

Mortality Portfolio Risk Management

By

Samuel H. Cox, Yijia Lin, Ruilin Tian, and Luis F. Zuluaga

Submitted to the

2009 American Risk and Insurance Association Annual Meeting

Please address correspondence to:

Yijia Lin
Department of Finance
University of Nebraska
P.O. Box 880488
Lincoln, NE 68588 USA
Tel: (402) 472-0093
Email: ylin@unlnotes.unl.edu

Samuel H. Cox
University of Manitoba
Winnipeg, Manitoba
Tel: (204) 474-7426
scox@cc.umanitoba.ca

Yijia Lin
University of Nebraska
Lincoln, NE 68588 USA
Tel: (402) 472-0093
ylin@unlnotes.unl.edu

Ruilin Tian
North Dakota State University
Fargo, ND 58108
Tel: (701) 231-6544
ruilin.tian@ndsu.edu

Luis F. Zuluaga
University of New Brunswick
Fredericton, NB Canada
Tel: (506) 458-7956
lzuluaga@unb.ca

MORTALITY PORTFOLIO RISK MANAGEMENT

ABSTRACT

In this paper, we offer a new method of managing the risk of unexpected changes in mortality underlying annuities and life insurance. This method maximizes the insurer's profit margin, subject to constraints on its downside mortality risk. We also show how to determine bounds on mortality margins when information on the moments of the distributions is known. We provide numerical examples to illustrate how to apply these methods and to use them to take advantage of natural hedging effects of annuity and life insurance mortality.

1. INTRODUCTION

Life insurance companies sell a wide variety of life insurance and annuity products. The insurer's liability for these products depends on future interest and mortality rates, which explains why company managers and regulators focus on these risks. Unanticipated mortality improvement (as well as a decline in interest rates) can be very serious. For example, unanticipated mortality improvement was a factor in the failure of Equitable Life, once a highly-regarded United Kingdom life insurer (see Ombudsman (2008)).

During recent years, economic and policy changes have made mortality projection and risk management more important than ever. On the one hand, pension plans and annuity providers underestimated life expectancy for ages 60 and older in the past decade (Cowling and Dales, 2008). Cowling and Dales (2008) find that companies in the UK's FTSE100 index underestimated their aggregate pension liabilities by more than £40 billion. On the other hand, the recent findings of genetic analysts spur fears of a worldwide epidemic by confirming that today's "bird flu" is similar to the 1918 "Spanish flu" which killed more than 40 million people (Juckett, 2006). According to Toole (2007), losses due to a severe pandemic could amount to 25% of the US life insurance industry's statutory capital. While the great majority of US life insurance companies would weather such a pandemic, it is clear that these companies should be interested in mitigating the risk.

We describe a method life insurance companies can use to adjust their portfolio of life insurance and annuities to better manage mortality extreme outcomes while maintaining a relatively efficient risk-return relationship. This method is called the “mean-variance with conditional value at risk constraint”, denoted “MV+CVaR”. The MV+CVaR approach combines Markowitz portfolio theory and conditional value-at-risk (CVaR) by optimizing the tradeoff between mean and variance subject to a lower bound on CVaR. Although the MV+CVaR portfolios are suboptimal relative to Markowitz counterparts in terms of the mean-variance efficiency, they have lower downside risk. Therefore this approach may be more appealing to insurance companies that are required to meet various solvency requirements. Our approach is empirical with all calculations based on the company’s mortality experience.

Once a mortality portfolio has been determined, the corresponding empirical returns can be used to estimate moments of the returns. We show how such moment information can be used to determine semi-parametric upper and lower bounds on the actual underlying portfolio return. Knowledge about the underlying mortality distribution may be limited. The literature on mortality forecasts uses a variety of stochastic processes to model mortality or longevity risk. The semi-parametric approach is based solely on moments, not a distributional assumption. This method gives bounds on portfolio return that will be satisfied by any distribution with the same moments. This method provides a mechanism for checking the downside risk of a MV+CVaR efficient mortality portfolio valid for any distribution with the same moments.

The remainder of this paper is organized as follows. The next section describes how to calculate profit margins of mortality portfolios. Section 3 discusses the MV+CVaR mortality optimization model. A numerical example illustrates how to implement the approach. Section 4 describes the semi-parametric bounds method. We show how to compute the semiparametric upper and lower bounds for MV+CVaR efficient portfolios and then perform the bound analysis on those portfolios. In Section 5, we study the natural hedging effect by adding an annuity to a portfolio with only life insurance. Section 6 extends the analysis to different MV+CVaR efficient mortality portfolios. Section 7 is our conclusion.

2. MORTALITY RISK PORTFOLIOS

Consider an insurer selling n lines of business at times $t = 0, \dots, T$. For simplicity, we assume, for each line of business i , the insurer collects a single premium at time t and pays a death or survival benefit at the end of the year.¹ These business lines include various life insurance and annuity products. The insurer calculates the present value of expected life insurance or annuity payments, so called *net premium*, based on its forecasted mortality rates.

Consider a contract based on a life (x) at time t . The company has forecasts of annual death and survival rates for (x) in future years, $q_{x+s,t+s}$ and $p_{x+s,t+s}$, for $s = 0, 1, \dots$. The j -year survival rate for (x) determined at t is ${}_j p_{x,t}$ where

$${}_j p_{x,t} = \begin{cases} 1 & \text{for } j = 0 \\ p_{x,t} \times p_{x+1,t+1} \times \dots \times p_{x+j-1,t+j-1} & \text{for } j \geq 1 \end{cases} \quad (1)$$

The net single premium for a k -year term life insurance on age (x) issued at time t is calculated as

$$A_{x:\overline{k}|,t}^1 = \sum_{j=0}^{k-1} v^{j+1} {}_j p_{x,t} q_{x+j,t+j} \quad (2)$$

where v is the 1-year discount factor. The calculation of the whole life insurance net single premium $A_{x,t}$ is similar; just let $k \rightarrow \infty$.

The net single premium for an annual payment immediate life annuity written on (y) at time t is calculated as

$$a_{y,t} = \sum_{j=1}^{\infty} v^j {}_j p_{y,t}. \quad (3)$$

Actual mortality results are observed in the years following t . For example, after k years, the company will have observed the actual death rates $\tilde{q}_{x+s,t+s}$ for $s = 0, 1, \dots, j-1$. Define the present

¹Our techniques can also be applied to policies with monthly or annual premiums. It is only to save space that we consider single premium policies.

value of the actual death benefits or survival payments, $\tilde{L}_{i,t}$, as

$$\tilde{L}_{i,t} = \begin{cases} \tilde{A}_{x:\bar{k}|,t}^1 = \sum_{j=0}^{k-1} v^{j+1} ({}_j\tilde{p}_{x,t}) (\tilde{q}_{x+j,t+j}) & \text{for life insurance;} \\ \tilde{a}_{y,t} = \sum_{j=1}^{\infty} v^j ({}_j\tilde{p}_{y,t}) & \text{for life annuity,} \end{cases} \quad (4)$$

where ${}_j\tilde{p}_{x,t}$ (or ${}_j\tilde{p}_{y,t}$) represents the actual j -year survival rate for the age x (or the age y) at time t , and $\tilde{q}_{x+j,t+j}$ is the actual one-year death rate for the age $x+j$ during year $t+j$.

For the business line i written in year t , in principle, we have $E[\tilde{L}_{i,t}] = P_{i,t}$, where the net single premium

$$P_{i,t} = \begin{cases} A_{x:\bar{k}|,t}^1 = \sum_{j=0}^{k-1} v^{j+1} ({}_j p_{x,t}) (q_{x+j,t+j}) & \text{for life insurance;} \\ a_{y,t} = \sum_{j=1}^{\infty} v^j ({}_j p_{y,t}) & \text{for life annuity.} \end{cases} \quad (5)$$

However, in most cases the realized $\tilde{L}_{i,t}$, differs from the expected amount $P_{i,t}$, the lump-sum premium charge by the insurance company at t when the policy i was written. To illustrate, we define

$$\tilde{r}_{i,t} = \frac{P_{i,t}}{\tilde{L}_{i,t}} - 1. \quad (6)$$

If $\tilde{r}_{i,t} > 0$, it indicates that the insurer pays a benefit lower than its expectation, and vice versa. We call $\tilde{r}_{i,t}$ the “pure margin” for line i , which is known at the expiration of the policy but is random at time t . Denote the weight vector of n lines of business as $w = [w_1, w_2, \dots, w_n]$. The total pure margin of the insurer across n lines of business written at time t equals

$$\tilde{r}_t(w) = \sum_{i=1}^n w_i \tilde{r}_{i,t}. \quad (7)$$

In general, the insurer charges a risk premium over and above the expected mortality or longevity payments. The risk premium covers the insurer’s costs of bankruptcy and information asymmetry associated with the assumed risk (Jean-Baptiste and Santomero, 2000). Assume the insurer charges a risk premium π_i on line i , which is a convex function of the random pure margin $\tilde{r}_{i,t}$,

$$\pi_i = a + \frac{c}{2} E[\tilde{r}_{i,t}^2], \quad (8)$$

where a is a positive constant. The random margin $\tilde{r}_{i,t}$ acts as a proxy for the probability of financial distress caused by line i , and the constant $c/2$ is the distress cost factor with $c > 0$. That is, all else equal, a higher business risk is associated with a higher cost of financial distress. As such, the insurer charges a higher risk premium.

Consequently, the profit margin $\tilde{m}_{i,t}$ of line i is the sum of the pure margin in equations (6) and the risk premium in (8),

$$\tilde{m}_{i,t} = \pi_i + \tilde{r}_{i,t} = a + \frac{c}{2} \mathbf{E} [\tilde{r}_{i,t}^2] + \tilde{r}_{i,t}. \quad (9)$$

Then it is straightforward to obtain the total profit margin $\tilde{m}_t(w)$ across n lines of business written at time t ,

$$\tilde{m}_t(w) = \pi(w) + \tilde{r}_t(w) = a + \frac{c}{2} \sum_{i=1}^n w_i \mathbf{E} [\tilde{r}_{i,t}^2] + \sum_{i=1}^n w_i \tilde{r}_{i,t}. \quad (10)$$

3. MORTALITY PORTFOLIO OPTIMIZATION

Markowitz portfolio optimization (Markowitz, 1952) originally applied to investors who are holding a portfolio and seeking to maximize expected return for a given level of risk, measured by the variance of return. More recently portfolio optimization techniques have been applied to corporation management, helping find optimal business strategies to maximize business profits. We show how to apply this approach to mortality and/or longevity risks subject to capital constraints. Life insurers are motivated to search for optimal business compositions that could maximize profits and minimize downside risks. In this section, we propose the MV+CVaR approach to achieve this goal.

