Longevity Risk in Fair Valuing Level-Three Assets in Securitized Portfolios*

Peter M. Mazonas, CPA Life Settlement Financial, LLC

P.J. Eric Stallard, ASA, MAAA, FCA, Advisor to Life Settlement Financial, LLC

Lynford Graham, CPA, PhD Bentley University

e-mail: peter.mazonas@LSF-llc.com

ABSTRACT

Fair value accounting aims to establish a three-level hierarchy that distinguishes (1) readily observable measurement inputs from (2) less readily observable measurement inputs and (3) unobservable measurement inputs. Level 3 longevity valued assets will pose unique valuation risks once securitized pools of these alternative asset classes come to market as investment vehicles for pension plans and individual retirement accounts. No uniform framework is available to assure consistent fair market valuation and transparency for investor decision-making. Applying existing international auditing standards and analytical procedures (IFRS 13) will offer a platform upon which fund managers, their auditors, and actuaries can agree upon uniform valuation and presentation guidelines. Application of these quasi-governmental standards will bring future liquidity to otherwise illiquid capital market instruments. This paper presents a valuation methodology consistent with fair value accounting and auditing standards. The methodology incorporates the longevity predictive modeling of Stallard (2007) in a form that is compatible with Bayes Factor weighted average valuation techniques based on Kass and Raftery (1995). The methodology is applicable to fair valuation of life settlement portfolios where the combination of too few large death benefit policies and large variances in individual life expectancy estimates currently challenge accurate valuation and periodic re-valuation.

Keywords: Fair value; mark-to-market; longevity risk; IFRS 13; longevity cost calculator; level 3 asset valuation

INTRODUCTION

Level 3 longevity valued assets will pose unique valuation risks once securitized pools of these alternative asset classes come to market as investment vehicles for pension plans and individual retirement accounts. No uniform framework has yet been established to assure consistent fair market valuation and transparency for investor

decision-making. On May 12, 2011, the International Accounting Standards Board (IASB) and the U.S. Financial Accounting Standards Board (FASB) jointly issued International Financial Reporting Standard 13, (IFRS 13) requiring these assets to be accounted for and reported using fair value accounting.[1] IFRS 13 follows exposure drafts issued in May of 2010 which contain examples.[2][3] Auditors have also developed auditing standards that apply to assets that are fair valued. Application of these clarified auditing and accounting standards (IFRS 13 and AU §328[4]), analytical procedures (IFRS 13 and AU §329), and Accounting Standards Codification (ASC §820), will offer a platform upon which fund managers, their auditors, and actuaries can agree upon uniform valuation and presentation guidelines. Reconciling the variations in longevity estimates and applying these professional standards will increase confidence in the valuation of these securities, bringing further liquidity to these rapidly expanding capital market instruments.

FASB ASC §820.10 (superseding FAS 157 and 159) lays out the framework for the fair value hierarchy and prioritizes the inputs to valuation techniques. The measurement inputs (variables, assumptions, and methods used) are classified according to a three-level hierarchy that distinguishes (1) readily observable inputs from (2) less readily

---

*Presented at the Sixth International Longevity Risk and Capital Markets Solutions Conference, Sydney, Australia.

[1] On 12 May 2011 the International Accounting Standards Board (IASB) and the Financial Accounting Standards Board (FASB) issued new guidance on fair value measurement and disclosure requirements for International Financial Reporting Standards (IFRSs) and US generally accepted accounting principles (GAAP). The guidance, set out in IFRS 13 *Fair Value Measurement* and an update to Topic 820 in the FASB's Accounting Standards Codification® (formerly referred to as SFAS 157), completes a major project of the boards' joint work to improve IFRSs and US GAAP and to bring about their convergence. A summary is available at: http://www.iasplus.com/standard/ifrs13.htm

[2] On May 26, 2010, the Financial Accounting Standards Board (FASB) issued an Exposure Draft: *Proposed Accounting Standards Update—Accounting for Financial Instruments and Revisions to the Accounting for Derivative Instruments and Hedging Activities—Financial Instruments (Topic 825) and Derivatives and Hedging (Topic 815)*. The topic sections refer to the Board's new Accounting Standards Codification (ASC) and are included in ASC section 820.10, *Fair Value Measurement*. The exposure draft is available at: http://www.fasb.org/cs/ContentServer?site=FASB&c=Page&pagename=FASB%2FPage%2FSectionPage&cid=1176157086783

[3] Board Comment number 40 in the May 26, 2010 Exposure Draft states (p. 119): "The Board decided that life settlement contracts should be included in the scope of the proposed guidance. The Board observed that requiring fair value measurement would, in effect, eliminate the option to use the investment method described in Subtopic 325-30."

[4] AU 328 *Auditing Fair Value Measurement and Disclosures* superceded AU 342 *Auditing Accounting Estimates*. AU 328 expanded the auditor's responsibilities when fair values were involved. See Menelaides et al. (2003).

observable inputs and (3) unobservable inputs, with the overall level of the measurement (Levels 1 through 3) determined by the lowest level (i.e., highest numbered) input. Life settlements, annuities, and reverse mortgages are classified as Level 3 assets because their fair value is determined by an unobservable future event – the death of the insured, annuitant, or borrower.

The relevant AICPA/PCAOB rule-based standards (known as SAS 101) have been in effect since June 2003. The IASB standards are principle-based, but the thrust of the accounting and auditing standards vary only slightly and have the same intent of routinely marking portfolios to fair value while providing clear, understandable, and transparent investor reporting. The open question is one of implementation in the murky world of individual level life expectancy estimates. This issue is relevant for life settlements, life settlement derivative and leveraged debt products, annuities, and reverse mortgages.

Life expectancy estimates for individual persons are based on refinements of cohort life expectancy estimates, which require projections of future survival, with life expectancies defined as the areas under the respective survival curves. The cohorts are generally defined on the basis of sex and year of birth (or current age); further restrictions may be imposed, e.g., only including freshly underwritten lives with no impairments.

Life expectancy estimates for individual persons are subject to two types of error: (1) systematic error due to failure of the projections to accurately predict future survival rates in the respective cohorts; and (2) idiosyncratic error due to failure of the refinements of the cohort life expectancy estimates to accurately reflect the different mortality risks faced by different individuals.

Two properties will be useful in the following discussion. (1) Under the law of large numbers (LLN), given a sequence of independent, randomly selected samples of increasing size in a given cohort, the corresponding sequence of average life expectancies will converge to the cohort life expectancy. (2) Under the assumption that the sampling is random, the expected value of the individual life expectancies in each sample will be the same value, viz., the cohort life expectancy.

The relevant life expectancies for the full cohort and the various samples could be computed using the Kaplan-Meier estimator[5] of the survival curve if one were willing to wait until the last member of each sample and ultimately the entire cohort (assuming that it is sufficiently large) had died. For valuation purposes, this would be impractical;

---

[5] Kaplan and Meier (1958).

hence, it is necessary to use estimates of cohort and individual life expectancies based on the information available at the time of valuation.

This means that the systematic errors noted above will be present and that one cannot rely on the LLN to support the assumption that the average life expectancy values in increasingly large samples will converge to the cohort life expectancy estimates; they will converge to some value, but it may be very different from the one assumed[6]. This was the main point of the paper by Milevsky et al.[7] who showed how this type of error can be accommodated in a stochastic pricing model based on the Sharpe ratio. Bauer et al.[8] considered alternative approaches to this problem, comparing the Sharpe ratio with the Wang transform and option-type longevity derivatives, which they favored.

Importantly, Milevsky et al.[9] emphasized that stochastic mortality risk is different from model specification risk; for the purposes of exposition, they assumed that the mortality model was specified precisely, without error. This assumption, however, was not recommended for use in applied practice.

Stallard[10] argued that, due to the very large number of mortality risk factors and the variations in the timing of their effects, it is highly unlikely that any mortality model can be specified without error and the potential impact of model misspecification due to the omission of one or more salient factors is substantial.

Wang and Preston[11] showed that the omission of just one factor, cigarette smoking, was sufficient to generate a 22.5% underestimate in the projected probability of surviving from age 50 to age 85 for U.S. males based on two versions of the Lee-Carter model, one with and the other without smoking. Similarly, Reither et al.[12] argued that the omission of birth cohort in the standard Lee-Carter model fails to capture the effects of obesity which may lead to slow-downs or reversals in the projected rates of mortality improvement.

Purushotham et al.[13] provided best-estimate mortality improvement rates by age (in the age range 25–100+ years), sex, smoking status, and type of insurance (individual life

---

[6] Olivieri (2001).
[7] Milevsky et al. (2006).
[8] Bauer et al. (2010).
[9] Milevsky et al. (2005).
[10] Stallard (2006).
[11] Wang and Preston (2009).
[12] Reither et al. (2011).
[13] Purushotham et al. (2011).

vs. annuity), but without adjustments for obesity. The differences were substantial, e.g., ranging from 1.0% to 2.0% per year at age 75–84.

Granieri[14] compared the 10-year survival of approximately 52,000 life settlement applicants (LSAs) in the 21st Services, LLC database (aged 70+ years; with an average age of 78 years at the time of application) with the corresponding expected insurance experience based on the 2008 Valuation Basic Table (VBT)[15], with the latter tables matched by age, sex, and smoking status. The LSA survival curve was substantially lower than the VBT survival curve, falling close to mid-way between the VBT and the matched U.S. general population survival curves. For example, the LSA survival curve reached the 80% level at 65 months, compared to 81 months for the VBT and 43 months for the U.S. general population survival curves.

The LSA population differs from the life settlement transaction (LST) population in that the latter is a select nonrandom subset of the former, with the selection based, in part, on the buyer's assessment of the insured's survival prospects. Thus, one would expect the LST survival curve to fall materially below the LSA curve, so that the use of the LSA curve could be a serious model misspecification error, with the use of the VBT curve constituting an even more egregious error. The LST population is the appropriate reference population for life settlement valuations, and it is the source of the cohorts whose survival will be the targets of convergence for the small-sized life settlement portfolios under present consideration.

The remainder of the paper focuses on idiosyncratic errors due to inaccurate representations of the individual mortality risks and survival curves. This is a critical problem in the valuation of life settlement portfolios—and one which has received little attention in the actuarial literature. Resolution of the issues related to idiosyncratic errors is an essential step in the development of a fully integrated model that simultaneously addresses all of the sources of error identified above.

