Milliman White Paper

Prepared by:

Stephen Conwill, FSA, MAAA

January 2016

Milliman

Life insurance risks

Observations on Solvency II and the modeling of capital needs

TABLE OF CONTENTS

HOW TO USE THIS WHITE PAPER 2

I. INTRODUCTION 3

II. OBSERVATIONS ON INSURANCE RISK 5

Correlations, tails, and discontinuities 5

III. MORTALITY 6

Standard formula calibration 6

Internal model enhancements 6

IV. LONGEVITY 8

Standard formula calibration 8

Internal model enhancements 8

Causal mortality models 10

V. DISABILITY/MORBIDITY 11

Standard formula calibration 11

Internal model enhancements 11

A Bayesian approach to morbidity forecasts 11

VI. LIFE EXPENSE AND OPERATIONAL RISK 14

Standard formula calibration 14

Internal model implications 15

VII. LAPSE RISK 16

Standard formula calibration 16

Internal model enhancements 16

A combination of techniques 17

VIII. EXAMPLE: DYNAMIC LAPSE ANALYTICS FOR SINGLE-PREMIUM WHOLE LIFE 22

Premise 22

Traditional analysis 22

Predictive analytics 23

Bridging techniques 25

Cognitive mapping, shape of distribution, and enhancing domain knowledge 25

IX. FURTHER COMMENTARY ON RISK CORRELATIONS 26

X. INNOVATIONS AND FUTURE DIRECTIONS 28

HOW TO USE THIS WHITE PAPER

Over the past decade, Milliman has conducted research on a variety of topics related to economic capital modeling. This research includes analyses of investment risk, operational risk, insurance risk, correlation and combination of risks, and other aspects of risk and capital management. During the course of recent conversations with our clients, it has become clear that there is a dearth of material related to life insurance risks. While excellent papers exist on some topics, such as management of longevity risk, the results of many interesting avenues of research have not yet been widely disseminated. Our goal in presenting this white paper is to describe and discuss a variety of promising and practical new techniques.

We emphasize that this white paper is not intended as a comprehensive or definitive treatise. Nonetheless, we are confident that readers with an interest in life insurance risk and capital management will find a wealth of engaging material herein. Some of the exposition, of course, will be familiar. In particular, in each section, we ground the discussion with a reference to the Solvency II standard formula. This will be “old hat” to some readers: you won’t offend us if you skip a section or two! Indeed, we offer the following advice: ***Don’t hesitate to “flip through” this white paper, skip sections that are familiar, scan the diagrams, and skim the text before returning to the sections you find the most intriguing and relevant to your work.***

Please send email comments to stephen.conwill@milliman.com. We will respond to all comments or inquiries.

I. INTRODUCTION

Solvency II moved towards completion on the heels of the financial crisis. It reflects the input of many constituents over a long period of deliberation. Given a very difficult task and very difficult circumstances, the framers crafted what should be a workable and comprehensive framework for risk and capital management. However, it is important to appreciate that many of the core ideas of Solvency II are drawn from the realms of finance and investing. These may or may not extend naturally to the domain of insurance risk. In some cases, given the paucity of historical data and the difficulty of measuring correlations and tail-risk behaviors, a healthy dose of skepticism is warranted. How viable is value at risk (VaR)? How credibly can we measure the impact of a one-in-200 event? Risk management will always require an interplay of analytics and common sense. Effective capital management is as much an art as it is a science.

Yet tools are increasingly available that enhance our ability to develop forecasts and to critique these forecasts as part of the capital management, business planning and decision-making process. In particular, we are seeing an explosive growth in potential applications of data analytics, predictive algorithms, and behavioral modeling to challenges related to insurance risk. This white paper describes several key ideas that can serve as building blocks for a new analytical framework.

Solvency II life insurance risk submodules include mortality, disability-morbidity, longevity, life expense, revision risk, lapse risk, and catastrophe risk. In principle, Solvency II focuses on the impact on net asset value (NAV) that may emerge due to errors in the estimation of the level,[1] trend, and volatility of insurance events. The analysis underlying the development of the standard formula focused on level and trend risk.[2]

Conventional statistical techniques were employed in parameterizing the standard formula. These included analysis of large data sets to construct appropriate stress scenarios and to evaluate levels of correlation among risks. In the course of their work, the Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS)[3] drew data from across the EU, including Polish and U.K. lapse data, Swedish disability data, and longevity data from nine jurisdictions. Approaches and parameters for stress testing were refined throughout the Quantitative Impact Study (QIS) process.

With the possible exceptions of longevity and disability risks, life insurers tend to view insurance risks as relatively stable and easily managed. Risk managers have tended to focus on investment risk. This is not unreasonable. The perils associated with market and credit risks were brought into clear focus at the time of the financial crisis. Indeed, many of the life insurer insolvencies that have occurred over the past 30 years have stemmed from problems on the asset side or flawed product design. Insurers are typically confident when it comes to underwriting, in-force management, and the mitigation of insurance risk through reinsurance and financial market mechanisms.

The nature of the life insurance business is changing. As a result of globalization, low interest rates, and aging demographics, we are likely to see new waves of product development over the next decade. Risk products may come to the fore. As the number of retirees grows, insurers will increasingly assume risks related to longevity, post-retiree medical needs, and long-term care. As we see increases in the retirement age in the developed world and a trend towards meaningful part-time work among the aged, there will be a need for new and innovative mortality and disability coverages.

In this environment, to the extent that investment and protection elements continue to be bundled, insurers may face significant challenges in evaluating the cost of embedded options. The exercise of investment-oriented options will often have implications for mortality or morbidity experience. The management of protection risks may be deeply intertwined with the management of investment risks; lapse risk will have both protection and investment elements. Understanding the correlation among risks is essential and non-trivial.

1 In some literature, level risk is referred to as "basis risk." We use these terms interchangeably in this document.
2 To the extent that volatility risk is not diversified through reinsurance, it may be valued by stochastic modeling of the individual incidence of insured contingencies (e.g., death, disability, illness).
3 CEIOPS was established by the European Commission in 2003 and was replaced by the European Insurance and Occupational Pensions Authority (EIOPA) in 2011.

Innovations in technology and data analytics have the potential to drive major changes in the insurance business. Increasingly, technology will allow us to:

- Collect larger and larger sets of data
- Merge or apply analytic techniques to heterogeneous data sets
- Gain a deeper understanding of the relationships among variables describing attributes of our data
- Gain a better understanding of time series data
- Make and test hypotheses about future results
- Analyze data in real time
- Combine quantitative and qualitative information when gaining business understanding
- Systematically update models and hypotheses as new information becomes available

Analytic techniques such as those outlined above will allow us to develop better forecasts and to better understand the limitations of our forecasts, thereby improving our capital and risk management capabilities. While sophisticated technologies have been available to support asset side analytics for at least a decade, it is only now that innovations in data analytics, predictive algorithms, and behavioral modeling are beginning to emerge as critical features of capital management in the realm of insurance risk and asset-liability management (ALM).

Milliman is a leader in research that facilitates the integration of new technologies with the economic capital measurement and strategic decision-making process. As part of an ongoing effort to disseminate this research, we offer this paper focusing on the management of insurance risk under Solvency II and similar regimes. For key risks, and with respect to the critical issue of correlations, we look at current practice. We reflect on the suitability of current practice as our environment evolves, and hope to facilitate a pragmatic discussion of key management questions. Does current practice offer an adequate framework for the evaluation of risk? Does it provide the management insight required to efficiently deploy and fine-tune capital management? What emerging techniques can offer management greater ability to deal with these risks?

The sections in this paper correspond reasonably closely to the Solvency II underwriting risk modules. We have excluded catastrophe risk from consideration—it is certainly worthy of a separate white paper. Also, our definitions of risk differ somewhat from those under Solvency II. In particular, insurance lapse risk is so inseparable from investment risk that we find it hard to isolate components tied to non-financial market considerations.

A primary thesis underlying our analysis is as follows:

> ***The understanding of one-in-200 insurance risk events is impossible without a deep understanding of the drivers of risk events and the implications of risk correlations. Correlations among insurance risks (such as lapse, mortality, and morbidity) and correlations among insurance, investment, and operational risks, may be more significant than is generally recognized. An understanding of these drivers and correlations will facilitate appropriate interventions; as such, risk measurement is highly intertwined with the sophistication of management response.***

Over the past decade, Milliman has led the industry in research into emerging risks and risk correlations. We have developed important techniques applicable to specific risks, such as longevity, disability, and medical risk. We highlight some of our research throughout this white paper, and encourage readers to contact us for further details.

II. OBSERVATIONS ON INSURANCE RISK

What is a one-in-200 event? Given available data, the complex interactions among a multitude of risks, changing circumstances, and a forever evolving environment, how can we credibly model a one-in-200 hit to NAV? On the one hand—if we are honest with ourselves—we may reasonably view it is an impossible task. Counterintuitively, perhaps, this does not negate the value of the framework. The value is in the process. The value is in the insights gained when trying to answer a fundamentally unanswerable question.

Insurance risks envisioned under Solvency II are easy to enumerate:

1. Mortality
2. Longevity
3. Disability morbidity
4. Lapse
5. Expense
6. Revision
7. Catastrophe

A one-in-200 event combines risks associated with level, trend, and volatility. Because volatility is viewed as subsumed within level and trend, the focus is on level and trends, where level risk refers to the risk that our best estimate at time zero is incorrect, and trend refers to the risk that our expectations for future evolution in a particular risk is incorrect.

Level risk can be evaluated by (a) looking back at our track record and, (b) taking account of changes that have occurred to our markets, underwriting, and management. In looking back, we can assess how accurate we have typically been when estimating time zero experience. A measurement of our historical success will reflect both error in estimation and random fluctuations. In this sense, it combines level and volatility. Looking forward, we need to adjust our evaluation of level risk to account for changes in our markets and strategies. Have we recently entered new markets? Are we underwriting new risks with limited experience? Have we changed our approach to underwriting or claims management? If so, we might expect greater level risk than we have historically faced. Are we writing larger volumes of business? Have we tightened underwriting standards? If so, level risk may have declined.

