.33
30	-9.91	-7.20	-5.71	-3.14	-2.54	-2.18
35	-7.46	-5.31	-4.24	-3.26	-2.55	-2.13
40	-5.47	-3.92	-3.22	-3.21	-2.42	-2.02
45	-4.00	-2.90	-2.48	-2.86	-2.10	-1.79
	Coefficient of Kurtosis					
25	157.96	102.91	74.47	31.19	22.93	18.77
30	114.76	65.62	44.96	27.92	19.40	15.36
35	66.63	36.98	26.01	24.55	16.17	12.65
40	36.66	21.01	15.80	19.68	12.61	10.08
45	20.41	12.39	10.11	14.12	9.15	7.71

Table 4 (contd.)
Temporary Annuity for 20, 30, and 40 Years
Using the I.I.D. Model with δ = 7%

	$\sigma = 0.05$			$\sigma = 0.07$		
Age	20 yrs	30 yrs	40 yrs	20 yrs	30 yrs	40 yrs
			Net Single Premium			
25	10.431	12.128	12.910	10.535	12.283	13.099
30	10.405	12.054	12.770	10.509	12.208	12.955
35	10.342	11.909	12.527	10.445	12.059	12.703
40	10.222	11.659	12.147	10.323	11.803	12.312
45	10.017	11.264	11.602	10.114	11.399	11.751
			Standard Deviation			
25	1.3463	1.7504	1.9818	1.8128	2.3592	2.6667
30	1.3735	1.8021	2.0444	1.8316	2.3921	2.6995
35	1.4625	1.9314	2.1818	1.8966	2.4815	2.7839
40	1.6283	2.1450	2.3885	2.0229	2.6369	2.9188
45	1.8749	2.4290	2.6389	2.2196	2.8520	3.0880
			Coefficient of Skewness			
25	-0.81	-0.60	-0.48	0.06	0.21	0.30
30	-0.86	-0.66	-0.53	0.03	0.16	0.25
35	-1.15	-0.89	-0.72	-0.16	0.00	0.10
40	-1.50	-1.12	-0.91	-0.46	-0.23	-0.10
45	-1.67	-1.20	-0.99	-0.74	-0.43	-0.28
			Coefficient of Kurtosis			
25	9.59	7.83	7.02	5.32	4.94	4.82
30	9.30	7.46	6.60	5.27	4.86	4.72
35	9.72	7.46	6.48	5.55	4.94	4.73
40	9.83	7.18	6.21	5.88	4.98	4.70
45	8.80	6.30	5.58	5.84	4.77	4.52

Table 5
Whole Life Annuity
Using the I.I.D. Model with δ = 7%

Age	$\sigma = 0.0$	$\sigma = 0.03$	$\sigma = 0.05$	$\sigma = 0.07$
		Net Single Premium		
25	13.097	13.174	13.312	13.525
30	12.860	12.933	13.064	13.265
35	12.526	12.593	12.715	12.902
40	12.074	12.136	12.247	12.417
45	11.489	11.544	11.643	11.794
		Standard Deviation		
25	1.093	1.536	2.135	2.863
30	1.239	1.625	2.175	2.860
35	1.488	1.804	2.283	2.902
40	1.806	2.055	2.454	2.994
45	2.157	2.350	2.672	3.125
		Coefficient of Skewness		
25	-6.242	-2.033	-0.380	0.366
30	-4.809	-1.914	-0.437	0.313
35	-3.769	-1.927	-0.626	0.162
40	-3.009	-1.889	-0.833	-0.047
45	-2.399	-1.732	-0.951	-0.244
		Coefficient of Kurtosis		
25	55.440	15.873	6.496	4.793
30	34.979	13.205	6.166	4.696
35	22.083	11.305	6.125	4.689
40	14.505	9.441	5.981	4.670
45	9.765	7.493	5.485	4.505

Tables 1 to 3 show that the values of skewness and kurtosis are both large, reflecting the skew and sharply peaked shape of the density functions. They also show that an increase in σ is associated with decreases in the absolute values of the skewness and kurtosis because when σ increases, the density function will tend to be spread more evenly over its range. Tables 4 and 5 show similar features, although the skewness is negative for many cases and becomes more positive as σ is increased.

3 The Moving Average Process MA(q)

As noted by Frees (1990), the assumption that the sequence Δ_k consists of i.i.d. random variables is a useful advance on the traditional deterministic assumption that allows the volatility of interest rates to be incorporated in the model.

In practice, however, the i.i.d (normal or not) assumption will rarely be satisfied. Investment returns often feature underlying patterns, including characteristics such as dependency, trend, seasonality, and cyclic fluctuations. It is more reasonable to assume that successive interest rates are stochastic and dependent. Thus we assume that the force of interest follows a moving average model of order q (MA(q)) and generalize some of the ideas of Frees (1990), who used the simpler MA(1) model.

The MA(q) model accounts for some correlation between the rates and is tractable (in the mathematical sense) in terms of the calculation of insurance functions. Moving average time series models are more tractable than autoregressive time series models—as exemplified by comparison of our closed-form results with those of Panjer and Bellhouse (1980). Further, as noted by Mills (1999), the autocorrelation functions and partial autocorrelation functions of these two structurally different time-series models often appear similar.

Although there are many examples in the literature of successful empirical studies using $AR(p)$ models, there are fewer case studies using MA(q) models. Examples include:

- MA(2) model fitted to the differences in the U.K. de Zoete Equity Index from 1919 to 1978 (Maturity Guarantees Working Party, 1980)

- MA(1) model fitted to Salomon Brothers U.S. Bond Index from 1926 to 1985 (Frees, 1990)

- MA(1) model fitted to nominal returns on the U.K. Financial Times - Actuaries All Share Index from 1965 to 1995 (Mills, 1999).

The MA(q) model also has been applied to pension funding problems by Haberman and Wong (1997) and Bédard and Dufresne (1998).

3.1 Basic Results

We will now derive expressions for $E[v_k]$ and $E[v_r^m v_s^n]$ for an MA(q) process.

Proposition 3. *Suppose Δ_k follows a MA(q) process, i.e.,*

$$\Delta_k = \delta + a_0\varepsilon_k + a_1\varepsilon_{k-1} + a_2\varepsilon_{k-2} + \dots a_q\varepsilon_{k-q}.$$

for $q = 1, 2, \dots$ with $a_0 = 1$. It follows that

$$\sum_{r=1}^{k} \Delta_r = k\delta + \sum_{h=1-q}^{k} \alpha_h(k)\varepsilon_h \tag{21}$$

and

$$E[v_k] = e^{-k\delta} \prod_{h=1-q}^{k} M(-\alpha_h(k)) \tag{22}$$

for $k = 1, 2, \dots$, where $M(\tau) = E\left[e^{\tau\varepsilon_k}\right]$, and for (i) $q \geq k$,

$$\alpha_h(k) = \begin{cases} \sum_{j=1-h}^{q} a_j & \text{for } h = 1-q, 2-q, \dots, k-q-1 \\ \sum_{j=1-h}^{k-h} a_j & \text{for } h = k-q, k-q+1, \dots, 0 \\ \sum_{j=0}^{k-h} a_j & \text{for } j = 1, 2, \dots, k \\ 0 & \text{for } h \geq k+1 \end{cases} \tag{23}$$

and, (ii) $q < k$,

$$\alpha_h(k) = \begin{cases} \sum_{j=1-h}^{q} a_j & \text{for } h = 1-q, 2-q, \dots, 0 \\ \sum_{j=0}^{q} a_j & \text{for } h = 1, 2, \dots, k-q-1 \\ \sum_{j=0}^{k-h} a_j & \text{for } h = k-q, k-q+1, \dots, k \\ 0 & \text{for } h \geq k+1. \end{cases} \tag{24}$$

Proof: As

$$\Delta_r = \delta + \sum_{j=0}^{q} a_j \varepsilon_{r-j} \quad \text{for } r = \ldots, -1, 0, 1, \ldots$$

it follows that

$$\sum_{r=1}^{k} \Delta_r = k\delta + \sum_{r=1}^{k} \sum_{j=0}^{q} a_j \varepsilon_{r-j}.$$

Substituting $h = r - j$ yields

$$\sum_{r=1}^{k} \Delta_r = k\delta + \sum_{j=0}^{q} \sum_{h=1-j}^{h=k-j} a_j \varepsilon_h.$$

We change the order of summation and write

$$\sum_{r=1}^{k} \Delta_r = k\delta + \sum_{h=1-q}^{k} \alpha_h(k) \varepsilon_h$$

where the $\alpha_h(k)$s are constants to be determined.

When changing the order of summation, there are then two cases to consider (i) $q \geq k$ and (ii) $q \leq k - 1$.

Case 1: $q \geq k$.

The second term of equation (21) corresponds to summing over the area represented by the parallelogram ACFE in Figure 1. There are three distinct regions: the triangles ACB and DFE and the parallelogram BCDE. Changing the order of summation corresponds to changing from summing over horizontal strips to summing over vertical strips.

It is then straightforward to show that the form of $\alpha_h(k)$ is given by equation (23).

Case 2: $q \leq k - 1$.

The second term of equation (21) now corresponds to summing over the area represented by the parallelogram ADFB in Figure 2. The distinct regions are the triangles ACB and DFE and the rectangle BCDE. It is then straightforward to show that the form of $\alpha_h(k)$ is given by equation (24). The proposition is thus proved.

Figure 1
Regions of Summation for the Case $q \geq k$

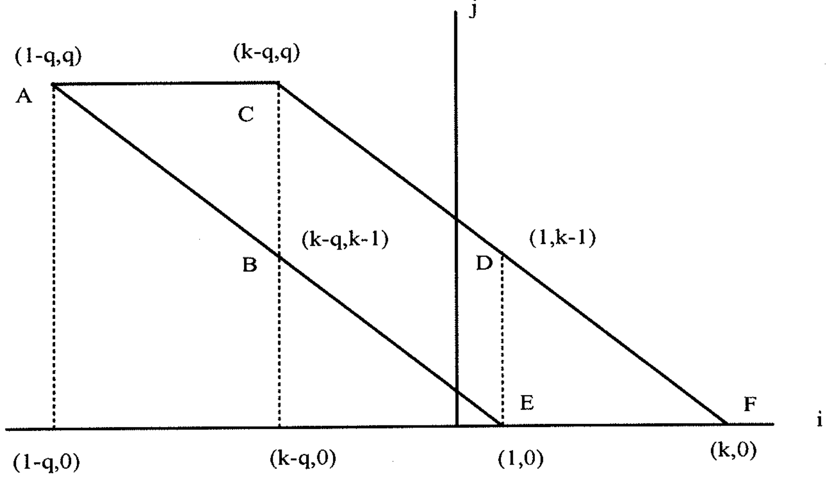
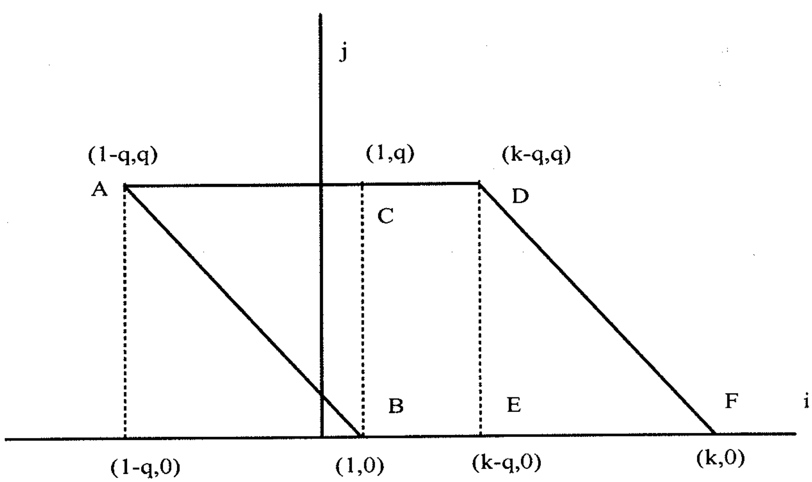


Figure 2
Regions of Summation for the Case $q \leq k - 1$



We can easily use Proposition 1 to prove the following:

$$E\left[v_t^n\right] = e^{-nk\delta} \prod_{h=1-q}^{t} M(-n\alpha_h(t)) \quad \text{for } t, n = 1, 2, \ldots \tag{25}$$

and, for $s, t, m, n = 1, 2, \ldots$

$$\begin{aligned} E\left[v_s^m v_t^n\right] &= e^{-(ms+nt)\delta} E\left[\exp\left(-m \sum_{h=1-q}^{Max(s,t)} (m\alpha_h(s) + n\alpha_h(t))\,\varepsilon_h\right)\right] \\ &= e^{-(ms+nt)\delta} \prod_{h=1-q}^{Max(s,t)} M\left(-(m\alpha_h(s) + n\alpha_h(t))\right) \text{ for } s \neq k. \end{aligned} \tag{26}$$

Under the simplifying assumption that $\varepsilon \sim N(0, \sigma^2)$,

$$M(\tau) = e^{\frac{1}{2}\tau^2\sigma^2}$$

and equation (25) yields

$$\begin{aligned} E\left[v_k^n\right] &= e^{-nk\delta} \prod_{h=1-q}^{k} e^{\frac{1}{2}\sigma^2(n\alpha_h(k))^2} \\ &= e^{-nk\delta + \frac{1}{2}n^2\sigma^2 \sum_{h=1-q}^{k}(\alpha_h(k))^2}. \end{aligned} \tag{27}$$

It is straightforward to show that the particular case $q = 1$ leads to the results obtained by Frees (1990).

