Predictive Modeling for Life Insurers

Application of Predictive Modeling Techniques in Measuring Policyholder Behavior in Variable Annuity Contracts

April 30, 2010

Guillaume Briere-Giroux, FSA, MAAA, CFA
Jean-Felix Huet, FSA
Robert Spaul, FIA
Andy Staudt, FCAS, MAAA
David Weinsier, FSA, MAAA

TOWERS WATSON

ABSTRACT

Since the end of 2003, variable annuity (VA) assets have exceeded $1 trillion (source: Towers Watson VALUE™ Survey). The appeal of this product to the market can be attributed to the combination of a direct investment in the equity markets with the guarantees commonly offered with these products (GMDB, GMIB, GMAB and GMWB)[1]. Policyholder behavior is a primary profit and risk driver for VA business, and accurate modeling of this assumption is therefore critical for pricing, reserving, hedging and risk management. Traditional modeling approaches have attempted to reflect VA policyholder behavior patterns based on product design, policy characteristics and policy performance. However, traditional approaches typically consider only a limited number of variables and fail to adequately capture certain correlations and interactions between variables.

This paper describes how predictive modeling techniques commonly used within the property and casualty (P&C) insurance industry can be applied to more effectively model certain policyholder behaviors, specifically VA lapse rates. VA lapse rates are driven by numerous interrelated factors, many of which are not modeled under traditional approaches. A VA is a complex financial product that provides an effective platform to illustrate how a predictive modeling approach can improve upon traditional modeling techniques. The key goals of this paper are to:

- Articulate to life insurance practitioners the potential benefits of applying predictive modeling techniques in measuring complex life insurance behavior;
- Describe how a predictive modeling approach can improve upon traditional methods used to model VA lapse behavior;
- Illustrate through a case study how predictive modeling techniques can be applied in practice.

[1] GMDB: guaranteed minimum death benefit; GMIB: guaranteed minimum income benefit; GMAB: guaranteed minimum accumulation benefit; GMWB: guaranteed minimum withdrawal benefit

Table of Contents

1. Predictive Modeling Background

1.1 What is predictive modeling?

Generally, predictive modeling can be thought of as the application of certain algorithms and statistical techniques to a data set to better understand the behavior of a target variable based on the co-relationships of several explanatory variables. Simply put, it involves analyzing data in order to understand risk, which is what actuaries have been doing for many years. The difference is the use of more advanced mathematics, algorithms and larger quantities of data, which makes the analysis computationally tractable. Rather than relying on a simple understanding of basic risk elements, predictive modeling enables the user to consider many confounding factors simultaneously by mining across a set of scenarios. This analysis will allow for making more informed decisions and limit the amount of subjective judgment required.

Examples of predictive modeling exist in other industries — baseball teams are drafting talent more effectively based on analyzing leading indicators of performance, online sales sites optimize expected profitability by tailoring recommendations (and sometimes prices) to each user, and hospitals employ predictive models in the diagnostic phase, resulting in more efficient and cost-effective patient care.

1.2 Applications in the P&C industry

In the context of insurance, predictive modeling techniques have primarily been used within the P&C industry to enhance understanding of current and/or future insured risks. This knowledge has led to improved risk segmentation, underwriting, pricing and marketing decisions. For example, auto insurance premiums reflect the fact that younger drivers are poorer risks than middle-aged and older drivers, and males are poorer risks than females. However, data also show a clear interaction between age and gender, i.e., the difference in relative risk between male and female drivers is much less pronounced at older ages than at younger ages. Quantifying this interaction between risk parameters was challenging under traditional pricing approaches, but a predictive model will recognize this and other interactions, enabling the insurer to develop premiums that accurately reflect the relative risk characteristics of the pool of underlying policyholders.

1.3 Potential uses of predictive modeling in life insurance and annuity modeling

Traditional assumption setting techniques are often restricted to a relatively small set of risk factors to maintain the credibility of the results. Consequently, these methods can only take limited consideration of the correlations in the data and of interactions between factors. With use of predictive modeling techniques, however, we are not only able to consider all risk factors simultaneously but can also allow for many interactions without significantly reducing the credibility of results. This allows for a macro view while at the same time also facilitating focus on the subtle, micro interactions between risk factors. Specifically, predictive modeling enables improved understanding of the factors influencing policyholder behavior, the interaction of such factors, and the potential impact on profitability and risk.

As seen in the P&C market, innovative carriers can potentially realize competitive advantages by developing more accurate estimates of key profit and risk drivers. Examples include:

- Identifying more/less profitable lines of business or distribution channels
- Improving profitability, either by offering more competitive rates or by avoiding selection effects through pricing insurance risks more accurately
- Improving risk management via more accurate asset/liability management or hedging
- Developing more accurate estimates of economic reserves and capital
- Easing compliance with certain regulatory, rating agency and reporting requirements (e.g., Solvency II, MCEV Principles).

2. Traditional Approaches to Modeling VA Lapse Behavior

This section briefly describes the primary drivers of VA lapse behavior and the traditional approaches commonly used to model this behavior. It then discusses certain shortcomings of traditional approaches that can be addressed by the use of predictive models. It should be noted that throughout this paper we use the standard industry term “lapse” to mean full surrender.

2.1 Drivers of VA lapse behavior

Primary factors that drive VA lapse behavior are described in **Table 2.1**. The factors can be generally categorized into four groups: product and guarantee design, distribution-related factors, policy characteristics, and policy performance. The table also indicates whether traditional modeling approaches typically reflect each factor.

TABLE 2.1
Factors that Drive VA Lapse Behavior

Category	Factor	Traditional Industry Modeling Practices
Product/Guarantee Design	Surrender charge length and strength	Reflected via grading up base rate, shock, shock + 1 and ultimate lapse rates
	Share class (A-share, B, C, L)	Reflected with specific surrender charge schedule
	Presence and nature of living benefits	Reflected, but highly approximate and somewhat speculative
Policy Characteristics	Policy duration	Reflected during surrender charge schedule
	Policy size	Typically not reflected
	Policyholder age and sex	Typically not reflected
	Life stage (accumulation/ income stage)	Reflected (occasionally)
Policy Performance	Guaranteed benefits in-the-moneyness	Reflected via deterministic formulas applied uniformly to base lapse rates
	Recent fund performance	Typically not reflected
Distribution	Commission structures (heaped vs. trail)	Typically not reflected for in-force modeling (often reflected in pricing, however)
	Distribution channel/ target market	Typically not reflected beyond what is captured in the aggregate experience

A typical VA writer will develop base lapse schedules, varying rates by policy duration, to reflect the product's design and structure (e.g., surrender charge, share class) and the presence and nature of living benefits. Additional base lapse schedules are occasionally generated to capture the impact of commission structures and life stage.

In addition to a base lapse table, a dynamic lapse component is typically applied to adjust expected lapse behavior up or down to reflect the in-the-moneyness of the guarantees. The dynamic lapse piece will typically vary by the type of guarantee rider.

