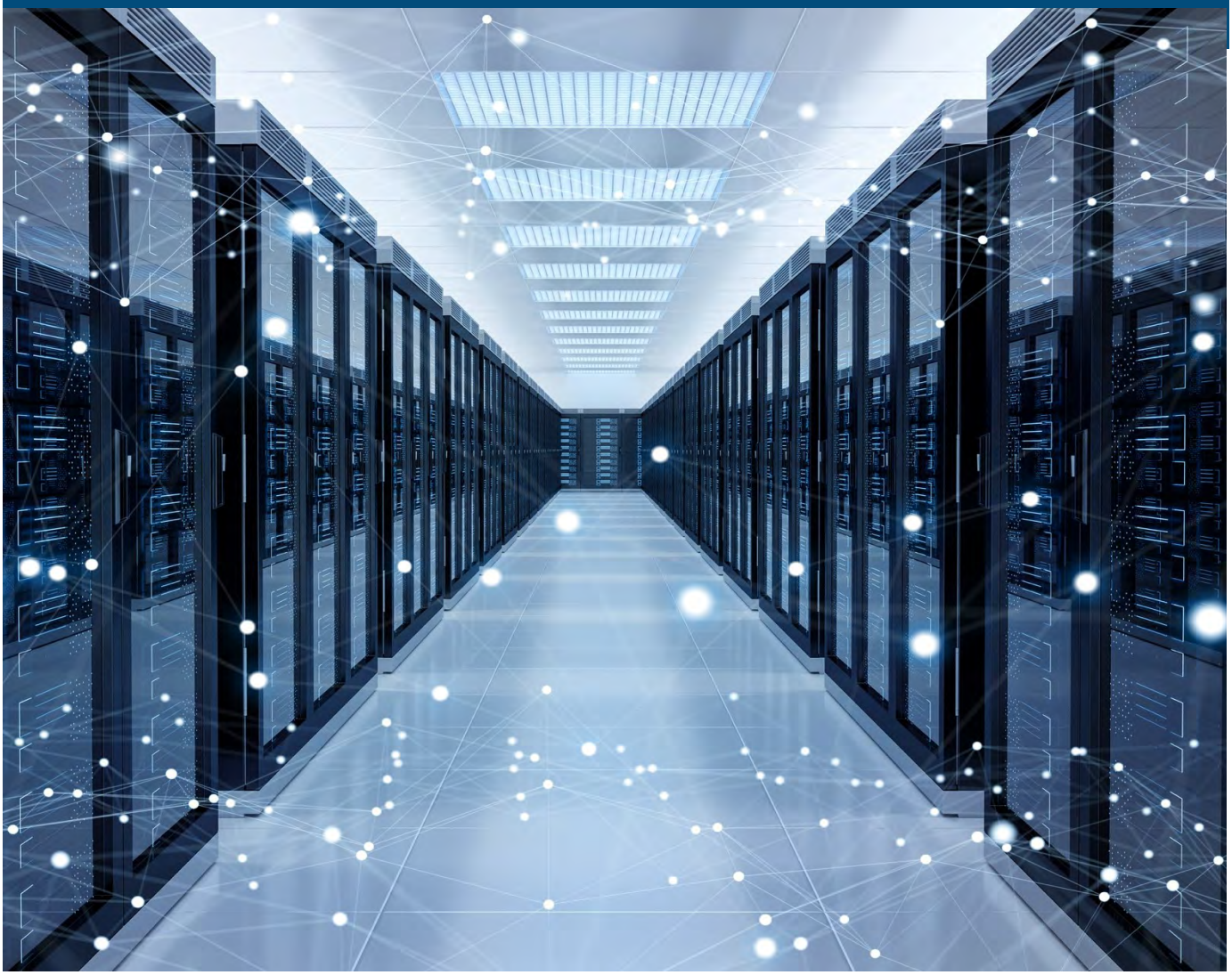
SOCIETY OF ACTUARIES

Innovation and Technology

Insurance Regulatory Issues in the United States



May 2019

Insurance Regulatory Issues in the United States

AUTHORS

Kemi Akinyemi, FSA, EA, MAAA
Actuarial Consultant
Risk & Regulatory Consulting

Lisa Parker, ASA, MAAA
Actuarial Consultant
Risk & Regulatory Consulting

Benjamin Leiser, FSA, MAAA
Consulting Actuary
Risk & Regulatory Consulting

Patricia Matson, FSA, MAAA
Partner
Risk & Regulatory Consulting

SPONSOR

Society of Actuaries
Actuarial Innovation & Technology
Program Steering Committee

Caveat and Disclaimer

The opinions expressed and conclusions reached by the authors are their own and do not represent any official position or opinion of the Society of Actuaries or its members. The Society of Actuaries makes no representation or warranty to the accuracy of the information

Insurance Regulatory Issues in the United States

CONTENTS

Insurance Regulatory Issues in the United States

Executive Summary

The Society of Actuaries Actuarial Innovation & Technology Program Steering Committee engaged Risk & Regulatory Consulting LLC ("RRC" or "the researchers") to conduct research on the intersection of U.S. insurance ratemaking and analytics techniques in advanced modern rating systems and associated regulatory considerations in the U.S.