Recall equations (6) and (7), which give the corresponding result for the i.i.d. case:

$$E[v_k^n] = e^{-nk\delta + \frac{k}{2}n^2\sigma^2}.$$

It is not clear which of these two expected values (in equations (6) and (27)) is the larger and, hence, which would lead to larger single premiums or reserves. The answer depends on whether or not the term

$$\sum_{h=1-q}^{k} (\alpha_h(k))^2 \gtreqless k.$$

3.2 Life Insurance Actuarial Functions

If we consider the general insurance and annuity contracts given in equations (14) and (15), respectively, we can easily calculate their first two moments.

For convenience, we will define the quantities

$$D_k(n) = \prod_{h=1-q}^{k} M\left(-n\alpha_h(k)\right) \tag{28}$$

and

$$G_{s,t}(m,n) = \prod_{h=1-q}^{Max(s,t)} M\left(-\left(m\alpha_h(s) + n\alpha_h(t)\right)\right). \tag{29}$$

For Z, the nth moment is given by

$$\begin{aligned} E[Z^n] &= E[E[v_{K+1}^n b_{K+1}^n | K]] \\ &= E[e^{-n\delta(K+1)} D_{K+1}(n) b_{K+1}^n] \\ &= \sum_{k=0}^{\infty} e^{-n(k+1)\delta} D_{k+1}(n) b_{k+1}^n {}_{k|1}q_x. \end{aligned} \tag{30}$$

For the case of general annuity contracts, the moments are more difficult to determine. The first two moments are:

$$\begin{aligned} E[Y] &= E[E[\sum_{s=0}^{K} v_s c_s | K]] \\ &= E[\sum_{s=0}^{K} c_s e^{-s\delta} D_s(1)] \\ &= \sum_{k=0}^{\infty} \sum_{s=0}^{k} c_s e^{-s\delta} D_s(1) {}_{k|1}q_x \\ &= \sum_{s=0}^{\infty} c_s e^{-s\delta} D_s(1) {}_s p_x \end{aligned} \tag{31}$$

and

$$
\begin{aligned}
E[Y^2] &= E[E[(\sum_{s=0}^{K} v_s c_s)^2 | K]] \\
&= E[E[\sum_{s=0}^{K} v_s^2 c_s^2 + 2 \sum_{r=1}^{K} \sum_{s=0}^{r-1} v_s c_s v_r c_r | K]] \\
&= E[\sum_{s=0}^{K} e^{-2s\delta} D_s(2) c_s^2] + 2E[\sum_{r=1}^{K} \sum_{s=0}^{r-1} e^{-2(r+s)\delta} G_{r,s}(1,1) c_s c_r] \\
&= \sum_{k=0}^{\infty} {}_{k|1}q_x \sum_{s=0}^{k} e^{-2s\delta} D_s(2) c_s^2 \\
&\quad + 2 \sum_{k=1}^{\infty} {}_{k|1}q_x \sum_{r=1}^{k} \sum_{s=0}^{r-1} e^{-2(r+s)\delta} G_{r,s}(1,1) c_r c_s. \qquad (32)
\end{aligned}
$$

3.3 Examples

Tables 6 to 8 present the calculated values for the moments in the cases of an endowment, temporary insurance, and whole life insurances in the $MA(1)$ case (with $a_1 = 0.5$) with $\delta = 0.07$ for different choices of x, n (where appropriate) and σ. The tables show that an $MA(1)$ model with $\varepsilon \sim N(0, \sigma^2)$ yields a higher standard deviation for Z than an i.i.d. model with $\Delta_k \sim N(\delta, \sigma^2)$. This is because Δ_k has a higher variance.

Table 6
Endowment Insurance for 20, 30, and 40 Years
Using An $MA(1)$ Model with $a_1 = 0.5$ and $\delta = 7\%$

	$\sigma = 0.00$			$\sigma = 0.03$		
Age	20 yrs	30 yrs	40 yrs	20 yrs	30 yrs	40 yrs
			Net Single Premium			
25	0.0088	0.0158	0.0257	0.0089	0.0160	0.0260
30	0.0131	0.0253	0.0405	0.0132	0.0255	0.0411
35	0.0225	0.0425	0.0645	0.0226	0.0429	0.0653
40	0.0395	0.0705	0.0991	0.0397	0.0712	0.1004
45	0.0675	0.1125	0.1448	0.0680	0.1135	0.1465
			Standard Deviation			
25	0.0688	0.0763	0.0794	0.0694	0.0773	0.0807
30	0.0800	0.0901	0.0922	0.0808	0.0914	0.0939
35	0.1022	0.1129	0.1114	0.1032	0.1147	0.1136
40	0.1332	0.1414	0.1334	0.1347	0.1436	0.1360
45	0.1705	0.1711	0.1549	0.1723	0.1737	0.1578
			Coefficient of Skewness			
25	8.95	6.76	5.72	8.99	6.74	5.67
30	6.97	4.92	4.16	7.01	4.92	4.13
35	5.10	3.56	3.15	5.13	3.57	3.12
40	3.68	2.60	2.50	3.71	2.61	2.48
45	2.63	1.90	2.06	2.65	1.91	2.04
			Coefficient of Kurtosis			
25	91.00	58.89	47.46	91.58	58.51	46.49
30	56.77	33.44	27.65	57.24	33.26	27.07
35	31.21	18.68	17.00	31.54	18.61	16.68
40	17.07	11.08	11.54	17.27	11.04	11.33
45	9.52	7.05	8.39	9.64	7.01	8.24

Table 6 (contd.)
Endowment Insurance for 20, 30, and 40 Years
Using An $MA(1)$ Model with $a_1 = 0.5$ and $\delta = 7\%$

	$\sigma = 0.05$			$\sigma = 0.07$		
Age	20 yrs	30 yrs	40 yrs	20 yrs	30 yrs	40 yrs
			Net Single Premium			
25	0.0089	0.0162	0.0266	0.0091	0.0166	0.0276
30	0.0133	0.0260	0.0421	0.0136	0.0267	0.0437
35	0.0229	0.0437	0.0669	0.0233	0.0448	0.0693
40	0.0402	0.0724	0.1026	0.0408	0.0743	0.1061
45	0.0687	0.1154	0.1495	0.0699	0.1183	0.1542
			Standard Deviation			
25	0.0705	0.0790	0.0831	0.0722	0.0819	0.0870
30	0.0823	0.0939	0.0971	0.0846	0.0978	0.1023
35	0.1052	0.1179	0.1176	0.1083	0.1230	0.1241
40	0.1373	0.1476	0.1407	0.1414	0.1541	0.1484
45	0.1757	0.1786	0.1632	0.1809	0.1863	0.1719
			Coefficient of Skewness			
25	9.05	6.73	5.58	9.18	6.75	5.48
30	7.08	4.93	4.07	7.21	4.98	4.03
35	5.20	3.58	3.08	5.31	3.64	3.06
40	3.76	2.63	2.44	3.86	2.67	2.42
45	2.69	1.92	2.00	2.77	1.96	1.98
			Coefficient of Kurtosis			
25	92.83	58.02	44.93	95.40	57.86	43.07
30	58.28	33.09	26.18	60.42	33.30	25.27
35	32.25	18.60	16.19	33.68	18.87	15.74
40	17.69	11.04	11.01	18.53	11.22	10.72
45	9.88	7.00	8.01	10.37	7.09	7.81

Table 7
Term (Temporary) Insurance for 20, 30, and 40 Years
Using An $MA(1)$ Model with $a_1 = 0.5$ and $\delta = 7\%$

	$\sigma = 0.05$			$\sigma = 0.07$		
Age	20 yrs	30 yrs	40 yrs	20 yrs	30 yrs	40 yrs
			Net Single Premium			
25	0.2507	0.1310	0.0759	0.2535	0.1331	0.0774
30	0.2522	0.1354	0.0846	0.2550	0.1375	0.0861
35	0.2560	0.1444	0.0999	0.2588	0.1465	0.1015
40	0.2632	0.1599	0.1240	0.2660	0.1622	0.1258
45	0.2757	0.1848	0.1591	0.2785	0.1871	0.1611
			Standard Deviation			
25	0.0401	0.0550	0.0652	0.0551	0.0600	0.0673
30	0.0438	0.0625	0.0753	0.0579	0.0673	0.0776
35	0.0547	0.0787	0.0936	0.0667	0.0829	0.0960
40	0.0721	0.1017	0.1173	0.0817	0.1054	0.1199
45	0.0946	0.1289	0.1433	0.1023	0.1324	0.1461
			Coefficient of Skewness			
25	12.17	9.48	7.74	4.88	7.47	7.24
30	10.28	7.41	5.82	4.64	6.15	5.52
35	7.77	5.46	4.32	4.45	4.84	4.16
40	5.70	4.03	3.26	4.03	3.74	3.18
45	4.17	2.98	2.50	3.38	2.84	2.45
			Coefficient of Kurtosis			
25	165.99	106.91	76.52	49.26	78.08	69.89
30	122.68	68.83	46.31	42.80	53.88	43.09
35	71.70	38.79	26.67	34.77	33.26	25.39
40	39.42	21.94	16.08	25.50	20.06	15.55
45	21.87	12.83	10.21	17.03	12.18	9.96

Table 7 (contd.)
Term (Temporary) Insurance for 20, 30, and 40 Years
Using An $MA(1)$ Model with $a_1 = 0.5$ and $\delta = 7\%$

	$\sigma = 0.05$			$\sigma = 0.07$		
Age	20 yrs	30 yrs	40 yrs	20 yrs	30 yrs	40 yrs
			Net Single Premium			
25	0.2585	0.1369	0.0801	0.2663	0.1429	0.0844
30	0.2600	0.1414	0.0890	0.2678	0.1474	0.0934
35	0.2638	0.1504	0.1045	0.2715	0.1565	0.1093
40	0.2710	0.1662	0.1291	0.2787	0.1724	0.1343
45	0.2835	0.1912	0.1647	0.2912	0.1976	0.1703
			Standard Deviation			
25	0.0762	0.0691	0.0715	0.0722	0.0819	0.0870
30	0.0784	0.0759	0.0820	0.0846	0.0978	0.1023
35	0.0852	0.0907	0.1006	0.1083	0.1230	0.1241
40	0.0976	0.1125	0.1249	0.1414	0.1541	0.1484
45	0.1157	0.1390	0.1515	0.1809	0.1863	0.1719
			Coefficient of Skewness			
25	2.28	5.26	6.43	1.62	3.64	5.42
30	2.29	4.64	5.04	1.64	3.43	4.45
35	2.49	4.01	3.92	1.75	3.22	3.61
40	2.62	3.32	3.05	1.89	2.87	2.90
45	2.52	2.63	2.37	1.92	2.40	2.29
			Coefficient of Kurtosis			
25	16.28	48.08	59.24	8.51	27.25	46.18
30	15.59	36.63	37.90	8.46	23.16	31.52
35	15.59	25.85	23.32	8.84	18.89	20.76
40	14.55	17.26	14.70	9.04	14.24	13.69
45	11.91	11.17	9.57	8.39	10.00	9.16

Table 8
Whole Life Insurance
Using An $MA(1)$ Model with $a_1 = 0.5$ and $\delta = 7\%$

Age	$\sigma = 0.0$	$\sigma = 0.03$	$\sigma = 0.05$	$\sigma = 0.07$
		Net Single Premium		
25	0.0469	0.0479	0.0496	0.0524
30	0.0630	0.0641	0.0663	0.0697
35	0.0856	0.0870	0.0896	0.0936
40	0.1161	0.1178	0.1208	0.1256
45	0.1556	0.1576	0.1611	0.1665
		Standard Deviation		
25	0.0739	0.0753	0.0781	0.0827
30	0.0837	0.0856	0.0890	0.0948
35	0.1006	0.1028	0.1069	0.1138
40	0.1221	0.1246	0.1294	0.1372
45	0.1459	0.1487	0.1540	0.1627
		Coefficient of Skewness		
25	6.24	6.10	5.86	5.53
30	4.81	4.70	4.53	4.31
35	3.77	3.70	3.58	3.43
40	3.01	2.96	2.87	2.78
45	2.40	2.36	2.30	2.23
		Coefficient of Kurtosis		
25	55.44	53.50	50.16	45.54
30	34.98	33.74	31.69	29.08
35	22.08	21.41	20.32	19.01
40	14.50	14.13	13.54	12.87
45	9.77	9.55	9.22	8.88

Thus, for the $MA(q)$ model it is straightforward to show that Δ_k follows a normal distribution with mean δ and variance σ_q^2 where

$$\sigma_q^2 = \sigma^2(1 + a_1^2 + \cdots + a_q^2), \tag{33}$$

which is greater than σ^2. So the variance in $MA(q)$ is higher than for the equivalent $N(\delta, \sigma^2)$.We then note that increasing q leads to an increase in the variance of Δ_k.

4 Concluding Comments

As noted by Frees (1990) and Dufresne (1992), moving average processes often lead to tractable results. They are simpler to manipulate than the full ARMA processes, but still incorporate dependence over time because of the relatively simple form of the covariance structure.

In this paper, we demonstrate the tractability and convenience in the case of standard present value calculations in a life insurance context. There is a duality between the standard AR and MA models that often makes it difficult to distinguish the two models when fitting models to observational data (Frees, 1990). Any lack of fit with actual data from $MA(q)$ models may be offset by the simplifications arising from their use.

References

Baxter, M. and Rennie, A. *Financial Calculus: An Introduction to Derivative Pricing.* Cambridge, England: Cambridge University Press, 1996.

Bédard, D. and Dufresne, D. "Pension Funding with Moving Average Rates of Return." Centre for Actuarial Studies, University of Melbourne, Research Paper No. 67 (1998).

Bellhouse, D.R. and Panjer, H.H. "Stochastic Modelling of Interest Rates with Applications to Life Contingencies, Part II." *Journal of Risk and Insurance* 48 (1981): 628–637.

Box, G.E.P. and Jenkins, G.M. *Time Series Analysis: Forecasting and Control,* revised edition. San Francisco, Calif: Holden Day, 1976.

Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A. and Nesbitt, C.J. *Actuarial Mathematics,* 2nd Edition. Schaumburg, Ill.: Society of Actuaries, 1997.