2.2 Shortcomings of traditional approaches

Traditional approaches to modeling VA lapse behavior have the following shortcomings:

- Historical data will show a single lapse rate, which is a function of both ‘base behavior’ and ‘dynamic behavior;’ however, traditional approaches do not easily allow for identifying which component of the single aggregate rate is base and which is dynamic. When attempts are made to separate these impacts, the credibility of the resulting groups decreases. Thus, the impact of these separate pieces cannot be precisely validated.
- Traditional approaches make suboptimal use of historical experience data. In a typical experience study, the data is 1) categorized, 2) aggregated and 3) analyzed. By splitting the data into categories, the exposure bases available to analyze a given relationship (e.g., policy year effect for a particular product) become smaller, which results in a loss of credibility. Aggregating the experience for a given variable does not control for the contribution of other variables influencing the experience for that group. This creates ‘noise’ that increases the amount of data required to extract a credible relationship when analyzing a single variable at the time.
- As shown in Table 2.1, traditional approaches typically consider a limited number of explanatory variables to account for a complex behavior. Many of these variables are readily available (e.g., age, gender, asset allocation, past withdrawals), but others could be categorized as ‘exotic’ variables that could also be collected and analyzed to help predict VA lapse behavior (e.g., indicators of financial sophistication such as credit score, education levels, profession/industry).
- Traditional approaches typically ignore interactions between variables, whereby the effect of one variable is influenced by a second variable.
- Traditional approaches do not fully account for correlations between explanatory variables, which can result in double counting effects or not attributing an effect to the right variable.

A predictive modeling approach can address these shortcomings. Through the use of a case study, Section 3 explains and demonstrates how the application of predictive modeling techniques can enhance the modeling of VA lapse behavior.

3. Case Study: Application of Predictive Modeling Techniques to VA Lapse Behavior

3.1 Case study methodology

Underlying data

Using a case study, this section illustrates how the application of a predictive model to modeling VA lapse rates can improve upon traditional approaches. The underlying analysis was performed on a large sample of hypothetical but representative data. This data was developed based on actual industry experience, with certain adjustments, resulting in an exposure base and product mix representative of a typical VA writer. The resulting data set featured a typical age, share class, fund allocation, commission type and rider mix by year of issue. More than 10 issue years are included in the data. The in-the-moneyness (ITM) for the living benefit riders (e.g., GMWB, GMIB) was representative of actual historical market conditions, including actual experience in the tumultuous years 2008 and 2009.

Traditional model

The traditional model employs a typical industry approach to modeling VA lapse rates, reflecting the following factors:

- Base rate varying by policy duration
- Surrender charge length and strength
- Shock lapse at the end of the surrender charge period
- Commission structure
- Presence and nature of living benefits

- ITM of living benefits, defined as [1–(account value / benefit base)]

Predictive model

The final predictive model, derived as a generalized linear model (GLM) as described in detail in the Appendix, is based on the following variables present in the case study data:

- Base rate varying by policy duration
- Surrender charge length and strength
- Proximity to end of surrender charge
- Commission structure
- Presence and nature of living benefits
- ITM of living benefits
- Premium (i.e., policy size)
- Fund value
- Portfolio mix (aggressive, balanced, conservative, cash)
- Attained age.

We note that the case study data did not encompass all the potential variables that could influence VA lapse rates. The study included only variables for which data was robust and credible enough to build a representative model of VA lapse rates.

3.2 Results

This section analyzes the performance of the predictive model against the traditional model.

Model validation

The data set was randomly split into two distinct groups in order to facilitate an objective model validation. The first group, made up of 70% of the aggregate data set, was used to set the model parameters. The second group, the remaining 30% of the aggregate data set, was then used to test how effectively the model predicted actual lapse behavior. That is, the first group of data was used to fit the models. These models then project an expected set of lapse rates for the policies in the second group (the “E” in an actual-to-expected study). The actual lapse experience in the second group was then designated as the “A” to see how well the models predicted actual results. In practice, once the final model form is selected, the model is refit to 100% of the data set in order to estimate the parameters of the final model.

Actual/expected analysis

Actual-to-expected results by duration and ITM are presented in **Figure 1** and **Figure 2**, respectively.

Figure 1: Actual versus expected lapse rates by duration

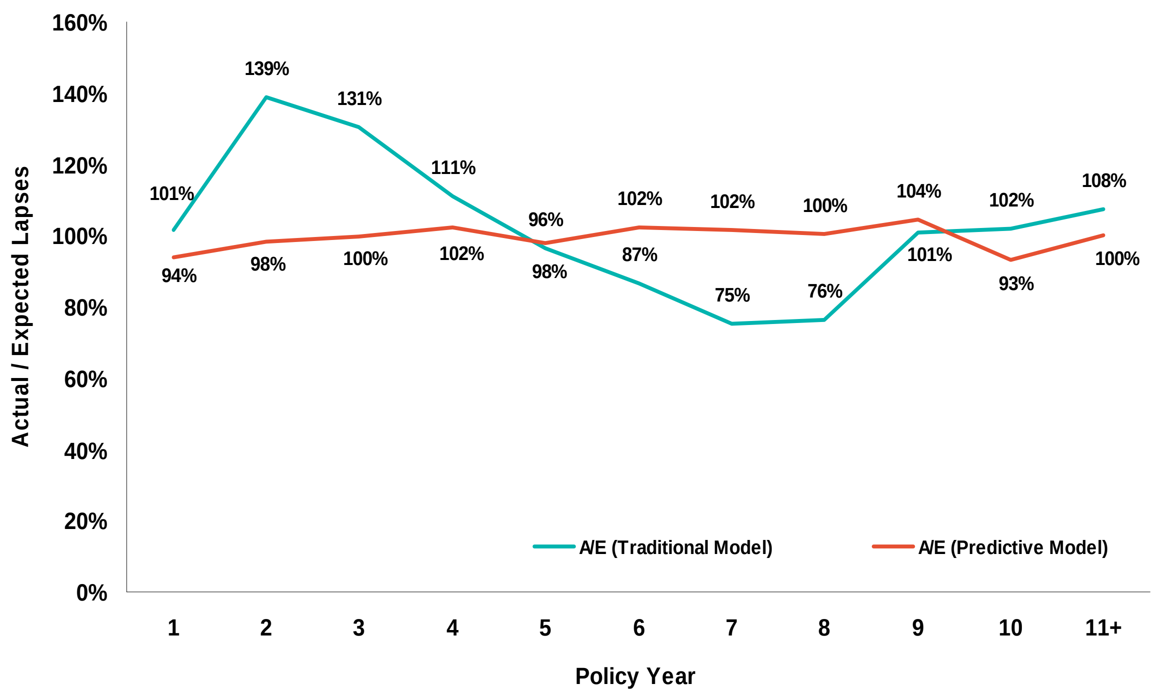
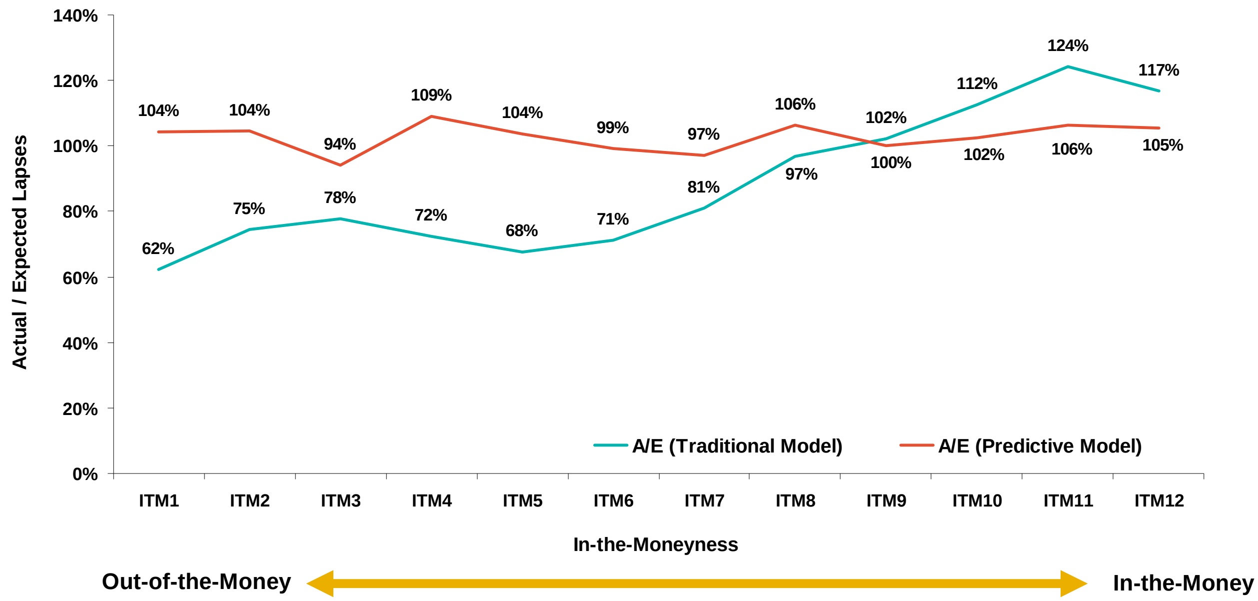