3.1. Portfolio Optimization with CVaR Constraints. As a measure of risk, conditional value-at-risk (CVaR) is defined as the expected loss/return exceeding a given value-at-risk (VaR). There are several applications of the CVaR methodology to portfolio optimizations. For example, Rockafellar and Uryasev (2000) consider minimizing CVaR, while requiring a minimum expected return. They generate an efficient CVaR_β -mean frontier by considering different expected returns. Krokmal et al. (2002) suggest minimizing the negative expected return subject to a CVaR constraint. However, the existing literature does not explicitly consider the tradeoff between mean

and variance subject to CVaR constraints. This paper fills the gap by adding one or more CVaR constraints to the traditional Markowitz problem. We call it the MV+CVaR approach.

Let σ_{ij} the covariance of margins of business lines i and j . Our MV+CVaR problem is to solve for portfolio weights $w = [w_1, w_2, \dots, w_n]$ in terms of the margins, so as to

$$\begin{aligned} \text{Minimize} \quad & \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \\ \text{subject to} \quad & \text{CVaR}_\beta(w) \geq \zeta \\ & w \in \mathbb{W}, \end{aligned} \tag{11}$$

where $\text{CVaR}_\beta(w)$ is the β -level CVaR of mortality portfolio margin calculated as

$$\text{CVaR}_\beta(w) = \text{E} [\tilde{m}(w) | \tilde{m}(w) \leq \text{VaR}_\beta(w)].$$

$\text{VaR}_\beta(w)$ is the β -level value at risk (VaR) that defines the minimal value such that the probability of portfolio margin not exceeding this value is β . Here, $\tilde{m}(w) = \sum_{i=1}^n w_i \tilde{m}_i$ is the random variable of portfolio margin. As the observed value of $\tilde{m}(w)$ in year t ($t = 1, \dots, s$), $\tilde{m}_t(w)$ equals $\sum_{i=1}^n w_i \tilde{m}_{i,t}$. We enforce a left-tail constraint to management downside risk by setting β at 0.05. The CVaR constraint in (11) ensures the tail expectation $\text{CVaR}_\beta(w)$ no lower than a pre-specified value ζ , thus reducing downside risk. \mathbb{W} is the subset of the mortality business composition feasible set. Specifically, any w in \mathbb{W} satisfies

$$\begin{aligned} \sum_{i=1}^n w_i \text{E}[\tilde{m}_i] &= m_0 \\ \sum_{i=1}^n w_i &= 1 \\ w_i &\geq 0, \quad i = 1, 2, \dots, n, \end{aligned} \tag{12}$$

where \tilde{m}_i is the random margin of line i and n is the number of business lines in the mortality portfolio and m_0 is a pre-specified level of profit margin. Assume no short selling for all lines, that is, $w_i \geq 0$ for each i .

Rockafellar and Uryasev (2000) show that the β -level CVaR can be realized as the solution to the problem

$$\max_{\alpha \in \mathbb{R}, w} G(\alpha, w) = \alpha - \frac{1}{\beta} \mathbb{E} [[\alpha - \tilde{m}(w)]^+]. \quad (13)$$

If a pair (α^*, w^*) achieves the maximization of (13), $G(\alpha^*, w^*)$ will return β -level CVaR and α^* will give the corresponding β -level VaR.

Theorem 1. *Problem (11) is equivalent to the following problem (14) in the sense that their objectives achieve the same minimum values.*

$$\begin{aligned} \min_{\alpha \in \mathbb{R}, w} \quad & \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \\ \text{subject to} \quad & \alpha - \frac{1}{\beta} \mathbb{E} [[\alpha - \tilde{m}(w)]^+] \geq \zeta \\ & w \in \mathbb{W}. \end{aligned} \quad (14)$$

The proof of Theorem 1 is provided in the Appendix A.

We can calculate $\mathbb{E} [[\alpha - \tilde{m}(w)]^+]$ by using the empirical values:

$$\mathbb{E} [[\alpha - \tilde{m}(w)]^+] = \frac{1}{s} \sum_{t=1}^s [\alpha - \tilde{m}_t(w)]^+,$$

where s is the number of observations. By using auxiliary variables v_t , $t = 1, \dots, s$, the constraint $\alpha - \frac{1}{\beta} \mathbb{E} [[\alpha - \tilde{m}(w)]^+] \geq \zeta$ can be realized as

$$\begin{aligned} \alpha - \frac{1}{\beta} \frac{1}{s} \sum_{t=1}^s v_t &\geq \zeta \\ v_t &\geq \alpha - \tilde{m}_t(w) \quad \forall t = 1, \dots, s \\ v_t &\geq 0 \quad \forall t = 1, \dots, s. \end{aligned} \quad (15)$$

Below we proceed with an example to illustrate our MV+CVaR approach.

3.2. Numerical Illustration. Assume an insurer sells three types of life insurance in year t : 5-year term life insurance ($i = 1$) on a male aged $x = 35$, 10-year term life insurance ($i = 2$) on male aged $x = 25$, and whole life insurance ($i = 3$) on a male aged $x = 40$. Further we assume (a) the

TABLE 1. Summary Statistics – Profit Margin of Three Types of Life Insurance

Lines	Mean	Variance	Skewness	Kurtosis
$\tilde{m}_{A_{35:\overline{5} }}^1$	0.0070	0.0078	0.3197	0.5024
$\tilde{m}_{A_{25:\overline{10} }}^1$	0.1114	0.0216	0.4906	-0.2144
$\tilde{m}_{A_{40}}$	0.0371	0.0003	0.5992	-0.8741

insurer applies the Renshaw et al. (1996) model to predict future mortality rates, with which it uses to determine the premium P_i ; and (b) the insurer has the same mortality experience as the US population, which determines its payment \tilde{L}_i .

Table 1 shows the descriptive statistics for the profit margins of these three lines of life insurance. Please refer to Appendix B for details on how we estimate those values. In particular, the 10-year term life insurance on male aged 25, $A_{25:\overline{10}|}^1$, has the highest expected profit margin and variance. The whole life insurance on male aged 40, A_{40} , has a higher expected margin but lower variance than the 5-year term life insurance on male aged 35, $A_{35:\overline{5}|}^1$. In practice, the insurer may not have the same profit margins as what we use in this paper. Nevertheless, the same techniques discussed below can be applied to a given insurer's situation.

Suppose originally these three types of life insurance generate the same amount of premiums. That is, 33.33% of total premium comes from the 5-year term life insurance, 33.33% from the 10-year term life insurance, and 33.33% from the whole life insurance ($w_0 = [33.33\%, 33.33\%, 33.33\%]$). beginning of year t equals

$$\tilde{m}_t(w_0) = \sum_{i=1}^3 \frac{1}{3} \tilde{m}_{i,t}, \quad (16)$$

with the summary statistics shown in the row called “Original” of Table 2.

TABLE 2. Summary Statistics of Profit Margins of Original and MV+CVaR 3-Line Portfolios

$\tilde{m}(w)$	Mean	Variance	Skewness	Kurtosis	Mode	CVaR _{5%}
Original	0.0518	0.0057	0.5292	0.1647	0.0264	-0.0749
MV+CVaR	0.0518	0.0008	0.2077	-0.5352	0.0696	0.0068

To search for an optimum business strategy, we solve the MV+CVaR optimization problem (14).

We specify

$$m_0 = \sum_{i=1}^3 w_i E[\tilde{m}_i] = 0.0518,$$

and

$$\zeta = \text{CVaR}_{0.05}(\tilde{m}(w_0)) + 0.05|\text{CVaR}_{0.05}(\tilde{m}(w_0))|, \quad (17)$$

where $\text{CVaR}_{0.05}(\tilde{m}(w_0))$ is the 5%-level CVaR of the original portfolio. Then we get the following weights for the MV+CVaR optimal business composition:

$$w_{\text{MV+CVaR}} = [0, 19.81\%, 80.19\%].$$

This means the insurance company should stop writing the 5-year term life insurance, put around 20% of its business on the 10-year life insurance, and underwrite the rest 80% on the whole life insurance. The row called ‘‘MV+CVaR’’ in Table 2 shows the summary statistics for the MV+CVaR efficient mortality portfolio. The MV+CVaR efficient portfolio significantly reduces the variance from 0.0057 to 0.0008. Notably, in Figure 1, the distribution of the MV+CVaR portfolio’s profit margin is tightly grouped around the mean relative to that of the original portfolios. Furthermore, Table 2 shows that, although the new portfolio’s skewness decreases a little bit, its 5%-CVaR increases dramatically from -0.0749 to 0.0068. In sum, we conclude the MV+CVaR portfolio has a better mean-variance tradeoff and a lower downside risk than the original one.

4. BOUNDS ANALYSIS WITH MOMENT METHODS

How well does the optimal business strategy suggested by the MV+CVaR approach in section 3 secure an insurer’s financial position? In this section, we apply the moment methods to address this question. Specifically, rather than managing only the first two moment of the portfolio with portfolio optimization techniques, we incorporate the higher moment information to the mortality risk management.