In many instances, it will be more difficult to estimate trend risk than level risk. Trends in insurance risk are influenced by complex factors including medical advances, societal trends, government actions, the economy, and a company's competitive situation.

What is a one-in-200 risk event? What does it mean to say that over a one-year time horizon, we have experienced a one-in-200 event? The notion is arguably more intuitive in the realm of investments than it is in the realm of insurance. A one-in-200 investment event will typically emerge in the here and now as a tangible market disturbance—a severe fall in equities, a severe change in credit spreads, or a major uptick in defaults. These are events whose impact can generally be measured over a short time horizon. On the other hand, many one-in-200 insurance risk "events" involve a change in view, a realization that future expectations must be revised. Examples would include emerging data that implies that longevity is improving at a rate greater than once envisioned, or a realization that disability recovery rates appear to have materially declined. Expectations emerge over the one-year time horizon—and the capital implication is immediate—but the actual cash-flow impact of revised expectations is felt over years or decades.

Correlations, tails, and discontinuities

The issue of correlation may be more difficult in the realm of insurance risk than it is in the case of investment risk. Yet it may be equally important. In the case of investment risk, there is substantial empirical data available to calibrate our correlation models, typically built on matrices or copulas. It is possible to define a theoretical approach, parameterize a formula, backtest the formula against historical eras, and monitor the viability of the formula as new experience emerges. While this is not altogether impossible in the case of insurance risk, it may be much more difficult. The risk drivers that influence insurance experience are often difficult to quantify. Standard statistical approaches may be well-suited to specific risks and relatively shorter time periods but may no longer be appropriate as the time periods lengthen, when we try to understand the tail of a distribution, and when we try to understand complex correlations, as among mortality, lapse, and investment. Further, it may be very difficult to evaluate the implications of discontinuities using common statistical approaches. For these reasons, we believe that Bayesian and behavioral techniques are essential, because they facilitate the modeling of insurance risk drivers, the understanding of severe risk events, and an analysis of correlations among such events.

III. MORTALITY

For any company with a significant book of in-force business and a relatively long history of underwriting mortality risks, an evaluation of level risk may be relatively straightforward. Analysis of experience over the recent past, for example, five to 10 years, will offer empirical insights into the degree to which actual experience has tracked forecasts. Because of the potential importance of tail risk, experience studies of 20 years or more may also provide useful information. A granular assessment by block of business can improve overall results, though caution needs to be exercised to ensure statistical credibility and comparability in the course of aggregation. Also, stochastic analysis may reveal offsetting effects from aggregating complimentary blocks of business.

An evaluation of trend risk will typically be more difficult than the estimation of level risk. Nonetheless, as described below, risk and capital management efforts can be enhanced by employing bespoke mortality forecasts that take account of the specific circumstances of any operation.

Standard formula calibration

In the course of the Solvency II development process, a consensus emerged that mortality risk should be reflected in the standard formula by measuring the impact on NAV of an annual uniform percentage increase in mortality rates. The final standard formula specifies a permanent annual increase of 15% above best estimate mortality. This increase is intended to reflect level, trend, and volatility risk.

Internal model enhancements

For any company with a large book of life insurance and extensive underwriting experience, the standard formula assumption may be conservative when compared with level and volatility risk. Companies employing an internal model will likely want to separate an analysis of level/volatility risks from a consideration of trend risk.

Level and volatility risks

Volatility risk will depend on the size of the book of business and is a relatively straightforward calculation. In the case of blocks with few insured lives, concentrations of risk can lead to material discontinuities in benefit cash flows. In this situation, a company may opt to use stochastic scenarios with randomized dates of death subject to survival curves that reflect the level and trend risks modeled for each scenario.

Level (or basis) risk may vary by type of business, and will depend on many considerations. Whether past experience can provide a fully credible estimate of level risk will depend on the extent to which future circumstances can be expected to follow the past. In particular, it is necessary to reflect on considerations such as:

- Has the average age or duration of the in-force book evolved significantly?
- Are new circumstances emerging for which prior data was not available—for example, a growing book of renewals on term business?
- Have lapse rates changed, with potential consequences for anti-selection?
- Have underwriting standards recently been revised?
- Have new products, new markets, or new distribution channels materially affected the mix of in-force business?
- Is the actual-to-expected ratio level or does it vary by duration?

If very little evolution in the in-force block is anticipated, an analysis of historical actual-to-expected ratios can support an estimation of level risk.

If considerations such as those described in the list above suggest that in-force drift[4] is likely, it becomes essential to segment the in-force book into blocks that accurately reflect variations in anticipated mortality experience. Level risk factors can be developed separately for each block. These separate factors can be incorporated into forecasts that measure impact on NAV for the purpose of evaluating an aggregate 99.5% impact. Especially when this segmentation leads to a large number of small blocks, Monte Carlo methods may facilitate the development of trend risk assumptions.

4 In this context, we are using in-force drift to mean changes in the demographic attributes of the block that will affect actual-to-expected ratios in absence of any change in the underlying mortality of a population.

A secondary consideration for level risk relates to what may be described as long-term underwriting risk. Underwriting places individual lives (select or substandard) in distinct mortality classes. Each underwritten block will have mortality attributes that differ from blocks that have not been underwritten. Data clearly demonstrate that this underwriting effect wears off over time. The selection period used in the United States for select and ultimate mortality has lengthened from five years in the not-so-distant past to 25 years in recent industry tables. Does a “preferred” status or a substandard rating fully wear off over time so that all lives ultimately grade to standard? While experience indicates that the underwriting impact will wear off to some extent over time, the degree and timing is not fully known.

Trend risk

While lifestyle and environmental factors could dampen a century-long trend of mortality improvement, very few experts anticipate deteriorating mortality in the future. A best-estimate forecast is likely, therefore, to include some assumption as to future improvements. Trend risk then measures the extent to which the actual emerging trend differs from that underlying the best estimate assumption.

The flat 15% factor employed in the standard formula to reflect level, trend, and volatility risk implicitly involves an average over time. In reality, actual mortality is likely to diverge from the best estimate assumption, with the divergence growing over time. Historical mortality improvements have ranged from 1% to 2% per annum in developed countries. If our best estimate liability assumes 1.5% per annum improvement and our stress test assumes half of this level (i.e., 0.75% per annum), it will take close to 20 years before actual mortality exceeds our forecast by 15%. Thus, for a company with a large block of business and little level/volatility risk, the 15% factor gives the appearance of conservatism. We need to be very careful, however, not to assume that trend risk is driven fully or even primarily by biometric considerations. For example, the competitive market situation—in particular, the extent of price competition—will affect policyholder behavior, lapse rates, anti-selection, and therefore mortality. In a market such as Japan, where margins have historically been high on mortality coverage, companies are likely to see less improvement in in-force block mortality than may be suggested by improvements in population mortality. Anti-selective lapses are likely to increase as pricing becomes more and more competitive. The current global low interest rate environment adds complexity to this situation. Even in markets where mortality risks have been competitively priced, if investment and insurance coverage is combined, some degree of anti-selection should be expected in a rising-rate environment. When rates rise, current policyholders will likely find attractive investment opportunities at competing institutions. Policyholders who are in better health are in a better position to seek risk protection elsewhere.

Stochastic and Bayesian methods

Stochastic methods may facilitate analysis of anti-selective lapses. One possible approach is to calibrate a mortality distribution that explicitly reflects correlations to the lapse assumption.

Historical analysis demonstrates that annual mortality improvement fluctuates from year to year, and over long-term periods, across both gender and attained age. A stochastic model should recognize and project long-term trends that affect both genders and multiple ages while still capturing annual volatility. As an example, Milliman’s Risk and Economic Volatility Evaluation of Annuitant Longevity (REVEAL)[5] software can be used to analyze historical mortality rates and generate random scenarios of future mortality improvement. Projections reflect long-term trends and annual volatility while still maintaining correlations across age and gender that are consistent with historical annual mortality rates. Each component may be modified with scalars to test variations that may fall outside of the range of historical results.

The interaction of policy design elements, policyholder behavior, the macroeconomic environment, the competitive environment, and mortality are nonlinear and difficult to model employing standard statistical techniques. However, stochastic mortality generators can be incorporated into a more comprehensive Bayesian model. As discussed in later sections of this white paper, Bayesian and other techniques allow us to integrate expert opinion, historical data, and accruing information to develop enhanced forecasts and monitoring techniques.

5 Milliman’s REVEAL is a system developed to analyze longevity risk. REVEAL generates stochastic projections on pension and annuity liabilities with volatile assumptions (i.e., baseline mortality, mortality improvement, extreme mortality and longevity events, and annuitant [or plan participant] behavior, such as retirement dates and benefit elections). Find more information about REVEAL at http://www.milliman.com/Solutions/Products/REVEAL/.

IV. LONGEVITY

Though longevity risk is tied to our ability to forecast survival, considerations differ significantly from those pertaining to mortality. Longevity risk is not simply the complement of mortality risk. It emerges primarily as a result of long-term guarantees. Volatility is rarely an issue, except in the case of small blocks where any discontinuity in cash flows may be significant. However, level risk and trend risk are highly relevant. For example, insurers pricing pension buyouts may be exposed to immediate and significant level risk, depending on the quality and quantity of pension experience data available. Further, the effect of trend risk may be amplified because it can linger for many years with ongoing and potentially growing financial impact.

When modeling business exposed to both mortality and longevity risk, it can be important to recognize the correlation of trend risk across age and gender. Because correlations will vary by age and gender, there will not be a direct offset, but stochastic models can capture and maintain the historical relationships. Models that directly capture this information can allow companies to better understand and respond to their risk exposures.

Standard formula calibration

The standard formula incorporates longevity risk through a 25% immediate and permanent decrease in mortality, irrespective of policy attained age or duration.

Internal model enhancements

In calibrating the standard formula, CEIOPS examined historical mortality improvement in nine countries. Mean annual rates of improvement and the standard deviation of these rates were calculated. Assuming that the mortality improvement rates were normally distributed, a stochastic model was used to determine 99.5th percentile survival probabilities by attained age and remaining coverage duration. The 25% shock employed in the standard formula represents an initial shock to mortality that aggregates the individual survival probabilities over age and policy duration.