Boyle, P.P. "Rates of Return as Random Variables." *Journal of Risk and Insurance* 43 (1976): 693–711.

Chadburn, R.G. and Wright, I.D. "The Sensitivity of Life Office Simulation Outcomes to Differences in Asset Model Structure." *Insurance: Mathematics and Economics* (2000), to appear.

De Jong, P. "The Mean Square Error of a Randomly Discounted Sequence of Uncertain Payments." In *Premium Calculation in Insurance.* (F. De Vylder Editor). Dordrecht, Holland: Reidel Publishing Co., 1984, 449–459.

Dhaene, J. "Stochastic Interest Rates and Autoregressive Integrated Moving Average Processes." *ASTIN Bulletin* (1989): 131–138.

Dufresne, D. "On Discounting when Rates of Return are Random." *Transactions of 24th International Congress of Actuaries* (1992): 27-44.

Frees, E.W. "Stochastic Life Contingencies with Solvency Considerations." *Transactions of Society of Actuaries* 42 (1990): 91–148.

Gerber, H.U. *Life Insurance Mathematics* 2nd Edition. Zurich, Switzerland: Springer and Swiss Association of Actuaries, 1995.

Giaccotto, C. "Stochastic Modelling of Interest Rates: Actuarial vs. Equilibrium Approach." *Journal of Risk and Insurance* 53 (1986): 435–453.

Haberman, S. and Wong L.Y.P. "Moving Average Rates of Return and the Variability of Pension Contributions and Fund Levels for a Defined Benefit Pension Scheme." *Insurance: Mathematics and Economics* 20 (1997): 115–135.

Marceau, E. and Gaillardetz, P. "On Life Insurance Reserves in Stochastic Mortality and Interest Rates Environment." *Insurance: Mathematics and Economics* 25, (1999): 261–280.

Maturity Guarantees Working Party. *Journal of the Institute of Actuaries* 107, Appendix C (1980): 116–133.

Mills T.C. Chapter 2 in *The Econometric Modelling of Financial Time Series*, 2nd Edition. Cambridge, England: Cambridge University Press, 1999.

Panjer, H.H and Bellhouse, D.R. "Stochastic Modelling of Interest Rates with Applications to Life Contingencies." *Journal of Risk and Insurance* 49 (1980): 91–110.

Pollard J.H. "On Fluctuating Interest Rates." *Bulletin de l'Association Royale des Actuaries Belges* 66 (1971): 68–94.

Waters, H.R. "The Moments and Distributions of Actuarial Functions." *Journal of the Institute of Actuaries* 105 (1978): 61–75.

Westcott, D.A. "Moments of Compound Interest Functions under Fluctuating Interest Rates." *Scandinavian Actuarial Journal* (1981): 237–244.

Wilkie, A.D. "The Rate of Interest as a Stochastic Process—Theory and Application." *Transactions of 20th International Congress of Actuaries* 1 (1976): 325–338.

Wright, I.D. "The Application of Stochastic Asset-Liability Modelling Techniques within a Pension Fund Environment." Ph.D. Thesis. Edinburgh: Heriot-Watt University, 1997.

Modeling Corporate Bond Default Risk: A Multiple Time Series Approach

Wai-Sum Chan*

Abstract†

A multiple time series approach is used to forecast the short-term U.S. corporate bond default level. These time series have two auxiliary economic variables: U.S. price inflation and U.S. GNP growth rate. Actual U.S. data from the turn of the century to the present are used to estimate the parameters of multivariate time series model. Diagnostic checks are performed to examine adequacy of the model. The model's forecast for the aggregate U.S. bond default level in 2000-2001 are 0.42% and 0.56%, respectively, while the forecast for the speculative-grade default rate in 2000 is 3.6%, which is more pessimistic than some other forecasts available in the market.

Key words and phrases: *autoregressive, moving average, stationary, forecasting, high-yield bonds, vector time series*

1 Introduction

A bond is said to be in default when the bond issuer has missed a payment of interest, filed for bankruptcy, or announced a distressed-creditor restructuring. The default rate is measured on an initial population of bonds for a finite period of time, such as one year.[1]

*Wai-Sum Chan, Ph.D., F.S.A., is an associate professor of actuarial science in the University of Hong Kong. Dr. Chan has a B.B.A. (Honours) in accounting from the Chinese University of Hong Kong and a Ph.D. in applied statistics from Temple University, U.S.A. He is a Fellow of both the Society of Actuaries and the Royal Statistical Society.

Dr. Chan's address is: Department of Statistics & Actuarial Science, The University of Hong Kong, Pokfulam Road, HONG KONG. Internet address: *chanws@hku.hk*

†This research is supported by a CRCG grant from the University of Hong Kong. The author is indebted to the anonymous referees for their helpful comments on an earlier version of this paper.

[1]For example, Caouette et al., (1998, p. 195) computed the annual default rates for all domestic corporate U.S. bonds from 1971 to 1997.

The incidence of default by U.S. corporate bond issuers is spread unevenly over this century, with high rates of default in the 1910s, the Great Depression of the 1930s, and again in the late 1980s and early 1990s. Figure 1 shows the aggregate default rate for all U.S. corporate domestic bond issuers as an annual time series from 1900 to 1999 and is derived using data in Vanderhoof et al., (1989), Altman and Kishore (1998), and Altman et al., (2000). Notice how aggregate corporate default risk has ebbed and flowed since 1900. Though the risk of asset default has not been a real threat to life insurance companies over the last 50 years, this can easily change. As most life insurance companies in U.S. hold a significant portion of corporate bonds[2] in their investment portfolios, it is important for actuaries to watch for movements of bond default levels.

This paper investigates the possibility of using a multiple (vector) time series model to provide short-term forecasts of the future level of aggregate bond defaults. In addition to the bond default rate, two other economic variables are incorporated into the vector model: the U.S. price inflation rate and the U.S. gross national product (GNP) growth rate. The inflation rate is the most important driving force of some commonly used actuarial stochastic models (Wilkie, 1995). On the other hand, the GNP growth rate is an important leading indicator of economic stability (Vanderhoof et al., 1989). The price inflation time series and the GNP growth rate time series are shown in Figures 2 and 3, respectively.

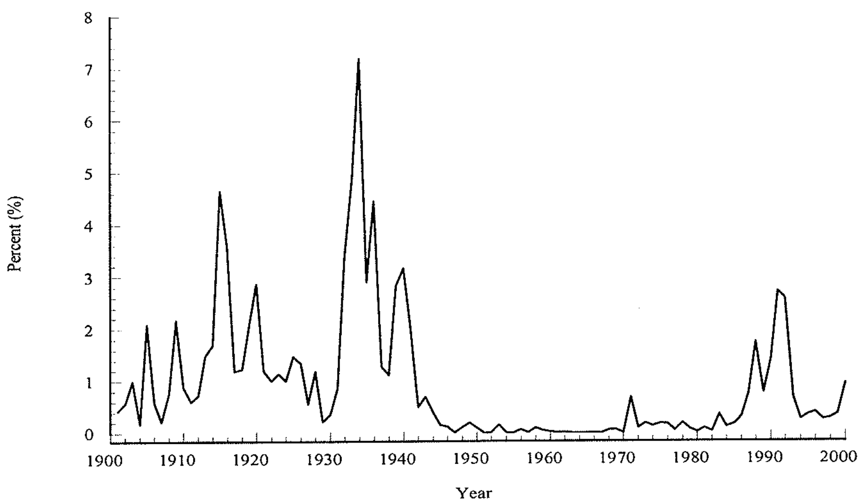
The procedure suggested by Tiao and Box (1981) is used to build a multivariate stochastic model for the three variables. This procedure has the advantage of being more direct and transparent, as compared with alternatives due to Granger and Newbold (1977) and Sims (1980). The sequential and iterative steps of tentative specification, estimation, and diagnostic checking parallel those of the well-known Box-Jenkins method in the univariate time series case. The model is completely determined by the data. Actuarial applications of the Tiao and Box approach can be found in Frees et al., (1997). Unfortunately, detailed model building information was not given in Frees' paper.

The main objectives of this paper are:

- To introduce actuaries to some of the advanced multiple time series analysis techniques used in building vector stochastic models;

[2]See, Vanderhoof et al. (1989, p.547).

Figure 1
Aggregate Default Rate for U.S. Corporate Domestic Bonds (B_t)

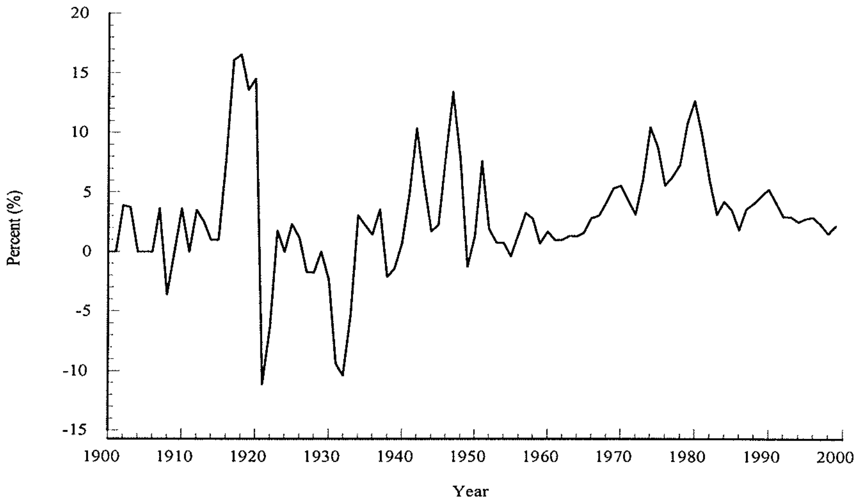


Data Sources: Vanderhoof et al., (1989), Altman and Kishore (1998), and Altman et al., (2000).

- To illustrate the Tiao and Box multiple time series model building procedure in a step-by-step manner so that actuaries who are not expert in this area can still follow the procedure;
- To provide actuaries with a tool for determining whether or not a block of business of an entire company has enough surplus to withstand a possible catastrophic event (Zurcher, 1993);
- To provide actuaries with a tool for determining whether or not a leading economic indicator may serve an an alarm signal for future possible jumps in the bond default levels.

The paper is organized as follows. Section 2 provides a review of the multiple time series modeling approach due to Tiao and Box (1981). Discussion is restricted to points necessary for describing the applications in this paper. Further details can be found in Tiao and Box (1981), Lütkepohl (1993), and Reinsel (1997). Section 3 describes the data while

Figure 2
U.S. Price Inflation (I_t)



Data Sources: Liesner (1989), ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt and http://www.bea.doc.gov/bea/ARTICLES/NATIONAL/NIPA/1998/0898nip3.pdf

Section 4 deals with the process of fitting the model. Section 5 provides an analysis of high-yield bonds. Section 6 concludes the paper.

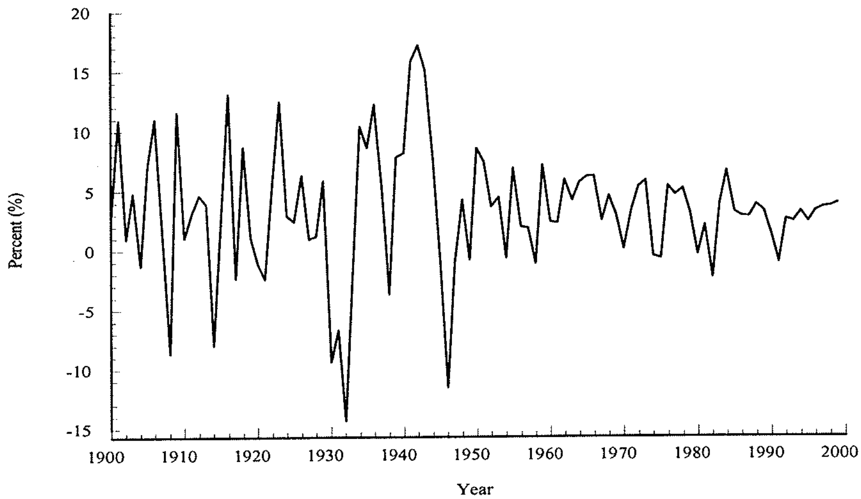
2 Multiple Time Series Analysis

Consider an m-element stationary column vector time series $\mathbf{Y}_t$ with mean $\boldsymbol{\mu}$ for $t = \ldots, -1, 0, 1, \ldots$. $\mathbf{Y}_t$ follows a vector autoregressive moving-average (ARMA) process of order p and q if

$$\mathbf{\Phi}(B)\mathbf{Y}_t = \mathbf{c} + \mathbf{\Theta}(B)\boldsymbol{\varepsilon}_t \tag{1}$$

where B is the backward shift operator such that $B\mathbf{Y}_t = \mathbf{Y}_{t-1}$, $\mathbf{\Phi}(B)$ and $\mathbf{\Theta}(B)$ are are $m \times m$ matrix polynomials in B of finite degrees p and q respectively,

Figure 3
U.S. Gross National Product (GNP) Growth Rate (G_t)



Data Sources: Liesner (1989), ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt and http://www.bea.doc.gov/bea/ARTICLES/NATIONAL/NIPA/1998/0898nip3.pdf

$$\boldsymbol{\Phi}(B) = \mathbf{I} - \boldsymbol{\varphi}_1 B - \cdots - \boldsymbol{\varphi}_p B^p$$
$$\boldsymbol{\Theta}(B) = \mathbf{I} - \boldsymbol{\theta}_1 B - \ldots - \boldsymbol{\theta}_q B^q$$

$\mathbf{c}$ is a m-dimensional constant column vector, and $\{\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{mt})'\}$ is a sequence of independent and identically distributed Gaussian random column vectors with mean zero and positive-definite variance-covariance matrix $\boldsymbol{\Sigma} = \{\sigma_{ij}\}$. The zeros of the determinantal polynomials $|\boldsymbol{\Phi}(B)|$ and $|\boldsymbol{\Theta}(B)|$ are all assumed to be on or outside the unit circle. It implies that the vector process is both stationary and invertible.