The moment problem was first studied by Tchebyshev, Markov and Stieltjes in the 1870’s. The classical Hausdorff moment problem, named after Felix Hausdorff, examines the necessary and

sufficient conditions that a given sequence $\mu_0, \mu_1, \mu_2, \dots$ be the sequence of moments of a probability distribution $f(z)$,

$$\int_{z=-\infty}^{+\infty} z^i f(z) dz = \int_{F(z)=0}^1 z^i dF(z) = \mu_i, \quad \forall i = 0, 1, \dots,$$

with the cumulative distribution function $F(z)$ constant outside the closed unit interval $[0, 1]$. So $dF(z)$ does not contribute any mass on the interval $\{(-\infty, 0) \cup (1, \infty)\}^2$. Since $\int_{z \in \mathbb{R}} f(z) dz = 1$ for all feasible distributions, $\mu_0 = 1$ in all moment problems. By setting $\phi(Z) = \mathbb{I}_{(-\infty, d]}(Z)$ as the indicator function for the event $Z \leq d$, $E[\phi(Z)] = \Pr(Z \leq d)$ produces the cumulative distribution function of Z . In our paper, Z is the profit margin $\tilde{m}(w)$ of a MV+CVaR efficient mortality portfolio. The solving the following problem yields the upper bound of $\Pr(Z \leq d)$:

$$\begin{aligned} \max \quad & E_F[\phi(Z)] \\ \text{subject to} \quad & E_F[Z^i] = \mu_i, \quad \forall i = 0, 1, \dots, n, \\ & F(z) \text{ is a probability distribution on } Q, \quad Q \subset \mathbb{R}. \end{aligned} \tag{18}$$

Similarly, the primal problem for the lower bound optimizes the following objective subject to the same constraints as (18):

$$\min \quad E_F[\phi(Z)]. \tag{19}$$

Notice that in the above upper and lower bounds problems, the only constraints given are the moment constraints. Therefore, these semiparametric bounds, called the ‘‘arbitrary bounds’’, are robust bounds that any feasible distribution with same moments must satisfy. If we set $d = \text{VaR}_\beta$, problems (18) and (19) are able to solve for a 100% confidence interval for the β -level VaR. This moment problem can be solved by some newly developed semidefinite programs such as SOSTOOLS (Prajna et al., 2004). In this paper, we will not go into details about the derivation of the dual problems and the application of the SOS programming solvers. Please refer to Parrillo (2000), Popescu (2005), Bertsimas and Popescu (2005) for details.

²Akhiezer (1965) notes: ‘‘two solutions are not taken to be distinct if their difference is a constant at all points at which it is continuous, and the moment problem is called determinate if it has a unique solution in this sense.’’

4.1. Unimodal Bounds. By adding some distribution assumption such as a unimodal distribution assumption to problems (18) and (19), we can narrow down the bounds on $\Pr(Z \leq d)$. A unimodal distribution is a distribution that has only one mode. As long as the distribution is unimodal, including this constraint can substantially narrow the range of the bounds. Many financial and insurance data have unimodal distributions, but some care must be exercised in applying this assumption. If the sample, is not in fact from a unimodal distribution, important tail behavior may be missed. To solve the moment problem with the unimodal constraint, which is called the “unimodal bounds” problem, we convert it to an arbitrary bounds problems (18) and (19) by applying Khintchine (1938)’s representation.³

The idea here is to transfer the unimodal bounds problem to its equivalent arbitrary bounds problem. Specifically, the objective function $\phi(Z)$ for the bounds on $\Pr(Z \leq d)$ is transferred to $\phi^*(Y)$ for the bounds on $\Pr(Y \leq d - m)$ ⁴. The unimodal bounds are obtained by solving the following problem:

$$\begin{aligned} \max(\text{ or min}) \quad & E_{F^*}[\phi^*(Y)] \\ \text{subject to} \quad & E_{F^*}[Y^i] = \mu_i^*, \quad \forall i = 0, 1, \dots, n, \\ & F^*(y) \text{ is a probability distribution on } Q^*, \quad Q^* \subset \mathbb{R}, \end{aligned} \tag{22}$$

where m is the mode of unimodal random variable $Z = m + UY$ in the Khintchine’s Representation. The moment of Y , μ_i^* , is calculated from the corresponding moments of Z as follows:

$$\mu_i^* = E[Y^i] = (i+1) \sum_{j=0}^i \binom{i}{j} E[Z^j] (-m)^{i-j}.$$

³Khintchine’s Representation: Z is unimodally distributed if and only if there are two independent random variables U and Y such that $Z = m + UY$, where U is uniformly distributed on $(0, 1)$ and m is the unique mode of the distribution.

⁴Considering the relationship between d and m , the function $\phi^*(Y)$ has the following two possible expressions:

(1) In this case of $d \geq m$,

$$\phi^*(Y) = \begin{cases} 1 & Y \leq d - m \\ \frac{d - m}{Y} & Y \geq d - m. \end{cases} \tag{20}$$

(2) In this case of $d < m$,

$$\phi^*(Y) = \begin{cases} 1 - \frac{d - m}{Y} & Y \leq d - m \\ 0 & Y \geq d - m. \end{cases} \tag{21}$$

The “max” in (22) finds the upper bound and the lower bound is obtained by applying “min”. For detail of the transferring procedure, please refer to Tian (2008)’s thesis.

4.2. Maximum-Entropy Distribution. Every probability distribution has some “uncertainty” associated with it. For example, an insurer may have some confidence in its estimates of moments of its margins, but not know exactly what the distribution is. The concept of “entropy” was introduced to provide a quantitative measure of this uncertainty. The maximum-entropy method has its theoretical basis in the work of Shannon (1948) in information theory and Jaynes (1957) in statistical physics.

As a special case of the general maximum entropy problem, we find the representative distribution given moments. The solution to the following problem, $f^*(z)$, is called the maximum-entropy distribution function.

$$\begin{aligned} \max_{f(z)} \quad & - \int_{b_1}^{b_2} f(z) \log f(z) dz \\ \text{subject to} \quad & \int_{b_1}^{b_2} z^i f(z) dz = \mu_i, \quad \forall i = 0, 1, \dots, n \\ & f(z) \geq 0, \end{aligned} \tag{23}$$

where $\mu_0, \mu_1, \dots, \mu_n$ are the given sequence of moments. The support $[b_1, b_2]$ is a subset of \mathbb{R} . The maximum-entropy distribution $f^*(z)$ is sensitive to the support interval, which is specified in advance. In practice, the trial-and-error experiments can help to generate an appropriate support.

The solution of problem (23) is a probability distribution, consistent with the given information, with maximal uncertainty. Accordingly, the maximum entropy probability is the “most likely”, “most unbiased”, “least prejudiced”, or “most uniform” distribution we can have. We regard it as an extension of our bound analysis because this problem incorporates only the moment information and yields a valid probability distribution, optimal in the sense that it uses as little information as possible. We will use this to illustrate our conclusions from the moment method on the risk analysis of MV+CVaR mortality portfolio.

4.3. Numerical Illustration. Now we extend the numerical example in Section 3.2 to include the moment method. Figure 1 shows the histograms of the original and the MV+CVaR 3-line

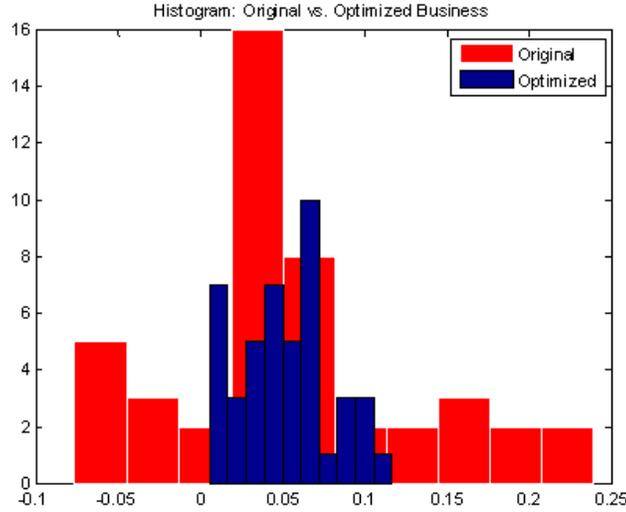


FIGURE 1. Histograms of the original and the MV+CVaR optimal life insurance portfolios.

portfolios based on 10 bins. From the histograms, it's not clear whether the original portfolio has a unimodal distribution. Adding the unimodal constraint are significantly narrows the bounds, as shown in Figure 2 for the case of 4-moment semi-parametric and the equally weighted 3-line mortality portfolio. For example, consider the 5%-VaR. Draw a horizontal line through the 0.05 level on the vertical axis. It intersects the solid curves at d values of -0.075 and -0.010. So if the original portfolio has unimodal distribution, the best we can say is that

$$-7.5\% \leq \text{VaR}_{0.05} \leq -1.0\%. \quad (24)$$

We should be very careful to make a unimodal assumption. If the actual data are not unimodal the 5%-VaR range is

$$-14.0\% \leq \text{VaR}_{0.05} \leq 2.5\%, \quad (25)$$

as determined by the intersection of the horizontal line with the $-o-$ curves. Theoretically, the normal distribution (the dotted lines in Figure 2) with the same mean and variance as the original mortality portfolio will fall between the 2-moment upper and lower arbitrary bounds. In our example, the normal curve is even within the 4-moment unimodal bounds. Under the normal assumption, the business's $\text{VaR}_{5\%}$ hits its lower bound, which is -7.5%.

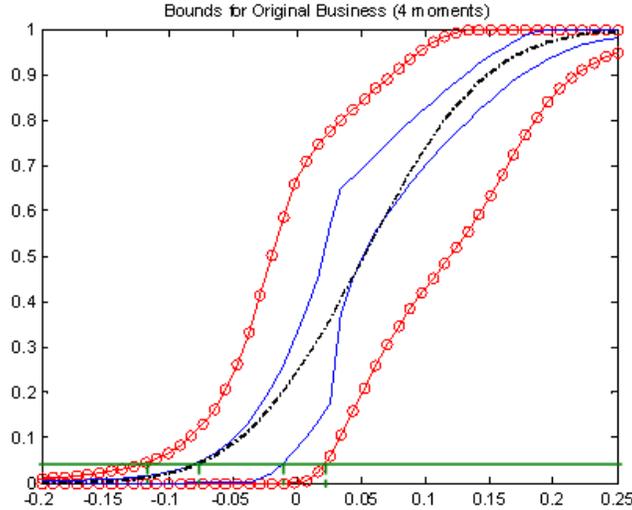


FIGURE 2. Bounds of the original 3-line life insurance portfolio. The lines with $-o-$ are the upper and lower arbitrary bounds with 4 moments. The solid lines show the unimodal bounds with 4 moments and the unimode 0.0264, and the dotted line in the middle represents the normal distribution with the same mean and variance as those of the original 3-line portfolio. The vertical axis stands for the cumulative probability $\Pr(\tilde{m}(w) \leq d)$, and the horizontal axis represents the portfolio's profit margin d .