Figure 2: Actual versus expected lapse rates by ITM



The predictive model shows an appreciably better fit than the traditional model when considering actual-to-expected ratios by policy duration and ITM. This result is primarily driven by correlations between policy duration and ITM that are captured in the predictive model but ignored by the traditional approach considered. As previously discussed, a known shortcoming of the traditional approach is that it does not easily allow one to distinguish between the base portion of lapse rates and the dynamic portion of lapse rates in the underlying experience data. Another driver is that the predictive model considers additional variables that the traditional model does not. Finally, predictive models consider the interactions between certain dependent variables. An example of interaction present in the data (and captured in the predictive model) was that the impact on lapse rates of ITM (a first explanatory variable) is influenced by policy size (a second explanatory variable). The data shows that higher premium policies exhibit a stronger link between ITM and lapse behavior (likely due to higher financial sophistication).

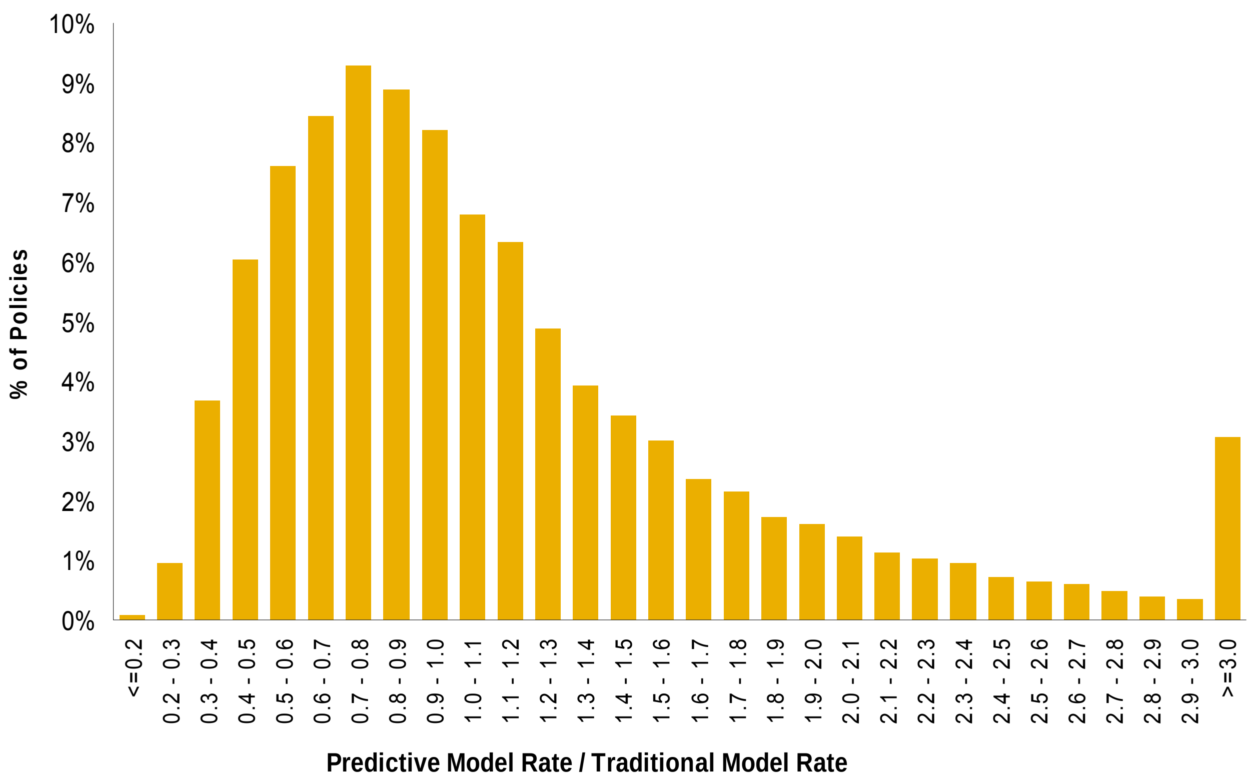
While these comparisons of actual-to-expected lapse rates by variable on an aggregate basis are useful, additional comparisons should be performed to verify the result. The next sections compare and validate the fit of the two models at more granular levels.

Policy-by-policy expected lapse rates: predictive vs. traditional

Figure 3 compares *expected* lapse rates emerging from the traditional model to the predictive model.

The x-axis is the ratio of the predictive model expected lapse rate divided by the traditional model expected lapse rate. A ratio of 1.0 indicates that the two models produce the same lapse rate for a given policy. A ratio under 1.0 indicates that the predictive modeling approach produces a lower lapse rate than the traditional model. Correspondingly, a ratio above 100% indicates that the predictive model produces a higher rate. In order to evaluate the distribution of the ratio, the ratios are grouped into categories and the relative frequency is shown on the y-axis.

Figure 3: Comparison of predictive model to traditional model expected lapse rates