The cross-covariance matrix of order k, $\boldsymbol{\Gamma}(k)$, is given by

$$\begin{aligned}\boldsymbol{\Gamma}(k) &= E[(\mathbf{Y}_t - \boldsymbol{\mu})(\mathbf{Y}_{t-k} - \boldsymbol{\mu})'] \\ &= \{\gamma_{ij}(k)\}, \qquad i, j = 1, \ldots, m \end{aligned} \tag{2}$$

for all integers k. Also,

$$\boldsymbol{\rho}(k) = \{\rho_{ij}(k)\} = \frac{\gamma_{ij}(k)}{\sigma_{ij}}$$

is defined as the corresponding cross-correlation matrix (CCM). Tiao and Box (1981) define the partial autoregression matrix (PAM) at lag k, denoted by $\boldsymbol{\Pi}(k)$, to be the last matrix coefficient when the data are fitted to a vector autoregressive process of order k. This is a direct extension of the Box and Jenkins (1976, page 64) definition of the partial autocorrelation function for univariate time series.

When $p = 0$, that is, $\mathbf{Y}_t$ is a vector MA(q) process, $\boldsymbol{\Gamma}(k)$ and $\boldsymbol{\rho}(k)$ are zero for $k > q$. On the other hand, the partial autoregression matrices $\boldsymbol{\Pi}(k)$ of a vector AR(p) process are zero for $k > p$. These cut off properties provide useful information for identifying the order of pure vector AR or MA models. However, both CCM and PAM are not useful when dealing with mixed vector ARMA processes (i.e., both $p > 0$ and $q > 0$). They do not exhibit cut off patterns. Simple inspection of the matrices $\boldsymbol{\rho}(k)$ and $\boldsymbol{\Pi}(k)$ would not, in general, give clear values of p and q for mixed models.

Tiao and Tsay (1983) proposed the extended cross-correlation matrix (ECCM) based on the concept of iterated least-squares regression. The asymptotic pattern of the ECCM for a vector ARMA(p, q) model is given in Table 1. There is a remarkable zero-triangle in the table and its vertex is in (p, q) position. Hence, the ECCM can be a useful tool in model specification, particularly for a mixed vector ARMA process.

Tiao and Box (1981) suggested an iterative modeling approach consisting of tentative specification, estimation, and diagnostic checking. For tentative specification the sample cross-correlation matrix (SCCM), denoted by $\hat{\boldsymbol{\rho}}(k) = \{\hat{\rho}_{ij}(k)\}$ is used. These statistics are particularly useful in spotting low order moving average models. If the series $\boldsymbol{\varepsilon}_t$ is a white noise, the standard error of each element of the SCCM is approximately $1/\sqrt{n}$. These statistics, however, provide a crude signal-to-noise ratio guide and are not meant to give formal significant tests.

Estimates of $\boldsymbol{\Pi}(k)$ and their standard errors can be obtained by fitting autoregressive models of successively higher order by least squares. Tiao and Box (1981) recommended using the likelihood ratio statistic to test the null hypothesis $\boldsymbol{\varphi}_k = 0$ against the alternative $\boldsymbol{\varphi}_k \neq 0$ if an AR(k) process is fitted. To conduct such a test, Bartlett's (1938) statistic, $M(k)$, is used. $M(k)$ is asymptotically χ^2 distributed with m^2 degrees of freedom if the null hypothesis is true.

Sample ECCM can be computed using iterated least-squares regressions. One can construct a two-way table from the sample matrices.

Table 1
The Asymptotic Pattern of the Extended Cross-Correlation Matrix for a Vector ARMA(p, q) Model

	MA order							
AR order	...	...	$q-1$	q	$q+1$	...	...	...
$\vdots$	X	...	X	X	X	X	...	X
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$p-1$	X	...	X	X	X	X	...	X
p	X	...	X	O	O	O	...	O
$p+1$	X	...	X	X	O	O	...	O
$p+2$	X	...	X	X	X	O	...	O
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\vdots$	X	...	X	X	X	X	...	O

Note: X represents nonzero matrix and O represents zero matrix.

The identification is carried out by visual searching the vertex of the zero-triangle inside the sample ECCM table. It is particularly useful in specifying the order of a mixed vector ARMA process.

After the order of the vector ARMA model is tentatively selected, asymptotically efficient estimates of the parameters can be determined using the maximum likelihood approach. Approximate standard errors of the estimates of the elements of $\boldsymbol{\varphi}_i$ for $i = 1, 2, \ldots, p$ and $\boldsymbol{\theta}_j$ for $j = 1, 2, \ldots, q$ can also be obtained and used to test for the significance of the parameters. Further gains in the efficiency of the estimates may be achieved by eliminating parameters that are found to be statistically insignificant. Interested readers may refer to Reinsel (1997, Chapter 5) for a detailed discussion of the maximum likelihood estimation for vector ARMA models.

The maximization of the likelihood function can be conducted by a conditional likelihood method or an exact likelihood method. The conditional likelihood method is computationally convenient, but may be inadequate if the sample size is not sufficiently large. Thus, in this paper we estimate the parameters initially using the conditional likelihood approach and eliminate parameters that are small relative to their standard error. The model is then re-estimated using the exact likelihood method. To guard against incorrectly specifying the model, a detailed diagnostic analysis of the residuals is required. This includes an examination of the plots of standardized residuals and the ECCM table of the residuals.

3 Preliminary Data Analysis

3.1 Data Transformation

The vector time series data consist of three key variables: the U.S. corporate bond default rate (B_t), the U.S. inflation rate (I_t), and the U.S. GNP growth rate (G_t). The data are available from 1900 to 1999. Summary statistics for the observed time series are given in Table 2.

The aggregate bond default rate, on the average over the past 100 years, is less than 1%. On the other hand, the average inflation rate and growth rate are around 3%. From Table 2, we also observe that the distributions of B_t and I_t are positively skewed while the distribution of G_t is negatively skewed. Furthermore, all the observed distributions are leptokurtosis (fat tail), with the default rate distribution having the thickest tail.[3]

This suggests that a transformation of the default rate might be called for, so the square-root transformation (a special case in the class of power transformations introduced by Box and Cox, 1964) of the default rate is used, i.e.,

$$D_t = \sqrt{B_t}.$$

The square-root transformation not only stabilizes the variance and the kurtosis of the default rate, but also prevents the default rate from being negative.

The summary statistics for the transformed variable, D_t, are also given in Table 2. The coefficients of skewness and kurtosis of D_t are significantly improved (in the sense of being closer to a Gaussian distribution). It is not unexpected that (D_t, I_t) and (D_t, G_t) are negatively correlated, while (I_t, G_t) is positively correlated. A view of the possible interrelationships of the economic variables using scatter-plot diagrams is given in Figures 4 through 7. These figures show some strong contemporaneous relationships among series. It justifies the use of multiple time series model for the variables.

[3]An observed distribution is called leptokurtosis if its sample coefficient of excess kurtosis is greater than zero.

Table 2
Summary Statistics in Percent (%)

	B_t	D_t	I_t	G_t
Sample Size	100	100	100	100
Mean	0.9173	0.7587	2.9955	3.3198
Median	0.3795	0.6159	2.6635	3.2705
Standard Deviation	1.2679	0.5875	4.7830	5.2751
Minimum	0.0000	0.0000	-11.0930	-14.3160
Maximum	7.1840	2.6803	16.5250	17.1430
Skewness	2.2613	0.8702	0.1943	-0.3606
Kurtosis	6.0373	0.2410	1.6398	1.5019
			Correlation Matrix	
	D_t	1.00	-0.25	-0.20
	I_t	-0.25	1.00	0.18
	G_t	-0.20	0.18	1.00

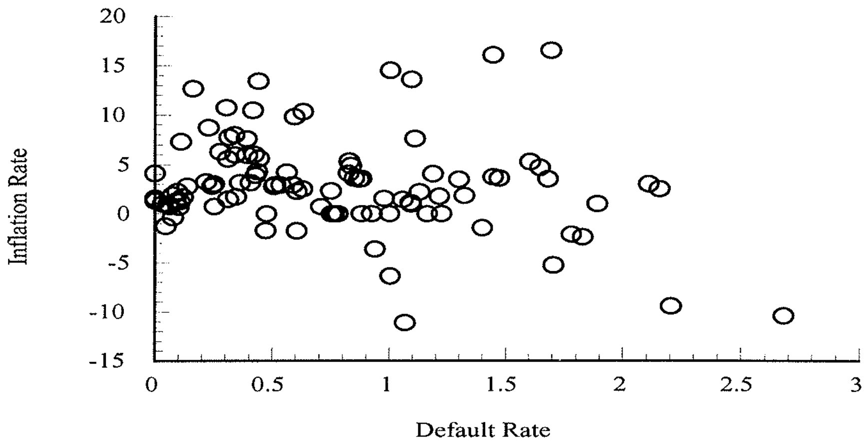
Note: B_t, I_t and G_t are expressed in percent.

For comparison purposes we fit univariate time series models to the economic variables following the orthodox Box and Jenkins (1976) approach. Table 3 gives the sample autocorrelation function (SACF) and the sample partial autocorrelation function (SPACF) of each individual variable up to order 8. The sample autocorrelation coefficients of D_t are exponentially decaying. On the other hand, the sample partial autocorrelations are cut off after lag one. It indicates an AR(1) model for the bond default series. Both the SACF and SPACF for the inflation series are decaying after lag one. It is likely that the underlying process for I_t is an ARMA(1,1). As both lag-1 and lag-4 autocorrelations are significant for the G_t series, an AR(4) model is appropriate.

3.2 Checks for Outliers

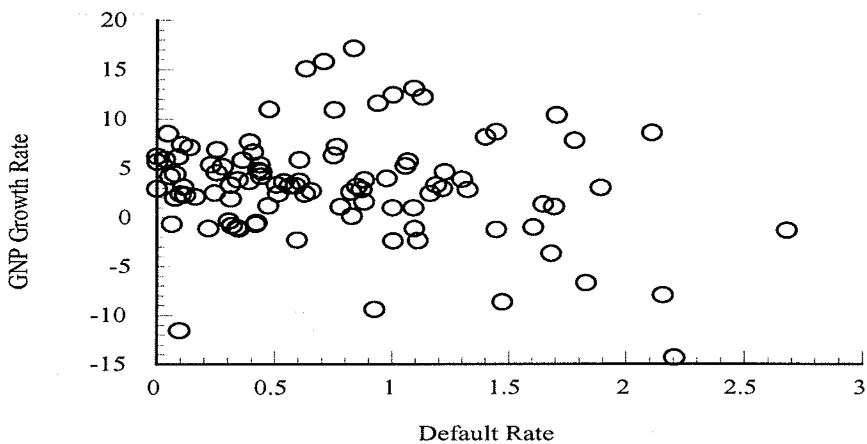
Time series observations are often influenced by interruptive events such as strikes, outbreaks of wars, sudden political or economic crises, or even unnoticed errors of typing and recording. The consequences of these interruptive events create spurious observations, which are inconsistent with the rest of the series. Such observations are usually referred to as outliers.

Figure 4
Scatter Plot of the Variables I_t and G_t



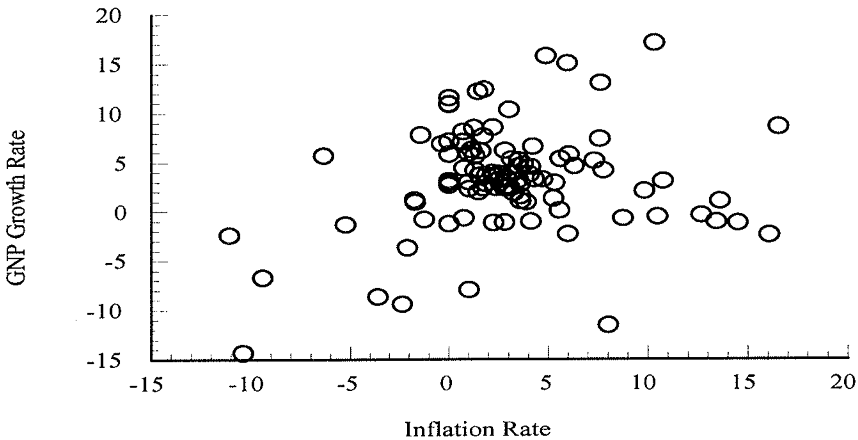
Data Sources: Liesner (1989), ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt and http://www.bea.doc.gov/bea/ARTICLES/NATIONAL/NIPA/1998/0898nip3.pdf

Figure 5
Scatter Plot of the Variables B_t and G_t



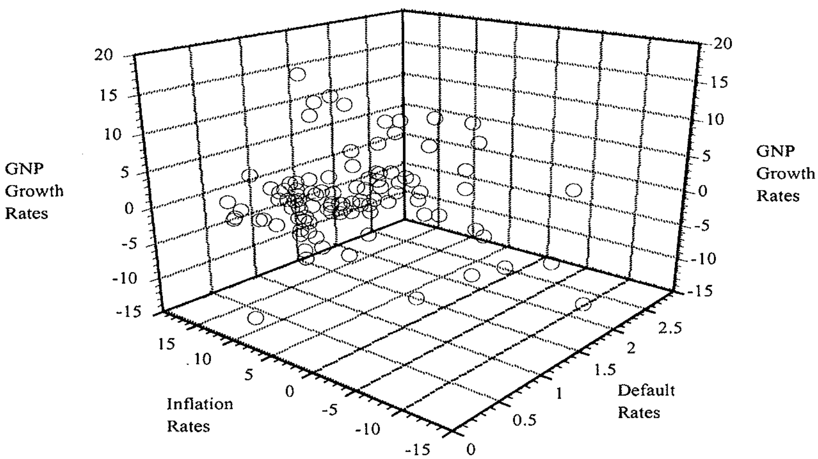
Data Sources: Liesner (1989), ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt and http://www.bea.doc.gov/bea/ARTICLES/NATIONAL/NIPA/1998/0898nip3.pdf

Figure 6
Scatter Plot of the Variables I_t and G_t



Data Sources: Liesner (1989), ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt and http://www.bea.doc.gov/bea/ARTICLES/NATIONAL/NIPA/1998/0898nip3.pdf

Figure 7
3D Scatter Plot of the Variables B_t, I_t and G_t



Data Sources: Liesner (1989), ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt and http://www.bea.doc.gov/bea/ARTICLES/NATIONAL/NIPA/1998/0898nip3.pdf

Table 3
Autocorrelation Coefficients
And Partial Autocorrelation Coefficients

	Lag Order							
	1	2	3	4	5	6	7	8
	(a) Autocorrelation Coefficients							
D_t	0.79	0.67	0.59	0.55	0.49	0.45	0.41	0.35
	(0.10)	(0.15)	(0.18)	(0.20)	(0.21)	(0.23)	(0.23)	(0.24)
I_t	0.62	0.26	0.14	.08	0.15	.17	0.10	0.01
	(0.10)	(0.13)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.15)
G_t	0.28	0.03	-0.09	-0.23	-0.16	0.05	.09	-0.02
	(0.10)	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)	(0.12)	(0.12)
	(b) Partial Autocorrelation Coefficients							
D_t	0.79	0.11	0.07	0.12	-0.01	0.02	0.02	-0.06
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
I_t	0.62	-0.21	0.12	-0.04	0.19	-0.04	-0.00	-0.08
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
G_t	0.28	-0.05	-0.09	-0.20	-0.05	0.11	0.03	-0.13
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)

Note: Standard errors of autocorrelations are given in parentheses.