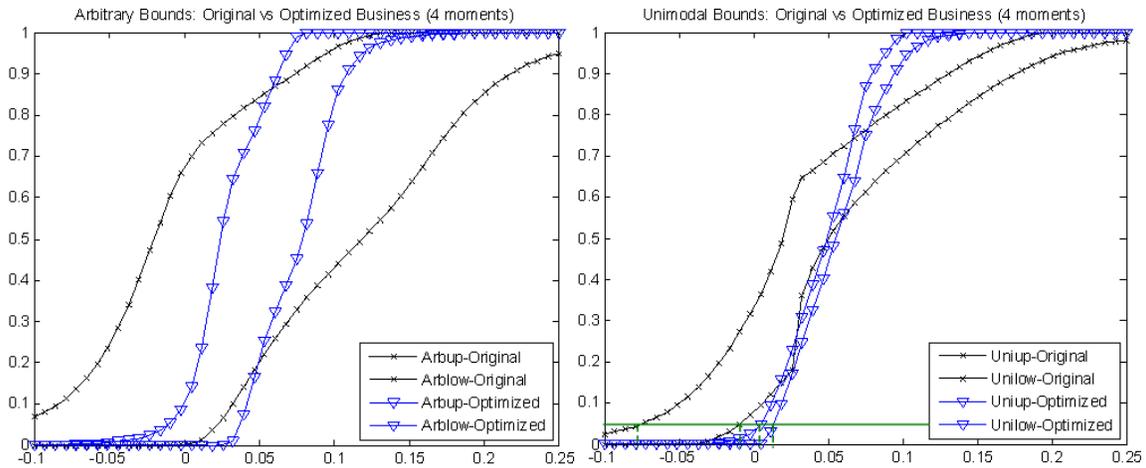


FIGURE 3. 4-moment Arbitrary and unimodal bounds of the original 3-line life insurance portfolio and its MV+CVaR optimum. The left plot draws the arbitrary bounds and the unimodal bounds are shown in the right plot. In both plots, the lines with $-x-$ represent the upper and lower bounds of the original business. Bounds on the optimal MV+CVaR business strategy are shown in the $-\Delta-$ curves. The vertical axis stands for the cumulative probability $\Pr(\tilde{m}(w) \leq d)$, and the horizontal axis represents the portfolio's profit margin d .

Figure 3 shows the 4-moment arbitrary and unimodal bounds of the original life insurance business with $w_0 = [33.33\%, 33.33\%, 33.33\%]$ and its MV+CVaR optimum with $w_{\text{MV+CVaR}} = [0, 19.81\%, 80.19\%]$. Notably, the bounds for the MV+CVaR life insurance business strategy are very narrow since its variance is low at 0.0008. No matter whether the arbitrary or the unimodal bounds are investigated, by applying the MV+CVaR approach, the downside risk is greatly decreased. To illustrate, similar to what we did in Figure 2, we draw a horizontal line through the 0.05 level on the vertical axis to analyze the 100% confidence bounds on 5%-VaR. If we assume both the original and optimized portfolios are unimodally distributed, the bounds of portfolio profit margin on $\text{VaR}_{0.05}$ rises to

$$0.60\% \leq \text{VaR}_{0.05} \leq 1.25\%, \quad (26)$$

by adding 5% CVaR constraint to the traditional Markowitz optimization. Thus the lower bound is greatly improved to a positive number, which means that there is only 5% probability that the business annual profit margin will fall below 0.5%. This is very attractive to insurance companies, especially those who are seriously worry about the downside risk during the economy recession.

To take our discussion one step further, we solve the maximum-entropy problem (23). The results are shown in Figure 4, which compares the maximum-entropy distributions of the original and the MV+CVaR mortality portfolios. In particular, the distribution of the MV+CVaR efficient 3-line portfolio shifts to the right of original portfolio distribution. The optimized portfolio has a higher mean and lower variance than the original portfolio, which is consistent with the conclusion of the moment method.

5. NATURAL HEDGING EFFECTS

Cox and Lin (2007) argue that an insurer selling both life insurance and annuities is exposed to lower one-directional changes in mortality. This effect is called “natural hedging”. As such, adding annuities to a portfolio that is only composed of life insurance may lower the portfolio’s mortality risk. In this section, we examine how natural hedging improves the MV+CVaR optimal business composition and decreases downside risk with the moment methods.

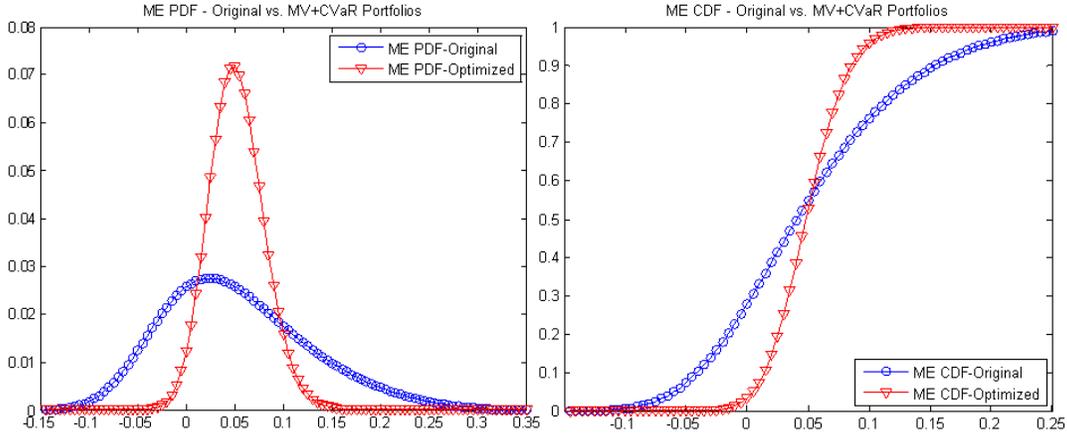


FIGURE 4. 4-moment maximum-entropy distributions of the original 3-line life insurance portfolio and its MV+CVaR optimum. The left plot graphs the density functions and the right one graphs the cumulative distribution functions. The curves $-o-$ are for the distributions of the original portfolio's profit margin and the curves $-\triangle-$ are for the distributions of the MV+CVaR portfolio's profit margin. The vertical axis stands for the probability, and the horizontal axis represents the portfolio's profit margin d .

TABLE 3. Summary Statistics – Profit Margin of Single Premium Immediate Life Annuity

Lines	Mean	Variance	Skewness	Kurtosis
$\tilde{m}_{a_{65}}$	0.0635	0.0002	-0.8590	0.4792

To illustrate the natural hedging effect, we add an annuity to the 3-line life insurance portfolio analyzed in Section 3.2. Specifically, in addition to selling the 5-year term life insurance on male aged 35, the 10-year term life insurance on male aged 25, and the whole life insurance on male aged 40, the insurer also underwrites a single-premium immediate life annuity on male aged 65. Following equations (3) and (4) and the same assumptions as what we use to calculate the profit margins of the three lines of life insurance, we obtain the summary statistics for $\tilde{m}_{a_{65}}$, as shown in Table 3.

Assume originally the insurer puts an equal weight (i.e. 25%) in each of these four lines of business. The expected margin of this equally weighted portfolio is 0.0547. To optimize this mortality portfolio, we specify a 5%-CVaR constraint with ζ in (11) determined by equation (17) and the objective profit margin $m_0 = 0.0547$. We get the MV+CVaR efficient 4-line mortality

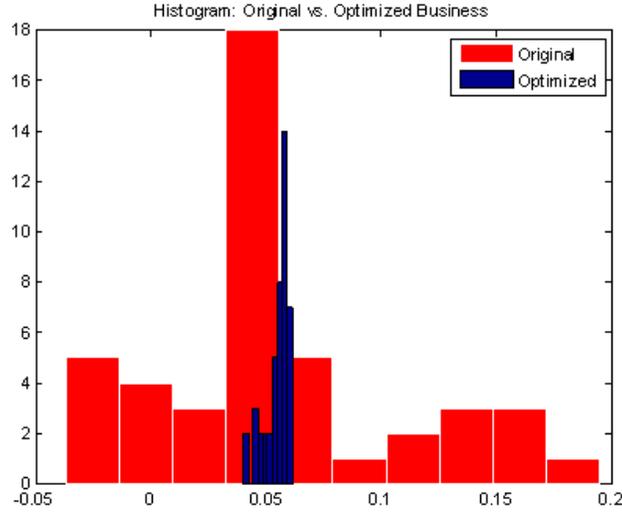


FIGURE 5. Histograms of the original and the MV+CVaR optimal 4-line portfolios.

portfolio with the weights,

$$w_{\text{MV+CVaR}} = [0, 2.52\%, 37.71\%, 59.77\%]. \quad (27)$$

That is, the insurer should stop selling the 5-year life insurance on male aged 35, and adjust its business to make 2.52% of its total premium from the 10-year life insurance on male aged 25, 37.71% from the whole life insurance on male aged 40, and 59.77% from the single premium immediate life annuity on male aged 65. Table 4 compares the original and MV+CVaR mortality portfolios. The MV+CVaR portfolio has a lower variance and skewness but a higher kurtosis, mode and 5%-CVaR. Figure 5 shows the histograms of the original and the MV+CVaR optimal 4-line portfolios based on 10 bins. According to the histograms, it seems that both the original and the optimized 4-line portfolios are unimodally distributed. So our following bounds analysis is performed based on the unimodal assumption.

TABLE 4. Summary Statistics of Profit Margins of Original and MV+CVaR 4-Line Portfolios

$\tilde{m}(w)$	Mean	Variance	Skewness	Kurtosis	Mode	CVaR _{5%}
Original	0.0547	0.0031	0.6107	0.1452	0.0385	-0.0363
MV+CVaR	0.0547	2.6E-05	-1.2334	0.7888	0.0575	0.0422

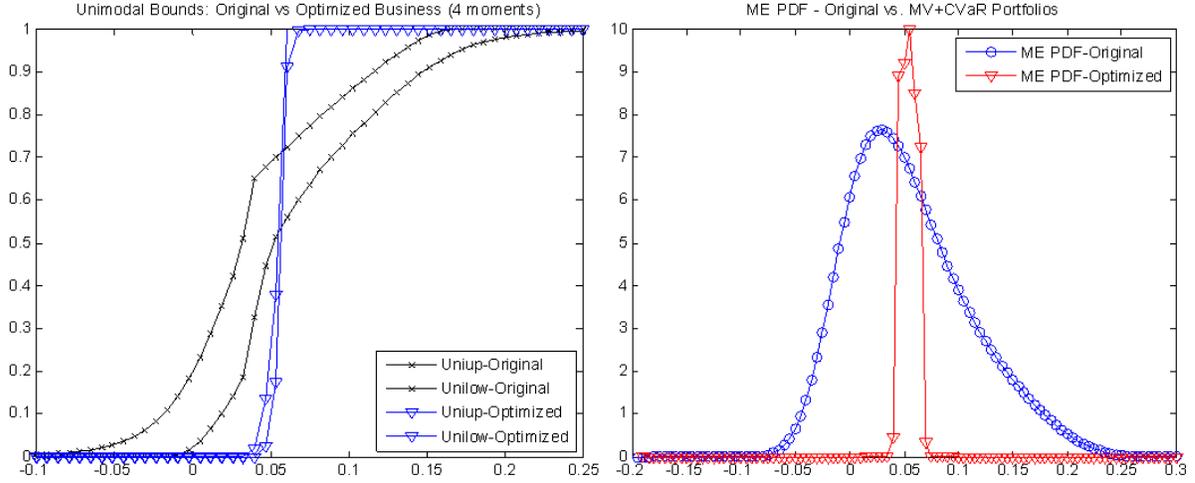


FIGURE 6. 4-moment unimodal bounds and maximum-entropy probability density of the original and MV+CVaR portfolios with 4 lines of business. The left plot shows the unimodal bounds of the original mortality portfolio and its MV+CVaR optimum. The right plot shows the maximum-entropy probability density of these two portfolios. The curves with $-x-$ in the left plot represent the bounds of the original portfolio. The $-o-$ curve in the right plot stands for the distribution of the original mortality portfolio. In both graphs, the unimodal bounds and the maximum-entropy probability distribution of the MV+CVaR portfolio are represented by the $-\triangle-$ curves.