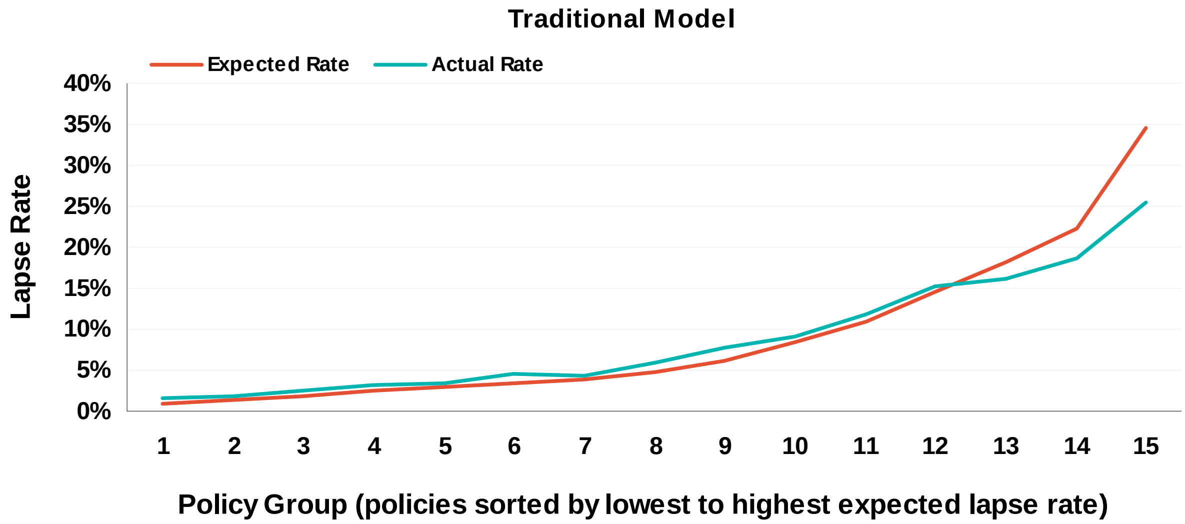
This graph tells us that, for a significant proportion of the policies, the two models produce radically different expected lapse rates. The absolute difference in the ratio is >=20% for 65% of the policies and >=60% for 23% of the policies. As shown on the far right side of Figure 3, this analysis also shows that for roughly 3% of policies, the predictive model produces a rate >=3.0 times the traditional rate, suggesting that the traditional model may have limitations in capturing the tails.

Lift Charts

A lift chart is a common tool used in predictive modeling to test the viability of the model at a granular level. To build a lift chart, policies are sorted by *expected* lapse rate in ascending order from the underlying model tested, then policies with similar expected rates are grouped together (e.g., 15 groups of equal exposure). A policy may show up in a particular group using the traditional model but a different group for the predictive model, because the models may assess the expected risk of lapse for any given policy differently. For example, a policy may fall into group 10 of 15 when policies are sorted by expected lapse rate using the traditional model. However, it may appear in group 12 when the policies are sorted by expected lapse using the predictive model, because the predictive model has assessed the policy as having a higher risk of lapse than the traditional model. The *actual* rate of lapses in each of the groups is then recorded and the results plotted. Thus, under this approach, the *actual* lapse rate for policies within a given group will vary by model.

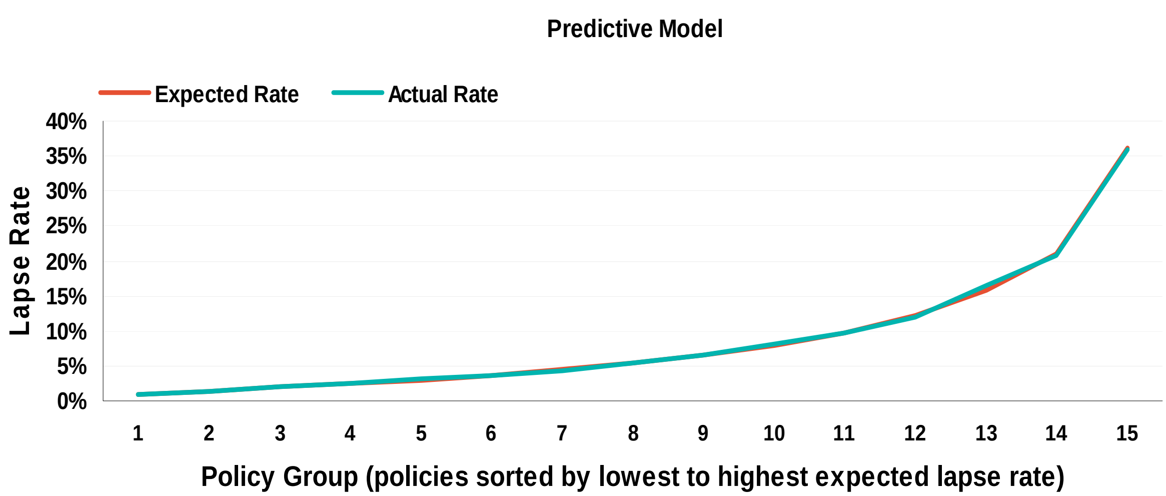
This process has been performed for the traditional model in **Figure 4a**. By definition, the expected lapse rate (red line) increases from group 1 to group 15, since group 1 contains the policies that have the lowest expected lapse rates, and group 15 contains the policies with the highest expected rates. The actual lapse rate for each of the groups (blue line) follows the expected rate fairly well.

Figure 4a: Lift chart for the traditional model showing actual and expected lapse rates



The process is repeated for the predictive model in **Figure 4b**.

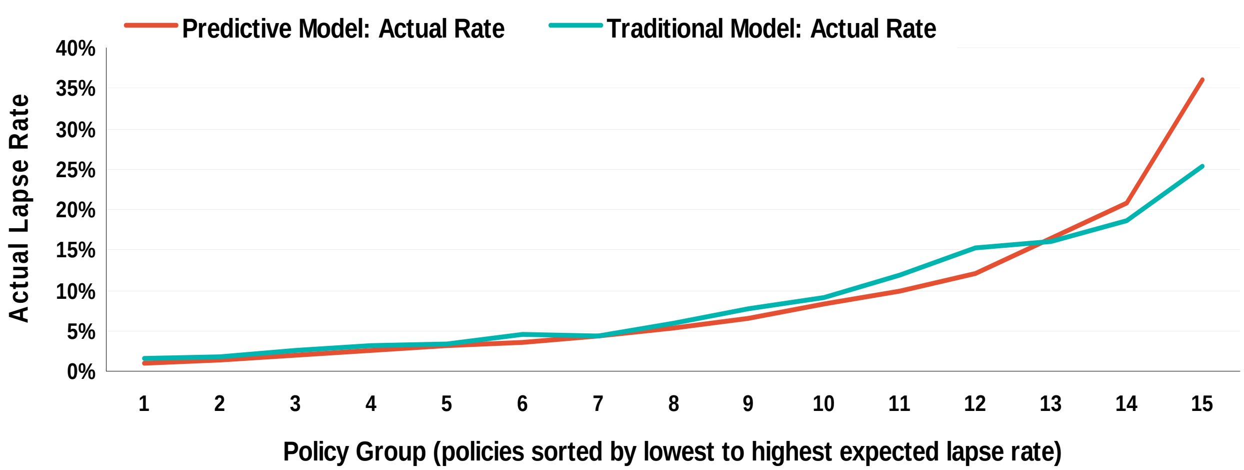
Figure 4b: Lift chart for the predictive model showing actual and expected lapse rates



Two effects are apparent: First, the actual and expected results are considerably closer across the range of groups, demonstrating the model's tighter fit to the validation data. Secondly, the range of actual lapse rates across the groups is wider with the predictive model, (1% to 36% compared to 2% to 25% in the traditional model). This demonstrates that the predictive model better differentiates between policies in the tails, indicating a better granular fit.

If actual lapse rates for each model are placed on the same graph, a lift chart can be used to compare two models, as shown in **Figure 5** following. The greater the difference between the high and low lapse rates (i.e., the "lift"), the better the model is at differentiating between policies by actual risk of lapsing. Here, the predictive model is clearly providing improved differentiation between the policies that are likely to lapse compared to those that are not, and is therefore a better model of VA lapse rates.