We begin by noting that the Kaplan-Meier estimator only works well for groups of individuals; for any single individual, the Kaplan-Meier estimator of life expectancy is equal to the individual's observed survival time, which is conceptually distinct from the individual's life expectancy at the time of valuation. The problem is that an individual's life expectancy is an inherently unobservable quantity, which makes it necessary to use some form of mathematical model in order to produce a valid estimator.

---

[14] Granieri (2011).
[15] SOA (2008).

A real world example illustrates the variations between commercial life expectancies (LEs) used in the life settlement marketplace to price purchase offers. These individual insured's LEs were the output from trained professionals from different firms all interpreting the same insured's medical records. They applied positive and negative factors and weightings to the items/issues in the medical records to develop a multiplier. This multiplier was applied to a single or multiple modified large population life tables based upon the insured's medical conditions to derive the estimated life expectancy for that single individual. These life expectancies and their related survival curves were then used to price the bundled policies.

From data compiled by the authors since January 2009, where three or more commercial LEs were prepared from each insured's medical records, the range between the low LE and the high LE per insured, relative to the low LE, averaged 31%[16]. For the same data, where the relative range between the low and the high LE was greater than 30%, the average size was 56%[16]. Therein lay the problem for individual policy pricing and portfolio valuation: the variability of the individual LEs was so large that their validity was highly suspect.

Our challenge was to craft a valid alternative fair valuation methodology that would satisfy the fund managers, auditors, and regulators and provide transparency to investors. In November 2009, one of the authors (Mazonas) was invited to testify before the SEC on the fair valuation of Level 3 assets, specifically life settlements. This request grew out of paper by Mazonas and Stallard, previously submitted to the SEC's Life Settlement Task Force[17]. The SEC's concern was the lack of standardized methodology for fair valuation and the historic lack of transparency. Their concern grew out of the proliferation of securitized life settlement financial instruments being sold or being designed to be sold to pension plans as rated and unrated portfolios of policies[18]. What was presented to the SEC and is described in the current paper is a careful melding of accounting and auditing standards, actuarial science, Bayesian statistics, and a relatively new, peer reviewed approach to predicting survival/mortality in single individuals. The goal was to reconcile the substantial variation in individual life

---

[16] Results are based upon data compiled from settlement brokers and tabulated by the lead author. EMSI, AVS and others have conducted studies that yield similar results.
[17] Available at
https://www.lifesettlementfinancial.com/pdf/Life%20Settlement%20Porfolio%20Valuation%20for%20Securitization%20LSF%202Nov09.pdf
[18] On July 22, 2010, the SEC's cross-divisional Life Settlement Task Force issued recommendations that the Commission recommend to Congress that it amend the Definition of Security under the Federal Securities Law to include life settlements. It further recommended Federal regulation of settlement brokers and provider to ensure legal standards of conduct are being met. It also called for consistent regulation of life expectancy underwriters.
http://www.sec.gov/news/studies/2010/lifesettlements-report.pdf

expectancy estimates to establish an ongoing basis for fair valuation of individual portfolios.

LONGEVITY COST CALCULATOR AS A POLICY AND PORTFOLIO VALUATION TOOL

Background

A life settlement is the purchase of an individual life insurance policy from the policy owner by a third party for a lump sum payment in cash. The insured in a typical transaction is over the age of 70 and has an average life expectancy of between 60 and 154 months. Most often these are universal life policies that are no longer needed and have grown too expensive to maintain the premium payments. The new policy owner must make all future policy premium payments to keep the policy in force and collects the death benefit at the death of the insured. The amount paid for the policy is the net present value of the stochastically determined future premium outflow subtracted from the net death benefit the new owner expects to receive.

In the summer of 2007, Life Settlement Financial (LSF) initiated work on a computer model for life settlement valuation which, in addition to provisions for accommodating three or more commercial LEs, also included LE calculations for the model described by Stallard[19]. LSF wrote and validated the new life settlement valuation model in robust web-based code and named it the Longevity Cost Calculator (LCC). In mid 2009, LSF launched this model as a settlement policy selection tool that provides low cost individual level survival functions based on "Grade of Membership" scores generated from data gathered in telephonic interviews with the insured or a proxy family member. The interviewee has no prior knowledge of the questions and experience has shown it is virtually impossible to game the system. Experience has shown that over thousands of LCC assessments seniors or their proxies tend to optimistic about their health, physical and mental conditions. These interviews are paid for by the life settlement brokers and represent a significant cost savings over how they have processed prospective settlement policies in the past. The LCC questionnaire-derived LE estimate remains attached to each policy irrespective of whether the funders purchase it or not. On-going this broad application of LCC assessments will facilitate the valuation of newly originated life settlement portfolios.

In addition to providing an LE estimate, the completed LCC questionnaire scores each insured using a four level Grade of Membership (GoM) system. The interrelationship of an insured's activity of daily living (ADL) impairments, instrumental activity of daily living

---

[19] Stallard (2007).

(IADL) impairments, and possible cognitive impairment affect those GoM scores and the trajectory of the individual's survival curve used to price a settlement offer. To quantify the health changes, the model generates both time varying and time invariant GoM scores, where the time invariant GoM scores characterize the predicted health status of each person at the time they are/were in the youngest age-group in the model, which for the current implementation was age-group 65–69. In both cases, the interpretation of the GoM scores is invariant over time. In the following, we describe the health status characteristics summarized by the GoM scores, and then provide a more formal description of the GoM model.

Characteristics of GoM Pure Types

GoM 1 (also referred to as "Pure Type I" or "Type I", with the Arabic or Roman numerals designating the rank ordering of the states by health status) refers to the healthiest component of the population. GoM 2, 3, and 4 capture a range of health problems that occur at different ages, with progressive and graded transitions from GoM 2 to GoM 3 and 4. GoM 2 refers to persons who have numerous medical problems, but few, if any ADL or other functional problems, or cognitive impairment.
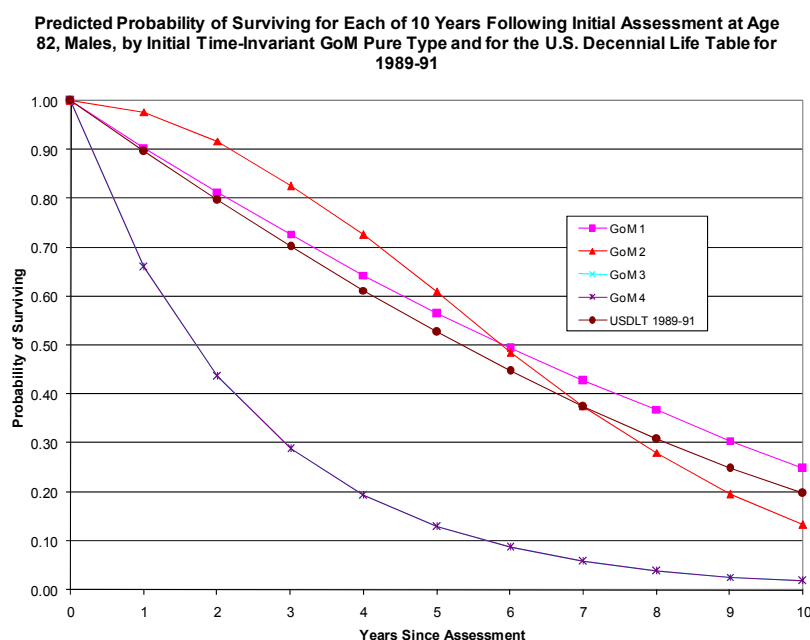
An important aspect of the model is that the individual GoM scores are fractional and sum to 1.0 when added over all four components. This is true for both the time varying and time invariant GoM scores, which implies that the deterioration in health status over age can only be represented by a reduction in GoM 1 and compensating increases in some combination of GoM 2, 3, and 4.

Persons with initial strong scores (i.e., close to 1.0, or 100%) on GoM 2 will live longer than traditional LE providers estimate; although this changes at older ages where these persons exhibit transitions to strong scores on GoM 4. Persons with initial strong scores on GoM 3 have minor medical problems, but mild/moderate cognitive impairments, usually not indicated in their medical records, although this also changes at older ages where these persons exhibit transitions to strong scores on GoM 4. Strong scores on GoM 4 identify people who have more serious medical problems, combined with serious ADL and/or cognitive impairments, also which are not usually indicated in their medical records. Persons with initial strong scores on GoM 3 and GoM 4 have shorter LE's than those issued by traditional LE underwriters.

Recall that the LE is the area under the relevant survival curve for the person or population for which the LE is being calculated. Thus, differences in LE between persons or groups of persons are best understood by examining the associated survival curves. This is illustrated in Figure 1 which displays plots of the predicted survival curves for males assessed at age 82 (i.e., age at last birthday is 82; the midpoint of the five year age interval 80–84 years used in model estimation) for the next 10 years

following the assessment, with a separate curve shown for each of the four time invariant GoM pure types and also for comparison the survival curve from the U.S. Decennial Life Table (USDLT) for 1989–91[20].