Figure 6 shows the 4-moment unimodal bounds and the maximum-entropy probability density of the original and MV+CVaR portfolios with 4 lines of business. The variance of the optimal 4-line portfolio is almost zero (see Table 4), so the upper and lower unimodal bounds of the MV+CVaR portfolio are very close to each other and its maximum-entropy probability density is approaching a dirac delta function.⁵

To explore the natural hedging effect, we compare the bounds of the MV+CVaR 3-line and 4-line portfolios. If we assume both optimized portfolios have the unimodal distributions, the right plot in Figure 7 shows how the 100% confidence interval of the 5%-level VaR improves by adding an annuity to the 3-line pure life insurance portfolio. Specifically,

$$0.6\% < \text{VaR}_{0.05}^{3L} < 1.25\% \Rightarrow 4.1\% < \text{VaR}_{0.05}^{4L} < 5.1\%. \quad (28)$$

⁵The dirac delta function is a generalized function representing an infinitely sharp peak bounding unit area: a “function” $\delta(x)$ that has the value zero everywhere except at $x = 0$ where its value is infinitely large in such a way that its total integral is 1.

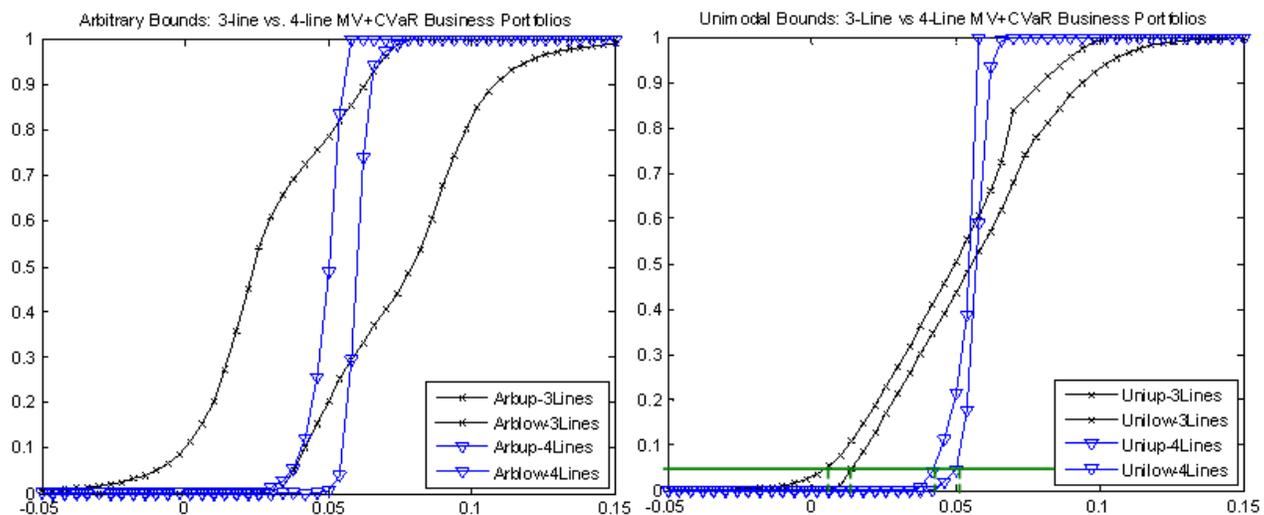


FIGURE 7. 4-moment Arbitrary and unimodal bounds of the MV+CVaR efficient 3-line and 4-line mortality portfolios. The left plot draws the arbitrary bounds and the unimodal bounds are shown in the right plot. In both plots, the curves with $-x-$ represent the upper and lower bounds of the MV+CVaR portfolio with three lines of life insurance. Bounds on the optimal 4-line mortality portfolio are shown as the $-\Delta-$ curves. The vertical axis stands for cumulative probability $\Pr(\tilde{m}(w) \leq d)$, and the horizontal axis represents the portfolio's profit margin d .

That is, the confidence interval of $\text{VaR}_{0.05}^{4L}$ stays at a much higher level of profit margin range than that of $\text{VaR}_{0.05}^{3L}$. It highlights the benefits of natural hedging: by including both annuity and life insurance in a portfolio, natural hedging increases profit but decreases mortality risk.

Discussion. In practice, it may be difficult for an insurer to directly accomplish an MV+CVaR optimal business composition, for example, the mortality portfolio weights in (27). Furthermore, for an insurer specializing in life insurance, entering the annuity business may not be practical (Cox and Lin, 2007). If an insurer is able to take advantage of the MV+CVaR approach and natural hedging at a low cost, it will gain from a higher profit margin. To achieve this goal, we recommend two possible solutions. First, the insurer can buy or sell reinsurance to rebalance its weights in various business lines. Second, the insurer can issue or purchase mortality or longevity securities from capital markets. The mortality-linked securities are new in the financial markets but have attracted a lot of attentions from insurers, investors, pension plans, and academia. As

the mortality-linked security markets develop, the insurer can cede or assume risk to realize the MV+CVaR efficient mortality portfolio at a lower cost.

6. FRONTIERS OF EFFICIENT MORTALITY PORTFOLIOS

So far our analysis has focused on improving an insurer's existing mortality portfolio with the MV+CVaR approach and used the moment method to examine whether and how this MV+CVaR portfolio could control downside risk. Can these techniques be applied to any insurance business composition? How superior is the MV+CVaR approach to the traditional Markowitz optimization method? How does the natural hedging reduce mortality risk for different MV+CVaR portfolios? To answer these questions, in this section, we extend our analysis to all efficient portfolios, not just a particular efficient portfolio given a level of profit margin. Specifically, we compare

- (1) the frontiers of the Markowitz and MV+CVaR efficient portfolios;
- (2) the frontiers of the 3-line and 4-line efficient portfolios.

6.1. Frontiers of Markowitz and MV+CVaR optimized Portfolios. Suppose an insurer sells the aforementioned four lines of business: $A_{35:\overline{5}}^1$, $A_{25:\overline{10}}^1$, A_{40} and a_{65} . We solve (a) the traditional Markowitz portfolio problem, and (b) the MV+CVaR problem (14) with a 5%-CVaR constraint, respectively, to obtain a set of efficient portfolios with different expected profit margins. Then, we compare their optimization results. To illustrate, we plot the mean-variance, skewness-variance, CVaR_{5%}-variance, and CVaR_{95%}-variance graphs. Each graph in Figure 8 is a piecewise linear interpolation based on 10 solved efficient portfolios.

While the top left graph of Figure 8 shows that the MV+CVaR efficient frontier somewhat deviates from the Markowitz efficient frontier in terms of the mean-variance tradeoff, the MV+CVaR approach effectively increases the skewness of low-variance portfolios shown in the skewness-variance graph. A higher skewness is desirable because it increases the likelihood of obtaining higher profit margins. The CVaR_{5%}-variance curves in the bottom left graph demonstrate that, for the same variance, the portfolios constructed from the MV+CVaR approach are able to reach a higher 5%-CVaR, implying a lower downside mortality risk. However, the impact of adding CVaR constraint to the Markowitz model on the high-variance portfolios is not as significant as that on

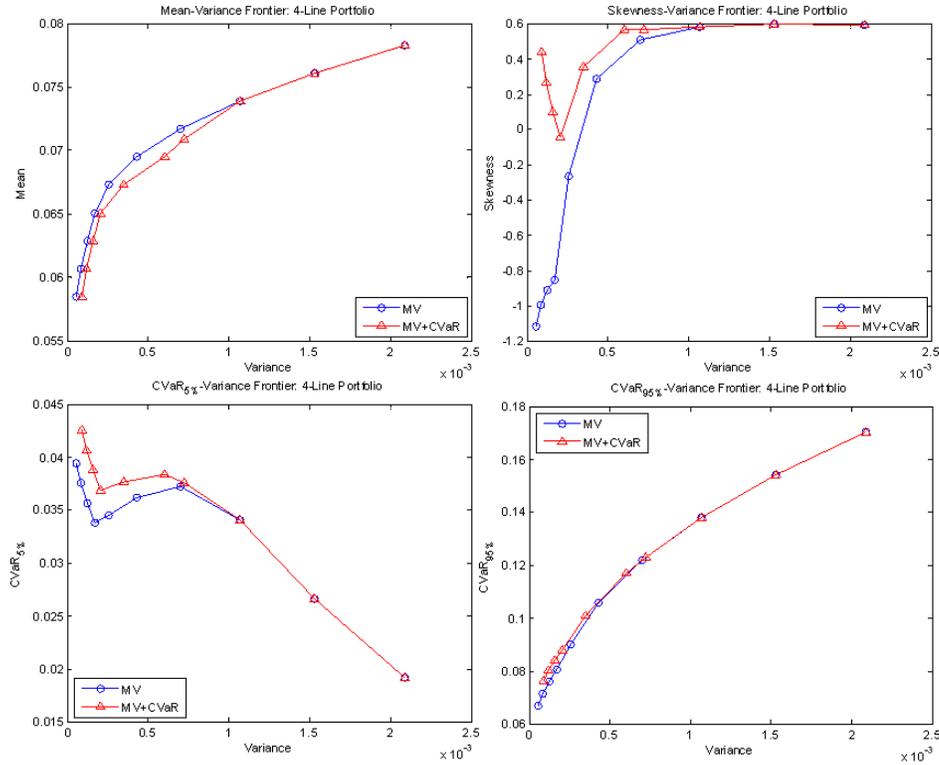


FIGURE 8. Efficient frontiers of optimal mortality portfolios with 4 lines of business. The mean-variance, skewness-variance, $\text{CVaR}_{5\%}$ -variance, and $\text{CVaR}_{95\%}$ -variance plots are shown in the top left, top right, bottom left, and bottom right graphs, respectively. Each graph is a piecewise linear interpolation based on 10 points. The curves with $-o-$ represent the traditional Markowitz frontiers. The $\text{MV}+\text{CVaR}$ efficient frontiers obtained by adding a 5%-level CVaR constraint are shown as $-\triangle-$ curves in all graphs.

the low variance portfolios. It is because the low-variance portfolio has more compact distribution, which makes it more sensitive to the tail-distribution management.