Figure 5: Comparison of traditional and predictive models using lift curves



Gains Charts

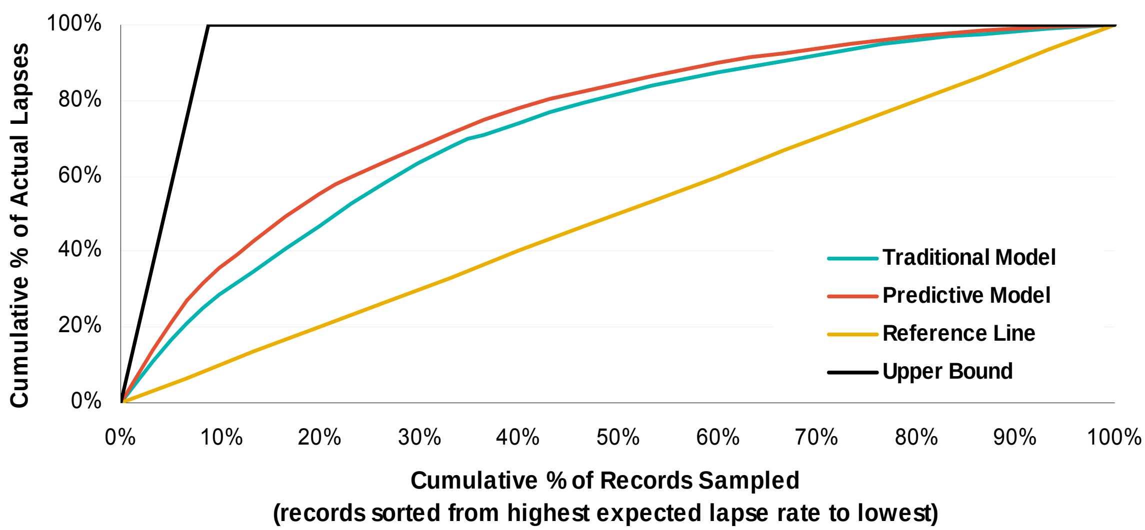
Gains charts are similar to lift charts but present the information in an alternative way. A gains chart sorts the policies by expected lapse rate in descending order. The cumulative lapse rate is then recorded as the data is stepped through policy by policy.

By definition, before the first record, the cumulative percentage of the total number of lapses will be 0%. At the end of the projection it will be 100%. If the model is no better than a random sort of the data, then we would expect a straight diagonal line that we label the Reference Line (yellow line) in **Figure 6**. In this case, 50% of the lapses have been found (y-axis) after sampling 50% of the records (x-axis). At the other extreme, a perfect model would have predicted 100% of the lapses in roughly the first 8% of records (8% is the average annual lapse rate). We label this the Upper Bound (black line).

Since the model is better than a random sort, we expect the cumulative percentage of lapses to increase more quickly than the cumulative percentage of records counted, and the line produced on the graph to be bowed to the left. The greater the area under the model line, the better the model is able to differentiate policies by risk of lapsing.

This figure shows that if the first 20% of policies are targeted, the predictive model (red line) would have predicted 55% of actual lapses, as compared to 47% for the traditional model (blue line), indicating a stronger model.

Figure 6: Comparison of traditional and predictive models using a gains chart



4. Conclusions

In this paper, we have described the potential benefits for life insurers of using predictive models, focusing on complex VA lapse rate behavior. Using a case study, we compared the performance of a traditional model to a predictive model in estimating VA lapse rates.

The predictive model had an appreciably better fit under a typical actual-to-expected analysis. To assess the models at a more granular level, we performed two additional tests. Lift charts and gain charts showed that the predictive model produces a more granular fit than the traditional model and better differentiates between policies with low and high risk of lapsing.

The overall assessment is that the predictive model can improve modeling of VA lapse behavior, relative to traditional approaches, due to the predictive model's ability to:

- Capture a greater number of risk factors (or variables) that drive VA lapse behavior
- Account for correlations between variables
- Capture interactions between variables, where the impact of one variable is impacted by another
- Use less data to achieve convergence by using seriatim data for the analysis.

The use of predictive modeling by life insurers can also lead to the following business and strategic benefits:

- More reliable pricing assumptions, less subjectivity and reduced assumption risk
- Identification of more profitable segments, distribution and target markets
- Product development based on more accurate estimates of policyholder behavior risk
- Improved risk mitigation (e.g., hedging, asset/liability management) by reducing policyholder behavior variances
- More streamlined models and better controlled model implementation, by replacing multiple tables and dynamic formulas by a single parameterized predictive model
- More accurate modeling of policyholder behavior in the tail, resulting in more accurate reserve and capital estimates.

5. References

Anderson et al. (2007), "A Practitioners Guide to Generalized Linear Models, Third Edition." http://www.watsonwyatt.com/europe/news/glmpaper/default.asp

Brieman, L., Friedman, J. H., Olshen, R.A., and Stone, C.J. (1984), Classification and Regression Trees. Chapman and Hall

Cherchiara, R. R., Edwards, M., Gambini, A. (2008), Generalized Linear Models in Life Insurance: Decrements and Risk Factor Analysis under Solvency II, 18th International AFIR Colloquium

McCullagh, P., Nelder, J. (1989), "Generalized Linear Models, Second Edition." Boca Raton: Chapman and Hall/CRC

Appendix: Developing the Predictive Model

This section describes many of the concepts involved in producing the predictive model used in the case study described in this paper. This is not necessarily intended to be a practitioner's guide, but it offers an introduction to some key concepts applied in developing a predictive model, specifically a generalized linear model. The topics covered in this appendix are outlined in **Table A.1**.

TABLE A.1
Topics for Developing a Predictive Model

Topic	Brief Description
A.1 Initial analysis	Initial analysis helps the user to understand the underlying data
A.2 Generalized linear modeling	Brief introduction to the theory of GLMs
A.3 Choice of GLM form	Decisions required for choosing the structure of the model
A.4 Choice of explanatory variables	Decisions required for choosing which variables to include in the model and which to exclude
A.5 Interpreting parameter estimate graphs	How to interpret the graphs that describe variable effects within a model
A.6 Modeling continuous variables	A method for including continuous variables in a model
A.7 Interactions	A method for including the effect of an interaction within a model