Figure 1

Predicted Probability of Surviving for Each of 10 Years Following Initial Assessment at Age 82, Males, by Initial Time-Invariant GoM Pure Type and for the U.S. Decennial Life Table for 1989-91

According to the USDLT, the male LE at age 82 was 6.2 years. This value was less than the LE's of 6.9 and 6.3 years for GoM 1 and 2, but was substantially higher than the LE's of 2.5 years each for GoM 3 and 4 (these LE-values were the same because an initial GoM 3 "converted" to GoM 4 prior to age 82).
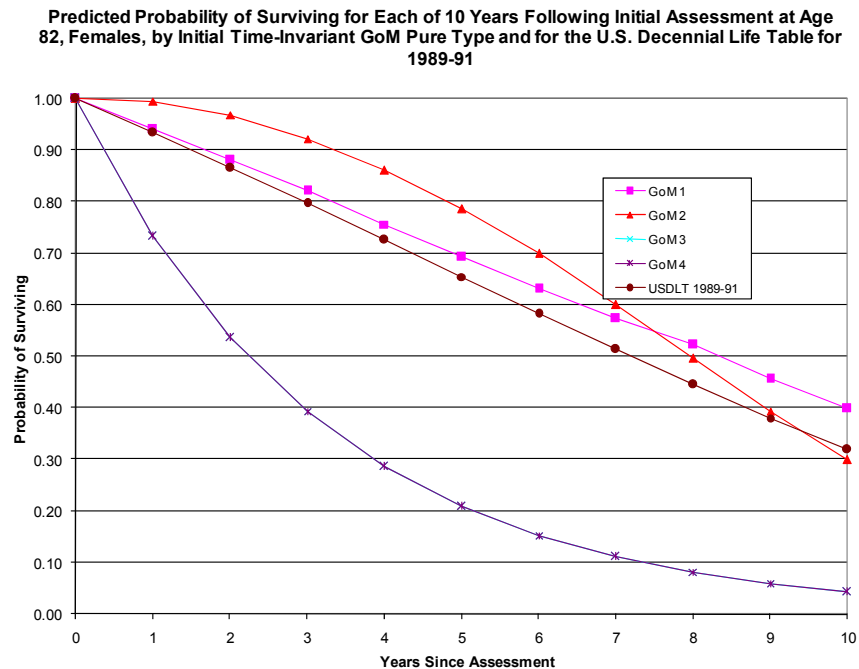
Given that both the survival curve and its slope are used to price an offer or value a component policy within a life settlement portfolio (the former for premium costs, the latter for death benefit offsets), it is important to accurately estimate these quantities. Armed with this knowledge, an underwriter could price an offer to outbid the competition, win the policy, but not pay the full price indicated by alternative commercially available proprietary pricing models. The same pricing differential would be manifest as a valuation error in the life settlement portfolio application.

In Figure 1, an initial GoM 2 had approximately a 16% greater likelihood of surviving through years 1 through 5 until the lines converged at year 6, near the 6.2 year LE. On the other hand, an initial GoM 3 or 4 had approximately a 75% lower likelihood of surviving through years 1 through 5, with corresponding reductions for persons who had initial *fractional* scores on GoM 3 or 4, with complementary *fractional* scores on GoM 1

---

and/or 2. Again, we note that the sum of all four fractional scores must equal 1.0 (100%), with the GoM scores for any given individual derived from his/her answers to up to 76 questions on the web-based or telephonic interview (selected from 95 questions in the original analysis; see Stallard)[19].

Figure 2

Predicted Probability of Surviving for Each of 10 Years Following Initial Assessment at Age 82, Females, by Initial Time-Invariant GoM Pure Type and for the U.S. Decennial Life Table for 1989-91

The corresponding plots (Fig. 2) for females aged 82 years at assessment display similar patterns, but with somewhat longer LE's, 7.8 years in the USDLT, which was less than the LE's of 9.2 and 8.2 years for GoM 1 and 2, but was substantially higher than the LE's of 3.2 years each for GoM 3 and 4 (as for males, the values were the same because an initial GoM 3 "converted" to GoM 4 prior to age 82).

For both sexes, the differences in the survival curves and their slopes illustrated the potential for substantial overpayment for a policy or for error in valuation without this additional knowledge regarding the level and slope of the relevant survival curves.

The 76 predictor variables in the LCC assessment questionnaire were ranked in order of importance by the ratio chi-squared/d.f. It was not until the 20th of the 76 predictors that one found a medical condition/diagnosis of the form typically reported on medical records[21]. The best predictors involved ADL/IADL impairments and functional limitations, which do not have ICD-9-CM codes and, hence, are typically not reported on medical records.

---

[21] See Stallard (2007, pp. 49-50).

The web-based LCC model replicates the original peer reviewed model of Stallard[19]. The mortality component of that model was calibrated using 120,832 male person-years and 196,270 female person-years of exposure data over an 18-year period 1984–2002, commencing with the 1984 National Long Term Care Survey (NLTCS). The annual mortality probabilities were based on 20,428 deaths (8,583 males and 11,845 females) among 32,389 participants in the survey (12,974 males and 19,415 females) in the 18-year period.

Formal Description

GoM provides a general approach to the analysis of longitudinal data with categorical variables, where the individual participants are subject to censoring (or loss) due to mortality (or other endpoint events). The approach can be understood by anyone who is familiar with matrix multiplication including the dot product multiplication of vectors.

We begin by defining the K-element vector $\mathbf{g}_i$ to be the column vector of initial GoM scores for the ith individual in the study sample, which is assumed to contain I such individuals. As noted above, $K=4$ in the current application. By convention, we use lower case k to index the elements g_{ik} of $\mathbf{g}_i$, which are subject to the constraints that $g_{ik} \geq 0$ for every k where the sum over k of g_{ik} is exactly 1.0. We use boldface regular font to denote vectors and matrices, with $\mathbf{g}_i^T$ denoting the K-element transpose of $\mathbf{g}_i$.

The initial, time invariant GoM scores are used in our application to characterize the predicted health status of each person at the time they are/were aged 65–69, the youngest age-group in our model.

To represent the effects of aging, we define the K-element vector $\mathbf{g}_{it}$ to be the column vector of time varying GoM scores for the ith individual, and $\mathbf{V}_t$ to be the $K \times K$ upper triangular transition matrix that maps $\mathbf{g}_i$ to $\mathbf{g}_{it}$, where t is used to denote age (or time):

$$\mathbf{g}_{it}^T = \mathbf{g}_i^T \mathbf{V}_t . \tag{1}$$

The form of $\mathbf{V}_t$ was restricted to upper triangular in order to force the GoM scores for the lower numbered pure types to decline over age with compensating increases for the higher numbered pure types. As a result, the most common pure type at the oldest ages was GoM 4 (e.g., see Stallard,[19] Figures 7 and 8). Note, however, that except for the global constraints of nonnegativity and summation to 1 we placed no further restrictions on the values of the initial GoM scores for individuals. As a result, the most common pure type at age 65-69 was GoM 1, with 40% scoring 1 on this type. However, nearly 12% of the sample scored 1 on GoM 3, 2% scored 1 on GoM 2, and 0.5% scored 1 on GoM 4; with the rest of the sample having fraction scores on the various pure types (Stallard, 2007, Table 4).[19]

The form of $\mathbf{V}_t$ was further restricted so that the changes in $\mathbf{g}_{it}$ from one time to the next were represented as:

$$\mathbf{g}_{it+1}^T = \mathbf{g}_{it}^T \mathbf{U}_t , \qquad (2)$$

where $\mathbf{U}_t$ is the $K{\times}K$ upper triangular transition matrix that maps $\mathbf{g}_{it}$ to $\mathbf{g}_{it+1}$.

Equations (1) and (2) are written without an error term. Thus, the GoM score transitions from one time to the next are deterministic and the model can be described as a "fixed-effect" model.

The initial time invariant GoM scores, $\mathbf{g}_i$, are determined jointly with the $\mathbf{V}_t$ matrices using maximum likelihood estimation procedures based on the selected predictor variables in the longitudinal data files used for model calibration. Once the model has been calibrated, the initial time invariant GoM scores for new individuals can be estimated conditional on the $\mathbf{V}_t$ matrices. Both procedures are described in Stallard[19].

In the current application, we use the GoM model to estimate individual mortality risks and survival curves. Let p_{it} denote the conditional probability that an individual who is alive at age t continues to survive to at least age $t + 1$; and let $q_{it} = 1 - p_{it}$ denote the complementary conditional probability of death. Let π_{kt} and μ_{kt} denote the corresponding conditional survival and mortality probabilities which are defined, not for individuals, but for each of the K pure types. Let $\boldsymbol{\pi}_t$ and $\boldsymbol{\mu}_t$ denote the K-element vectors containing respectively the elements π_{kt} and μ_{kt} for $k = 1, \ldots , K$. The GoM survival model can be written as

$$p_{it} = \mathbf{g}_{it}^T \boldsymbol{\pi}_t \qquad (3)$$

and the GoM mortality model as

$$q_{it} = \mathbf{g}_{it}^T \boldsymbol{\mu}_t . \qquad (4)$$

The estimates of the π_{kt} and μ_{kt} can also be obtained using the maximum likelihood estimation procedures in Stallard (2007). One additional manipulation substitutes equation (1) into (3) and (4), yielding

$$p_{it} = \mathbf{g}_i^T \mathbf{V}_t \boldsymbol{\pi}_t = \mathbf{g}_i^T \boldsymbol{\pi}_t^* \qquad (5)$$

$$q_{it} = \mathbf{g}_i^T \mathbf{V}_t \boldsymbol{\mu}_t = \mathbf{g}_i^T \boldsymbol{\mu}_t^* , \qquad (6)$$

where the composite vectors $\boldsymbol{\pi}_t^*$ and $\boldsymbol{\mu}_t^*$ are K-element vectors of "adjusted" conditional survival and mortality probabilities, respectively. This adjustment yields the identical predicted values for the individual probabilities. However, it makes the predicted

probabilities functions of the initial time invariant GoM scores, instead of the time varying GoM scores.

Individual marginal (cumulative) survival functions are obtained by multiplying the individual conditional survival probabilities over the relevant age range. For example, the function for the marginal survival of individual i beyond age t_0, to any later age t, is generated as:

$$S_{it} = \prod_{\tau=t_0}^{t-1} p_{i\tau}, \qquad\qquad (S_{it_0} = 1) \qquad\qquad (7)$$

where p_{it} is generated from (3) or (5). These calculations were illustrated in Figures 1 and 2 with $t_0 \approx 82.5$ years and $p_{it} = \pi_{kt}^*$, for $k = 1, \dots, K$.

Results

Table 1 presents the estimated age-specific conditional probabilities of death for the four GoM pure types, separately by sex, for use with the time varying and time invariant GoM scores, denoted by the headings "unadjusted" and "adjusted" probabilities, respectively. The unadjusted and adjusted parameters for GoM 4 (Type IV) were identical because GoM 4 is an absorbing state which, from Figures 1 and 2, can be seen to have the poorest survival. The unadjusted probabilities for GoM 1, 2, and 3 were, pairwise, lower than the corresponding adjusted probabilities.

The entries in the "total" rows for the unadjusted probabilities were the estimates of the single parameter that best summarized (in the sense of maximum likelihood estimation) the age-specific values in the columns above each entry. The unadjusted probabilities were relatively flat over the age dimension, indicating that the age increases in mortality risk were represented by the increases in the GoM scores over age. The instability in parameter estimates for the unadjusted probabilities beginning at ages 95–99 and older was attributable to the substantially smaller sample sizes at these ages (Stallard, 2007, Table 7).[19] The entries in the "total" rows for the adjusted probabilities were the exposure-weighted averages of the age-specific parameter estimates in the columns above each entry.