In construction of an internal model, a variety of enhancements are clearly possible, including:

- Varying mortality improvement by block of business; in particular, regional and other demographic factors can be reflected
- Explicit reflection of in-force block attributes; in particular, studies in the United Kingdom and elsewhere have uncovered significant cohort effects
- Explicit reflection of product design attributes; for example, deferral periods, certain periods, and joint benefit provisions may materially impact the 99.5th percentile calculation
- Causal modeling of mortality; rather than simply assuming a normal distribution, it is possible to examine and explicitly model future mortality based on assumptions as to public health, lifestyle, and medical innovation

We noted in earlier research that internal models developed to comply with Solvency II may produce excessive asset requirements compared with true economic-based internal models. As such, appropriate economic-based management tools may help companies manage their capital positions.[6]

We offer further comments on each of these aspects below.

Regional and demographic factors

Even among developed countries, significant differences exist in mortality rates and life expectancies, including life expectancy at age 65 or above. When developing nations are included in a comparison, variation is, of course, even greater. It is particularly interesting that even among developed nations, there have been significant differences in mortality improvement. These observations are illustrated in Figure 1, which shows female life expectancy at age 65 for various countries, along with the increase in life expectancy that emerged from 1960 through 2000.

6 Silverman, S. & Simpson, P. (October 2011), “Modelling Longevity Risk for Solvency II,” Milliman research report. Available at http://us.milliman.com/uploadedFiles/insight/life-published/modelling-longevity-risk.pdf.

In 1960, Japan and Sweden were developed nations with a high standard of living. Female life expectancy at age 65 was 16.7 years for both nations. In the 40 years that followed, Japanese life expectancy improved to a materially greater degree than it did in Sweden, so that by 2000, Japanese female life expectancy at age 65 was 24.0 years, compared to 21.0 years in the case of Sweden.

FIGURE 1: AGE 65 FEMALE LIFE EXPECTANCY FOR SELECTED COUNTRIES IN 2000

COUNTRY	LIFE EXPECTANCY	YEARS GAINED, 1960 - 2000
Japan	24.0	7.3
Sweden	21.0	4.3
Germany	20.8	5.4
OECD	20.5	4.4
United States	20.0	4.5
Mexico	18.0	2.8
Russia	16.5	n/a

Source: OECD (2011), "Life expectancy and healthy life expectancy at age 65," Health at a Glance 2011: OECD Indicators, OECD Publishing.

A company that operates in multiple countries needs to reflect the longevity in each country not only when choosing a best estimate of current mortality but also when constructing scenarios for improvement.

Within a single country, there will be material differences in longevity among different demographic groups, especially depending on socioeconomic class and type of benefit (e.g., private annuity contract, public pension, private pension, or group pension). For example, while U.S. overall life expectancy significantly lags behind that of Japan, mortality rates among upper middle class Americans are much closer to the Japanese average. In addition, differences exist between geographic regions in the United States due to different cultural, lifestyle, and healthcare delivery factors. These will affect not only current mortality levels but also anticipated trends.

Cohort effects

Mortality studies conducted in any country or subgroup of a population will reveal differing mortality improvement rates by age and gender. A more subtle but highly relevant phenomenon is the cohort effect. Mortality improvement rates can vary significantly by birth year. These differences can persist throughout a lifetime—in particular, far into old age. Extensive analysis of the cohort effect has been carried out in the United Kingdom. Important work has been completed in Japan and many other countries.

In the United Kingdom, the cohort born between 1925 and 1944 has experienced significantly greater mortality improvement than the cohort born before or after. This effect has been attributed to a large number of factors, including childhood diet, public health measures during childhood, and the effects of smoking.

Data from the U.K. Institute of Actuaries (The Cohort Effect: Insights and Explanations by R.C. Willets) is reproduced in Figure 2:

FIGURE 2: AVERAGE ANNUAL POPULATION MORTALITY IMPROVEMENT RATES BY DECADE MALES, ENGLAND AND WALES

AGE GROUP	1990	1995	2000
40 – 44	0.1%	0.3%	0.9%
45 – 49	**3.1**	0.4	0.2
50 – 54	**2.3**	**2.9**	1.7
55 – 59	**3.1**	**2.4**	**3.5**
60 – 64	**3.6**	**3.6**	**3.2**
65 – 69	2.1	**3.8**	**4.5**
70 – 74	2.0	2.5	**4.6**
75 – 79	2.2	2.6	2.8

The remarkable aspect of this data, as well as data from other countries, is that an effect that may have causes in childhood, or even in the life circumstances of one's parents, can persist into old age. The differentials shown in Figure 2—for example, improvements in a range of 3% to 4% for one cohort, versus zero to 1% for another—are clearly material and are highly relevant to pricing and capital management for annuity business, reinsurance, and portfolio transfers. It is possible that the improvement seen for some cohorts in the middle ages may not continue as they pass into older ages and the causes of mortality begin to change (e.g., the onset of degenerative diseases).

Policy provisions

Policy provisions obviously have a direct impact on the degree of risk assumed by the insurer or other entity. To the extent that selection effects are involved, there are material secondary effects as well—that is, the evolution of experience will depend on the nature of the benefits offered. For example, in the case of deferred annuities, we need to consider the impact of issues such as: Are annuity guarantees offered from the time of policy issue? Does the underlying pricing or reserving table assume improving mortality rates? In the payout stage, is there a certain period? Are there liquidity options that may promote anti-selection?

Causal mortality models

Through aggregate analyses of historical mortality improvement, it is possible to gain a broad sense of a range of likely future results. However, in order to optimize pricing, reserving, and capital management, it is often essential to construct a causal mortality model and to link that model to the specific attributes of a portfolio in question. In the context of Solvency II, a sophisticated stochastic projection model such as Milliman's REVEAL allows professionals to gain insight into the tail of a distribution, facilitating construction of a 99.5% stress scenario.

The sources of mortality improvement have varied dramatically over time. There has been a complex interaction between mortality rates by cause and factors such as scientific advances, culture, public policy, and approach to medical delivery.

By combining an analysis of historical experience with medical expert opinion, it is possible to craft credible assumptions about the potential mortality improvement in a particular population. This allows us not only to forecast average improvements but also to gain an understanding of the degree of uncertainty associated with each variable. By introducing an appropriate distribution (to reflect, for example, anticipated skewness and the likely properties of the tail of the distribution), we can develop stochastic analyses that yield insights into deviations over time and the emergence of extreme scenarios.

As an example, historical experience did not provide a blueprint for the rapid increase in deaths attributable to AIDS and the subsequent decline that resulted from new treatments. By conducting stochastic simulations of shifts and discontinuities in causal mortality, we can gain insight into attributes of experience that are hard to capture solely through retrospective analyses of mortality improvement.

Consideration should also be given to how mortality rate volatility may arise from additional causes:

1) Extreme long-term events: Beyond the trend risk associated with historical levels of mortality improvement, it is conceivable that events may cause mortality rates to change faster and more abruptly than anticipated. These changes could result in higher or lower experienced mortality. For example, a medical breakthrough can lead to a quick and long-term reduction in death rates related to a specific condition or disease. Alternatively, a new drug-resistant virus could cause an immediate and long-term increase in deaths by infection. Tools such as REVEAL allow users to project extreme long-term mortality events based on assumptions as to underlying individual causes of death, including but not limited to infections, neoplasms, circulatory diseases, and external causes.

2) Catastrophic short-term events: Unlike the extreme longevity occurrences, some deviations in mortality trends may have a significant temporary impact before experience reverts to the norm. Examples of these types of events include pandemics and acts of terrorism.

These additional sources of volatility tend to have effects that move mortality in one direction, producing potentially large and asymmetric results. Stochastic tools can facilitate our understanding of low-probability/high-impact events and how to mitigate the financial risk associated with them.

V. DISABILITY/MORBIDITY

Forecasts of morbidity frequently entail a greater degree of uncertainty than do those of mortality or survival. This is in part because medical advances may have a more immediate and significant impact on medical/surgical/disability benefits than they do on mortality. Also, macroeconomic and societal/cultural factors come into play more dramatically. Trends in medical delivery, government reimbursement policy, provision of long-term care, economic growth, and unemployment can all have material and sometimes unexpected consequences for insured benefits. As a result, morbidity models need to carefully reflect context and take into account more than just biometric factors in the evaluation of likely future claims and severe stress scenarios.

Standard formula calibration

This greater uncertainly is reflected in the standard formula. In contrast to the 15% mortality stress, the standard formula incorporates a stress of 50% of first year morbidity, followed by a permanent increase of 25% in all future years. Where policy benefits depend on a concept of recovery (as in some disability, long-term care, or daily hospitalization plans), the standard formula stress is based on a 20% permanent decrease in recovery rates.

Internal model enhancements

Morbidity is a very broad term, encompassing a wide number of policy types.

In almost all cases, companies defining internal model stresses will want to reflect differences by benefit type, country, and the demographic profile of in-force policyholders.

It is interesting to reflect on the impact of policyholder behavior in the context of claims and benefit utilization. In the case of life insurance, benefits will depend on policyholder choices such as election of a renewal option, surrender benefit, or paid-up option. Annuity benefits will vary depending on a policyholder's option to annuitize and may be affected by policy liquidity features.

In the case of morbidity, there is a very wide range of potential policyholder behaviors. For example:

- Disability incidence and recovery will be affected by the state of the macro-economy and a person's occupation and "work ethic."
- Long-term care utilization will depend on economic status and home care options.
- Surgical benefits may depend on a patient's own preference for invasive versus less invasive procedures.

The future level of morbidity benefits will usually correlate with a jurisdiction's approach to medical delivery and financing. Changes in medical delivery may lead to lower costs in one benefit but higher costs in another. For example, an increasing prevalence of hip replacement surgeries may increase hospital/surgical costs in the short term but lead to a reduction in disability or long-term care benefits.

Government policy may lead to increased or decreased costs in the private sector. For example, government initiatives that place pressure on hospitals to reduce the length of medical stays should lead to a reduction in payments on daily hospital plans. Cost-shifting efforts, on the other hand, may simply move costs from the public to the private sector.