As retaining outliers can lead to erroneous model specification and biased predictions (Chan, 1998), an outlier analysis is performed on the specified models for the time series D_t, I_t, and G_t. No outliers were found for the D_t and G_t series; a switch outlier[4] was detected in I_t at $t = 1921$. The magnitude of the outlier is estimated as 9.00. See de Jong and Penzer (1998) for more on time series outlier detection and switch outliers. The analysis in this paper is based on the outlier-adjusted series. Finally, the fitted univariate time series models for each series are summarized in Table 4.

[4]A switch outlier occurs where there are consecutive extreme values on either side of the current level of the series.

Table 4

Univariate Time Series Models for D_t, I_t and G_t

Variable	Model	Equation	$\hat{\sigma}^2$	Q_{12}
Bond Default	AR(1)	$D_t = \underbrace{0.16}_{(0.06)} + \underbrace{0.80}_{(0.06)} D_{t-1} + \varepsilon_t$	0.13	7.2
Inflation	ARMA(1,1)	$I_t = \underbrace{1.56}_{(0.55)} + \underbrace{0.52}_{(0.11)} I_{t-1} + \underbrace{0.49}_{(0.11)} \varepsilon_{t-1} + \varepsilon_t$	8.55	7.6
GNP Growth	AR(4)	$G_t = \underbrace{3.06}_{(0.68)} + \underbrace{0.27}_{(0.10)} G_{t-1} - \underbrace{0.21}_{(0.10)} G_{t-4} + \varepsilon_t$	24.38	7.4

Notes: Standard errors of estimates are given in parentheses; $\hat{\sigma}^2$ denotes the estimates of residual variance; and Q_{12} is the Box-Pierce portmanteau statistic of the residuals with lag order up to 12. Note that Q_{12} is asymptotically distributed as a χ^2 with degrees of freedom equal to 15 minus the number of parameters estimated. None of the Q_{12} statistics reported is significant at the 5% level.

3.3 Check for Cointegration

Cointegration analysis has attracted considerable research interest in recent years. Engle and Granger (1991) and Rao (1994) have described large growth in the business and economic applications of this area. A vector time series is said to be cointegrated if each of the series taken individually is nonstationary with a unit root, while some linear combination of the series is stationary. Cointegration of two (or more) time series suggests that there is a long-run, or equilibrium, relationship between them. The error correction mechanism (ECM) developed by Engle and Granger (1987) can reconcile the short-run behavior of an economic variable with its long-run behavior.

It should be noted that if the vector $(D_t, I_t, G_t)'$ process is cointegrated, then it is not correct to fit a vector ARMA model to the differenced data. Therefore, it is important to check for cointegration using the observed series. The first requirement for cointegration is that D_t, I_t, and G_t are each individually nonstationary with a unit root. We employ the augmented Dickey-Fuller (ADF) test to examine each series. For a stochastic variable Y_t, Dickey and Fuller (1981) considered the following regression model:

$$(1 - B)Y_t = \alpha_0 + \alpha_1 t + \delta Y_{t-1} + \sum_{j=1}^{m} \beta_j \left\{(1 - B)Y_{t-j}\right\} + \varepsilon_t.$$

The null hypothesis is that $\delta = 0$; that is, a unit root exists in Y (i.e., Y is is nonstationary with a unit root). The ADF test is applied to the three observed series with $m = 2$, the results are given in Table 5. The ADF tests indicate that not all the series are nonstationary with a unit root, and hence the vector process $(D_t, I_t, G_t)'$ is not cointegrated.

4 The Fitted Model

The multiple time series modeling procedures mentioned in Section 2 of this paper can be efficiently performed using matrix-based computer languages such as S-Plus, GAUSS, MATLAB, and SCA. The computations performed in this section are carried out using the SCA system (Liu and Hudak, 1994).

Model Specification: The sample cross-correlation matrix (SCCM) and the partial autoregression matrix (PAM) are first examined. Tiao and Box (1981) suggested summarizing the SCCM and PAM using indicator symbols +, −, and ·, where + denotes a value greater than twice

Table 5
Augmented Dickey-Fuller (ADF) Tests

	ADF Test Statistic	Critical Value (at 5%)	Conclusion
D_t	-2.71	-3.45	Nonstationary with unit root
I_t	-3.99	-3.45	Stationary
G_t	-5.61	-3.45	Stationary

Notes: The ADF test statistic is simply the t-ratio for $\delta = 0$ and the critical values are obtained from MacKinnon (1991, Chapter 13).

the estimated standard error, $-$ denotes a value less than twice the estimated standard error, and $\cdot$ denotes an insignificant value based on the above criteria. The resulting indicator matrices for SCCM and PAM are given in Table 6.

Both SCCM and PAM do not provide a cut off pattern. This suggests that the underlying process could be a mixed process. Therefore, the sample ECCM table is computed. The results are presented using indicator symbols in Table 7. We find a zero-triangle in the table and its vertex is in (1,1) position. Hence, we tentatively specify a vector ARMA(1,1) model for the data.

Model Estimation: The specified ARMA(1,1) model is first estimated using conditional likelihood method. All parameters in the model are computed. We call this model a "full model" in Table 8. Imposing zero restrictions on the coefficients that are insignificant, we re-estimate the model by exact likelihood method. The final model is given in Table 8.

It should be noted that only stationary and invertible vector time series models were considered for the process $(D_t, I_t, G_t)'$. That is, it was assumed that the zeros of the determinantal polynomials $|\Phi(B)|$ and $|\Theta(B)|$ are all on or outside the unit circle. It is important to check these assumptions for the final fitted model in Table 8. For the vector ARMA(1,1) model, the stationarity and invertiblity conditions are equivalent to restricting all the eigenvalues of $\boldsymbol{\varphi}_1$ and $\boldsymbol{\theta}_1$ inside the unit circle (Wei, 1990, p. 345). The eigenvalues of $\hat{\boldsymbol{\varphi}}_1$ and $\hat{\boldsymbol{\theta}}_1$ are $(0.8602, 0.3688, 0.4810)$ and $(0.3660, -0.5240, 0.0000)$, respectively, which suggests that the final fitted model in Table 8 satisfies the basic assumptions.

Table 6
Indicator Matrices for SCCM and PAM

	lag (k) 1	2	3	4	5
	(a) Sample Cross-Correlation Matrix (SCCM)				
	$\begin{pmatrix} + & - & - \\ \cdot & + & + \\ \cdot & \cdot & + \end{pmatrix}$	$\begin{pmatrix} + & - & \cdot \\ \cdot & + & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} + & - & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} + & - & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & - \end{pmatrix}$	$\begin{pmatrix} + & - & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$
	(b) Partial Autoregression Matrix (PAM)				
	$\begin{pmatrix} + & \cdot & \cdot \\ \cdot & + & \cdot \\ \cdot & \cdot & + \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & - & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & + \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$
$M(k)$	193.34	25.04	10.71	15.06	7.53

Notes: Critical values for $M(k)$: 16.9 for 5% level; 21.7 for 1% level.

Table 7
Indicator Matrices for ECCM

AR Order (p)	MA Order (q) 0	1	2	3	4
0	$\begin{pmatrix} + & - & \cdot \\ \cdot & + & + \\ \cdot & \cdot & + \end{pmatrix}$	$\begin{pmatrix} + & - & \cdot \\ \cdot & + & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} + & - & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} + & - & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & - \end{pmatrix}$	$\begin{pmatrix} + & - & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$
1	$\begin{pmatrix} \cdot & \cdot & \cdot \\ - & + & + \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$
2	$\begin{pmatrix} - & \cdot & + \\ - & + & + \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & + & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$
3	$\begin{pmatrix} \cdot & + & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & + & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & - & \cdot \\ \cdot & - & \cdot \end{pmatrix}$	$\begin{pmatrix} - & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ - & \cdot & + \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$
4	$\begin{pmatrix} + & \cdot & - \\ - & \cdot & + \\ \cdot & - & \cdot \end{pmatrix}$	$\begin{pmatrix} + & + & \cdot \\ \cdot & \cdot & \cdot \\ - & \cdot & + \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$

The full estimated model can be re-written as follows:

$$\begin{aligned}\hat{D}_t &= 0.126 + 0.853\, D_{t-1} - 0.012\, I_{t-1} + 0.007\, G_{t-1} + \varepsilon_{D,t}\\ &\quad - 0.358\, \varepsilon_{D,t-1} + 0.019\, \varepsilon_{I,t-1} - 0.024\, \varepsilon_{G,t-1} \qquad (3)\\ \hat{I}_t &= 1.789 + 0.716\, D_{t-1} + 0.519\, I_{t-1} - 0.266\, G_{t-1} + \varepsilon_{I,t}\\ &\quad - 0.887\, \varepsilon_{D,t-1} + 0.480\, \varepsilon_{I,t-1} + 0.336\, \varepsilon_{G,t-1} \qquad (4)\\ \hat{G}_t &= 1.367 + 1.825\, D_{t-1} + 0.003\, I_{t-1} + 0.160\, G_{t-1} + \varepsilon_{G,t}\\ &\quad - 0.790\, \varepsilon_{D,t-1} - 0.148\, \varepsilon_{I,t-1} + 0.158\, \varepsilon_{G,t-1}. \qquad (5)\end{aligned}$$

The final estimated model can be re-written as follows:

$$\begin{aligned}\hat{D}_t &= 0.124 + 0.898\, D_{t-1} - 0.014\, G_{t-1} + \varepsilon_{D,t} - 0.366\, \varepsilon_{D,t-1} \qquad (6)\\ \hat{I}_t &= 1.569 + 0.481\, I_{t-1} + \varepsilon_{I,t} + 0.524\, \varepsilon_{I,t-1} + 0.110\, \varepsilon_{G,t-1} \qquad (7)\\ \hat{G}_t &= 1.141 + 1.429\, D_{t-1} + 0.331\, G_{t-1} + \varepsilon_{G,t} \qquad (8)\end{aligned}$$

with

$$\hat{\Sigma} = \begin{pmatrix} 0.123 & -0.280 & -0.844 \\ -0.280 & 8.096 & 4.455 \\ -0.844 & 4.455 & 25.045 \end{pmatrix}. \qquad (9)$$

Diagnostic Checking: The indicator matrices of the residual ECCM are given in Table 9. The zero-triangle is pointing at the (0,0) position. It indicates that there is no significant serial correlation information left in the residuals. The portmanteau test of McLeod and Li (1983) is based on the squared residuals of a time series model and is a test for homoscedasticity of the residuals. The test statistics for the residuals from the fitted models (6) to (8) are 1.4799, 1.2413, and 1.6902, respectively. They should be compared with a χ_1^2 variate (critical value at the 5% level is 3.841). We conclude that the residuals are homoscedastic and the fitted model is adequate for the series.

Table 8
Estimation Results

c	$\hat{\varphi}_1$			$\hat{\theta}_1$			$\hat{\Sigma}$		
(a) Full Model									
0.126	0.853	−0.012	0.007	0.358	−0.019	0.024	0.120	−0.298	−0.847
(0.077)	(0.068)	(0.009)	(0.018)	(0.121)	(0.016)	(0.021)			
1.789	0.716	0.519	−0.266	0.887	−0.480	−0.336	−0.298	7.884	4.480
(1.222)	(1.070)	(0.127)	(0.213)	(1.197)	(0.117)	(0.188)			
1.367	1.825	0.003	0.160	0.790	0.148	−0.158	−0.847	4.480	25.094
(1.373)	(1.227)	(0.139)	(0.241)	(1.624)	(0.182)	(0.232)			
(b) Final Model									
0.124	0.898	0	−0.014	0.366	0	0	0.123	−0.280	−0.844
(0.045)	(0.046)		(0.006)	(0.109)					
1.569	0	0.481	0	0	−0.524	−0.110	−0.280	8.096	4.455
(0.541)		(0.098)			(0.103)	(0.050)			
1.141	1.429	0	0.331	0	0	0	−0.844	4.455	25.045
(0.875)	(0.797)		(0.089)						

Notes: Standard errors of estimates are given in parentheses.