Table 1

Unadjusted and Adjusted Probabilities of Death within One Year in Four Pure-Type GoM Model, by Sex and Attained Age at Time of Exposure

Exposure Age	Males Unadjusted Probability by Type				Males Adjusted Probability by Type			
	I	II	III	IV	I	II	III	IV
65-69	0.000	0.000	0.132	0.138	0.000	0.000	0.132	0.138
70-74	0.002	0.005	0.194	0.246	0.002	0.005	0.194	0.246
75-79	0.000	0.013	0.178	0.319	0.038	0.013	0.319	0.319
80-84	0.000	0.023	0.184	0.340	0.095	0.023	0.340	0.340
85-89	0.000	0.000	0.192	0.330	0.127	0.202	0.330	0.330
90-94	0.000	0.204	0.205	0.323	0.198	0.323	0.323	0.323
95-99	0.503	1.000	0.076	0.434	0.226	0.434	0.434	0.434
100-104	—	0.000	0.325	0.528	0.372	0.528	0.528	0.528
Total	0.001	0.007	0.180	0.314	0.041	0.033	0.253	0.271

Exposure Age	Females Unadjusted Probability by Type				Females Adjusted Probability by Type			
	I	II	III	IV	I	II	III	IV
65-69	0.000	0.000	0.081	0.140	0.000	0.000	0.081	0.140
70-74	0.001	0.003	0.108	0.223	0.001	0.003	0.108	0.223
75-79	0.000	0.003	0.111	0.249	0.018	0.003	0.249	0.249
80-84	0.000	0.005	0.116	0.267	0.059	0.005	0.267	0.267
85-89	0.000	0.000	0.128	0.271	0.089	0.110	0.271	0.271
90-94	0.000	0.000	0.182	0.272	0.127	0.272	0.272	0.272
95-99	—	0.446	0.000	0.388	0.168	0.388	0.388	0.388
100-104	—	0.000	0.148	0.499	0.274	0.499	0.499	0.499
Total	0.001	0.002	0.117	0.280	0.036	0.037	0.201	0.239

Note: "—" denotes cells for which probablilites could not be estimated; see text for discussion of
Source: Stallard (2007)

Figure 3

Observed and Average Predicted Probabilities of Death, Males, by 5-Year Age Groups

Figure 4

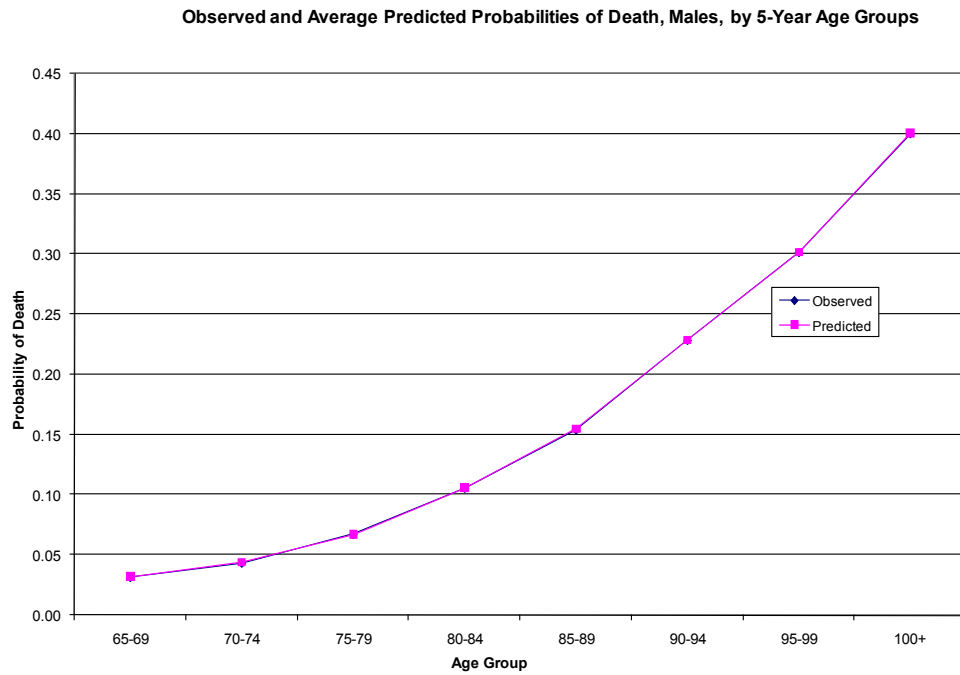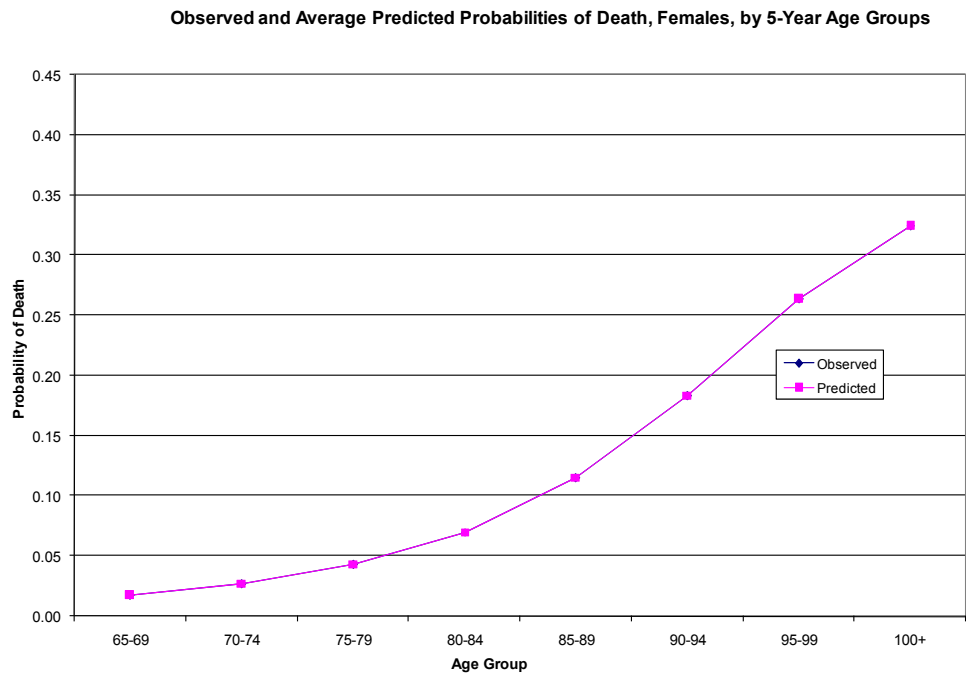**Observed and Average Predicted Probabilities of Death, Females, by 5-Year Age Groups**

Figures 3 and 4 display the plots of the observed and average predicted age-specific probabilities of death for males and females, respectively. The average predicted probabilities were the mean values of the predicted probabilities for the individual

respondents to the NLTCS based on equation (3). The observed probabilities were the ratios of the counts of the age-specific number of deaths divided by the corresponding number of respondents alive at the beginning of each 1-year observation period. The differences between the observed and average predicted probabilities were very small and were statistically nonsignificant, with chi-squared values of 1.36 and 0.29, respectively, each with 6 d.f. using the approximations in Hosmer and Lemeshow[22] to count degrees of freedom (reference chi-squared values were 12.59 and 16.81 at the conventional 5% and 1% significance levels).

Credibility

Before presenting additional measures of accuracy, it will be useful to consider the random statistical fluctuations expected from estimates based on the numbers of events likely to be observed in different sized samples, assuming fixed underlying event rates. Table 2 extends the results of Longley-Cook[23] to show the minimum number of expected events needed to meet various credibility standards for accuracy and reliability, with relative accuracy defined by the ratio:

R = (observed no. of deaths − expected no. of deaths) / expected no. of deaths.   (8)

Table 2

Credibility and Event Counts

Maximum Acceptable Departure from the Expected Count	Probability of Observed Count Falling Within the Acceptable Range		
	90%	95%	99%
	Minimum Required Expected Count		
+/−2.5%	4,329	6,146	**10,616**
+/−5.0%	**1,082**	1,537	2,654
+/−7.5%	481	683	1,180
+/−10%	271	384	663
+/−20%	68	**96**	166
+/−30%	30	43	74
+/−40%	17	24	41
+/−50%	**11**	15	27

Source: Based on Longley-Cook (1962).

To be 99% confident that the maximum relative error is less than 2.5%, the sample size needs to be large enough to produce 10,616 deaths (boldface font in Table 2). To be 90% confident that the maximum relative error is less than 5%, the sample size needs to be large enough to produce 1,082 deaths, which is the standard size used for full

---

[22] Hosmer and Lemeshow (1980).
[23] Longley-Cook (1962).

credibility in the actuarial literature. To be 95% confident that the maximum relative error is less than 20%, the sample size needs to be large enough to produce 96 deaths. To be 90% confident that the maximum relative error is less than 50%, the sample size needs to be large enough to produce 11 deaths. Thus, as the expected number of deaths falls, after rounding, from 10,000 to 1,000 to 100 to 10, the relative error increases from about 2.5% to 50%.

Practical considerations often dictate sample sizes less than that needed for full actuarial credibility. Table 2 indicates that sample sizes of 271 and 384 can yield relative errors of ±10% at the 90% and 95% probability levels, respectively, consistent with A.M. Best's recommendation that the collateral pool for life settlement portfolios consist of at least 300 lives[24].

Additional Measures of Accuracy

Random statistical fluctuations are inherently unpredictable. Hence, our assessment must focus on our ability to generate accurate values for the expected number of deaths among any selected set of insured lives. These values are computable as the sums of the individual predicted probabilities of death for the individuals comprising each set.

Figures 5 and 6 display the plots of the observed and average predicted probabilities of death with the study sample grouped into 10 categories according to the individual predicted probabilities of death, using fixed cutpoints at multiples of .05 (5%), separately for males and females. Chi-squared statistical tests of fit of the models are presented separately in Tables 3 and 4.

Visually, one can see that the observed probabilities increased across the 10 categories for both sexes, except for category 9 for females. However, the chi-squared test of the deviation for that one point indicated that the difference was statistically nonsignificant, with a chi-squared value of 3.44 with 1 d.f. (Table 4; reference values are 3.84 and 6.63 at the conventional 5% and 1% significance levels).