A Bayesian approach to morbidity forecasts

In order to cope with complex dynamics such as those described above, Milliman has developed a Bayesian approach coupled with Monte Carlo simulation to construct forecasts of morbidity experience.

Because our ultimate goal is to develop a distribution associated with expected claims, we incorporate information about policy benefit structures and company claims management practices into our model.

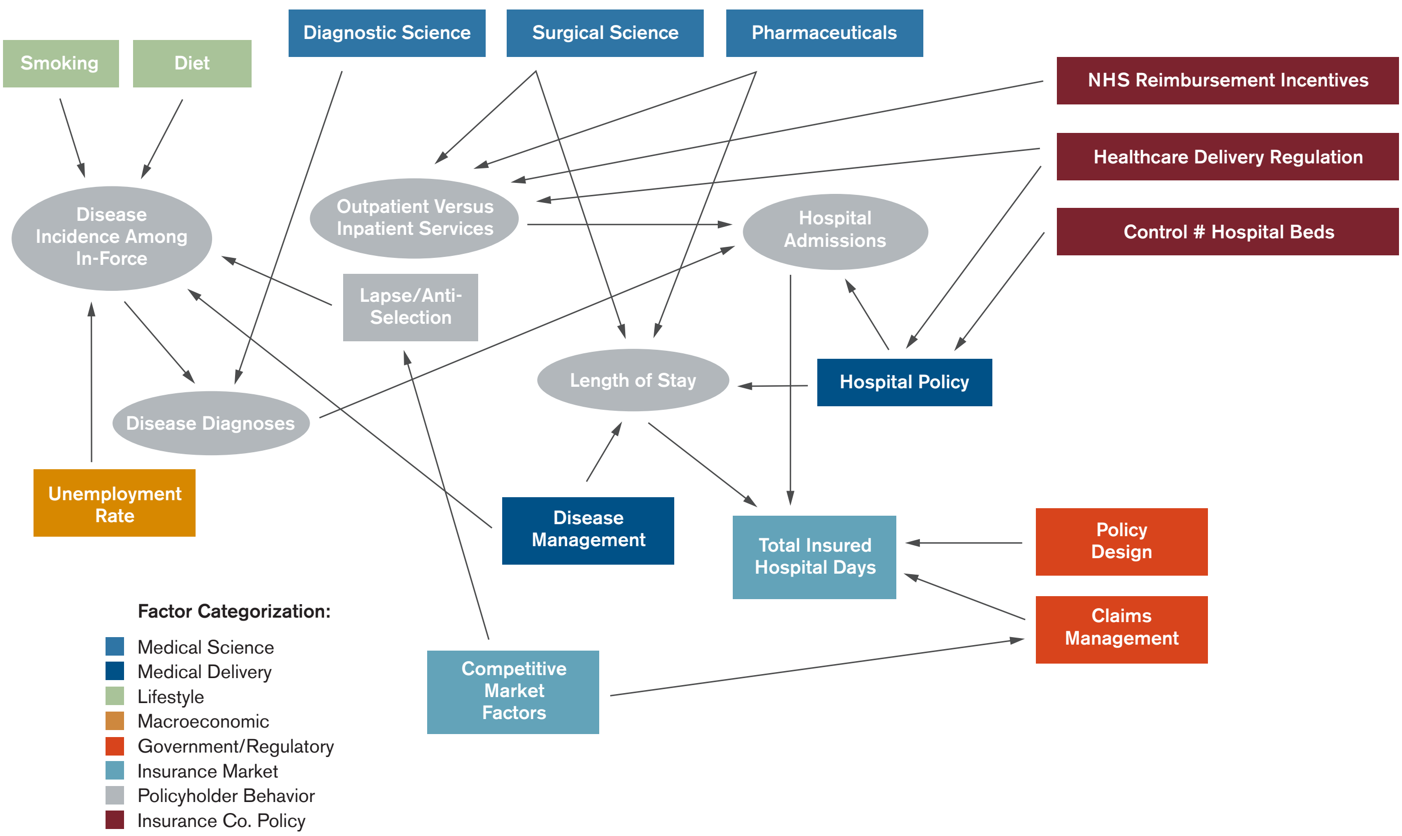
Our model typically reflects the following factors:

- Medical science
- Medical delivery
- Lifestyle
- Macroeconomic factors
- Government/regulatory factors
- Insurance market
- Policyholder behavior
- Insurance company policy

A key point is that many factors beyond the underlying prevalence of disease in a given population will affect trends in morbidity claims. When developing forecasts, insurance companies often introduce broad trend factors based on historical data. However, because such a large number of biometric and societal factors are involved, future improvement may differ materially from that experienced in the past. More elaborate models may incorporate forecasts tied to underlying incidence of disease–cancer, heart disease, stroke, etc. As a first step in the modeling process, disease incidence is viewed as primarily biometric in nature and as the primary driver of insurance claims. Even with disease incidence, however, it is hard to separate biometric factors from those driven by medical practice. Diagnostic techniques and definitions of disease change over time. In addition, for an in-force block at time zero, claims incidence will relate to factors such as underwriting, anti-selection, approach to medical delivery, regulation, the macro-economy, and the competitive environment. When companies offer products with long-term guarantees (e.g., as in whole life hospital, critical illness plans, or long-term disability), additional model sophistication can offer essential insights. For example, a Milliman analysis of available data coupled with expert opinion suggests that the standard deviation of benefit payments tied to hospital lengths of stay, surgeries, and outpatient treatments may be high. For any company writing substantial volumes of this business, it is critical that it understand the degree of uncertainly when developing pricing, reserving, capital management, and reinsurance strategies.

The Bayesian diagram in Figure 3 describes relationships among many of the factors that will ultimately drive insurance claims on an apparently straightforward daily hospitalization plan. Using Monte Carlo methods and the Bayesian relationships described in the diagram, it is possible to develop a distribution of future claims consistent with a well-defined set of assumptions about the drivers of future experience. This approach allows management to gain insight into the importance of specific risk drivers as they pertain to claim forecasts and the uncertainty inherent in these forecasts. After careful vetting, the resulting claims distributions can be incorporated into internal models that are used for planning, capital management, and risk management purposes.

FIGURE 3: BAYESIAN PROJECTION: JAPANESE MARKET DAILY HOSPITAL PAYMENTS



In the Bayesian diagram, each of the eight key factors described above is decomposed into sub-factors that are subject to independent analyses. For example, two critical lifestyle factors are diet and smoking. A distribution describes the potential percentage of smokers in the population by broad demographic category, at a future point in time. Similarly, diet can reflect changes in basic parameters such as fat and carbohydrate intake, or a more direct measure of diet/exercise, such as a future distribution of the population by body mass index. These lifestyle factors, along with macroeconomic factors (in this example, unemployment) and medical delivery factors (in this example, disease management) are the key drivers of projected future disease incidence. Assumptions as to lapse and anti-selection allow us to move from a model of the general population to a model of disease diagnoses on a specific block of insured lives. Hypotheses as to the evolution of medical science affect both disease diagnoses and the course of treatment following diagnosis. Various other interventions and impacts of government and insurance company policy are introduced.

The design and parameterization of a Bayesian model like the one illustrated in Figure 3 is admittedly difficult. In designing such a model, we need to take into consideration the availability of data and the potential to incorporate expert opinion. Calibration, testing, and model use require creativity and appropriate caution. However, companies will gain important insights into risk dynamics through the model-building process and through the process of interpreting results. In particular, the Bayesian approach allows us to gain a better understanding of contingencies and relationships among our variables. The model illustrated in Figure 3, developed for the Japanese market, allows us to gain insights into questions such as: What will happen if the Japanese population continues to adopt a Western diet? How could changes to government reimbursement policies affect claims in the private sector? Could growing competition expose us to material lapse-driven anti-selection? How could medical advances shift delivery of services between inpatient and outpatients, leading to materially differing claim patterns? Questions broadly of this nature will be of interest to risk managers regardless of jurisdiction.

VI. LIFE EXPENSE AND OPERATIONAL RISK

At first glance, calibration of capital requirements pertaining to life company expense may not generate a passionate discussion or high degree of controversy. Nonetheless, this is arguably an area where the standard formula is inadequate. On the one hand, the stress scenarios prescribed under the standard formula appear mild—or at least are based on a clear assumption that the future will not look like the past. Also, it is very hard to separate considerations of life company expense from those that pertain to operational risk. Though operational risk is developed independently of insurance risks under Solvency II, we offer a few comments in this section. The Bayesian approach described in other sections of this white paper is naturally applied in the realm of operational risk, and can provide important insights into expense risk dynamics.

Standard formula calibration

The standard formula expense risk stress is based on the following two components:

- An increase of 10% in future expense, relative to best estimate, in all years
- An increase of 1% per year above a best estimate inflation assumption

EIOPA described the primary drivers of expense risk as:

- Staffing costs
- Commission costs
- IT costs
- Costs of land and buildings occupied

Risks, of course, will vary significantly depending on individual company circumstances.

The following critical assumption underlies the selection of these stress assumptions:

> ***The undertaking operates in a macroeconomic environment where inflation, though subject to fluctuations, is broadly under control (i.e., inflation targeting)***[7]

Especially with the reference to inflation targeting, there is a clear assumption that public policymakers—legislators, administrators, central bankers—will be more judicious than has been true in the past. A look at history over the past century, as shown in Figure 4, suggests that the standard formula assumption is less severe than may be implied by this history:

FIGURE 4: PERIODS OF INFLATION IN DEVELOPED ECONOMIES

COUNTRY	YEARS	AVERAGE ANNUAL INFLATION
United Kingdom	1973 – 1983	14.0 %
Germany	1923 – 1924	hyperinflation
Japan	1972 – 1977	12.6 %
United States	1977 – 1982	9.7 %

If our modeling pertains to a developed economy, a hyperinflation event is something we can reasonably exclude from consideration. In a given developed economy, these events appear to occur with a frequency of less than one in 200. In any case, capital levels become essentially irrelevant in the face of a disruption of this magnitude.