Table 9
Indicator Matrices for ECCM of the Residuals

AR Order (p)	MA Order (q): 0	1	2	3	4
0	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$
1	$\begin{pmatrix} + & \cdot & - \\ - & \cdot & + \\ \cdot & \cdot & + \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$
2	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & - \\ \cdot & \cdot & - \\ + & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$
3	$\begin{pmatrix} + & \cdot & \cdot \\ + & \cdot & \cdot \\ + & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} + & \cdot & \cdot \\ + & - & \cdot \\ + & - & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ - & \cdot & \cdot \\ - & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$

5 An Extension

From 1900 through 1945 significant increases in the default rate were typically preceded by weakness in the overall economy as reflected in the GNP growth. Since 1945, it has more often been the case that increases in the default rate occur in advance of a weakening in the general economy. For example, in the worst episode of the post-war era, the default rate began to rise in 1985 rising from 0.315% to its peak of 2.715% in 1990. The GNP growth, on the other hand, peaked in 1984 but did not fall below the zero mark until the year of 1991. The results in equations (6) to (9) are able to describe such lead-lag relationship among the variables explicitly. Furthermore, the final model also captures the correlation momentum among innovations (residuals) implicitly through $\hat{\Sigma}$.

The U.S. high-yield bond market has been developing rapidly since 1980. Many investment managers now consider high-yield bonds a separate and distinct asset class. By the end of 1996, it was estimated that insurance companies and pension funds owned more than 40% of the high-yield debt market. It is important to study the historical default rate on high-yield bonds. Unfortunately, the history of high-yield market is short. Only 40 quarterly default figures on high-yield bonds are available from 1990 to 1999 (Altman et al., 2000).

The quarterly high-yield default rate, QD_t, as well as its corresponding quarterly GNP growth rate, QG_t, are plotted in Figure 8. Using the multiple time series modeling approach as described in Section 2 yields the following model for the series:

$$\sqrt{\widehat{QD}_t} = 1.156 + .279\sqrt{QD}_{t-1} - 0.846\, QG_{t-1} + \varepsilon_{D,t} - 0.732\, \varepsilon_{G,t-1} \tag{10}$$

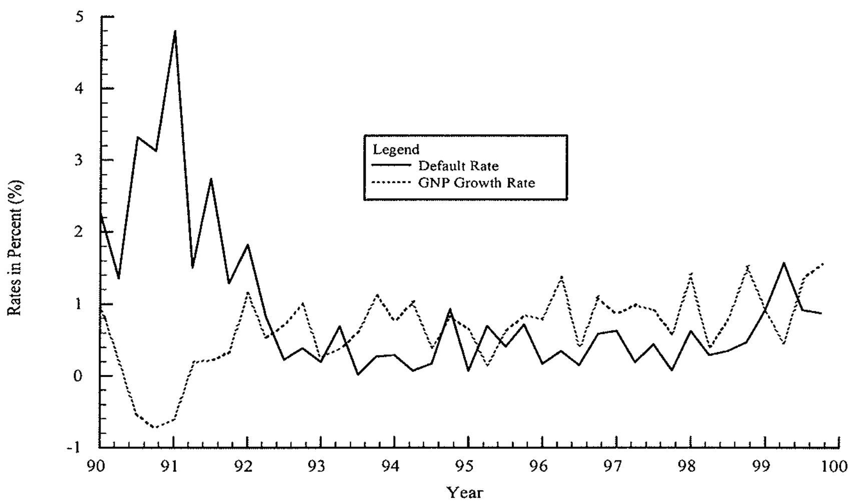
$$\widehat{QG}_t = 1.197 - 0.624\sqrt{QD}_{t-1} + \varepsilon_{G,t} \tag{11}$$

with

$$\hat{\Sigma} = \begin{pmatrix} 0.082 & -0.045 \\ -0.045 & 0.204 \end{pmatrix}. \tag{12}$$

The final model shows a strong first-order contemporary lead-lag relationship between the quarterly high-yield default level and the quarterly growth rate.

Figure 8
Quarterly High-Yield Default Rates and GNP Growth Rates, 1990–1999



Data Sources: Liesner (1989), ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt and http://www.bea.doc.gov/bea/ARTICLES/NATIONAL/NIPA/1998/0898nip3.pdf

6 Closing Comments

The aggregate bond default rates were below average in 1998 and 1999. Based on the fitted equations (3) to (6), the forecast of the default rates for 2000, 2001, and 2002 are 0.341%, 0.417%, and 0.562% respectively.

Based on the equations (7) to (9), the model's forecast for quarterly high-yield default rates for 2000 are 0.896%, 0.928%, 0.899%, and 0.908% respectively. These figures imply an annual forecast of 3.6% high-yield default rate in 2000, which is somewhat more pessimistic than the forecast of 2.8% produced by Altman et al. (2000).

In this paper we have illustrated multiple time series modeling techniques through the analysis of U.S. corporate bond default data. This method has the advantage of being simple to use. The iterative cycles of tentative specification, estimation, and diagnostic checking parallel those of the well-known Box-Jenkins (1976) method. The methodology has been implemented by some time series computer packages, such as SCA (Liu and Hudak, 1994). Vector time series models might be useful to other actuarial applications, say, stochastic asset modeling (Wilkie, 1995), pension simulation (Knox, 1993), and solvency assessment (Hardy, 1993). Research in some of these topics is in process.

There are many books and research papers related to other aspects of default risk or credit risk. Interested readers may refer to Duffie and Huang (1996), Duffie and Singleton (1998), Jarrow (1998), Altman (1999), and Jarrow and Turnbull (2000).

References

Altman, E.I. "Market Dynamics and Investment Performance of Distressed Securities." In *Workouts and Turnarounds II: Global Restructuring for the Next Century*, edited by Cushman, R., Napoli, D. and Sigoloff, S. Homewood, Ill.: Business One Irwin, 1999.

Altman, E.I. and Kishore, V.M. "Defaults and Returns on High Yield Bonds." In *The Handbook of Corporate Debt Instruments.* (ed. F.J. Fabozzi). New Hope, PA: Frank J. Associates, 1998: pp. 255-272.

Altman, E.I., Cooke, D. and Kishore, V.M. "Defaults and Returns on High Yield Bonds: Analysis Through 1998 and Default Outlook for 1999-2001." *Working Paper Series S-99-10.* New York University Salomon Center, 1999.

Altman, E.I., Hukkawala, N. and Kishore, V.M. "Defaults and Returns on High Yield Bonds: Analysis Through 1999 and Default Outlook for 2000-2002." *Working Paper Series S-00-10.* New York University Salomon Center, 2000.

Bartlett, M.S. "Further Aspects of the Theory of Multiple Regression." *Proceedings of Cambridge Philosophical Society* 34 (1938): 33-40.

Box, G.E.P. and Cox, D.R. "An Analysis of Transformations." *Journal of the Royal Statistical Society* B26 (1964): 211-252.

Box, G.E.P and Jenkins, G.M. *Time Series Analysis: Forecasting and Control.* San Francisco, CA: Holden-Day, 1976.

Caouette, J.B., Altman, E.I. and Narayanan, P. *Managing Credit Risk: The Next Great Financial Challenge.* New York, N.Y.: John Wiley & Sons, 1998.

Chan, W.S. "Outlier Analysis of Annual Retail Price Inflation: A Cross-Country Study." *Journal of Actuarial Practice* 6 (1998): 149-172.

De Jong, P. and Penzer, J. "Diagnosing Shocks in Time Series." *Journal of the American Statistical Association* 93 (1998): 796–806.

Dickey, D.A. and Fuller, W.A. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." *Econometrica* 49 (1981): 1057–1072.

Duffie, D. and Huang, M. "Swap Rates and Credit Quality." *Journal of Finance* 51 (1996): 921–950.

Duffie, D. and Singleton, K. "Simulating Correlated Defaults." *Working Paper.* Graduate School of Business, Stanford University, 1998.

Engle, R.F. and Granger, C.W.J. "Co-integration and Error Correction: Representation, Estimation and Testing." *Econometrica,* 55 (1987): 251–276.

Engle, R.F. and Granger, C.W.J. *Long-Run Economic Relationships: Readings in Cointegration.* Oxford, England: Oxford University Press, 1991.

Frees, E., Kung, Y.C., Rosenberg, M., Young, V. and Lai, S.W. "Forecasting Social Security Assumptions." *North American Actuarial Journal* 1 (1997): 49–82.

Hardy, M.R. "Stochastic Simulation in Life Office Solvency Assessment." *Journal of the Institute of Actuaries* 120 (1993): 131–151.

Jarrow, R. "Current Advances in the Modelling of Credit Risk." *Derivative: Tax, Regulation, Finance* May/June (1998): 1-12.

Jarrow, R. "The Intersection of Market and Credit Risk." *Journal of Banking and Finance* 12 (2000): 271–299.

Knox, D.M. "A Critique of Defined Contribution Plans Using a Simulation Approach." *Journal of Actuarial Practice* 1 (1993): 49–68.

Liesner, T. *One Hundred Years of Economic Statistics.* New York, N.Y.: Facts on File, 1989.

Liu, L.M. and Hudak, G.B. *Forecasting and Time Series Analysis Using the SCA Statistical System.* Chicago, Ill.: Scientific Computing Associates, 1994.

Lütkepohl, H. *Introduction to Multiple Time Series Analysis.* Berlin, Germany: Springer-Verlag, 1993.

MacKinnon, J.G. "Critical Values of Cointegration Tests." In *Long-Run Economic Relationships: Readings in Cointegration,* Engle, R.F. and Granger, C.W.J. (Editors). Oxford, England: Oxford University Press, 1991.

McLeod, A.I. and Li, W.K. "Diagnostic Checking ARMA Time Series Models using Squared-Residual Autocorrelations." *Journal of Time Series Analysis* 1983: 269–273.

Rao, B.B. *Cointegration for the Applied Economist.* New York, N.Y.: St. Martin's Press, 1994.

Reinsel, G.C. *Elements of Multivariate Time Series Analysis.* New York, N.Y.: Springer-Verlag, 1997.

Tiao, G.C. and Box, G.E.P. "Modeling Multiple Time Series with Applications." *Journal of the American Statistical Association* 76 (1981): 802–816.

Tiao, G.C. and Tsay, R.S. "Multiple Time Series Modeling and Extended Sample Cross-Correlations." *Journal of Business and Economic Statistics* 1 (1983): 43–56.

Vanderhoof, I.T., Albert, F., Tenenbein, A. and Verni, R. "The Risk of Asset Default — Report of the Society of Actuaries C-1 Risk Task Force of the Committee on Valuation and Related Areas." *Transactions of the Society of Actuaries* 46 (1989):547–591.

Wilkie, A.D. "More on a Stochastic Investment Model for Actuarial Use." *British Actuarial Journal* 1 (1995): 777–964.

Zurcher, M.L. "C-1 Bond Risk Analysis." *ARCH* 3 (1993): 627–639.

Independent Claim Report Lags and Bias in Forecasts Using Age-to-Age Factor Methodology

Stewart Gleason*

Abstract

This paper finds that when report lags are assumed to be independent, the age-to-age factor method produces biased estimates when applied to claim count development data. Two distributions are considered as models for the ultimate number of claims for an accident period: (i) a Poisson distribution, and (ii) a negative binomial distribution. In the Poisson case, the assumption of independent report lags implies the independence of the total number of claims reported in any two periods. In the negative binomial case, however, assuming that report lags are independent does not imply that increments are independent, and a somewhat different argument is required. Finally, it is proved that weighted average forecasts exhibit a smaller bias than do straight average estimates.

Key words and phrases: *loss development, Poisson, negative binomial, report lag*

1 Introduction

Stanard (1985) observes an apparent bias in forecasts of ultimate claims when commonly used reserving methods are applied to simulated data. His approach is to specify a stochastic model of the emergence of claims over time and use it to generate data to be used as input

*Stewart Gleason, Ph.D., A.C.A.S., A.S.A., is an actuary with Guy Carpenter & Company, Inc. He received his Ph.D. in mathematics (harmonic analysis) from the Courant Institute at New York University. Dr. Gleason's background is in reinsurance and professional liability with responsibilities including both pricing and loss reserving. His current research interests include simulation modeling and loss variation analysis. Prior to joining Guy Carpenter & Company, Inc., he was a consultant in the actuarial services group of Ernst & Young LLP. He has also been employed with the Prudential Reinsurance Company, CNA Insurance Companies, and Odyssey Reinsurance Corporation.

Dr. Gleason's address is: Guy Carpenter & Company, Inc., Two World Trade Center, New York, NY 10048, USA. Internet address: *stewart.gleason@guycarp.com*

to various reserving methods. One of the methods he selects is the familiar age-to-age factor method—he finds that the method produces overstated forecasts of ultimate claims in certain cases.

Stanard's simulation model assumes that the report lag of each claim is independent. This assumption has been presented in other work, particularly that of Weissner (1978, 1981). As I will prove, however, when report lags are assumed to be independent, the age-to-age factor method is biased when applied to claim count development data. Two models are considered for the ultimate number of claims in an accident period: (i) the Poisson distribution and (ii) the negative binomial distribution.

In the Poisson model, the assumption of independent report lags implies the independence of the total number of claims reported in any two periods and provides an example of an emergence process with independent increments. A general argument may be made to show that the age-to-age factor methodology gives biased results when the underlying process is known to have independent development increments. In the negative binomial model, which is the model specified by Stanard, assuming that report lags are independent does not imply that increments are independent, and a somewhat different argument is required.

The arguments presented here will use Jensen's inequality. Stanard notes in Appendix A of his paper that the observed bias is likely due to the fact that the expected value of the ratio of two non-constant random variables is not necessarily equal to the ratio of their expected values, i.e., in general

$$\frac{E[X]}{E[Y]} \neq E[\frac{X}{Y}].$$

Jensen's inequality may be used to show that, under certain conditions,

$$E\left[\frac{X}{Y}\right] > \frac{E[X]}{E[Y]}.$$

These ratios will arise as the usual claims development or age-to-age factors.