The Hosmer-Lemeshow chi-squared test produced statistically significant total chi-squared statistics of 155.65 and 62.99, respectively, for males and females, each with 8 d.f. (reference values are 15.51 and 20.09 at the conventional 5% and 1% significance levels).

Several comments are in order:

1. The tests indicated that the models displayed in Figures 5 and 6 did *not* fit the data. This means that at least some of the deviations of the observed from predicted numbers of deaths were larger than expected by chance.

---

[24] Modu (2008).

Figure 5

Observed and Average Predicted Probabilities of Death, Males, by Predicted-Probability Class Intervals with Cutpoints at Multiples of 5%
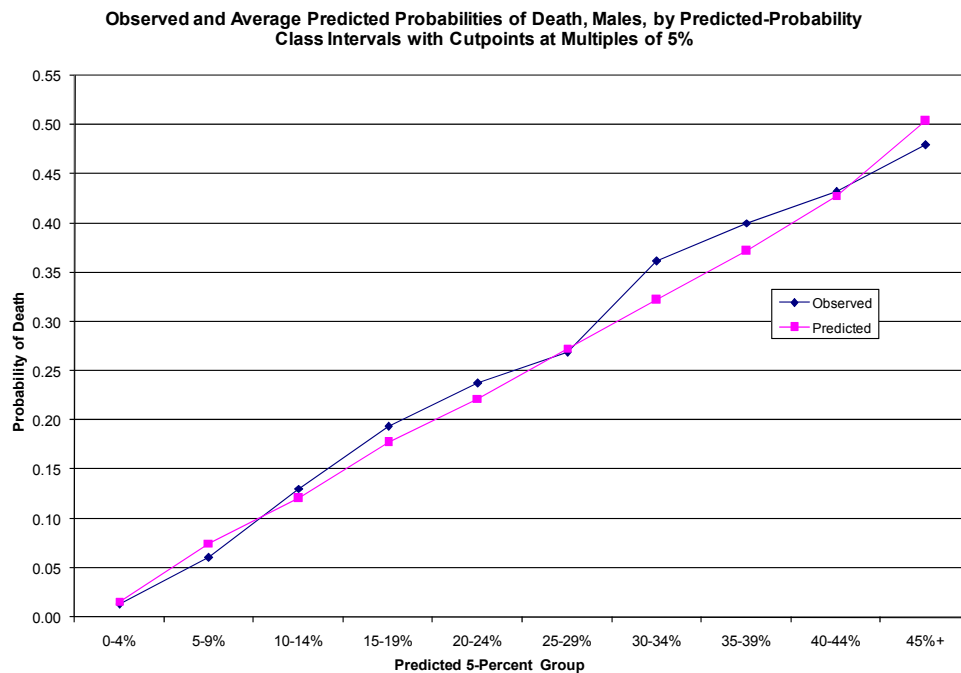

Figure 6

Observed and Average Predicted Probabilities of Death, Females, by Predicted-Probability Class Intervals with Cutpoints at Multiples of 5%
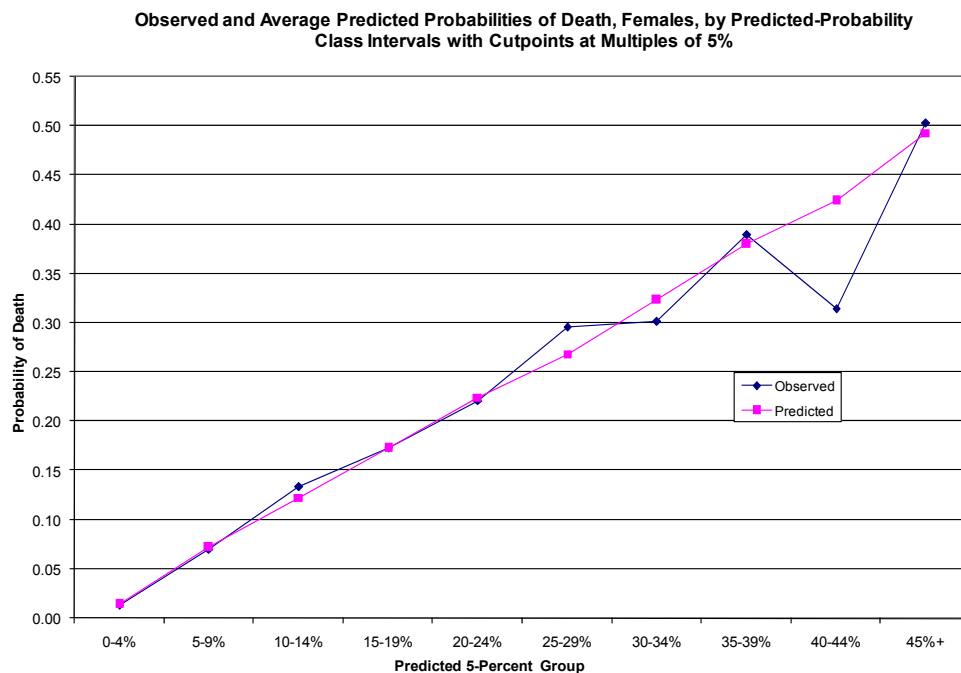
Table 3

Observed and Average Predicted Probabilities of Death, Males, by Predicted-Probability Class Intervals with Cutpoints at Multiples of 5%

Percentage Group	Number of Person-Years at Risk	Observed Number of Deaths	Expected Number of Deaths	Observed Probability	Average Predicted Probability	Hosmer-Lemeshow Chi-Squared
0-4%	61,463	801	981	0.013	0.016	**33.60**
5-9%	23,256	1,407	1,747	0.061	0.075	**71.73**
10-14%	20,100	2,622	2,420	0.130	0.120	**19.20**
15-19%	7,705	1,490	1,363	0.193	0.177	**14.29**
20-24%	4,557	1,082	1,011	0.237	0.222	6.44
25-29%	2,095	563	571	0.269	0.273	0.16
30-34%	1,361	492	438	0.361	0.322	**9.73**
35-39%	115	46	43	0.400	0.372	0.39
40-44%	132	57	56	0.432	0.428	0.01
45%+	48	23	24	0.479	0.504	0.11
Total	120,832	8,583	8,655	0.071	0.072	**155.65**

Table 4

Observed and Average Predicted Probabilities of Death, Females, by Predicted-Probability Class Intervals with Cutpoints at Multiples of 5%

Percentage Group	Number of Person-Years at Risk	Observed Number of Deaths	Expected Number of Deaths	Observed Probability	Average Predicted Probability	Hosmer-Lemeshow Chi-Squared
0-4%	111,425	1,434	1,614	0.013	0.014	**20.33**
5-9%	44,837	3,124	3,239	0.070	0.072	4.42
10-14%	18,004	2,396	2,202	0.133	0.122	**19.46**
15-19%	10,245	1,770	1,770	0.173	0.173	0.00
20-24%	6,536	1,443	1,465	0.221	0.224	0.43
25-29%	3,413	1,009	915	0.296	0.268	**13.17**
30-34%	617	186	199	0.301	0.323	1.32
35-39%	914	356	347	0.389	0.380	0.35
40-44%	70	22	30	0.314	0.424	3.44
45%+	209	105	103	0.502	0.493	0.08
Total	196,270	11,845	11,885	0.060	0.061	**62.99**

2. These were identified by the boldface font in the rightmost columns of Tables 3 and 4 using a cutpoint equal to the critical value of 6.63 using the conventional 1% significance level.
3. For males, the five significant deviations were for the four lowest probability groups in the range 0–19%, and 30–34%; for females, the three significant deviations were for the 0–4%, 10–14%, and 25–29% probability groups.
4. The expected counts for the eight groups with significant deviations ranged from 438 to 2,420, and five of the eight exceeded the 1,082 cutpoint for highly credible data in Table 2; all eight exceeded the minimum level of 300 lives recommended by A.M. Best[24] for life settlement collateral pool sizes.

We concluded that there were factors operating in these data that were not represented in our model. This should not be surprising given that the model used four GoM scores to summarize data on the original set of 95 questions concerning medical conditions, activities of daily living (ADLs), cognitive and behavioral impairments. Moreover, there may be other influential factors not included in the set of 95 questions. Given a sufficiently large sample one would expect to identify significant deviations from *any model* using the statistical procedures described above.

Two additional comments provide additional perspective:

1. Only two groups had significant deviations for both sexes: the 0–4% and 10–14% probability groups. This suggested that the deviations from the models were not predictable, at least at the higher probability levels which are of greatest concern to life settlement providers.
2. Nonrandom deviations in the GoM models can be tolerated if they are sufficiently small, relative to the errors that would occur in the absence of the GoM models.

To quantify the size of the nonrandom deviations, we applied linear regression analysis with the observed probabilities regressed on the average predicted probabilities, which are shown in Tables 3 and 4, obtaining R-squared values of 0.985 and 0.942, respectively.

The R-squared value for the combined data was 0.961 which may be interpreted as a measure of the accuracy of the GoM models: 96% of the variance of the observed probabilities was accounted for by the expected probabilities produced by the GoM models. The remaining 4% of the variance constituted a tolerable level of nonrandom deviation in the GoM models.

We considered the possibility that the linear regression analysis may not fully represent the impact of small deviations at the lower probability levels in Tables 3 and 4. This was motivated in part by the chi-squared tests which indicated that most of these deviations were statistically significant. To deal with this issue, we generated a second set of regressions with the logarithms of the observed probabilities regressed on the logarithms of the average predicted probabilities, obtaining R-squared values of 0.994 and 0.989, respectively.

The R-squared value for the combined data was 0.990 which may be interpreted as an alternative measure of the accuracy of the GoM models: 99% of the variance of the logarithm of the observed probabilities was accounted for by the logarithm of the expected probabilities produced by the GoM models. The remaining 1% of the variance constituted an even more tolerable level of nonrandom deviation in the GoM models.

Two additional questions were important to our assessment of the accuracy of the model.

The *first* question was whether the GoM scores added any significant information beyond the information already available using the age-specific mortality probabilities displayed in Figures 3 and 4; and if so, how much? This question can be directly addressed using the log–likelihood-ratios for the four sex-specific models listed in Table 5 to generate the corresponding AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) statistics typically used for model assessment.