A.1 Initial analysis

Introduction

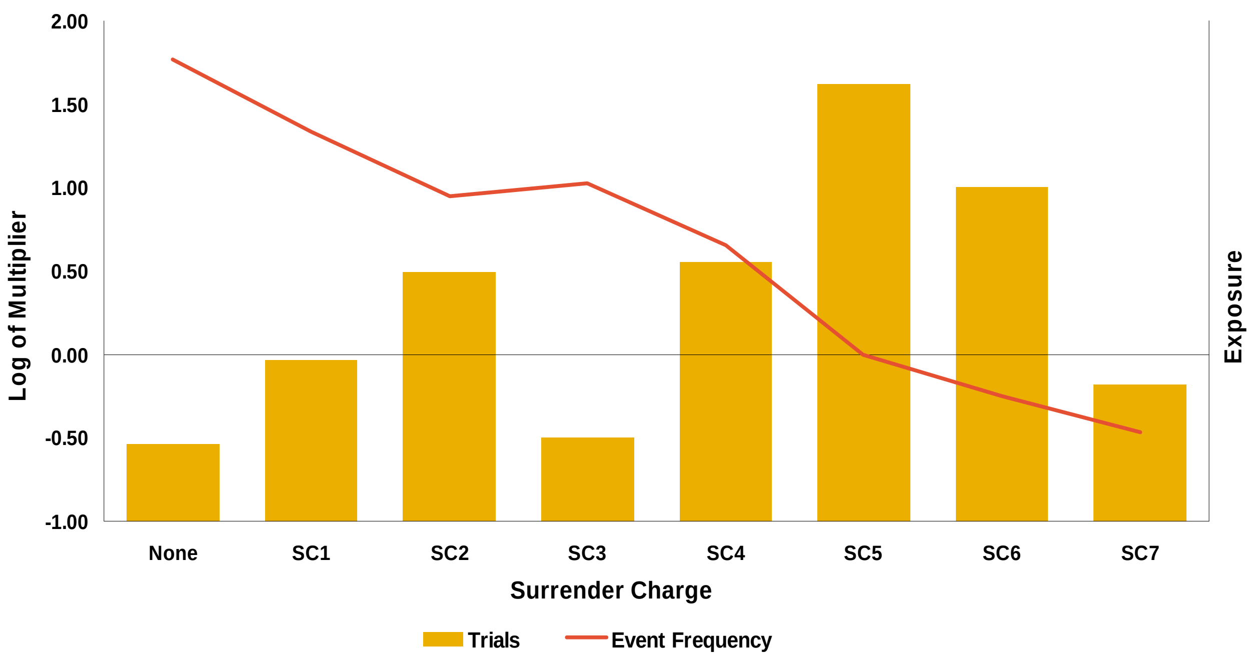
As with any experience analysis, the first step is to check and clean the data and decide on the business in scope for the analysis (e.g., in terms of product coverage and time length, frequency of study or "time step" used for the analysis). It is important to note changes in policy classifications over time, as well as make sure that the various in-force statistics are appropriately captured, etc. This step was straightforward for the case study described in this paper because of the use of hypothetical data.

It is sensible for any multivariate analysis to start with steps to help the modeler to get a sense of the data that they are working with and to check for errors that may have been missed during data cleaning. This will typically include univariate analysis (e.g., the study of lapses by one variable at a time), as well as correlations between the explanatory variables and bivariate analyses (e.g., study of lapses by a combination of two variables).

Univariate analysis

Figure A.1 shows the effect of surrender charge levels on lapse rates in the case study data. Results are shown relative to a particular base level, often the level with the most exposure.

Figure A.1: Univariate results by surrender charge



Such analysis allows us to see if the grouping of numeric variables is sensible; in this example the modeler may decide to split the SC5 level to improve precision across this range. Univariate results also highlight potential significant variables, but we must be wary because correlations in the underlying data may be distorting the results. At this stage we don't know whether surrender charge is a key driver of lapse rates or whether the effect is caused by a variable correlated to surrender charge (e.g., policy year). Further analysis is required to determine the true effect.

Correlations

Correlations are calculated between the surrender charge and the other explanatory variables using the Cramer's V statistic, which takes a number between 0 and 1 where 1 indicates either perfect positive or negative correlation. A typical rule of thumb for interpreting Cramer's V statistic is that values less than 5% imply very weak dependence and values greater than 30% imply very strong dependence. **Table A.2** indicates that the type of benefit and the policy duration are both strongly associated with surrender charge. This makes sense, given that surrender charges vary by policy duration and that VA living benefits are still mostly in the surrender charge period. These results should be kept in mind when considering whether the effect shown in the univariate analysis is driving lapse rates or whether the effect of these other correlated variables is impacting the results. To answer these questions, we need to perform a multivariate analysis that considers many variables simultaneously and allows for correlations in the data.

TABLE A.2
Correlations of Variables with Surrender Charge Using the Cramer's V Statistic

Variable	Cramer's V Statistic
Commission structure	6%
Size of fund	5%
Major asset class	13%
Type of benefit	29%
Age	12%
In-the-moneyness	19%
Policy duration	34%
Proximity to end of surrender period?	14%
Premium	6%
Intensity of withdrawals from the fund	6%

A.2 Generalized linear modeling

Generalized linear models are just one type of a vast family of multivariate regression models. They are standard models for many risks in the P&C industry as they provide a framework that:

- Is relatively easy to understand
- Allows for a wide range of statistical diagnostics
- Is extremely flexible to a wide range of distributions, input variables and applications.

Explanations of GLM theory that detail the underlying assumptions of the approach used are widely available for the interested reader. What follows is a basic description to help the reader understand some of the implementation considerations and results.

For this study, the lapse rate is assumed to have a distribution from the exponential family, and the expected rate for a given policy is assumed to have the following form:

$$E[L_i] \equiv \mu_i = g^{-1}(\eta_i)$$

where μ_i is the mean lapse rate, g^{-1} is known as the inverse link function and η_i is called the linear predictor.

The linear predictor consists of an intercept term, α , plus a set of covariates, $\underline{X}_i$ (that explain the details of each policy, *i*) and parameters that are applied to them, $\underline{\beta}$ called betas.

$$\eta_i = \alpha + \underline{X}_i . \underline{\beta}$$

In advance, a base policy is determined where values for all the explanatory variables are defined (e.g., male, 60 years old, five-year policy duration, policy with a GMWB benefit that is 10% in-the-money). Often, variable values are chosen that provide the maximum exposure. The intercept term is the value for the linear predictor for this base policy. The combination of betas and covariates for every other policy determines how the linear predictor is adjusted from the intercept term (e.g., female, 65 years old, etc).

This linear predictor will produce a number for each policy that lies on the real line (-∞, +∞). The inverse link function, g^{-1}, transforms this value to an annual lapse rate on the real line (0, 1).

The problem becomes how to find the parameters and the intercept term. Values are chosen to maximize the log-likelihood for the observed lapses in our data (using the assumed distribution) and these are found using iterative numerical techniques.

To present these results in an easy-to-understand format, it is possible in some cases to turn the results of the GLM into a set of multiplicative relativities and a base level. The example in **Figure A.2** has a base lapse rate of 10% corresponding to a policy in its 10th year which is neither in nor out of the money. The same policy in its eighth year has a lapse rate 150% higher. In this approach, the relative effect of the levels becomes clear.

Figure A.2: Example of a multiplicative approach to setting lapse rates

Policy Lapse Rate = Base Level × Duration Multiplier × ITM Multiplier

Base Level
10%

X

Duration	Multiplier
3	0.3
5	0.5
8	2.5
10	1
:	:

X

ITM	Multiplier
-30%	1.2
-10%	1.2
0	1
20%	0.8
40%	0.5
:	:

A.3 Choice of GLM form

Within the framework of generalized linear modeling, the form of the model must be chosen. The main choices will include, but are not limited to, the link function (g^{-1}), the error distribution (ε) and the combination of explanatory variables.

When modeling VA lapse behavior, several choices are available; some of the most common are shown in **Table A.3**.

TABLE A.3
Model Structures for Modeling VA Lapse Behavior

What we are trying to predict	Link Function	Error Term
The probability of lapse	Logit	Binomial
The account value given that a lapse has occurred	Log	Gamma
The expected value of a lapse (combination of the two previous models)	Log	Tweedie
The expected time until lapse	Inverse	Inverse Gaussian

A.4 Choice of explanatory variables

The modeler has a number of choices to make in selecting the variables to include in the model. Some variables will have a significant impact on the lapse rate while others will not. Insignificant variables are removed so as to avoid overfitting to the data and to reduce the complexity of the final model. For certain variables there will be limited scope to set the number of levels (e.g., gender). However, the modeler has discretion for numerical variables that have been grouped into bands (e.g., age split in five-year groups) or where particular groups show similar effects on the results. Principle components or other advanced clustering algorithms can be used to further enhance variable selection.