It is worth pausing to ask ourselves whether the era of double-digit inflation is well beyond us. In the last century, there have been three significant nearly global episodes of double-digit inflation: those associated with the two world wars and those that emerged during and immediately after the two oil crises of the 1970s.

7 EIOPA (July 25, 2014), "The underlying assumptions in the standard formula for the Solvency Capital Requirement calculation."

Over the past 30 years, inflationary pressures have receded due to a large number of factors, including:

- Globalization; in particular, increased access to resources and inexpensive labor
- Technological advances
- Aging demographics in the developed world

It is highly unclear whether the stable prices of the past several decades are a reflection of sage public policy or simply the result of a substantial receding of inflationary pressures due to factors such as those outlined above. In this environment of low growth, low inflation, and low interest rates, many governments, including the United States, Japan, and European economies large and small, have increased borrowing to levels rarely seen in peacetime. Given this accrual of debt[8] and the fact that countries have often used inflation as a means of reducing the real value of debt, it is unclear whether the macroeconomic assumption implicit in the standard formula is prudent and in keeping with a one-in-200 philosophy.

We are certainly not suggesting that companies allocate large amounts of new capital to protect against the possible emergence of severe but inherently unpredictable geopolitical events. We would caution against a Panglossian view as well. A protocol of stress testing based on geopolitical hypotheses should be viewed as one essential part of the risk management process.

Internal model implications

A key point is that life expense risk cannot be viewed in isolation. A scenario that leads to expense inflation beyond that envisioned by the standard formula will have implications for most other risks; certainly for investment risks, but also mortality, morbidity, lapse, and operational risk.

As an example of one approach to the analysis of these risks, we refer the reader to Milliman's operational risk modeling framework.[9]

In this document, we discuss a rigorous framework that can be applied to develop quantitative insights into operational risk and related macroeconomic and insurance risks.

This modeling framework can be used to develop insights into questions such as:

1. What factors may combine to produce extreme outcomes?
2. What countermeasures may be effective in the case of an inflationary environment?
3. How can organizational structure be changed to assure a more resilient operation?
4. What are key relationships between product design, inflation, expense risk, and policyholder behavior?

We encourage companies to reflect on these and other issues in the course of model development and use.

8 It is also worth noting the growing benefit obligations/declining workforces in many countries.
9 J. Corrigan, P. Luraschi, & N. Cantle (February 2013), "Operational risk modeling framework," Milliman research report. Available at http://us.milliman.com/uploadedFiles/insight/life-published/operational-risk-modelling-framework.pdf.

VII. LAPSE RISK

Lapse risk emerges through a complex interaction of macroeconomic, financial market, competitive, and policyholder behavior factors. It is very hard to measure and mitigate using traditional statistical and financial techniques. Milliman has pioneered the use of cognitive mapping, policyholder surveys, predictive analytics, and Bayesian statistical techniques to develop insights into lapse risk that are unachievable with traditional data analytics. Some of the techniques are described in this section; an illustrative case study is presented in Section VIII.

Standard formula calibration

In calibrating the standard formula, the intent was to separate lapse risk emerging from non-financial market considerations from that tied to experience in the financial markets. Thus, the standard formula shocks are intended to reflect changes to lapse experience that may emerge due to company reputation and competitive market conditions. It is worth noting that such factors may lead to either higher or lower lapses. A concern over company finances may lead to high lapses, especially on investment-intensive products. On the other hand, a flight to quality could lead to lower lapses at well reputed companies.

The standard formula requires the effect of three shocks to be calculated:

- Up shock: The impact of a permanent, increase in lapse equal to 50% of the base rate
- Down shock: The larger impact between (shock lapse rate = 50% times base lapse rate) and (shock lapse rate = base lapse rate – 20%)
- One-time shock: The impact of a mass lapse event of 40% or 70%, depending on the type of business

The 99.5% scenario represents the largest impact on NAV of these three shocks.

Internal model enhancements

Lapse risk is often so interrelated with other risks—including, for example, market risk, mortality risk, and operational risk—that it is very difficult to separate the impacts of financial and non-financial risks, as envisioned in the standard formula.

Some of the factors that influence policyholder behavior include:

- Product type
- Distribution channel
- Company reputation
- Competitive market situation
- Company retention strategies
- Demographic variables, including age, gender, occupation, and income
- Macro-economy, in particular, growth rates and unemployment rates
- Level of market interest rates, especially as compared to policy credited rates
- “Moneyness” of guarantees, especially in the context of variable annuities
- Broader financial market variables and performance, including exchange rates and equity indices
- Policyholder health
- Policyholder financial situation
- Policyholder alternative protection and investment opportunities
- The degree to which a policyholder may be viewed as “rational” from a financial market perspective

To date, companies have primarily employed traditional statistical techniques to develop assumptions regarding policyholder behavior and dynamic lapsation. In particular, historical lapse performance is tracked according to a few key variables—such as product type, distribution, and policy duration—and the relationship of lapse experience to financial market performance is assessed.

This approach has yielded some valuable insights, especially in the realm of variable annuities, where it is clear that policyholders pay reasonably close attention to embedded options. Lapse rates on variable annuity business are often low when policy guarantees are "in-the-money." They tend to increase rapidly as policies move from in-the-money to out-of-the-money (due typically to favorable equity market experience), but decline again if equity performance remains favorable.

With respect to general account products, experience such as that from the United States in the late 1970s and early 1980s offers some insight into experience that may emerge when rates begin to rise. There is abundant evidence to support a belief that policyholders pay attention to policy guarantees and market interest rates. When market interest rates materially exceed guarantees, lapse rates increase. More precise statements—in particular, formulas linking market rates, credited rates, and lapse rates—are difficult to derive from available data. Indeed, it is likely that a simple formula will not adequately capture the complexity of behavior and the evolution of behavioral patterns over time. Experience from the United States and other markets suggests that even in the case of rapid rise in interest rates, single-year lapse rates rarely rise above an approximately 30% shock, though the cumulative effect over several years may be more severe. Practitioners have experimented with a variety of parametric dynamic lapse formulae, including linear, step, and arctangent approaches. Given the evolution of products and markets in recent years and the long period of declining and low interest rates, it is very difficult to parameterize these formulas with a high degree of confidence. They are a useful starting point, but need to be supplemented with new techniques.

A combination of techniques

In absence of policyholder behavior data for a single line of policies covering a wide range of macroeconomic/financial eras, it is difficult to precisely parameterize a dynamic lapse formula.

As a result, in order to develop credible forecasts, it is essential to combine a range of techniques. In addition, it is necessary to develop "bridging" techniques—that is, ways to use experience on one block to make credible hypotheses about likely experience on business having broadly similar attributes.

Milliman's approach is application driven, but typically includes some or all of the following:

- Traditional data analytics
- Predictive analytics techniques
- Cognitive mapping/expert opinion
- Policyholder surveys and simulations
- "Bridging" techniques
- Bayesian models
- Iterative model refinement

A key point is that where complex human behavior is involved, it will not be possible to develop a fully precise and definitive formula. We are faced with challenges that at first glance may seem insurmountable—for example, sparse data, heterogeneous data, and constantly changing circumstances. Nonetheless, techniques are available to significantly increase our understanding. These techniques allow us to bring actionable information to management. We point our readers, in particular, to the Milliman VALUES study, which analyzes more than 100,000,000 quarterly observations on variable annuity contract holder behavior.[10]

10 Milliman (2015), "Milliman VALUES 2014: Variable annuity industry lapse experience study overview." Available at http://us.milliman.com/uploadedFiles/insight/2015/milliman-values-14.pdf.

Each technique is described in further detail below.

Traditional data analytics

The starting point for any modeling exercise is almost always a simple review of available data. What do we know about lapse rates by policy type, distribution channel, issue age, gender, and duration? Clear patterns that can be expected to persist should be present. Lapse rates will typically be higher in the early durations or when there is a step-up in premiums. Paid-up plans may have lower longer-term lapse rates. Lapse rates may be more volatile on business written through bank channels than on business sold by tied agents. Age effects are frequently observed. Anomalies can be examined and efforts made to explain them.

Simple regression analysis may facilitate an understanding of trends and relationships.

Predictive analytics techniques

Predictive analytics techniques allow refinements in analysis that are often not possible with traditional techniques, due to factors such as:

- Large volumes of potentially heterogeneous data
- A large number of data attributes
- Complex and unexpected interactions of attributes
- Time series phenomena

A very wide range of predictive analytics techniques are available. Different techniques can be applied to different types of problems. Also, the techniques differ in terms of their explanatory potential. For example, neural networks may offer a very powerful approach to uncovering unexpected relationships among variables. However, these and other techniques can appear to the user as somewhat in the nature of a "black box." It is difficult to understand why the relationship exists—why the apparent driver may be predictive. This makes it more difficult to explain results to management and to implement actions that might lead to better future results. Also, if the predictive technique does not offer insights into the "why" of the relationship, we should have a concern over whether the relationship may persist.

Other techniques, such as decision trees, are more intuitive. Most predictive analytics projects will require empirical testing of a variety of techniques to determine which is the most appropriate in a given situation. If we can find a predictive relationship with simpler, more intuitive approaches, this is typically the best choice.

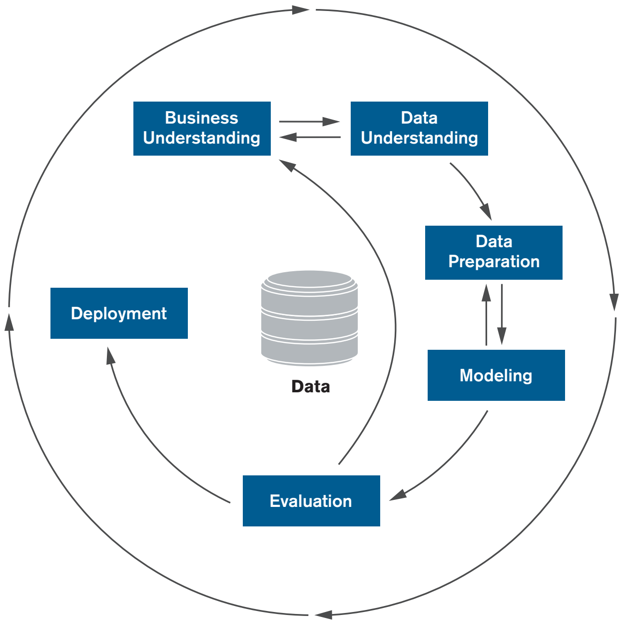
While there is a great promise for predictive analytic techniques in the realm of insurance risk management, their power should not be oversold. Predictive algorithms form an essential part of our toolkit, but are only one of many.