Table 5

Log-Likelihood and Related Statistics for Testing the Joint and Separate Contributions of the Age-Specific and the GoM-Specific Probabilities of Death, by Sex

# Model Description	Log-Likelihood-Ratio	d.f.	AIC	BIC	ΔAIC	ΔBIC
Males						
1 Constant Probability	0.00	1	2.00	9.06	10,878.66	10,659.87
2 Age-Specific Probabilities (no GoM)	1,632.10	8	-3,248.19	-3,191.73	7,628.46	7,459.08
3 GoM-Specific Probabilities (no Age)	5,278.87	4	-10,549.75	-10,521.52	326.91	129.30
4 Age&GoM-Specific Probabilities	5,470.33	32	**-10,876.66**	**-10,650.82**	0.00	0.00
Females						
1 Constant Probability	0.00	1	2.00	9.38	16,276.04	16,047.27
2 Age-Specific Probabilities (no GoM)	3,776.30	8	-7,536.59	-7,477.56	8,737.45	8,560.34
3 GoM-Specific Probabilities (no Age)	7,873.29	4	-15,738.58	-15,709.06	535.46	328.83
4 Age&GoM-Specific Probabilities	8,169.02	32	**-16,274.04**	**-16,037.89**	0.00	0.00

Model 1 was the simplest model. It assumed that the sex-specific mortality probability was constant over age and GoM scores. Model 2 assumed that the sex-specific mortality probabilities increased over age but not over GoM scores, following the observed values displayed in Figures 3 and 4. Model 3 assumed that the sex-specific mortality probabilities increased over GoM scores but not over age, following the unadjusted values displayed in the Totals row of Table 2. Model 4 assumed that the sex-specific mortality probabilities increased over GoM scores and over age, following the unadjusted values displayed in the age-specific rows of Table 2.

The log–likelihood-ratios were generated as the difference in the value of the log–likelihood for each model and the log–likelihood for Model 1. The degrees of freedom (d.f.) were defined as the number of parameters in each model. The AIC (Akaike Information Criterion) was calculated as the log–likelihood-ratio plus 2 times the d.f. The BIC (Bayesian Information Criterion) was calculated as the log–likelihood-ratio plus the product of the d.f and the logarithm of the number of deaths. The best model was the one that had the minimum value of AIC or BIC (indicated with boldface font in Table 5); the relative performance of each model was assessed using the difference between its value of AIC or BIC and the minimum value of these statistics (labeled ΔAIC or ΔBIC in

Table 5). Differences of 10 or more points were regarded as strong evidence in support of the model with the lower AIC or BIC value.

For both sexes and both criteria, Model 4 was overwhelmingly selected as the best model.

To determine whether the GoM scores added significant information beyond the information already available using the age-specific mortality probabilities, we needed to compare the value of ΔAIC or ΔBIC for Model 2 with the reference value of 10. For both sexes and both criteria the values exceeded the reference values by a factor of 746–874, indicating that the additional information provided by the GoM scores was huge.

Comparison of Models 2 and 3 provided additional confirmation of the power of the GoM model. The ΔAIC and ΔBIC for Model 2 were each about 30.0% smaller than the corresponding value for Model 1 for males and about 46.5% smaller for females. The ΔAIC and ΔBIC for Model 3 were each about 97.0% smaller than the corresponding value for Model 1 for males and 95.0% smaller for females. Thus, if one were forced to choose between Models 2 and 3, then Model 3 would be selected as the better model and the improvement offered by Model 3 would be huge. Model 3 would offer 95–97% of the improvement over Model 1 that could ultimately be obtained using Model 4. This would be far in excess of the 30–46% improvement offered by Model 2.

The *second* question was whether the excellent calibration displayed in Figures 5 and 6 continued when the predictions were stratified by age-groups. The results of these stratifications are displayed in Figures 7 and 8 for males and females aged 65–99 years old. The death counts for males at age 100+ fell below the standard CMS cutoff of 11 events and hence were suppressed.

For comparability, females were restricted to the same age range. The aberrant point in Figure 6 for the 40–44% group turned out to be solely for females aged 100+ which meant that this point was excluded from Figure 8.

The labeling of the groups (Pct.Age) in Figures 7 and 8 combined the lower bounds of the 5-Percent labels in Figures 5 and 6 with the lower bounds of the age-groups in Figures 3 and 4. Thus, "0.65" identifies persons aged 65–69 years with predicted probabilities in the range 0–4%; similarly "35.95" identifies persons aged 95–99 years with predicted probabilities in the range 35–39%. The groups were ordered by increasing predicted probabilities, and within each probability group, by increasing age.
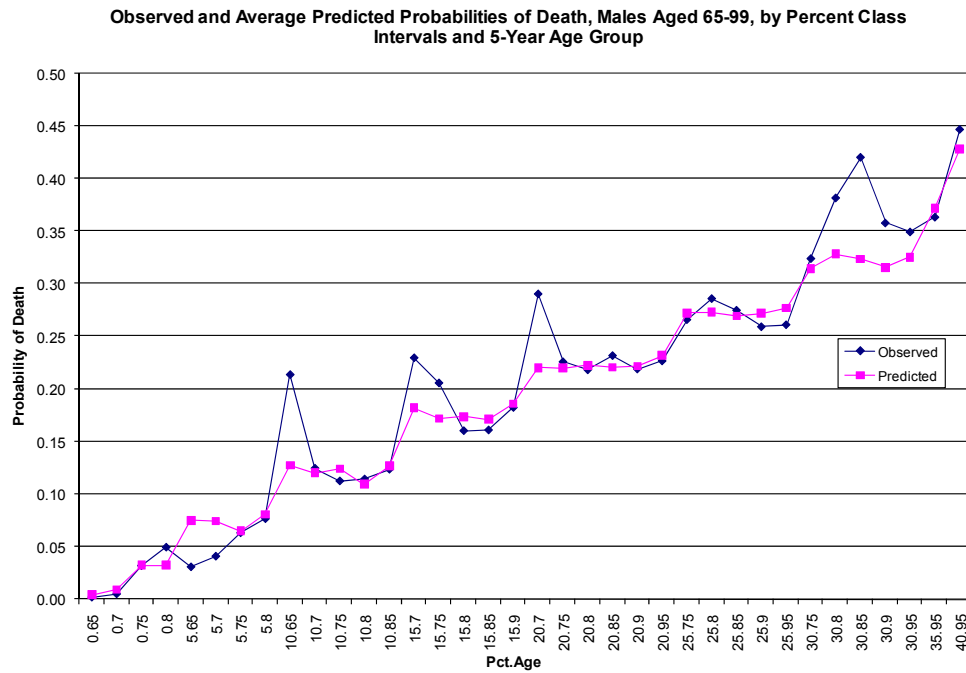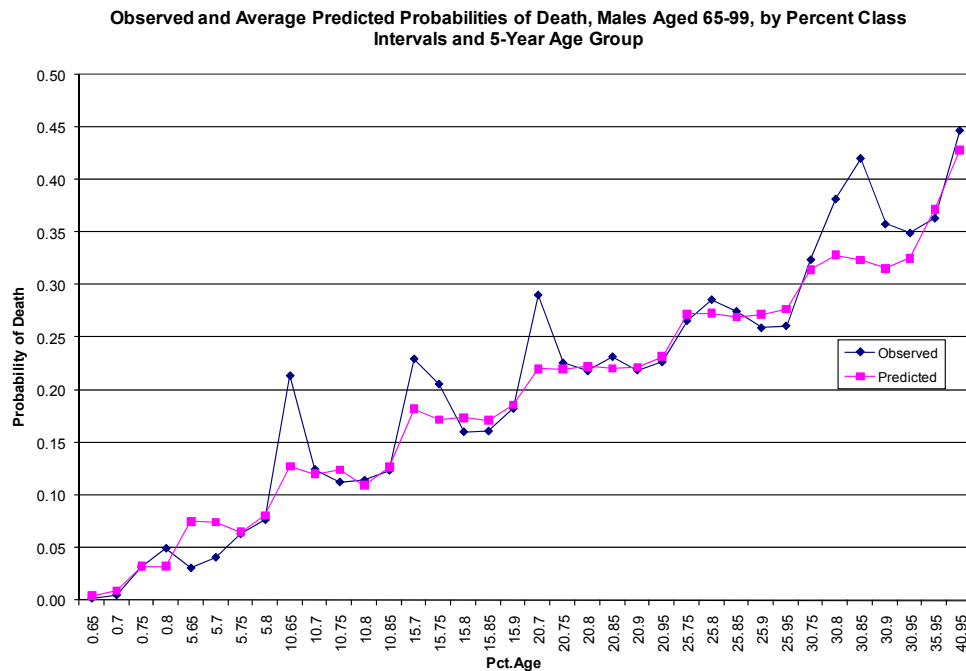
Figure 7

Observed and Average Predicted Probabilities of Death, Males Aged 65-99, by Percent Class
Intervals and 5-Year Age Group

Figure 8

Observed and Average Predicted Probabilities of Death, Males Aged 65-99, by Percent Class
Intervals and 5-Year Age Group

Figures 7 and 8 (males and females, respectively) showed that the largest deviations
were for males aged 65–69 (with offsetting deviations for 5–9% and 10–14%) and aged

70–74 (with offsetting deviations for 5–9% and 15–19%), but these same deviations were not replicated for females.

One possible explanation for the male result involved the design of the NLTCS in which persons who did not have ADL or IADL impairments at the time of the survey did not receive the in-person assessment; such persons "screened out" without answering any of the detailed health questions. Thus the estimates of their GoM scores have substantially larger errors than would be the case for persons who answered most or all of the 95 health-related questions in the original model (or the 76 questions in the LCC implementation).

To quantify the size of the deviations, we applied linear regression analysis with the observed probabilities regressed on the predicted probabilities, as done for the data in Figures 5 and 6, obtaining R-squared values of 0.943 and 0.960, respectively, for males and females, with a combined R-squared value of 0.950. We generated a second set of regressions with the logarithms of the observed probabilities regressed on the logarithms of the predicted probabilities, obtaining R-squared values of 0.958 and 0.984, respectively, with a combined R-squared value of 0.968. The implied accuracy was thus in the range 95–97%, depending on the form of the regression.