It is worth noting that the $\text{MV}+\text{CVaR}$ approach aims at reshaping the left tail of the profit margin distribution, which corresponds to high losses. The approach does not account for the right tail representing high profits. This is confirmed by the bottom right graph in Figure 8, which shows that the $\text{CVaR}_{95\%}$ curve of the $\text{MV}+\text{CVaR}$ portfolios just barely differs from that of the Markowitz counterparts. It suggests that the $\text{MV}+\text{CVaR}$ approach has no significant effect on changing the right tail of the distribution.

6.2. Natural Hedging for Efficient Portfolios. We investigate the natural hedging effect for various $\text{MV}+\text{CVaR}$ efficient portfolios to extend our analysis in Section 5. With the same setup as

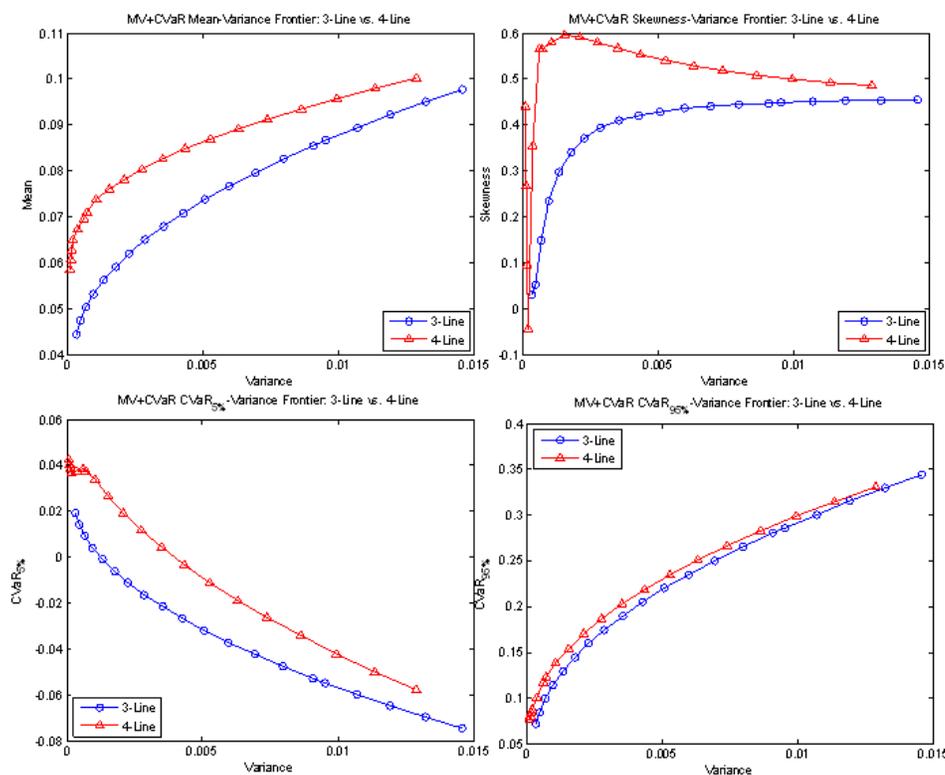


FIGURE 9. Efficient frontiers of MV+CVaR 3-line and 4-line portfolios with a 5%-CVaR constraint. The mean-variance, skewness-variance, CVaR_{5%}-variance, and CVaR_{95%}-variance plots are shown in the top left, top right, bottom left, and bottom right graphs, respectively. Each graph is a piecewise linear interpolation based on 20 points. The curves with $-o-$ represent the 3-line frontiers and the $-\triangle-$ curves stand for the frontiers of the 4-line efficient portfolios.

the example in Section 5 but setting a series of different expected profit margins, Figure 9 shows that the 4-line portfolios ($-\triangle-$ curves) including both life insurance and annuity lines outperform the portfolios that only contain life insurance ($-o-$ curves). Given the same variance, the 4-line efficient portfolio achieves higher profit margin, higher skewness, higher CVaR_{5%}, and higher CVaR_{95%} than their 3-line counterparts. Therefore, the inclusion of annuities in the mortality portfolio offsets the effect of longevity risk on the life insurance policies. This provides a new evidence to support the benefits of natural hedging. Specifically in our example, the improvement of frontier by adding an annuity can be explained by the negative correlation between the annuity a_{65} and the three lines of life insurance, i.e. $A_{35:\overline{5}}^1$, $A_{25:\overline{10}}^1$, and A_{40} , as shown in Table 5.

TABLE 5. Correlation of of Annuity and Life Insurance Profit Margins

	$\tilde{m}_{A^1_{35:\overline{5} }}$	$\tilde{m}_{A^1_{25:\overline{10} }}$	$\tilde{m}_{A_{40}}$
$\tilde{m}_{a_{65}}$	-0.16094	-0.12171	-0.68109

7. CONCLUSION

This paper incorporates the moment methods and the portfolio theory for mortality risk management. It contributes to the mortality risk management literature in two ways. First, we propose the MV+CVaR approach to manage mortality portfolio risk with a reasonable sacrifice of the mean-variance efficiency. Specifically, the MV+CVaR approach controls the tail risk by specifying the mean value of the lowest $\beta\%$ profit margin to be no less than some pre-specified value. At the same time, it optimizes the tradeoff between mean and variance. This method, in general, is effective in obtaining an optimal mortality risk portfolio while controlling its downside risk.

Second, we apply the moment methods in mortality risk management. Without assuming any arbitrary distribution, we calculate the semiparametric upper and lower bounds on the cumulative probability of profit margins for mortality portfolios. The bounds are used to illustrate the 100% confidence interval of the downside risk, which is measured by β -level VaR. In particular, we propose how to use the moment methods to investigate downside risk of MV+CVaR efficient mortality portfolios. In addition, as an extension to the moment methods, we derive the maximum-entropy distribution of mortality portfolios. We use the maximum-entropy approach to provide a robustness check for the moment methods because the maximum-entropy approach provides us a representative distribution that is the most unbiased one given moment information.

Although we illustrate our MV+CVaR and moment methods with up to four lines of business and based on the annual observations, these methods can be easily extended to more lines of business and high frequency data. Furthermore, our approaches can be refined by incorporating asset risks to examine an life insurer's overall expected shortfall. We leave this for future research.

APPENDIX A

Proof of Theorem 1. For a given time horizon, β -level VaR is defined as

$$\text{VaR}_\beta(w) = \alpha_\beta(w) = \min\{\alpha \in \mathbb{R} : \Pr(\tilde{m}(w) \leq \alpha) \geq \beta\},$$

where, $\tilde{m}(w) = \sum_{i=1}^n \mathbb{E}[\tilde{m}_i]w_i$, is the random variable of mortality portfolio margin. The corresponding β -level CVaR is the expected value of profit margin given the margin being not greater than the β -level VaR, $\alpha_\beta(w)$.

$$\text{CVaR}_\beta(w) = \theta_\beta(w) = \mathbb{E}[\tilde{m}(w) | \tilde{m}(w) \leq \alpha_\beta(w)],$$

Define G_β as a function on $\mathbb{W} \times \mathbb{R}$ that satisfies:

$$G_\beta(w, \alpha) = \alpha - \frac{1}{\beta} \mathbb{E}[[\alpha - \tilde{m}(w)]^+],$$

where $[g]^+ = \max\{g, 0\}$. Rockafellar and Uryasev (2000) proved that $\alpha_\beta(w)$ is determined by maximizing $G_\beta(w, \alpha)$:

$$\alpha_\beta(w) = \max_{\alpha \in \mathbb{R}} G_\beta(w, \alpha).$$

Therefore, if the set consisting of $\alpha_\beta(w)$ is nonempty, one always has

$$\alpha_\beta(w) \in \arg \max_{\alpha \in \mathbb{R}} G_\beta(w, \alpha) \text{ and } \theta_\beta(w) = G_\beta(w, \alpha_\beta(w)).$$

With the notations and equivalency discussed above, we have

$$\begin{aligned} & \min_{w \in \mathbb{W}} [\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j], \quad \text{s.t.} \quad \text{CVaR}_\beta(x) \geq \zeta \\ & \quad \quad \quad \Downarrow \\ & \min_{w \in \mathbb{W}} [\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j], \quad \text{s.t.} \quad \zeta - \theta_\beta(w) \leq 0 \\ & \quad \quad \quad \Downarrow \\ & \min_{w \in \mathbb{W}} [\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j], \quad \text{s.t.} \quad \zeta - \max_{\alpha \in \mathbb{R}} [G_\beta(w, \alpha)] \leq 0 \\ & \quad \quad \quad \Downarrow \\ & \min_{w \in \mathbb{W}} [\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j], \quad \text{s.t.} \quad \min_{\alpha \in \mathbb{R}} [\zeta - G_\beta(w, \alpha)] \leq 0, \end{aligned} \tag{P1}$$

where $\mathbb{W} = \left\{ \sum_{i=1}^n \mathbb{E}[\tilde{m}_i]w_i = m_0 \right\} \cup \left\{ \sum_{i=1}^n w_i = 1 \right\} \cup \{w_i \geq 0, i = 1, 2, \dots, n\}$ is a subset of the mortality business composition feasible set.