Continuous variables can be treated as categorical (e.g., policy size grouped into 10 levels), modeled directly (policy size as a number [*x*]), or as a function of the variable (typically polynomials [e.g., x^3], logs [e.g. ln(*x*)], or using splines). Where the lapse rate is dependent on a particular combination of variables interaction terms can be used.

The modeler has a range of tools to assist in making these decisions, a selection of which is shown in **Table A.4**.

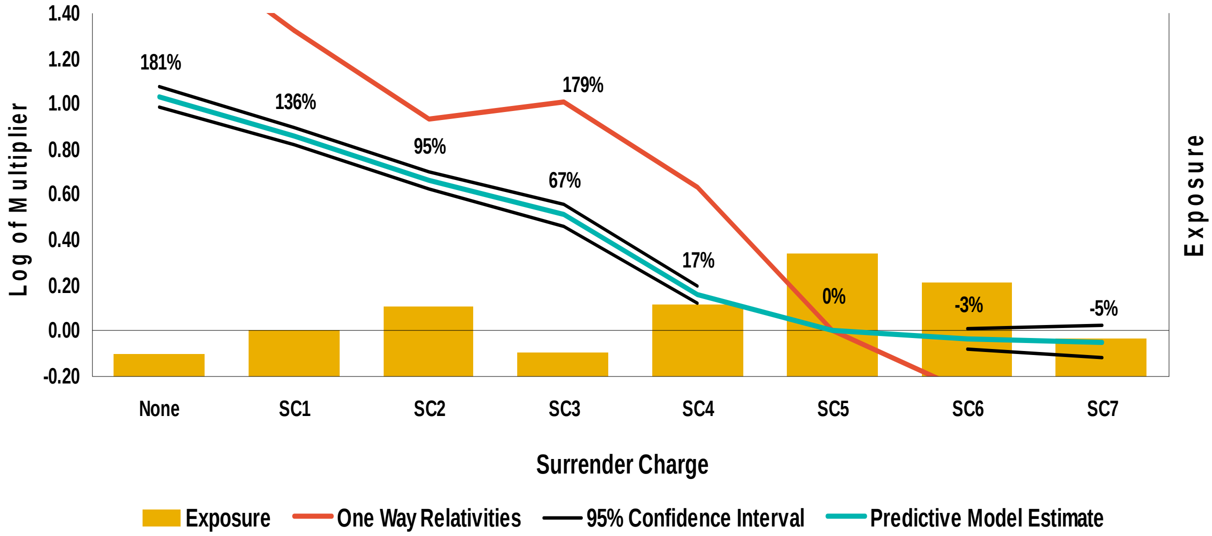
TABLE A.4
Certain Tools Available in Assessing Model Fit

Parameter estimate graphs	See section below
P-value tests	The significance of each variable can be determined through tests of significance by comparing a model with and without the variable in question
Model validation	As demonstrated in Section A.3, a model's performance can be tested against a separate sample of data using a range
Stepwise macros	Many software packages are capable of selecting models automatically through iterative model fitting. These methods should be used with care and should not replace the modeler's judgment.
AIC and BIC	The Akaike's information criterion (AIC) and Bayes information criterion (BIC) are two single figure measures of goodness of fit for a model useful in comparing similar models

A.5 Interpreting parameter estimate graphs

Of all the diagnostics available to assess fitted parameters, graphical output is often the most helpful to convey information that can be studied quickly and effectively. In this paper parameter estimate graphs are used to show the *relative* impact of levels in variables. The intercept term (the base lapse rate) is not usually a major concern when deciding on the model form.

Figure A.3: Relative effect of surrender charge within a GLM



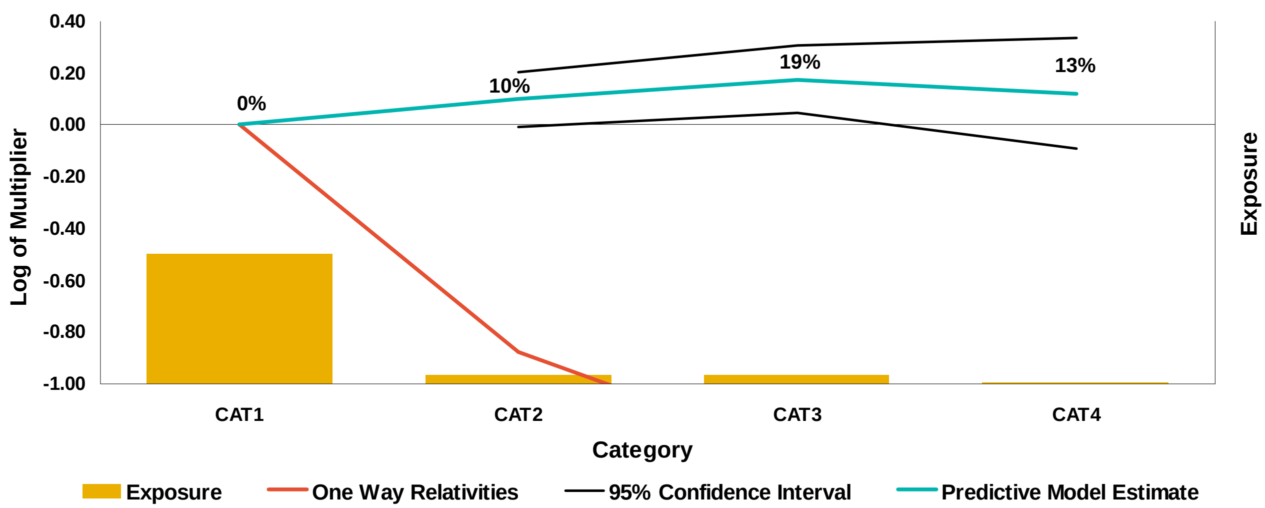
One such graph would be shown for each potential variable to include in the model. In this example, **Figure A.3**, the variable in question is surrender charge, grouped into eight levels as in the univariate analysis. The Predictive Model Estimate (blue line) shows the multiplicative impact of each level relative to the level with the most exposure from the multivariate analysis. The effect of a policy having a surrender charge class of SC3 is 67% higher than the best case surrender charge of SC5, all else being equal.

The One Way Relativities (red line) show the relativities implied by a simple univariate analysis. These relativities make no allowance for the fact that the difference in experience may be explained in part by other correlated factors. The comparison of the univariate and parameter estimates shows how the univariate effect of surrender charge is actually spread across several correlated variables. The univariate relativity for the SC3 case is 179% (i.e., the lapse rate for this category is 279% of SC5 lapses), compared to only 67% for the predictive model estimate (167% of SC5 lapses), when accounting for the effect of other variables.

The 95% Confidence Interval (black lines) indicate two standard errors on either side of the parameter estimate – the confidence interval around our estimate. The interpretation of these "confidence intervals" is not as straightforward as the name might imply, however, suffice it to say, smaller confidence intervals suggest a greater degree of certainty in where the true relativity might lie. The confidence interval will be wider when there is greater uncertainty in the parameter estimate due to either low exposure volume, where other correlated variables also explain the risk, or where the underlying experience is excessively random.