Ongoing Model Validation and Recalibration

LSF has been conducting retrospective validation tests using data sets with hundreds of assessments collected over the last seven years in assisted living communities. The questions in these assessments have been mapped to the 76 questions in the LCC.

Complementing these retrospective validation tests, LSF is in the process of collecting new prospective data, for use in additional validation activities, from the answers to the 76 questions on the LCC applications provided by cooperating life settlement brokers, providers and funders.

Prospective data collection by life settlement providers will allow comparisons of the LE and survival predictions based on the LCC model to the corresponding predictions of commercial LE providers, and comparisons to the actual outcomes, based on the tracking of each insured until death. Additional questions have been added to the interview and will be included in future recalibrations. As sufficiently credible experience is accumulated, the parameters of the model will be recalibrated to better capture unmeasured aspects of insured survival that are unique to life settlement participants.

The 32,389 individuals observed in consecutive assessments over the term of the original NLTCS calibration data were a random sample of Medicare enrollees who participated in the study. Assessments began either in 1982 when the NLTCS started, or at a later date when the participants were aged 65–69. Because life insurance data

were not collected, we do not know how many of these seniors either had life insurance in-force or may not have been medically insurable at age 65. What we do know from Census Bureau data is that more than 80% of U.S. seniors at age 65 are homeowners. Moreover, there is nearly a one-to-one correlation in the U.S. between homeownership and the purchase of life insurance.

Thus it was reasonable to expect that the variability of the general population experience represented in the NLTCS would be sufficiently broad that some subset of its participants would be reasonably representative of the experience of life insurance policy holders given that the answers to most or all of the 76 health-related questions were known. The conditioning on the 76 questions is important because it controls for most of the health-related selection biases that may distinguish the insured from the non-insured populations.

Application of the same model to the life settlement population must consider that life settlement applicants and transaction completers may exhibit further health-related selection biases that distinguish them from each other and from the non-settlement insured population,[14] given that the life settlement participants are a small fraction of the total insured population. Again, it is reasonable to expect that most of these biases will be controlled if the answers to most or all of the 76 health-related questions are known.

The purpose of the prospective data collection is to assess the size and direction of biases that may exist, and when necessary to recalibrate the model parameters as needed to remove them. Thus, periodic assessment and recalibration are important parts of the proposed methodology.

VALUATION METHODOLOGY FOR SETTLEMENT PORTFOLIO MANAGERS

A settlement portfolio's management is required to prepare fair valuation estimates and have these audited by independent auditors. The standards require that the valuation be determined using methodologies to corroborate and reconcile the results. The LCC tool is a low cost, peer reviewed, published methodology that can be incorporated in a defined process of reconciling the life expectancies as part of the portfolio fair valuation process.

The proposed methodology for annually establishing fair value is compliant with ASC §820.10 and is straightforward and drawn from ISA §540 / AU §328 and applies accepted Bayesian statistical methodology to arrive at a weighted average fair value per policy and thus the sum will be a fair value of the portfolio. Annually, the difference between the purchase value and the newly established value will be clearly documented, transparent and available to be consistently used year after year in fair valuation. Policy and portfolio cash flows can then be stress tested based upon this reconciliation of life expectancy to arrive at the portfolio's current fair value.

Proposed steps at underwriting and portfolio origination:

1. Each settlement policy at the point of underwriting will have three (3) commercial LE's, ideally from the same three commercial LE providers for the entire portfolio. LE providers will each provide a table of calculated survival functions for each insured life.
2. The settlement application will contain the LCC assessment questions or a telephonic assessment will be completed for automated LCC scoring. When the responses to the initial application questions deviate from the telephonic assessment responses an in person interview is scheduled and conducted through a registered nursing service to resolve differences.
3. Each policy will be priced independently for each commercial LE and the LCC LE using the same probabilistic pricing model. Typically, the funder will require that the offer be priced based upon commercial LE's only, no matter how many LE's may have been discarded at the time of pricing.
4. If the offer is accepted, the policy and its four (4) related LE data will become part of the portfolio valuation data base.
5. Initially at the time of portfolio formulation, equal weighting (1/4 to each) will be assigned to the four life expectancy providers (i.e., the three commercial LEs plus the LCC LE).

The importance of a consistent valuation methodology can most clearly be seen in the positive impact it has on the risk premium component of the discount rate used to calculate the present value of the net future periodic portfolio cash flows that establish the portfolio's current market value. A relatively homogeneous and actuarially credible portfolio with a reduced standard deviation of life expectancy estimates supports the argument for a lower risk premium. The discount rate is made up of the cost of funds for the underlying financial instrument, e.g. average 10-year treasury rate for a 10-year bond or similar instrument or average fund life, plus the risk premium. The risk premium component is made up of four factors:

1. The probability of a forced sale of the portfolio because of inadequate cash reserves or where a liquidity facility is not maintained;
2. The delta in the standard deviation of probable cash flows because of inadequate actuarial credibility introduced by a small number of policies or a lumpy portfolio;
3. The delta in the standard deviation of probable cash flows introduced by a wide ranging variance in LEs per policy that has not been challenged and mathematically tightened up by a process such as the one described in this paper and quantified below.
4. The delta in the standard deviation of probable cash flows introduced by irreducible systematic risks relating to errors in projecting average future survival

rates in the respective cohorts due to various combinations of model misspecification errors, errors in estimating the model parameters, and inherent stochasticity in future mortality rates.

At this writing the average 10-year U.S. Treasury Bond coupon rate is 3.27%. A risk premium equal to an additional 4.98% would yield a portfolio current valuation approximately equal to the investment in the portfolio. However a risk premium equal to only an additional 3.27% would add a 32% increase to the market value of the portfolio. Conversely, a risk premium of an additional 9.81% would discount the portfolio to 58% less than cost. This highlights the need for strict and consistent underwriting and valuation methodologies.

Proposed steps at quarterly and annual fair valuation:

1. Compare the actual to expected mortality results for each of the four LE providers over the examination period. The expected survival counts will be generated by summing the survival functions (S_{it}) by LE provider for that period for each policy in the portfolio for which an LE value was provided. The change from one period to the next in the sum of S_{it} will constitute the expected number of deaths for the period as predicted by each of the four LE providers.

Table 6
Identical Actual-to-Expected Analysis Performed for Each LE Provider

	EXPECTED SURVIVAL AND DEATHS						ACTUAL SURVIVAL AND DEATHS					
	Expected Survival			Expected Deaths			Actual Survival			Actual Deaths		
Portfolio Member Code	Year-End 1	Year-End 2	Year-End 3	Year-End 1	Year-End 2	Year-End 3	Year-End 1	Year-End 2	Year-End 3	Year-End 1	Year-End 2	Year-End 3
0001-F	0.9577	0.9095	0.8564	0.0423	0.0905	0.1436	1.0000			0.0000		
0002-F	0.7824	0.6428	0.4987	0.2176	0.3572	0.5013	1.0000			0.0000		
0003-F	0.8974	0.8021	0.7895	0.1026	0.1979	0.2105	1.0000			0.0000		
0004-F	0.8547	0.7952	0.6587	0.1453	0.2048	0.3413	1.0000			0.0000		
⁓	⁓	⁓	⁓	⁓	⁓	⁓	⁓	⁓	⁓	⁓	⁓	⁓
1000-F	0.9211	0.9033	0.8524	0.0789	0.0967	0.1476	1.0000			0.0000		
Totals	883	811	731	117	189	269	852			148		Actual
Less Expected Deaths										117		Expected

2. Compare the actual number of deaths for the period against the expected number to ascertain the accuracy of each LE provider relative to their prediction.

3. The portfolio valuation for each of the four LE providers will define the bounds of the valuation. For example, assume the independently determined portfolio

values are $255 million, $274 million, $316 million and $329 million[25]. Absent better precision the portfolio has a value between $255 million and $329 million.

Table 7
NPV Valuation of Portfolio at Origination

LE Provider	NPV of Cash Flows Value	Assumed Probability	Probable Weighted Cash Flows	Probable NPV of Cash Flows
LCC	$329 Million	25%	$82.3 Million +	
1	$316 Million	25%	$79.0 Million +	
2	$274 Million	25%	$68.5 Million +	
3	$255 Million	25%	$63.8 Million =	$293.6 Million

4. Use Bayesian analysis as detailed by Kass and Raftery[26] to determine the weighted average total value of the portfolio.

Table 8
Results Oriented Weighted Valuation after Portfolio Year-One

LE Provider	NPV of Cash Flows Value	Assumed Probability	Probable Weighted Cash Flows	Probable NPV of Cash Flows
LCC	$329 Million	42%	$138.2 Million +	
1	$316 Million	16%	$50.6 Million +	
2	$274 Million	28%	$76.7 Million +	
3	$255 Million	14%	$35.7 Million =	$301.2 Million

5. This will allow the actual versus expected values for each policy's LE's to be evaluated and the LE providers to be ranked accordingly. The LE evaluations can be conducted via standard Bayesian methods or, more simply, by using Bayesian Information Criterion (BIC) measures of goodness of fit of the actual-to-expected probabilities of death computed separately for each LE model using the same pool of policies. Kass and Raftery[27] showed how BIC values[28] can be used (1) to rank the various LE models and (2) to generate optimal weighted averages of the outputs of the various LE models, where each of the four weights is interpretable as the Bayesian posterior probability that the corresponding LE model is correct (assuming that one of them is correct). Weighted averaging can then be applied to each individual policy, and to the

---

[25] Assumes a portfolio of $1.7 billion of face amount with future net cash flows to maturity discounted at 8% to an NPV.
[26] Kass and Raftery (1995, Eq. 18).
[27] Kass and Raftery (1995, Eqs. 9, 16 and 18).
[28] Note that BIC values can be generated using various approaches to measuring goodness of fit, including chi-squared statistics and regression-based R^2-statistics.

aggregate of all policies, in annually revaluing each portfolio. The weights can be updated each year (more frequently if the portfolio is large) as additional information on the actual number of deaths in that year becomes available. Over time, this will give greater weight to the better performing models. Disclosure of the chosen methodology at portfolio formulation and annual review will provide transparency to investors with respect to these longevity valued asset transactions.

6. Present the weighted average portfolio value as well as the individual portfolio values, and their associated weights, derived by using the LE values and survival curves from each of the four LE providers.