Therefore, to prove Theorem 1 is to prove the last equation in (P1) is equivalent to

$$\min_{w \in \mathbb{W}, \alpha \in \mathbb{R}} \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \right], \quad \text{s.t. } [\zeta - G_\beta(w, \alpha)] \leq 0, \quad (P2)$$

in the sense that their objectives achieve the same minimum values.

According to the Karush-Kuhn-Tucker conditions, the necessary conditions for the problem (P2) are stated as follows:

$$\begin{aligned} \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i^* w_j^* \right] + \lambda [\zeta - G_\beta(w^*, \alpha^*)] &\leq \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \right] + \lambda [\zeta - G_\beta(w, \alpha)] \\ &\Downarrow \\ \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i^* w_j^* \right] - \lambda G_\beta(w^*, \alpha^*) &\leq \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \right] - \lambda G_\beta(w, \alpha), \end{aligned} \quad (i)$$

and

$$\lambda [\zeta - G_\beta(w^*, \alpha^*)] = 0, \quad \lambda \geq 0, \quad w \in \mathbb{W}. \quad (ii)$$

To make conditions (i) and (ii) be the sufficient condition as well, we should prove

$$h_1(w) = \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \right] \quad \text{and} \quad h_2(w, \alpha) = \zeta - G_\beta(w, \alpha)$$

are convex functions, and

$$h_3(w) = \sum_{i=1}^n \mathbb{E}[\tilde{m}_i]w_i - m_0 \quad \text{and} \quad h_4(w) = \sum_{i=1}^n w_i - 1$$

are affine functions. It is obvious that $h_3(w)$ and $h_4(w)$ are affine functions. As for $h_1(w)$ and $h_2(w)$, given $\rho \in [0, 1]$,

$$\begin{aligned} h_1(\rho w + (1 - \rho)w') &= \mathbb{E}[(\rho w + (1 - \rho)w')^2] - [\mathbb{E}[\rho w + (1 - \rho)w']]^2 \\ &= \rho^2 \mathbb{E}(w^2) + (1 - \rho)^2 \mathbb{E}(w'^2) + 2\rho(1 - \rho)\mathbb{E}(ww') \\ &\quad - \rho^2 [\mathbb{E}(w)]^2 - (1 - \rho)^2 [\mathbb{E}(w')]^2 - 2\rho(1 - \rho)\mathbb{E}(w)\mathbb{E}(w'), \end{aligned} \quad (29)$$

and

$$\rho h_1(w) + (1 - \rho)h_1(w') = \rho E(w^2) - \rho [E(w)]^2 + (1 - \rho)E(w'^2) - (1 - \rho)[E(w')]^2. \quad (30)$$

Then we have

$$\begin{aligned} h_1(\rho w + (1 - \rho)w') - [\rho h_1(w) + (1 - \rho)h_1(w')] &= \rho(1 - \rho) [[E(w) - E(w')]^2 - E[(w - w')^2]] \\ &\leq \rho(1 - \rho) [[E(w) - E(w')]^2 - [E(w - w')]^2] \\ &= 0. \end{aligned} \quad (31)$$

So $h_1(w)$ is convex on w . The inequality above follows from the Jensen's inequality: $E(w^2) \geq [E(w)]^2$.

Let the vector $(w, \alpha) = g$, for $\rho \in [0, 1]$,

$$G_\beta(\rho g + (1 - \rho)g') = \rho \alpha + (1 - \rho)\alpha' - \frac{1}{\beta} E [[\rho \alpha + (1 - \rho)\alpha' - \tilde{m}(\rho w + (1 - \rho)w')]^+], \quad (32)$$

and

$$\begin{aligned} G_\beta(g) + (1 - \rho)G_\beta(g') &= \rho \alpha - \frac{1}{\beta} E [[\rho \alpha - \tilde{m}(\rho w)]^+] \\ &\quad + (1 - \rho)\alpha' - \frac{1}{\beta} E [[(1 - \rho)\alpha' - \tilde{m}((1 - \rho)w')]^+]. \end{aligned} \quad (33)$$

So we get

$$\begin{aligned} &G_\beta(\rho g + (1 - \rho)g') - [G_\beta(g) + (1 - \rho)G_\beta(g')] \\ &= \frac{1}{\beta} E [[\rho \alpha - \tilde{m}(\rho w)]^+] + E [[(1 - \rho)\alpha' - \tilde{m}((1 - \rho)w')]^+] \\ &\quad - \frac{1}{\beta} E [[\rho \alpha + (1 - \rho)\alpha' - \tilde{m}(\rho w + (1 - \rho)w')]^+]. \end{aligned} \quad (34)$$

Since $\tilde{m}(w) = \sum_{i=1}^n E[\tilde{m}_i]w_i$ is a linear transformation on w ,

$$\tilde{m}(\rho w + (1 - \rho)w') = \tilde{m}(\rho w) + \tilde{m}((1 - \rho)w').$$

Following $[a + b]^+ \leq [a]^+ + [b]^+$, we have

$$\begin{aligned} & \mathbb{E} \left[[\rho\alpha + (1-\rho)\alpha' - \tilde{m}(\rho w) - \tilde{m}((1-\rho)w')]^+ \right] \\ & \leq \mathbb{E} \left[[\rho\alpha - \tilde{m}(\rho w)]^+ \right] + \mathbb{E} \left[[(1-\rho)\alpha' - \tilde{m}((1-\rho)w')]^+ \right]. \end{aligned} \quad (35)$$

Therefore,

$$\begin{aligned} & G_\beta(\rho g + (1-\rho)g') - [G_\beta(g) + (1-\rho)G_\beta(g')] \\ & = \frac{1}{\beta} \left\{ \mathbb{E} \left[[\rho\alpha - \tilde{m}(\rho w)]^+ \right] + \mathbb{E} \left[[(1-\rho)\alpha' - \tilde{m}((1-\rho)w')]^+ \right] \right\} \\ & \quad - \frac{1}{\beta} \left\{ \mathbb{E} \left[[\rho\alpha + (1-\rho)\alpha' - \tilde{m}(\rho w) - \tilde{m}((1-\rho)w')]^+ \right] \right\} \\ & \geq \frac{1}{\beta} \left\{ \mathbb{E} \left[[\rho\alpha - \tilde{m}(\rho w)]^+ \right] + \mathbb{E} \left[[(1-\rho)\alpha' - \tilde{m}((1-\rho)w')]^+ \right] \right\} \\ & \quad - \frac{1}{\beta} \left\{ \mathbb{E} \left[[\rho\alpha - \tilde{m}(\rho w)]^+ \right] + \mathbb{E} \left[[(1-\rho)\alpha' - \tilde{m}((1-\rho)w')]^+ \right] \right\} \\ & = 0. \end{aligned} \quad (36)$$

So $G_\beta(w, \alpha)$ is concave on (w, α) and $h_2(w, \alpha) = \zeta - G_\beta(w, \alpha)$ is a convex function on (w, α) .

Thus (i) and (ii) are not only the necessary and but also the sufficient conditions of problem (P2).

First suppose that w^* is a solution of (P1) and $\alpha^* = \arg \max_{\alpha \in \mathbb{R}} G_\beta(w^*, \alpha)$. Next us show that (w^*, α^*) is a solution of (P2).

$$\begin{aligned} & \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i^* w_j^* \right] - \lambda G_\beta(w^*, \alpha^*) = \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i^* w_j^* \right] - \lambda \theta_\beta(w^*) \\ & \leq \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \right] - \lambda \theta_\beta(w) = \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \right] - \lambda \max_{\alpha \in \mathbb{R}} G_\beta(w, \alpha) \\ & \leq \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \right] - \lambda G_\beta(w, \alpha), \end{aligned} \quad (37)$$

and

$$\lambda [\zeta - G_\beta(w^*, \alpha^*)] = \lambda [\zeta - \theta_\beta(w^*)] = 0, \quad \lambda \geq 0, \quad w \in \mathbb{W}.$$

So conditions (i) and (ii) are satisfied and (w^*, α^*) is a solution of (P2).

Second, suppose the (w^*, α^*) achieves the minimum of (P2) and $\lambda \geq 0$. For fixed w^* , the point α^* minimizes the function $[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i^* w_j^*] - \lambda G_\beta(w^*, \alpha^*)$, and, consequently, the function $G_\beta(w^*, \alpha^*)$ is maximized. Then $\alpha^* = \arg \max_{\alpha \in \mathbb{R}} G_\beta(w^*, \alpha)$. We have

$$\begin{aligned} & \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i^* w_j^* \right] - \lambda \theta_\beta(w^*) = \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i^* w_j^* \right] - \lambda G_\beta(w^*, \alpha^*) \\ & \leq \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \right] - \lambda G_\beta(w, \alpha_\beta(w)) = \left[\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} w_i w_j \right] - \lambda \theta_\beta(w), \end{aligned} \quad (38)$$

and

$$\lambda[\zeta - \theta_\beta(w^*)] = \lambda[\zeta - G_\beta(w^*, \alpha^*)] = 0, \quad \lambda \geq 0, \quad w \in \mathbb{W}.$$

Therefore conditions (i) and (ii) are satisfied, and w^* is a solution of (P1). If a pair (w^*, α^*) achieves the maximization of equation (P2), $G_\beta(w^*, \alpha^*)$ returns β -level CVaR and α^* gives the corresponding β -level VaR.

APPENDIX B

Estimating Profit Margins Summarized in Table 1. Assume the insurer uses the Renshaw et al. (1996) model to forecast its future mortality rates. The Renshaw et al. (1996) model incorporates both the age variation and the underlying time trend of the force of mortality $\mu_{x,t}$. The force of mortality $\mu_{x,t}$ of age x in year t in Renshaw et al. (1996) is modeled as

$$\mu_{x,t} = \exp \left\{ \beta_0 + \sum_{j=1}^h \beta_j L_j(x') + \sum_{g=1}^r \alpha_g t'^g + \sum_{g=1}^r \sum_{l=1}^v \gamma_{gl} L_l(x') t'^g \right\}, \quad (39)$$

where t' and x' are the transformed times and ages mapped onto the interval $[-1, +1]$. For example, if we have mortality experience data for individual ages x ranging from 65 to 100 years for the period $t = 0$ to $t = 104$ years, t' and x' have the values,

$$t' = \frac{t - 52}{52} \quad \text{and} \quad x' = \frac{x - 82.5}{17.5}.$$

$L_j(z)$ and $L_l(z)$ in model (39) are the Legendre polynomials defined as follows:

$$L_0(z) = 1$$

$$L_1(z) = z$$

$$L_2(z) = (3z^2 - 1)/2$$

$$\vdots$$

$$(m+1)L_{m+1}(z) = (2m+1)zL_m(z) - mL_{m-1}(z),$$

where j and $l = 0, 1, 2, \dots, m+1$, and $z = t'$ or x' .