Finally, I will prove that weighted average forecasts exhibit a smaller bias than straight average estimates.

2 Preliminaries

2.1 Notation and Assumptions

For simplicity, the claims activity is divided into n consecutive and disjoint time periods of equal length, such as weeks, months, quarters, years, etc. For $i, j = 1, 2, \ldots$, let $X_{i,j}$ denote the number of incidents occurring in period i that are reported as claims in period $i + j - 1$ (i.e., with lag $j - 1$). The incremental development triangle at the end of the nth period is displayed in Table 1.

Table 1
Number of Accident Period i Claims Reported with Lag $j - 1$

	j						
i	1	2	...	n-$i+1$	...	$n-1$	n
1	$X_{1,1}$	$X_{1,2}$	...	$X_{1,n-i+1}$	...	$X_{1,n-1}$	$X_{1,n}$
2	$X_{2,1}$	$X_{2,2}$	...	$X_{2,n-i+1}$	...	$X_{2,n-1}$	
$\vdots$	$\vdots$	$\vdots$	...	...			
i	$X_{i,1}$	$\vdots$	$\vdots$	$X_{i,n-i+1}$			
$\vdots$	$\vdots$	$\vdots$	...				
$n-1$	$X_{n-1,1}$	$X_{n-1,2}$					
n	$X_{n,1}$						

These data are more commonly summarized as a cumulative development triangle (as shown in Table 2), where

$$S_{i,j} = \sum_{k=1}^{j} X_{i,k}.$$

The assumptions, however, will be stated in terms of the $X_{i,j}$s.

The basic problem for data given in this format is to estimate the total number of claims arising from each accident period from the number reported through the end of period n and from the claim reporting pattern. It is sufficient for our purposes to consider only the problem of forecasting the next reporting increment.

then it follows that

$$\Omega_{i,n-i+2} > E[X_{i,n-i+2} \mid S_{i,n-i+1} > 0] \equiv E[X_{i,n-i+2}].$$

Proof: Observe that, due to the independence of accident periods,

$$\begin{aligned}
\Omega_{i,n-i+2} &= E\left[S_{i,n-i+1}\frac{\sum_{k=1}^{i-1} X_{k,n-i+2}}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0,\, S_{i,n-i+1} > 0\right] \\
&= E\left[S_{i,n-i+1} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0,\, S_{i,n-i+1} > 0\right] \\
&\quad \times E\left[\frac{\sum_{k=1}^{i-1} X_{k,n-i+2}}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0,\, S_{i,n-i+1} > 0\right] \\
&= E\left[S_{i,n-i+1} \mid S_{i,n-i+1} > 0\right] \\
&\quad \times E\left[\frac{\sum_{k=1}^{i-1} X_{k,n-i+2}}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right].
\end{aligned} \tag{4}$$

Because of the independence of increments, it is also true that

$$\begin{aligned}
&E\left[\frac{\sum_{k=1}^{i-1} X_{k,n-i+2}}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right] \\
&\qquad = (i-1)E[X_{k,n-i+2}]\,E\left[\frac{1}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right].
\end{aligned} \tag{5}$$

Using Jensen's inequality, with $g(\mathbf{x}) = x_1 + \cdots + x_{i-1}$, one deduces that

$$\begin{aligned}
&E\left[\frac{1}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right] \\
&\qquad > \frac{1}{E\left[\sum_{k=1}^{i-1} S_{k,n-i+1} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right]}
\end{aligned}$$

and this inequality may be strengthened by noting that

$$
\begin{aligned}
E&\left[\sum_{k=1}^{i-1} S_{k,n-i+1} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right] \\
&\leq E\left[\sum_{k=1}^{i-1} S_{k,n-i+1} \mid S_{k,n-i+1} > 0, k = 1, 2, \ldots, i-1\right] \\
&= (i-1)E\left[S_{i,n-i+1} \mid S_{i,n-i+1} > 0\right].
\end{aligned}
$$

Thus

$$
\begin{aligned}
E&\left[\frac{1}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right] \\
&> \frac{1}{(i-1)E\left[S_{i,n-i+1} \mid S_{i,n-i+1} > 0\right]}. \qquad (6)
\end{aligned}
$$

Substituting equations (5) and (6) into (4) completes the proof. □

Readers will observe at this point that Theorem 1 is a statement of fact regarding ratios of independent random variables; it does not rely on the specific distribution of the underlying process. This should not be surprising because the age-to-age factor methodology also does not rely on the specific distribution of the underlying process. Intuition is the main guide in the construction of forecasts relying on identical distributions by lag. The conclusion is not that the age-to-age factor method is biased absolutely, but that it is not compatible with a claims process assumed to have independent increments.

3.2 Independent Increments from Independent Claims Lags: The Poisson Case

I will now prove that when the report lags are independent and the distribution of ultimate accident period claims is Poisson with constant mean λ then assumption (2) holds. The proof relies on two well-known properties of Poisson processes: the number of claims reported with lag $j-1$ is Poisson with mean λp_j,[1] where p_j is the probability that a claim from accident period i is reported in period $i+j-1$. In addition, the number of claims reported with lag $j-1$ and with lag $k-1$ are also independent.[2] Formally, this may be stated as:

[1]See, for example, Karlin and Taylor (1994, Chapter 5, Theorem 5.2, page 243).

In either case, the mean of M is,

$$E[M] = \frac{\alpha}{\beta}. \tag{9}$$

Proposition 5. *When N (the distribution of ultimate claims) is negative binomial with parameters $\alpha > 0$ and $\beta > 0$ and the report lags are independent, then N_j (the number of claims reported with lag $j - 1$) is negative binomial with parameters α and β_j where $\beta_j = \beta/p_j$.*

Proof: Again, $[N_j = k \mid N = n]$ has a binomial distribution when the report lags are independent and $n \geq k$.

$$\begin{aligned}
\Pr[N_j = k] &= \sum_{n=0}^{\infty} \Pr[N_j = k \mid N = n] \Pr[N = n] \\
&= \sum_{n=k}^{\infty} \frac{n!}{k!(n-k)!} p_j^k (1 - p_j)^{n-k} \frac{\Gamma(\alpha + n)}{n!\Gamma(\alpha)} \left(\frac{\beta}{1+\beta}\right)^{\alpha} \left(\frac{1}{1+\beta}\right)^{n}.
\end{aligned}$$

Changing the summation variable to $r = n - k$ produces

$$\begin{aligned}
\Pr[N_j = k] &= \frac{(p_j)^k}{k!\Gamma(\alpha)} \left(\frac{\beta}{1+\beta}\right)^{\alpha} \left(\frac{1}{1+\beta}\right)^{k} \sum_{r=0}^{\infty} \frac{\Gamma(\alpha + k + r)}{r!} \left(\frac{1 - p_j}{1+\beta}\right)^{r} \\
&= \frac{\Gamma(\alpha + k)}{k!\Gamma(\alpha)} \left(\frac{\beta}{1+\beta}\right)^{\alpha} \left(\frac{p_j}{1+\beta}\right)^{k} \left(1 - \frac{1 - p_j}{1+\beta}\right)^{-(\alpha+k)} \\
&\quad \times \sum_{r=0}^{\infty} \frac{\Gamma(\alpha + k + r)}{r!\Gamma(\alpha + k)} \left(\frac{1 - p_j}{1+\beta}\right)^{r} \left(1 - \frac{1 - p_j}{1+\beta}\right)^{\alpha+k} \\
&= \frac{\Gamma(\alpha + k)}{k!\Gamma(\alpha)} \left(\frac{\beta}{1+\beta}\right)^{\alpha} \left(\frac{p_j}{1+\beta}\right)^{k} \left(1 - \frac{1 - p_j}{1+\beta}\right)^{-(\alpha+k)} \\
&= \frac{\Gamma(\alpha + k)}{k!\Gamma(\alpha)} \left(\frac{\beta}{p_j + \beta}\right)^{\alpha} \left(\frac{p_j}{p_j + \beta}\right)^{k}
\end{aligned}$$

and the proposition is proved. □

Proposition 6. *When the distribution of ultimate claims, N, is negative binomial with parameters $\alpha > 0$ and $\beta > 0$ and the report lags are independent, $[N_j \mid N_k = s]$ has a negative binomial with parameters $\alpha_{jk} = \alpha + s$ and $\beta_{jk} = (\beta + p_k)/p_j$ provided $p_j > 0$. In particular,*

$$E[N_j \mid N_k] = \frac{(\alpha + N_k)p_j}{\beta + p_k}.$$

Proof: Again, $[N_j = k \mid N = n]$ is binomial when the report lags are independent and $n \geq k$.

$$\Pr[N_j = r \mid N_k = s] = \frac{\Pr[N_j = r, N_k = s]}{\Pr[N_k = s]}$$
$$= \frac{\sum_{n=0}^{\infty} \Pr[N_j = r, N_k = s \mid N = n] \Pr[N = n]}{\sum_{n=0}^{\infty} \Pr[N_k = s \mid N = n] \Pr[N = n]}.$$

Again, $[N_j = r, N_k = s \mid N = n]$ is multinomial, so the numerator may be rewritten as

$$\sum_{n=0}^{\infty} \Pr[N_j = r, N_k = s \mid N = n] \Pr[N = n]$$
$$= \sum_{n=r+s}^{\infty} \left[\frac{n!}{r!s!(n-r-s)!} p_j^r p_k^s (1 - p_j - p_k)^{n-r-s} \times \frac{\Gamma(\alpha + n)}{n!\Gamma(\alpha)} \left(\frac{\beta}{1+\beta}\right)^{\alpha} \left(\frac{1}{1+\beta}\right)^{n} \right]$$

and the denominator may be rewritten as

$$\sum_{n=0}^{\infty} \Pr[N_k = s \mid N = n] \Pr[N = n]$$
$$= \sum_{n=s}^{\infty} \frac{n!}{s!(n-s)!} p_k^s (1 - p_k)^{n-s} \frac{\Gamma(\alpha + n)}{n!\Gamma(\alpha)} \left(\frac{\beta}{1+\beta}\right)^{\alpha} \left(\frac{1}{1+\beta}\right)^{n}.$$

The numerator and denominator can be summed separately and reduced to give

$$\Pr[N_j = r \mid N_k = s] = \frac{\Gamma(\alpha_{jk} + r)}{r!\Gamma(\alpha_{jk})} \left(\frac{\beta_{jk}}{1+\beta_{jk}}\right)^{\alpha} \left(\frac{1}{1+\beta_{jk}}\right)^{r}$$

thus proving the proposition. □

One implication of Proposition 3 is that the increments are no longer independent. The following fact is also required:

Proposition 7. *When the distribution of ultimate claims, N, is negative binomial with parameters $\alpha > 0$ and $\beta > 0$ and the report lags are independent, then for $p_j > 0$ and $p_k > 0$,*

$$E[N_j \mid N_k > 0] = \frac{\alpha p_j}{\beta} \left(\frac{1 - \left(\frac{\beta}{\beta + p_k}\right)^{\alpha+1}}{1 - \left(\frac{\beta}{\beta + p_k}\right)^{\alpha}} \right). \tag{10}$$

Proof: Proceeding in a now familiar fashion but using the convenient, alternative form of the negative binomial probabilities, one sees that:

$$\Pr\left[N_j = r \mid N_k > 0\right] = \frac{\sum_{s=1}^{\infty} \sum_{n=0}^{\infty} \Pr\left[N_j = r, N_k = s \mid N = n\right] \Pr\left[N = n\right]}{\Pr\left[N_k > 0\right]}.$$

The denominator on the right side of this equation is

$$\Pr\left[N_k > 0\right] = 1 - \left(\frac{\beta}{\beta + p_k}\right)^{\alpha}.$$

The numerator on the right side of this equation is

$$
\begin{aligned}
&\sum_{s=1}^{\infty}\sum_{n=0}^{\infty}\Pr\left[N_j = r, N_k = s \mid N = n\right]\Pr\left[N = n\right] \\
&= \sum_{s=1}^{\infty}\sum_{n=r+s}^{\infty}\frac{n!}{r!s!\left(n-r-s\right)!}p_j^r p_k^s\left(1-p_j-p_k\right)^{n-r-s} \\
&\quad\times\frac{\beta^\alpha}{\Gamma\left(\alpha\right)}\int_0^\infty\frac{e^{-\lambda}\lambda^n}{n!}\lambda^{\alpha-1}e^{-\beta\lambda}d\lambda \\
&= \frac{\beta^\alpha p_j^r}{\Gamma\left(\alpha\right)r!}\int_0^\infty\left\{\sum_{s=1}^{\infty}\frac{\left(\lambda p_k\right)^s}{s!}\sum_{n=0}^{\infty}\frac{\left[\left(1-p_j-p_k\right)\lambda\right]^n}{n!}\right\} \\
&\quad\times\lambda^{\alpha+r-1}e^{-(\beta+1)\lambda}d\lambda \\
&= \frac{\beta^\alpha p_j^r}{\Gamma\left(\alpha\right)r!}\left\{\int_0^\infty\lambda^{\alpha+r-1}e^{-(\beta+p_j)\lambda}d\lambda\right. \\
&\quad\left.-\int_0^\infty\lambda^{\alpha+r-1}e^{-(\beta+p_j+p_k)\lambda}d\lambda\right\} \\
&= \frac{\Gamma\left(\alpha+r\right)}{\Gamma\left(\alpha\right)r!}\beta^\alpha p_j^r\left\{\frac{1}{\left(\beta+p_j\right)^{\alpha+r}}-\frac{1}{\left(\beta+p_j+p_k\right)^{\alpha+r}}\right\}.
\end{aligned}
$$