PREPARING FOR THE INDEPENDENT AUDIT

AU §328 directs independent auditors to first understand the process of developing the fair value estimates and the controls instituted by the entity to ensure the completeness, accuracy, and consistency of the methodology and data used in the computations. It will often be necessary for the auditor to verify that the data in the models are accurate and to evaluate the reasonableness of the assumptions used in the modeling.

The approach outlined in this paper will help provide documentation and support helpful to the auditor in assessing that a "reasonable basis" exists for the fair valuations. Another option open to auditors is to apply a different methodology to relevant fair value data and compare the results with the entity estimates. However, the estimates of the entity are generally only open to challenge when they fall outside of a "reasonable range" of outcomes as assessed by the independent auditor.

Because of the specialty nature of the application of fair value concepts to insurance valuations, the auditor is likely to engage a specialist to review the approach and computations underlying the entity estimates and to identify benchmark data that might be used to assess the reasonableness of the assumptions used in developing the estimate.

Recognizing these requirements, the entity can minimize its audit costs by creating clear and transparent documentation of the development of its estimates, establishing oversight and internal controls to ensure the quality of its estimates on a quarterly and annual basis, and having internal resources available to answer questions and explain the process during the audit.

Because consistency is an important accounting concept, fair value methodologies should be applied consistently from year-to-year, and thus should be chosen wisely. Departures in methodologies may indicate that an accounting change has occurred; the effects of the change should be quantified in financial reporting. Nonetheless, as more

and better proven methodologies arise, they should be considered for possible implementation.

<u>EXISTING ACCOUNTING FRAMEWORK</u>

The framework exists today to create uniformity in life settlement policy underwriting and related disclosure by applying existing GAAP at the time new policies are originated, existing policies are pooled for securitization, and annually at revaluation. These policy portfolios, whether rated or unrated, will require disclosure at the time they are pooled and sold, and subsequently when the pools are individually valued and audited because they are held for investment by issuers. For Level 3 assets where value is dependent upon a future unobserved outcome (the insured's death) GAAP, Accounting Standards Codification (ASC) 820.10.05–820.10.65 (previously FASB 157), requires the use of the best information available in the corroboration of the valuation methodology. The outcome of these methodologies must then be reconciled and disclosed. Specific relevant auditing standards include AU §328 (SAS 101) (*Auditing Fair Value Measurements and Disclosures*, effective June 2003).

AU §328.4 states "Management is responsible for making the fair value measurements and disclosures included in the financial statements. As part of fulfilling its responsibility, management needs to establish an accounting and financial reporting process for determining the fair value measurements and disclosures, select appropriate valuation methods, identify and adequately support any significant assumptions used, prepare the valuation, and ensure that the presentation and disclosure of the fair value measurements are in accordance with GAAP." We believe the straightforward rules-based methodology proposed above supports these objectives.

AU §328.40, under the heading *Developing Independent Fair Value Estimates for Corroborative Purposes*, states, "The auditor may make an independent estimate of fair value (for example, by using an auditor-developed model) to corroborate the entity's fair value measurement[29]. When developing an independent estimate using management's assumptions, the auditor evaluates those assumptions as discussed in paragraphs .28 to .37. Instead of using management's assumptions, the auditor may develop his or her own assumptions to make a comparison with management's assumptions. The auditor uses that understanding to ensure that his or her independent estimates take into consideration all significant variables and to evaluate any significant differences from management's estimates. The auditor also should test the data used to develop the fair value measurements and disclosures as discussed in paragraph .39." If a well documented and controlled process, supported by published research or established as an industry practice, is used to develop the entity fair value estimates, the auditor

---

[29] See AU §329, *Analytical Procedures.*

generally will "audit" the entity process, controls, and calculations rather than run alternative models that may need to be reconciled to the entity estimates[30].

Clear roadmaps that link to the data sources used and documented support for the key model assumptions will aid auditors in evaluating whether management has a "reasonable basis" for its estimates of fair value.

<u>CONCLUSION</u>

The proposed fair value methodology is credible, compliant with the accounting and auditing framework, and, most importantly, doable by management. Management is responsible for implementing a supportable assumptions-based valuation methodology that is transparent and controlled. Provided they do this they can present a completed valuation to the independent auditors to critique and avoid costly additional modeling. The proposed methodology for valuation should be consistently applied year-after-year. As the portfolio valuation improves, it adds income. Conversely, if the portfolio value were to decline in the future, consistently applying the methodology would identify the change in direction on a timely basis, and avoid lags in portfolio write-downs. This would mitigate the natural tendency to defer losses, pending the ugly last-minute, one-time write-down of asset values such as experienced recently, and expected to continue for the foreseeable future, in other collateralized asset classes.

---

[30] The testing requirements for fair value estimates are enumerated in paragraphs 23 to 42 of SAS 101 (AU 328).

References

Bauer, D., Boerger, M., and Russ, J., 2010. On the Pricing of Longevity-Linked Securities. *Insurance: Mathematics and Economics* 46:139–149.

Financial Accounting Standards Board (FASB), 2010. *Proposed Accounting Standards Update—Accounting for Financial Instruments and Revisions to the Accounting for Derivative Instruments and Hedging Activities—Financial Instruments (Topic 825) and Derivatives and Hedging (Topic 815)*. Financial Accounting Standards Board, Norwalk, CT, May 21, 2010.
http://www.fasb.org/cs/ContentServer?site=FASB&c=Page&pagename=FASB%2FPage%2FSectionPage&cid=1176157086783

Granieri, V., 2011. Mortality Experience of Three Senior Populations. Presented at Society of Actuaries Living to 100 Symposium IV, Orlando, FL, January 5–7.
http://livingto100.soa.org/pdf/2011-orlando-granieri.pdf

Hosmer, D.W. and Lemeshow, S., 1980. Goodness of Fit Tests for the Multiple Logistic Regression Model. *Communications in Statistics – Theory and Methods* 9(10):1043–1069.

International Accounting Standards Board (IASB). 2011. IFRS 13 Replaces the fair value measurement guidance dispersed across different IFRS standards. Kass, R.E. and Raftery, A.E., 1995. Bayes Factors. *Journal of the American Statistical Association* 90(430):773–795.

Kaplan, E.L. and Meier, P., 1958. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association* 53(282): 457-481.

Longley-Cook, L.H., 1962. *An Introduction to Credibility Theory*. Casualty Actuarial Society, New York.

Menelaides, S., Graham, L., and Fischbach, G., 2003. The Auditor's Approach to Fair Value. *Journal of Accountancy* 195(6):73–76.

Milevsky, M.A., Promislow, S.D., and Young, V.R., 2005. Financial Valuation of Mortality Risk via the Instantaneous Sharpe Ratio: Applications to Pricing Pure Endowments. *arXiv:0705.1302v1*.

Milevsky, M.A., Promislow, S.D., and Young, V.R., 2006. Killing the Law of Large Numbers: Mortality Risk Premiums and the Sharpe Ratio. *Journal of Risk and Insurance* 73(4):673-686.

Modu, E., 2008. Life Settlement Securitization, Best's Rating Methodology, A.M. Best Company, Inc., March 24, 2008.

www.ambest.com/debt/lifesettlement.pdf

National Center for Health Statistics (NCHS), 1997. U.S. Decennial Life Tables for 1989–91. Vol. 1, no. 1. National Center for Health Statistics, Hyattsville, MD.

Olivieri, A., 2001. Uncertainty in Mortality Projections: An Actuarial Perspective. *Insurance: Mathematics and Economics* 29:231-245.

Purushotham, M., Valdez, E., and Wu, H., 2011. *Global Mortality Improvement Experience and Projection Techniques*. Society of Actuaries, Schaumburg, IL. http://www.soa.org/files/pdf/research-global-mortality-improve-report.pdf

Reither, E.N., Olshansky, S.J., and Yang, Y., 2011. New Forecasting Methodology Indicates More Disease and Earlier Mortality ahead for Today's Younger Americans. Health Affairs 30(8):xxx—yyy (WEB FIRST) http://content.healthaffairs.org/content/early/2011/06/21/hlthaff.2011.0092.full.html

Society of Actuaries (SOA), 2008. 2008 Valuation Basic Table (VBT) Report & Tables. Society of Actuaries, Schaumburg, IL. http://www.soa.org/research/experience-study/ind-life/valuation/2008-vbt-report-tables.aspx

Stallard, E., 2006. Demographic Issues in Longevity Risk Analysis. *Journal of Risk and Insurance* 73(4):575-609.

Stallard, E., 2007. Trajectories of Morbidity, Disability, and Mortality among the U.S. Elderly Population: Evidence from the 1984-1999 NLTCS. *North American Actuarial Journal* 11(3):16–53. http://www.soa.org/library/journals/north-american-actuarial-journal/2007/july/naaj0703-2.pdf

Wang, H. and Preston, S.H., 2009. Forecasting United States Mortality using Cohort Smoking Histories. *Proceedings of the National Academy of Sciences USA* 106(2)393–398.

About the Authors

Peter Macrae Mazonas, CPA,MBA, is the Managing Partner of Life Settlement Financial (LSF) in San Francisco. LSF aggregates, manages and values longevity asset portfolios for clients. Prior to forming LSF in 2006 Mazonas was Director of Corporate Governance and Sarbanes Oxley for McKesson Corp and Invitrogen. Previously Mazonas created and managed the reverse mortgage arm of Transamerica; created the Private Banking and Executive Financial Planning divisions of Bank of America; and was a tax and audit associate at Price Waterhouse.

P.J. Eric Stallard, A.S.A., M.A.A.A., F.C.A., is a Research Professor in the Social Science Research Institute at Duke University, and Associate Director of the Center for Population Health and Aging, in the Duke Population Research Institute. His research interests include medical demography and health/LTC actuarial practice, with an emphasis on modeling longevity and mortality risks at the individual and aggregate levels. He won the Society of Actuaries' Edward A. Lew Award for the best paper in actuarial modeling for 2006.

Lynford Graham, CPA, Ph.D., CFE, is a Visiting Professor of Accountancy at Bentley University in Waltham MA. Dr. Graham is a past member of the AICPA's Auditing Standards Board. He edits the forthcoming 12th Edition of the Accountant's Handbook (Wiley), is a co-author of Information Technology Audits (CCH, 2011 Ed.), and is the sole author of 2 books on Internal Controls (Wiley, 2007, 2010). He holds an MBA and Ph.D. from the University of Pennsylvania, (Wharton School) and is a Certified Fraud Examiner.