The insurer uses the US population mortality tables, observed each year from 1901 to 2005 from the Human Life Table Database and the Human Mortality Database, to estimate model (39).⁶ The age range is $x = 25, 26, 27, \dots, 100$ for US males from $t = 0, 1, 2, \dots, 104$ (corresponding to calendar years 1901 – 2005). Younger ages have different mortality change patterns from those of older ages (Cox and Lin, 2007). Thus, to improve the model's goodness of fit, the insurer estimates model (39) for the age range 25-64 and the age range 65-100 separately. Following Lin and Cox (2005), the insurer fits model (39) with $h = 3$, $r = 1$ and $v = 1$. Details of the parameter estimates for two age ranges are given in Table 6.

Assume constant force of mortality within each year of age. After forecasting $\mu_{x,t}$ based on equation (39), the insurer predicts the one-year survival rate for each age x in year t as follows

$$p_{x,t} = \exp(-\mu_{x,t}).$$

As such, the probability that the age x at the beginning of year t will die within one year equals

$$q_{x,t} = 1 - p_{x,t} = 1 - \exp(-\mu_{x,t}).$$

⁶The tables for years 1901 to 1999 are from the Human Life Table Database and the tables for 2000 to 2005 are from the Human Mortality Database, published by the University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on June 8, 2008).

TABLE 6. The model (39) with $h = 3$, $r = 1$ and $v = 1$, fits the US Male Population data from 1901 to 2005. All of the coefficients are significant at the 1% level.

Panel A: Ages 25-64					
Coefficient	Estimate	Standard Error	Coefficient	Estimate	Standard Error
β_0	-4.8979	0.0023	β_3	-0.0831	0.0057
β_1	1.3029	0.0040	α_1	-0.6841	0.0040
β_2	0.1916	0.0050	γ_{11}	0.3240	0.0067
Adjusted R^2	0.9713				

Panel B: Ages 65-100					
Coefficient	Estimate	Standard Error	Coefficient	Estimate	Standard Error
β_0	-1.9971	0.0012	β_3	-0.0237	0.0030
β_1	1.3816	0.0021	α_1	-0.3055	0.0021
β_2	-0.0449	0.0026	γ_{11}	0.0751	0.0035
Adjusted R^2	0.9921				

Following equation (2), the net premiums (or the expected payments) of aforementioned three types of life insurance charged at the beginning of year t equal

$$\begin{aligned}
 P_{1,t} &= A_{35:\overline{5}|,t}^1 = \sum_{j=0}^{5-1} v^{j+1} ({}_j p_{35,t}) (q_{35+j,t+j}), \\
 P_{2,t} &= A_{25:\overline{10}|,t}^1 = \sum_{j=0}^{10-1} v^{j+1} ({}_j p_{25,t}) (q_{25+j,t+j}), \\
 P_{3,t} &= A_{40,t} = \sum_{j=0}^{\infty} v^{j+1} ({}_j p_{40,t}) (q_{40+j,t+j}).
 \end{aligned} \tag{40}$$

Following equation (1), the j -year survival rate for age x at the beginning of year t in equation (40), ${}_j p_{x,t}$, is calculated based on the projected mortality tables from year t to year $t + j - 1$. For example, the 3-year survival rate for age 35 at the beginning of year t ,

$${}_3 p_{35,t} = (p_{35,t}) \times (p_{36,t+1}) \times (p_{37,t+2}),$$

where the one-year survival rates $p_{35,t}$, $p_{36,t+1}$, and $p_{37,t+2}$ are read from the projected mortality tables in year t , $t + 1$, and $t + 2$, respectively. Similarly, $q_{x+j,t+j}$ is the one-year death rate for age $x + j$ based on the mortality table in year $t + j$. In sum, as the insured at the age of x in year t gets

j years older (i.e. $x + j$ years old in year $t + j$), the insurer shifts to or refers to the mortality table for year $t + j$ to locate $p_{x+j,t+j}$ and $q_{x+j,t+j}$.

Equation (4) suggests the present value of actual death payments of three types of life insurance sold in year t equal

$$\begin{aligned}\tilde{L}_{1,t} &= \tilde{A}_{35:\overline{5}|,t}^1 = \sum_{j=0}^{5-1} v^{j+1} ({}_j\tilde{p}_{35,t}) (\tilde{q}_{35+j,t+j}), \\ \tilde{L}_{2,t} &= \tilde{A}_{25:\overline{10}|,t}^1 = \sum_{j=0}^{10-1} v^{j+1} ({}_j\tilde{p}_{25,t}) (\tilde{q}_{25+j,t+j}), \\ \tilde{L}_{3,t} &= \tilde{A}_{40,t} = \sum_{j=0}^{\infty} v^{j+1} ({}_j\tilde{p}_{40,t}) (\tilde{q}_{40+j,t+j}).\end{aligned}\tag{41}$$

To illustrate the MV+CVaR method with reasonable margins, we assume the insurer has the same mortality experience as that of the U.S. population. Accordingly, $\tilde{L}_{1,t}$, $\tilde{L}_{2,t}$, and $\tilde{L}_{3,t}$ can be calculated from the actual U.S. mortality tables with the shifting method. When we use those historical mortality tables, we notice the 1918 worldwide flu epidemic pushed up the one-year death rates for different ages unevenly. The flu struck ages 0-50 more seriously than older ages, due to immunity that the older people had acquired as survivors of earlier flu epidemics (Cox and Lin, 2007). Since we have not experienced an epidemic for about a century, no one has immunity to epidemics, such as a deadly H1N1 flu. If an epidemic breaks out, it may have the same impact at all ages equally. To mimic a more realistic 1918-type flu in the future, when we calculate $\tilde{L}_{i,t}$, we adjust the 1917 and 1918 U.S. population mortality tables by setting one-year mortality rate changes in ages above 60 equal to the average mortality deterioration in ages 25-60. In addition, we suppose $a = 0.06$ and $c = 6$ in equation (9) and the interest rate 0.06. Then the total margin of life insurance i in year t equals

$$\tilde{m}_{i,t} = 0.06 + 3E[\tilde{r}_{i,t}^2] + \tilde{r}_{i,t},$$

where $\tilde{r}_{i,t} = \frac{P_{i,t}}{\tilde{L}_{i,t}} - 1$. As noted earlier, there are 105 realized U.S. population mortality tables from 1901 to 2005 ($t = 0, 1, 2, \dots, 104$). If we assume the terminal age is 100, we need 60 of those

mortality tables to calculate one $\tilde{A}_{40,t}$ observation and then get a $\tilde{m}_{A_{40,t}}$.⁷ Therefore, we have 45 observations for $\tilde{m}_{A_{40,t}}$ where $t = 0, 1, 2, \dots, 44$. To make consistent moment calculations, we also calculate $\tilde{m}_{A_{35:\overline{5}|}^1}^1$ and $\tilde{m}_{A_{25:\overline{10}|}^1}^1$ for $t = 0, 1, 2, \dots, 44$. Then we obtain the summary statistics in Table 1.

REFERENCES

- Akhiezer, N. I. (1965). *The Classical Moment Problem and Some Related Questions in Analysis* (1 ed.). New York: Hafner Publishing Company. Translated by N. Kemmer.
- Bertsimas, D. and I. Popescu (2005). Optimal inequalities in probability theory: A convex optimization approach. *SIAM J. Optimization* 15(3), 780–804.
- Cowling, C. and R. Dales (2008, January). The FTSE 100 and their pension disclosures. Technical report, Pension Capital Strategies.
- Cox, S. H. and Y. Lin (2007). Natural hedging of life and annuity mortality risks. *North American Actuarial Journal* 11(3), 1–15.
- Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Reviews* 106, 620–630.
- Jean-Baptiste, E. L. and A. M. Santomero (2000). The design of private reinsurance contracts. *Journal of Financial Intermediation* 9, 274–297.
- Juckett, G. (2006). Avian influenza: Preparing for a pandemic. *American Family Physician* 74(5), 783–790.
- Khintchine, A. Y. (1938). On unimodal distributions. *Izv. Nauchno Issled. Inst.Mat.Mech. Temsk. Gos. Univ.* 2(2), 1–7. In Russian.
- Krokhmal, P., J. Palmquist, and S. Uryasev (2002). Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of Risk* 4(2), 43–68.
- Lin, Y. and S. H. Cox (2005). Securitization of mortality risks in life annuities. *Journal of Risk and Insurance* 72(2), 227–252.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance* 7(1), 77–91.

⁷The terminal age in recent mortality tables is above 100. Here we just want to generate a reasonable time series of mortality margins for our optimization method. Our results will be qualitatively similar to that with a higher terminal age.

- Ombudsman (2008, July). Equitable life: a decade of regulatory failure part five: guide to the main report and summary of findings and recommendations. Technical report, The House of Commons, London.
- Parrillo, P. A. (2000, May). *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. Ph. D. thesis, Department of Control and Dynamical Systems, California Institute of Technology, Pasadena, CA.
- Popescu, I. (2005). A semidefinite programming approach to optimal-moment bounds for convex classes of distributions. *Mathematics of Operations Research* 30(3), 632–657.
- Prajna, S., A. Papachristodoulou, P. Seiler, and P. A. Parrilo (2004, June). *SOSTools Sum of Squares Optimization Toolbox for MATLAB Users Guide* (2.00 ed.). Pasadena CA 91125: Control and Dynamical Systems, California Institute of Technology. <http://www.cds.caltech.edu/sostools/>.
- Renshaw, A., S. Haberman, and P. Hatzopoulos (1996). The modeling of recent mortality trends in United Kingdom male assured lives. *British Actuarial Journal* 2(2), 449–477.
- Rockafellar, R. T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of Risk* 2(3), 21–41.
- Shannon, C. E. (1948, July, October). A mathematical theory of communication. *Bell System Technical Journal* 27, 379–423, 623–659.
- Tian, R. (May 2008). *Moment Problems with Applications to Value-at-Risk and Portfolio Management*. Ph. D. thesis, Georgia State University.
- Toole, J. (2007, May). Potential impact of pandemic influenza on the u.s. life insurance industry. Research report, Society of Actuaries.