It is worth observing that business applications to date have typically involved situations where vast volumes of data are available, and the goal of the predictive algorithm is relatively narrow. For example:

- Credit scoring: How likely is an individual to default?
- Fraud detection: What type of behavior is suggestive of fraud?
- Sales prospecting: Which individuals are most likely to buy our products?
- Product recommendations: Which books or movies will our customers want to buy?

Substantial effort is required to move the modeling process from one of discovery toward implementation; that is, the development of an efficient, well-defined application that is integrated into operations. Various efforts have been made to systematize the predictive analytics process, such as the CRISP-DM approach, described by the diagram in Figure 5.[11]

FIGURE 5: CRISP-DM PROCESS DIAGRAM



The CRISP-DM process diagram emphasizes that any predictive analytics exercise must begin with domain knowledge or business understanding. A successful modeling exercise is dependent on a solid understanding of the business and the type of problem we are trying to solve. As a next step, we examine data needs and availability. We can begin the exercise with data that is readily available. However, a key part of the process will be the accumulation of data going forward. It is highly unlikely that at the outset we will have an ideal database. Over time, through the CRISP-DM process, we will gain a better understanding of the data we need. We can accumulate value over time by gathering and maintaining this data.

Once we have analyzed and prepared available data, the modeling process can begin. Evaluation of model results will enhance our business understanding, which in turn allows us to refine our data-gathering approach and the models we deploy.

Cognitive mapping/expert opinion

A significant and typically untapped source of information about policyholder behavior is disbursed throughout almost any insurance organization.

Professionals who deal with policyholders—whether in prospecting, sales, benefit payments, or in-force management—will often have useful insights into behavior and trends.

Milliman conducts cognitive mapping exercises aimed at teasing out this knowledge and converting it to a form that can complement traditional and predictive data analytics. In the process of discovery, responses to an initial set of broad, open-ended questions allow a facilitator and the participants to progress toward a more and more focused line of inquiry. The facilitator will prepare questions in advance that are appropriate to the context and goals of the exercise, but will invariably develop additional lines of questioning depending on the topics and issues that emerge.

Cognitive mapping is valuable as an independent exercise, but is also valuable as the starting point for a more comprehensive process of Bayesian model construction.

11 Jensen, K., "CRISP-DM Process Diagram," own work, licensed under CC BY-SA 3.0 via Commons. Accessed December 15, 2015, at https://commons.wikimedia.org/wiki/File:CRISP-DM_Process_Diagram.png

Policyholder surveys and simulations

Companies rarely take advantage of one of the most reliable sources of information about policyholder behavior: the policyholders themselves. Carefully designed surveys can provide precise information about policyholder behavior in a very efficient manner. The survey information gathered can be used to develop hypotheses about policyholder behavior. As actual experience begins to emerge, predictive analytics techniques can be used to test hypotheses and calibrate models. New surveys can be developed to hone in on aspects of behavior that emerge empirically or through prior surveys.

In order to complete the process efficiently, a carefully selected representative sample of policyholders must be chosen. Questions must be very carefully developed to capture multiple aspects of behavior and to remove bias that can inadvertently be introduced during a survey process.

"Bridging" techniques

It is often the case that we have very little data available that is directly applicable to a block of business of interest. Nonetheless, through an analysis of experience on business that has certain similar attributes, we can gain important insights into the block we are primarily interested in. Various "bridging techniques" can be used–for example, when two blocks:

- Share a distribution channel
- Have some overlapping policy features
- Share other risk drivers
- Have overlapping policyholders

With respect to the final item, it is reasonably common for one company to issue multiple policies to a given individual. We may be able to use the prior lapse history for a given individual in our effort to develop dynamic lapse assumptions and to parameterize models. Or we can develop hypotheses about current policyholder behavior based on other transaction histories, such as inquiries, prior claims, complaints, etc. In this context, it is useful to introduce a concept of *propensity to lapse*. For a given company, this may be an attribute of an individual policyholder, or, it may be applied to a class of policyholders sharing certain attributes.

In the following paragraph, we introduce, informally, some notation to help illustrate the idea. We think in terms of a random variable that may describe lapse rates on granular blocks of business, or, for some seriatim calculation, a binary "lapse/did not lapse" for one interval of a projection. Where L is a random variable describing lapses, we model dependencies such as:

Lapse rate ~ $L(\Delta, \Sigma, E, \Pi)$, where Δ, Σ, E, Π are used to represent:

Δ = basic demographic variables such as age and gender

Σ = policy structural variables, describing, for example, benefits and premiums

E = macroeconomic information, capturing interest rates, inflation, equity values, etc.

Π = propensity to lapse

In concept, the propensity to lapse will give us information about the shape and parameters of a lapse function. Π may be viewed simply as an attribute of the individual policyholder, or may reflect other well-defined dependencies (for example, distribution channel). Although there will almost certainly be a dependence relationship between the propensity to lapse and other attributes that are of predictive value, by defining the concept, we are better equipped to use data from prior periods or from different blocks of business to develop hypotheses regarding the block of business in question. Considered in conjunction with our tools from predictive analytics, our cognitive mapping/Bayesian tools, and the bridging approach, the introduction of a general propensity to lapse can facilitate our efforts to distinguish predictive value from noise.

Bayesian models

Each of the techniques described above adds to our understanding of a complex and dynamic problem.

Bayesian modeling techniques, described previously in the context of morbidity and operational risk, offer a comprehensive way to assemble data and expert opinion in a single quantitative model.

In using the techniques described above to construct a comprehensive Bayesian model, we need to answer the following questions:

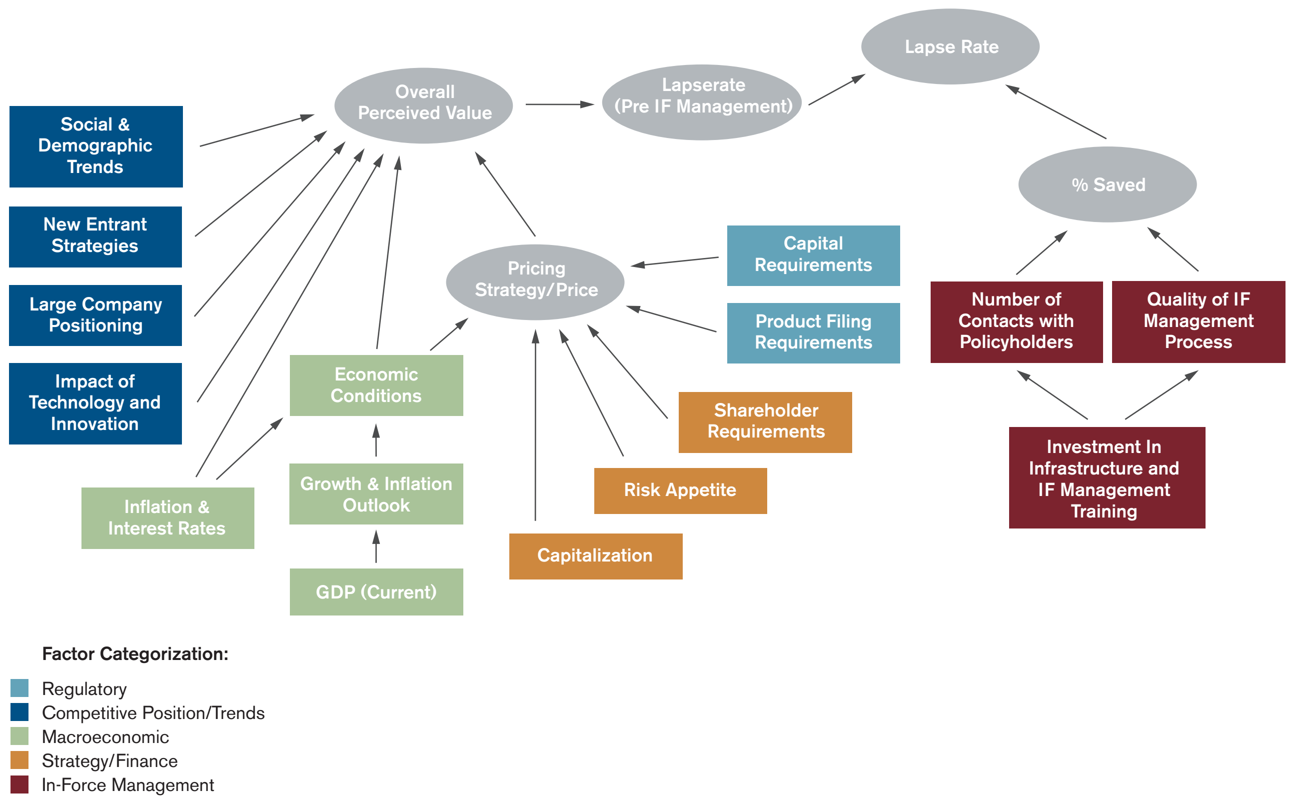
1. What policyholder attributes tie mostly closely with lapse behavior?
2. What are the drivers of this behavior?
3. How are the drivers of behavior linked?
4. Given available data and expert opinion, to what extent can we quantify the contingencies that link behaviors?

Item 3 above gets to the heart of how the Bayesian approach differs from a more traditional statistical approach. In a conventional approach, we think about correlations among variables and search for data that may give us insights into the nature of these correlations. In developing a Bayesian model, we focus more on drivers of behavior and causality. We are particularly interested in evaluating contingent probabilities. The design of a Bayesian model will need to take into consideration availability of data and expert opinion to construct, parameterize, and test the model.

A Bayesian projection of dynamic lapses

Figure 6 illustrates Bayesian relationships established with a goal of developing a lapse model (lapse distribution, reflecting expected mean and shape).