Even though the standard errors on the graph only indicate the estimated certainty of the parameter estimates relative to the base level, such graphs generally give a good intuitive feel for the significance of a factor. For example, in Figure A.3 it is clear that the factor is significant since the parameter estimates for the surrender charge levels are considerably larger than twice the corresponding standard errors for most levels. By contrast, **Figure A.4** following illustrates an example where a variable appears significant in the univariate analysis, but is not very significant when considering the impact of other variables. In this case there is only one parameter estimate more than two standard errors from zero, and we could consider removing this variable or collapsing some of the levels to improve the significance (perhaps two levels; CAT1 and CAT2, CAT3 and CAT4).

Figure A.4: Example of a factor with insignificant levels

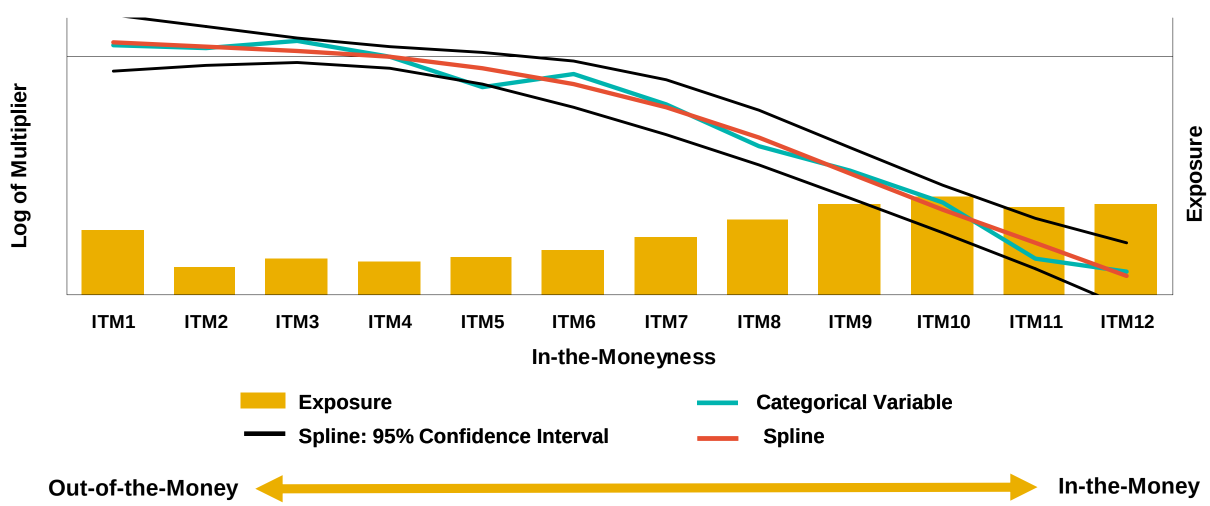


A.6 Modeling continuous variables

It is desirable to have smooth results, particularly for continuous variables where grouping can reduce the information contained, or if the model is used in pricing where smooth changes in prices are desired — across a range of ages for example. One method of achieving this is to use regression splines. The mathematics behind regression splines is beyond the scope of this paper, but this method can produce desirable effects.

Figure A.5 shows a spline used to fit a metric defining the ITM of policy guarantees (red line)[2]. The actual form of the spline is not easily graphed, but we can graph the weighted average parameter estimate that the spline would produce for each level of the categorical form of the variable. The parameter estimates found when fitting ITM as a categorical variable are also shown (blue line). It is clear that the spline has captured the effect of ITM over the range very effectively and removed some of the noise in the parameter estimates. In modeling lapse rates, different ITM metrics can be tested to see which one provides the most effective predictive model.

Figure A.5: Modeling in-the-moneyness as a continuous variable using splines



[2] For clarity, the univariate results and confidence intervals for the categorical form of the variable are not shown.

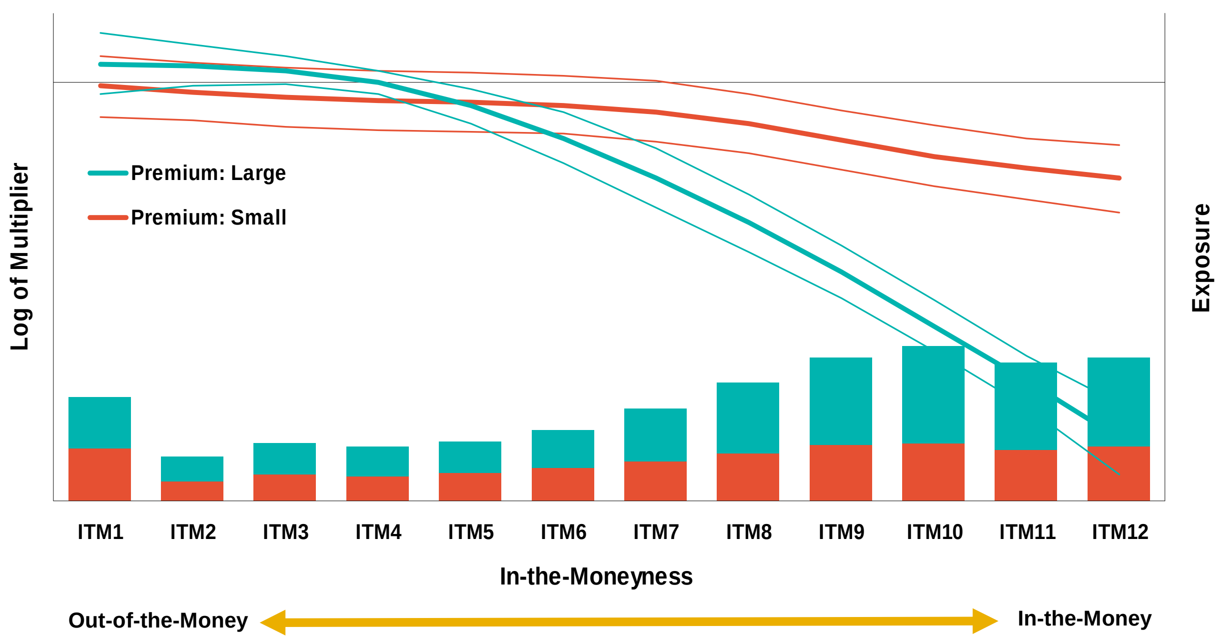
A.7 Interactions

The effect of some variables may be dependent on the values that other variables take. To use an example from auto insurance, it is widely known that young drivers are a poor risk compared to those in middle age and that males tend to have higher claim costs than females overall. What is interesting is that the gender effect on claim costs differs by age, such that young males are considerably worse risks than young females, but this effect is less significant as drivers get older.

Within the GLM framework we can allow the effect of combinations of variables through interactions.

Figure A.5 showed ITM where premium was a separate variable in the model with no interaction. In **Figure A.6**, the two separate variables have been replaced by an interaction of premium size with ITM. It is clear that the effect of ITM is very different depending on the size of the premium in this data.

Figure A.6: Modeling ITM as a spline with an interaction with premium size



TOWERS WATSON

For comments or questions, please call or e-mail:

Guillaume.Briere-Giroux@towerswatson.com	(860) 843-7083
JeanFelix.Huet@towerswatson.com	(212) 252-8061
Robert.Spaul@towerswatson.com	(516) 690-6696
Andy.Staudt@towerswatson.com	(415) 836-1328
David.Weinsier@towerswatson.com	(404) 365-1781

TOWERS WATSON