It follows that

$$
\Pr\left[N_j = r \mid N_k > 0\right] = \frac{\dfrac{\Gamma\left(\alpha+r\right)}{\Gamma\left(\alpha\right)r!}p_j^r\left\{\dfrac{1}{\left(\beta+p_j\right)^{\alpha+r}}-\dfrac{1}{\left(\beta+p_j+p_k\right)^{\alpha+r}}\right\}}{\dfrac{1}{\beta^\alpha}-\dfrac{1}{\left(\beta+p_k\right)^\alpha}}.
$$

It is straightforward to sum this expression to obtain the result:

$$\begin{aligned}
E\left[N_j \mid N_k > 0\right] &= \sum_{r=1}^{\infty} r \Pr\left[N_j = r \mid N_k > 0\right] \\
&= \sum_{r=1}^{\infty} \frac{\Gamma(\alpha + r)}{\Gamma(\alpha)(r-1)!} p_j^r \\
&\quad \times \frac{\left\{ \dfrac{1}{\left(\beta + p_j\right)^{\alpha+r}} - \dfrac{1}{\left(\beta + p_j + p_k\right)^{\alpha+r}} \right\}}{\dfrac{1}{\beta^\alpha} - \dfrac{1}{(\beta + p_k)^\alpha}} \\
&= \frac{\alpha p_j}{\dfrac{1}{\beta^\alpha} - \dfrac{1}{(\beta + p_k)^\alpha}} \sum_{r=0}^{\infty} \frac{\Gamma(\alpha + 1 + r)}{\Gamma(\alpha + 1)\, r!} p_j^r \\
&\quad \times \left\{ \frac{1}{\left(\beta + p_j\right)^{\alpha+1+r}} - \frac{1}{\left(\beta + p_j + p_k\right)^{\alpha+1+r}} \right\}
\end{aligned}$$

which is a difference of two negative binomial forms. This may be simplified as

$$\begin{aligned}
E\left[N_j \mid N_k > 0\right] &= \alpha p_j \left\{ \frac{\dfrac{1}{\beta^{\alpha+1}} - \dfrac{1}{(\beta + p_k)^{\alpha+1}}}{\dfrac{1}{\beta^\alpha} - \dfrac{1}{(\beta + p_k)^\alpha}} \right\} \\
&= \frac{\alpha p_j}{\beta} \left\{ \frac{1 - \left(\dfrac{\beta}{\beta + p_k}\right)^{\alpha+1}}{1 - \left(\dfrac{\beta}{\beta + p_k}\right)^{\alpha}} \right\}
\end{aligned}$$

and the proposition is proved. □

The key task may now be addressed, that is a theorem without the restrictions of Assumption 2.

Theorem 2. *When Assumption 1 holds and the distribution of ultimate claims is negative binomial, the expected value of the weighted average forecast is always greater than the expected value of the actual change, i.e.,*

$$\Omega_{i,n-i+2} > E[X_{i,n-i+2} \mid S_{i,n-i+1} > 0]$$

where $\Omega_{i,n-i+2}$ *is defined in equation (3).*

Proof: As in the proof of Theorem 1,

$$\begin{aligned}\Omega_{i,n-i+2} &= E\left[S_{i,n-i+1} \mid S_{i,n-i+1} > 0\right] \\ &\quad \times E\left[\frac{\sum_{k=1}^{i-1} X_{k,n-i+2}}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right]\end{aligned}$$

due to the independence of accident periods. In the proof of Theorem 1, it was possible to separate the expectation operator containing the ratio. As has been shown, however, independence of increments does not hold here—some other mechanism must be employed. To this end, one fixes the $S_{k,n-i+1}$ and computes the expectation in successive steps. But

$$\begin{aligned}E&\left[\frac{\sum_{k=1}^{i-1} X_{k,n-i+2}}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right] \\ &= E\left[\frac{E[\sum_{k=1}^{i-1} X_{k,n-i+2} \mid S_{1,n-i+1}, \ldots, S_{i-1,n-i+1}]}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right] \\ &= E\left[\frac{\sum_{k=1}^{i-1} E[X_{k,n-i+2} \mid S_{k,n-i+1}]}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right].\end{aligned}$$

Proposition 3 implies that this expression is equal to

$$\begin{aligned}&\frac{p_{n-i+2}}{\beta + \pi_{n-i+1}} E\left[\frac{\sum_{k=1}^{i-1} (S_{k,n-i+1} + \alpha)}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right] \\ &= \frac{p_{n-i+2}}{\beta + \pi_{n-i+1}}\left(1 + (i-1)\alpha E\left[\frac{1}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1} > 0\right]\right)\end{aligned}$$

where $\pi_k = p_1 + \ldots p_k$ for $k = 1, 2, \ldots$. Jensen's inequality implies that in turn

$$\frac{p_{n-i+2}}{\beta+\pi_{n-i+1}}\left(1+(i-1)\alpha E\left[\frac{1}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid \sum_{k=1}^{i-1} S_{k,n-i+1}>0\right]\right)$$
$$>\frac{p_{n-i+2}}{\beta+\pi_{n-i+1}}\left(1+\frac{(i-1)\alpha}{E\left[\sum_{k=1}^{i-1} S_{k,n-i+1} \mid \sum_{k=1}^{i-1} S_{k,n-i+1}>0\right]}\right)$$
$$\geq \frac{p_{n-i+2}}{\beta+\pi_{n-i+1}}\left(1+\frac{\alpha}{E\left[\sum_{k=1}^{i-1} S_{k,n-i+1} \mid \sum_{k=1}^{i-1} S_{k,n-i+1}>0\right]}\right).$$

From Proposition 2, one can see that

$$E\left[\sum_{k=1}^{i-1} S_{k,n-i+1} \mid \sum_{k=1}^{i-1} S_{k,n-i+1}>0\right] \leq (i-1)E\left[S_{k,n-i+1} \mid S_{k,n-i+1}>0\right]$$
$$=\frac{(i-1)\alpha\pi_{n-i+1}}{\beta\left(1-\left(\frac{\beta}{\beta+\pi_{n-i+1}}\right)^{\alpha}\right)}.$$

Substituting these results into the expression for $\Omega_{i,n-i+2}$ yields

$$\Omega_{i,n-i+2} > \frac{\alpha p_{n-i+2}}{\beta(\beta+\pi_{n-i+1})}\left(\frac{(\beta+\pi_{n-i+1})^{\alpha+1}-\beta^{\alpha+1}}{(\beta+\pi_{n-i+1})^{\alpha}-\beta^{\alpha}}\right)$$
$$= E\left[X_{k,n-i+2} \mid S_{k,n-i+1}>0\right]$$

from equation (10) of Proposition 4. The theorem is thus proved. □

5 Average Factors: Straight versus Weighted

There has been much discussion in actuarial circles regarding the merits of weighted average development factors as opposed to straight (unweighted) average development factors. In this section, it will be demonstrated that the straight average estimator,

$$\bar{X}_{i,n-i+2}=\frac{1}{i-1}\sum_{k=1}^{i-1} S_{i,n-i+1}\frac{X_{k,n-i+2}}{S_{k,n-i+1}},$$

cannot reduce or eliminate the bias seen in the weighted average estimator. For brevity, attention is restricted to the case of independent increments.

Before stating and proving the final theorem, one more proposition is needed:

Proposition 8. *For any set of positive m numbers $\{y_1, y_2, \ldots, y_m\}$,*

$$\sum_{k=1}^{m} \frac{1}{y_k} \geq \frac{m^2}{\sum_{k=1}^{m} y_k}.$$

Proof: Let Y denote the discrete random variable such that

$$\Pr[Y = y_k] = \frac{1}{m}, \qquad \text{for } k = 1, \ldots, m.$$

Applying Jensen's inequality yields

$$E\left[\frac{1}{Y}\right] = \frac{1}{m}\sum_{k=1}^{m} \frac{1}{y_k} \geq \frac{1}{E[Y]} = \frac{m}{\sum_{k=1}^{m} y_k}$$

and the proposition is proved. □

Theorem 3. *When both expectations are defined, the expected value of the straight average prediction is greater than the expected value of the weighted average prediction. That is,*

$$\begin{aligned} &E[\bar{X}_{i,n-i+2} \mid S_{k,n-i+1} > 0,\ k = 1, \ldots, i] \\ &\qquad \geq E[\hat{X}_{i,n-i+2} \mid S_{k,n-i+1} > 0,\ k = 1, \ldots, i]. \qquad (11) \end{aligned}$$

Proof: Again making use of the independence and symmetry between the periods, proving equation (11) is equivalent to proving the following inequalities:

$$\begin{aligned} &\frac{1}{(i-1)} \sum_{k=1}^{i-1} E\left[S_{i,n-i+1} \frac{X_{k,n-i+2}}{S_{k,n-i+1}} \mid S_{k,n-i+1} > 0\right] \\ &\quad \geq E\left[S_{i,n-i+1} \frac{\sum_{k=1}^{i-1} X_{k,n-i+2}}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid S_{1,n-i+1} > 0, \ldots, S_{i-1,n-i+1} > 0\right] \end{aligned}$$

or

$$E\left[\sum_{k=1}^{i-1}\left[\frac{1}{S_{k,n-i+1}} \mid S_{k,n-i+1} > 0\right]\right]$$
$$\geq (i-1)^2 E\left[\frac{1}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid S_{1,n-i+1} > 0, \ldots, S_{i-1,n-i+1} > 0\right]$$

or

$$E\left[\sum_{k=1}^{i-1}\frac{1}{S_{k,n-i+1}} - \frac{(i-1)^2}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \mid S_{1,n-i+1} > 0, \ldots, S_{i-1,n-i+1} > 0\right] \geq 0.$$

But as $S_{k,n-i+1} > 0$, $k = 1, \ldots, i-1$,

$$\sum_{k=1}^{i-1}\frac{1}{S_{k,n-i+1}} - \frac{(i-1)^2}{\sum_{k=1}^{i-1} S_{k,n-i+1}} \geq 0$$

from Proposition 5 and, therefore, is the expectation of this quantity. This proves the theorem. □

6 Beyond Claim Counts: Possible Extensions of the Results

To this point, only claim count development data have been considered. As actuaries are concerned also with the development of the amounts of paid and reported claims, it is natural to ask whether these results can be generalized to include the analysis of claim amounts.

The general case based on the assumption of independent increments may be adapted to include specific examples of paid claim development. Medical malpractice indemnity payments exhibit special behavior to which the work presented above may be applied. Such a claim typically is closed either with no payment or with a single payment of a final award or settlement. By defining the variable $X_{i,j}$ to be the total amount paid for claims closing in period j (not the number of claims reported in a given period), assumptions 1 and 2 of Section 3 will be met in this simplest of situations.

One requires that the closure lags of claims are independent and that, once determined, the amount of a claim cannot change in subsequent periods. In reality, there are subtle problems with this. One must tabulate amounts for incidents rather than claims: a medical incident may generate several claims whose closure lags are not only interdependent, they are the same. Settlements could be paid as periodic payments which means that the amounts are actually paid over many periods violating the independence assumption. In practice, annuities may be purchased at the time of closure to fund the payments and limit the payment to a single period.

Allocated loss adjustment expense (ALAE) payments for the same medical malpractice business do not exhibit these properties although they are often combined with indemnity for the purposes of development analysis. Payments of ALAE are made incrementally from the time of the claim report to the time of its closure. Partial payments related to the same claim or incident will appear in different periods. Hence, the increments $X_{i,j}$ cannot be expected to be independent when ALAE is included.

Although a bias argument similar to the negative binomial case might be constructed, the interdependence of the payment increments is more complicated than the claim count increments. An essential part of the negative binomial example is being able to specify the nature of the interdependence. For ALAE payments, it is not clear what the nature of this interdependence would be.

For reported loss amounts, a similar problem arises. In this case, the value of the claim may include not only numerous partial payments but also changing estimates of the unpaid portion of the claim. The case reserve estimates are included in order to stabilize the development and bring the initial value of the claim as close as possible to its final value. This introduces the possibility of negative increments and serves only to complicate their interdependence.

Extending the results to development of claim amounts is difficult. Perhaps the most promising approach would be to consider particular models presented by Stanard in his original paper where the claim amount structure is specified.

7 Closing Comments

It is not the purpose of this paper to advocate one set of assumptions regarding the independence of report lags over another. If one believes that expected development increments are directly proportional to the accumulated total claims at a given point in time, then one might conclude that methods based on independent increment assumptions produce understated results.

It is, however, apparent that Stanard's simulation test of the development method produces the correct observation. If one believes that individual report lags are independent, then the loss development methods will produce overstated results. One thing that the analytical work presented here does not show is the magnitude of the bias. Stanard's work produced measures of that in specific cases. The key point is that there is a fundamental incompatibility between loss development techniques and methods relying on independent report lags.

References

Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A., and Nesbitt, C.J. *Actuarial Mathematics,* Second Edition. Schaumburg, Ill.: Society of Actuaries, 1997.

Ross, S.M. *Introduction to Probability Models,* Fifth Edition. San Diego, Calif.: Academic Press, 1993.

Royden, H. *Real Analysis.* New York, N.Y.: Macmillan Publishing Co., 1968.

Stanard, J. "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques." *Proceedings of the Casualty Actuarial Society* 72 (1985): 124-148.

Taylor, H.M. and Karlin, S. *An Introduction to Stochastic Modeling.* San Diego, Calif.: Academic Press, 1994.

Weissner, E. "Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood." *Proceedings of the Casualty Actuarial Society* 65 (1978): 1-40.

Weissner, E. "Evaluation of IBNR on a Low Frequency Book Where the Report Development Pattern is Still Incomplete." In *Casualty Loss Reserve Seminar Transcripts.* Arlington, Va.: Casualty Actuarial Society, (1981).

Wheeden, R. and Zygmund, A. *Measure and Integral: An Introduction to Real Analysis.* New York, N.Y.: Marcel Dekker, Inc., 1977.