FIGURE 6: BAYESIAN PROJECTION: DYNAMIC LAPSE



VIII. EXAMPLE: DYNAMIC LAPSE ANALYTICS FOR SINGLE-PREMIUM WHOLE LIFE

We further illustrate the concepts described in the prior section by presenting an example of dynamic lapse analytics in the context of a block of single-premium whole life business, described below. The example that follows, though hypothetical, is a hybrid based on actual analytics undertaken by Milliman in various markets. The example below is described in the context of Japan, but the concepts generalize easily to other jurisdictions.

Premise

Over the past five to 10 years, in a very low interest rate environment, Japanese companies have sold substantial volumes of traditional single-premium whole life insurance. After paying a single premium, the insured will receive death protection for the whole of life. This product typically provides an underlying interest guarantee of about 1%, and may allow policyholders the option of book value withdrawals. The product poses a substantial ALM dilemma. Recent experience would suggest that lapse rates will be low, implying that policy duration is long. Even though average issue ages may be high at some companies, if current lapse rates continue, the duration of a block of newly issued policies is likely to exceed 20 years. In order to match this duration and achieve an adequate interest spread, companies generally chose to invest in long term obligations, especially Japanese Government Bonds. However, rising rates would undoubtedly lead to higher lapse rates on this business. The value of a book value withdrawal option, and the cost to companies writing this business, will depend on the extent to which policyholders actually exercise the option– i.e, the extent to which they lapse in a scenario of rising rates.

There is very little historical data directly applicable to facilitate the parameterization of a dynamic lapse formula. Because of this, most companies have made only cursory attempts to develop forecasts of policyholder behavior. In fact, by combining multiple sources of information and several analytic techniques, we can gain insights that will help us mitigate risk and improve the efficiency of capital deployment.

Traditional analysis

The starting point for our study is a traditional, static lapse study, as illustrated in Figure 7. We have simply reviewed company data for the past four years by age, gender, duration, and distribution channel. A simplified extract of results is presented below.

**FIGURE 7: ANNUAL LAPSE RATES BY POLICY DURATION, ISSUE AGE, AND GENDER
SINGLE PREMIUM WHOLE LIFE EXPERIENCE FOR 2011–2014**

Fiscal Year 2011

DURATION	FEMALE				MALE			
	AGE 35	AGE 45	AGE 55	AGE 65	AGE 35	AGE 45	AGE 55	AGE 65
1	7%	6%	4%	4%	9%	7%	5%	5%
2	5	4	4	3	9	5	4	3
3	5	4	3	2	5	4	3	2
4	4	4	2	1	4	4	2	2
5	4	3	1	1	4	3	2	1
6	3	2	1	1	4	3	2	1
7	3	2	1	1	4	3	2	1

Fiscal Year 2012

DURATION	FEMALE				MALE			
	AGE 35	AGE 45	AGE 55	AGE 65	AGE 35	AGE 45	AGE 55	AGE 65
1	7%	6%	4%	3%	9%	7%	5%	5%
2	5	5	6	6	8	5	4	3
3	5	6	5	7	6	4	3	2
4	4	4	2	3	5	4	2	2
5	4	3	3	3	4	3	2	1
6	3	2	2	3	4	3	2	1
7	3	2	2	2	4	3	2	1
8	2	2	2	2	3	3	1	1

Fiscal Year 2013

DURATION	FEMALE				MALE			
	AGE 35	AGE 45	AGE 55	AGE 65	AGE 35	AGE 45	AGE 55	AGE 65
1	7%	6%	4%	4%	9%	7%	5%	5%
2	5	4	4	3	9	5	4	3
3	5	4	5	3	5	4	3	2
4	4	4	2	3	4	4	2	2
5	4	3	1	2	4	3	2	1
6	3	2	1	1	4	3	2	1
7	3	2	1	1	4	3	2	1
8	2	2	1	1	3	3	2	2
9	2	1	1	1	3	2	2	1

Fiscal Year 2014

DURATION	FEMALE				MALE			
	AGE 35	AGE 45	AGE 55	AGE 65	AGE 35	AGE 45	AGE 55	AGE 65
1	7%	6%	4%	4%	9%	7%	5%	5%
2	5	4	4	3	9	5	4	3
3	5	4	3	2	5	4	3	2
4	4	4	2	1	4	4	2	2
5	4	3	1	1	4	3	2	1
6	3	2	1	1	4	3	2	1
7	3	2	1	1	4	3	2	1
8	3	2	1	1	3	2	1	1
9	2	1	1	1	2	1	1	1
10	2	1	1	1	2	1	1	1

Initial observations based on this traditional analysis are useful for modeling and forecasting, though are not particularly dramatic:

- Lapse rates on male lives are generally higher than lapse rates on female lives.
- Lapse rates are generally low, and are typically very low (1% to 3 %) after the fifth policy year.
- Lapse rates are higher at the younger issue ages.
- Results are quite consistent from year to year, though there is some unexpected volatility, especially among high age females in 2012.

Predictive analytics

We apply several predictive algorithms in an effort to uncover additional patterns. In particular, we are interested in trying to better understand the volatility of results experienced in 2012. Can we uncover a cause?

As a first step, we examine potential predictive attributes. Our goal is to find attributes that can be used to partition blocks so that key block attributes correlate with a differing tendency to lapse. Ideally, these correlations will be suggestive of causation. In order to gain insights into future results, we need to understand the drivers of past and likely future behavior. Some of these predictive attributes will be available as part of data that has already been captured. These include basic attributes about a policyholder—for example, age, gender, and prefecture of residence. Additional policyholder attributes that may be useful, if available, include marital status, number of children, type of employment, and income level. In addition, it may be possible to add useful predictive attributes that can be derived from the primary attributes. These secondary attributes may include, for example, information related to geography of residence, such as unemployment rates or public health data. Additional policyholder attributes could include information derived from other business in-force at the company.

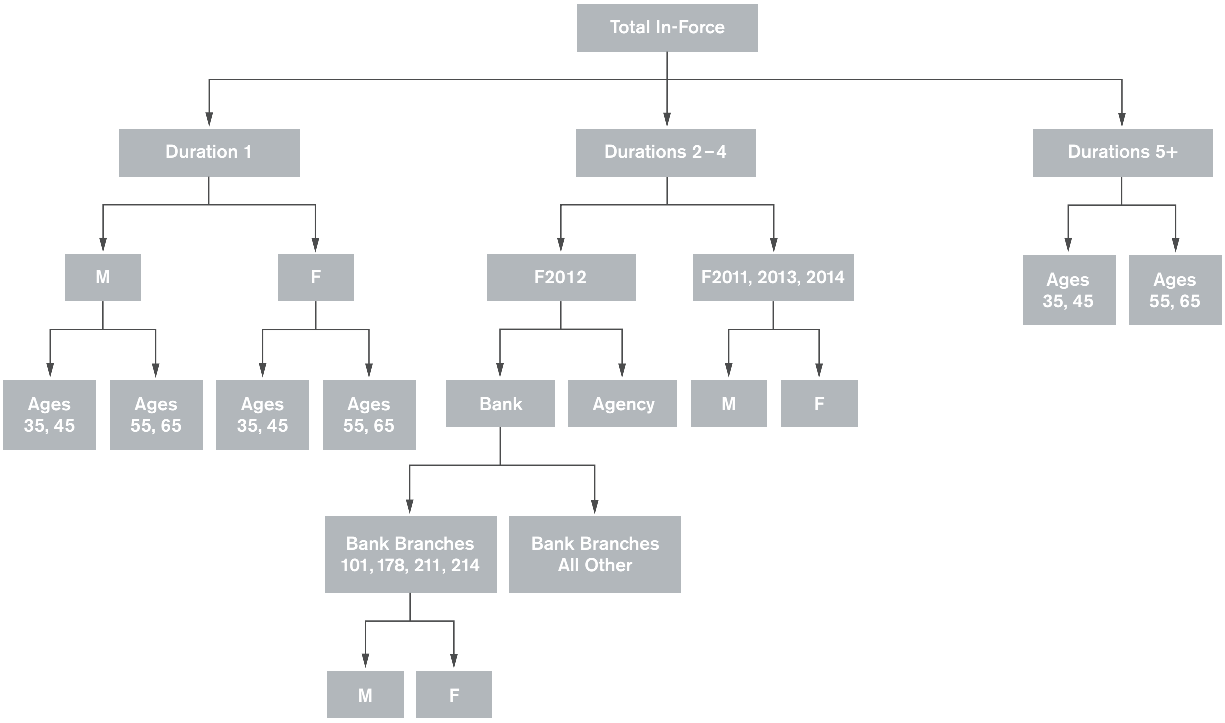
The following is a list of attributes selected for this analysis:

- Issue age
- Duration
- Gender
- Prefecture of residence
- Distribution channel (bank versus tied agency)
- Bank or agency code
- Marital status
- Income level category
- Number of children
- Number of policies
- Policy type attribute
- Prefectural unemployment rate
- Prefectural health status attribute
- Exposure year

After testing various algorithms, we chose to use a relatively simple decision tree algorithm. This algorithm iteratively partitions the in-force based on the attributes listed above, in a manner that maximizes the differential in lapse rates among blocks. The user can control various features of the algorithm, including the minimum size of each ultimate block and the depth of the "trees" that emerge in the course of application. The precise parameters chosen will be specific to a given application, but will typically err on the side of simplicity and will always aim to reduce the chance of "overfitting"– finding illusory results that are unlikely to persist over time or have clear causative explanation.

Results from one application of the decision tree algorithm are shown, in simplified form, in Figure 8.

FIGURE 8: DECISION TREE ANALYSIS OF PROPENSITY TO LAPSE ON A BLOCK OF SINGLE PREMIUM WHOLE LIFE BUSINESS



The algorithm used to split the in-force is based on the concept of information gain. In concept, we want to partition the business into blocks that have relatively similar lapse rates, with partitions based on some combination of the attributes we have selected. At each step in the construction of the decision tree, the algorithm chooses the partition under which blocks "look most similar" as it regards propensity to lapse. This similarity of state corresponds to a lower "entropy," as defined in information theory.[12]

12 The classic introduction to information theory is "The Mathematical Theory of Communication" by Claude E. Shannon. Shannon defines the entropy of an information system as $-\sum p \log p$, where the sum is over all system states and p is the probability that the system takes that state.

When adequate data is available and the algorithms are creatively applied, very powerful results can be obtained. In this example, the algorithm finds that lapse experience differs materially and reasonably consistently in the first year, mid-term, and longer term (Duration 1, Durations 2–4, and Durations 5+). This is not surprising, but things begin to get interesting in the second iteration:

- For Duration 1 business, the algorithm finds material differences by gender; after this, the most material driver is age.
- For the longer-duration business, however, gender differences are found to be immaterial; the algorithm immediately splits this block by age and terminates.
- Behavior is very different in the middle durations. The algorithm discovers surprising behavior during fiscal 2012. In this exposure year, there appear to be anomalies in the bank-issued business. As the algorithm digs deeper, it finds very high lapse rates occurring in blocks of business issued by four specific branches. In the next and final iteration of the algorithm, the high lapse rate issue seems to be most prevalent in business sold to female policyholders.

In this particular example, the predictive analysis would allow a company to introduce both operational changes and enhancements to its internal model. The year 2012 was significant because it coincided with a new product introduction. The bank branches in question were simply churning the business. New monitoring procedures could be introduced. Even with the best of controls, bank business was expected to experience higher lapse rates than those experienced on business written through its tied agents. While this assumption had previously been incorporated in the company's internal model, a more detailed analysis of the data suggested greater lapse propensity in bank sales than anticipated, even after adjusting for the more blatant churning activity that was uncovered.

Bridging techniques

After completing an initial study on a single block, as described above, techniques can be expanded to other blocks, beginning with material blocks having similar attributes. For example, variable annuity books or other savings business can be added to the study. As information is gathered, more definitive conclusions can be drawn–for example, about individual policyholders, distribution channels, individual agents and agencies, implications of product design, competition, and many other factors.

Cognitive mapping, shape of distribution, and enhancing domain knowledge

By combining advanced data analytic techniques with approaches to collecting expert opinion, we can begin to make more credible hypotheses about risk drivers, shapes of distributions, tail behaviors, and contingent probabilities. This information, in turn, can be used to help design our approach to predictive analytics. In particular, through cognitive mapping we can significantly enhance our domain knowledge, which is critical to the CRISP-DM or any predictive analytics process.

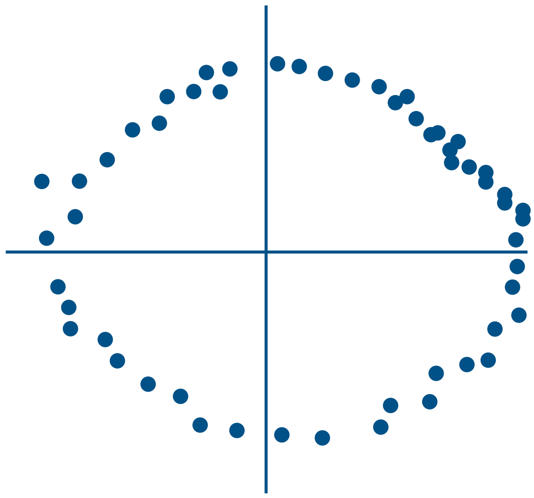
IX. FURTHER COMMENTARY ON RISK CORRELATIONS

Risk correlation has been studied extensively in the context of asset performance. For example, there is often extensive data available to design and calibrate correlation models such as those involving:

- Different classes of equities
- Equities and fixed income
- Asset classes across different geographies

Two primary approaches to modeling risk correlations involve correlation matrices and copulas. Correlation matrices in their most basic form express the pairwise correlations between a set of n random variables. If there is a broadly linear relationship among the variables in question, the correlation matrix will capture critical information about the nature of the relationship. When a complex, non-linear relationship exists between two variables, the correlation matrix will be of limited value. In particular, there are many examples of variables that have strong dependence relationships but zero correlation. A classic example is illustrated in the conceptual diagram in Figure 9. It represents trials (X,Y) from random variables X and Y that have correlation coefficient zero, but a clear dependence relationship.

FIGURE 9: TRIALS FROM DEPENDENT VARIABLES (X,Y) HAVING ZERO CORRELATION COEFFICIENT



Copulas allow aspects of a joint distribution describing correlation to be separated from elements describing the underlying shape of the distribution. Depending on the nature of tail correlation desired, practitioners may choose Gaussian, Student T, Clayton, or other copulas. Different distributions facilitate the modeling of "fatter tails" and higher degrees of correlation. Copulas may allow the modeling of relationships more complex than those captured by correlation matrices. However, calibration of copulas will often be difficult. Copula work well when there is a large volume of data available and when a reasonable degree of stability can be anticipated. As such, they are likely to work better in finance and investing than in the modeling of insurance risks.

This brings us back to the issue raised at the outset of this paper. Are the relatively simplified approaches described by the Solvency II standard formula generally adequate for reflecting life insurance risks, or is greater effort warranted? Of course, the answer will vary from company to company, but we make the following assertions:

- For many organizations, the capital level suggested by the standard formula for life risks may be well in excess of that which is required.
- However, due to risk correlations, some enterprises face risk levels significantly beyond that envisioned by the standard formula.

Let us reflect for a moment on one relatively small corner of the life risk edifice.

The relationship between mortality, lapse, and renewal rates has long been of interest to actuaries. Interesting work in this area dates to the late 1970s era of “term wars” in the United States and other markets. It was during this era that actuaries began to reflect on the drivers of anti-selection when developing mortality forecasts.[13] It was recognized that key factors tied to lapse rates and mortality anti-selection included:

- Gross premium rate patterns, in particular the slope of increasing premiums or degree of step up at renewal
- Agent incentives, primarily first-year and renewal commission rates
- Initial underwriting and policyholder health

Anti-selective lapsation and renewals caused severe stress for many U.S. companies during this era. At first glance, this appears to be a relatively clearly defined domain—projecting mortality rates on products that are relatively simple to describe. But the risk drivers quickly multiply in number. In fact, underwriting, pricing, and in-force management is difficult on these products. While mathematical methods do in fact increase our understanding and intuition, in many instances, even in this easy-to-describe domain, the drivers of experience quickly become complex. Certainly they may be hard to describe in terms of correlation matrices or copulas.

Yet if our goal is to allocate adequate capital to assure our survival in the face of a one-in-200 risk, it is exactly the correlations—more precisely, the nature of the dependences among risk drivers—that we need to understand. Perhaps we can be reasonably confident that we can understand a one-in-20 or one-in-50 insurance-type event. Examples might include the high lapses in the United States during the 1970s, the “run on the bank” that occurred at several Japanese insurers in the late 1990s, or longevity losses incurred by insurers and pension providers in the United Kingdom. Are these risks mostly independent? Do they truly represent one-in-200 contingencies? Of course, not all companies experienced dramatic elevations in lapse rates during this period, and we should be careful to distinguish a “one-in-200 event” from a “one-in-200-year event.” Nonetheless, tail risk can be exceptionally difficult to quantity. As we move further into the tail of a distribution that describes a single risk, we may find that not only are the tails fatter than we had anticipated, but also the correlations are more severe. If this is so, perhaps more capital is imperative—or perhaps it is indicative of a need for operational changes that would not have been apparent without a closer assessment of the key dependencies.

13 See, for example, Dukes and MacDonald (1980), “Pricing a Select and Ultimate Renewal Term Product,” *Transactions of the Society of Actuaries*, Volume 32.

X. INNOVATIONS AND FUTURE DIRECTIONS

Given a very difficult task and very difficult circumstances, the framers of Solvency II crafted what should be a workable and comprehensive framework for capital management—if the professional community effectively communicates the limitations inherent in any scheme dependent on long-term forecasts. Solvency II is only a start: Risk and capital management are ever-evolving disciplines. Given the remarkable pace of technological innovation, it would not be an exaggeration to say that Solvency II was conceived in a different era, or perhaps in a transitional era characterized by a migration from traditional to computationally intensive data analytics.

Innovation is always messy. When the textbooks are written, when the ideas are distilled, when we look back at what has emerged, it may look neat and clean and comprehensible. That is never the case when we are in the midst of discovery. And we are definitely now in an era of discovery. We are in the midst of an era where science and technology offer remarkable promise to insurance company customers and profit for companies with the capital and vision required to understand and adopt new technologies.

In this white paper, we have discussed innovations in a range of areas that are likely to move into the mainstream of risk management:

- Disease/cause of death models
- Predictive analytics
- Cognitive mapping
- Survey techniques integrated with risk management
- Bridging techniques
- Bayesian models/Monte Carlo simulations

As innovations in these areas move into the mainstream, companies that introduce them and collect the requisite data have the potential to enhance their rates of growth and earn superior risk-adjusted returns.

It is worth reemphasizing that much of the value to risk management is in the process itself. The purpose of a risk manager is to help an organization build value, in line with that organization's goals and risk appetite. If they are to achieve this, they must not be viewed as gate keepers or naysayers, but rather as partners and facilitators, as colleagues who help identify opportunities and manage the risk that opportunities inevitably entail.

New, emerging risk management technologies can help us move in that direction. They will help us to ask the right questions. They will help us to better frame our answers. They will help us explain risk analyses to senior management and assure that strategies and operations are consistent with them. They will help us to better understand our organizations—the risks we face, and our untapped potential.

What is a one-in-200 risk? Perhaps we can never know. Yet, ironically, we can enhance value through our efforts to answer this unanswerable question.

End note

Milliman is a leader in global research on the topics presented in this white paper and many additional areas that support risk management and value enhancement efforts. Some of our newer areas of research include applications of information theory, network theory, genetic algorithms, phylogenic approaches, and agent-based modeling. We are carrying out a variety of pilot projects in major jurisdictions around the world and invite you to join us in these efforts.

ABOUT MILLIMAN

Milliman is among the world's largest providers of actuarial and related products and services. The firm has consulting practices in life insurance and financial services, property & casualty insurance, healthcare, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

milliman.com

CONTACT

For further information, please contact your local Milliman office or:

Stephen Conwill, FSA, MAAA
stephen.conwill@milliman.com