RRC performed research on the current regulatory environment and emerging issues concerning the use of advanced analytics techniques for rating and summarized that information within this report. The researchers also developed and provided a questionnaire to some practitioners within the insurance industry, consulting, and regulatory environments to comment on current industry practices, the regulatory environment, and emerging issues around data analysis and predictive analytics for ratemaking. The responses to the questionnaire are not intended to represent all opinions; however, they provide a snapshot of some insurance practitioners' views on the use of predictive analytics in the insurance industry. It should be noted that the purpose of this report is not to promote the required use of advanced predictive analytics in the development of rates. Instead, the purpose is to explore the emergence of big data and predictive analytics in the insurance industry, as well as any associated issues that may arise from their use. Many companies are not using big data and advanced analytical techniques for various reasons, including resource constraints such as budget, time, and staff.

Companies use various basic and complex modeling techniques to determine the adequate rates for their products. Big data and advanced techniques are increasingly being used to determine rates, but there is limited regulatory guidance on the applicability of these advanced modeling techniques and tools on ratemaking. Actuaries currently rely on Actuarial Standards of Practice (ASOP) and laws and regulations such as state and federal statutes as their guidelines for developing advanced models, but they are looking to regulators to provide additional comprehensive guidance. Companies are looking to understand what documentation regulators need when rates are developed using innovative methods and advanced techniques. The property and casualty industry appears to be further along in their use of predictive analytics than the life and health industries.

Consultants assist companies in multiple capacities, including understanding the available data, model development, setting assumptions and choosing appropriate methods, model validation, external peer review, and more. Consultants also assist regulators in their review of the data and models used by companies for determining rates.

The regulators are tasked with ensuring that the rates for the insurance products are adequate and not unfairly discriminatory, among other responsibilities. Since the advanced modeling techniques could vary across companies, they are looking to companies to appropriately demonstrate that their use of big data and predictive analytics in determining rates are appropriate for the products and are not unfairly discriminatory. Some companies that are exploring advanced analytics are using them in ratemaking, while others are limiting their reliance on big data and advanced analytics until comprehensive regulatory guidance becomes available.

The models used for ratemaking could be developed internally by the company or with assistance from external vendors. Either way, the company must understand how the data is being used, how the model relies on the data for ratemaking, and be able to communicate this clearly to all key stakeholders, including policyholders, management, and the public.

This paper discusses a few modeling techniques that are used for ratemaking. Basic tools such as trending and linear regression are still in use, and models that are more complex are increasingly being used. One of the most common complex tools used is Generalized Linear Models (GLM), but Machine Learning is increasingly being considered. Generalized Linear Models are easier to explain to regulators and stakeholders than Machine Learning, while the results from Machine Learning are likely to be more accurate. Companies need to assess the risk and reward tradeoff between the tools that they consider or use.

Also discussed in this paper are current and emerging issues, recommended best practices, and implications of implementation of big data and advanced predictive analytics based on research and questionnaire responses. The risks to the insurer and policyholders, as well as the impacts of the use of advanced analytics in ratemaking, are also addressed.

As the use of predictive analytics evolves, the research and questionnaire responses show that models and methods used to determine rates must be defendable and explainable, and the companies must be able to demonstrate that the rates determined are not unfairly discriminatory.

Section 1: Introduction

The Society of Actuaries (SOA) Actuarial Innovation & Technology Program Steering Committee (AITPSC) sponsored this research study (hereafter "the Study") to investigate the intersection of U.S. insurance ratemaking and analytics techniques in advanced modern rating systems with associated regulatory considerations in the U.S. The researchers carried out the main objectives of this project, including background research, interviewing insurance practitioners (hereafter referred to as "panelists"), and developing this report. The objectives of the Study were the following:

1. Perform initial research with a goal of understanding industry activity around the following items:

    - Evolution of rate determination process in various practice areas
    - Regulatory requirements
    - Current techniques used in insurance ratemaking and consideration of statistical modeling techniques other than Generalized Linear Models
    - Types of variables allowed to be used in rating
    - Approach to review and validate the model results
    - Consultant interaction
    - Current and emerging issues during rate filing review
    - Monitoring actual experience versus pricing
    - Impacts on financial performance
    - Risks to the insurer and the policyholder

2. Identify and interview insurance industry, consulting, and regulatory panelists in the use of predictive analytics for ratemaking and/or the regulatory review of those rates, including across life, health, and property and casualty lines of business.

3. Supplement the research with responses from the surveys from, and interviews with, the panelists, including any practical issues noted. Responses are too varied to draw conclusions from the interviews alone because the use of predictive analytics and advanced modeling techniques is evolving and varies across lines of business.

4. Summarize the results of the research and interviews, including the approach, the information gathered, and the conclusions represented.

Based on the results of the research, surveys, and interviews, the researchers have summarized the approach, information gathered, and the resulting conclusions. The summary includes information that is responsive to each of the objectives outlined above. The results of the research were enhanced by the insurance panelists' written and verbal responses to the questionnaire. The responses to the questionnaire are not intended to represent all insurance practitioners' opinions; however, they provide a snapshot of some views on the use of predictive analytics in the insurance industry.

Risk & Regulatory Consulting (RRC) conducted this research. Information regarding RRC can be found on their website: www.riskreg.com.

Section 2: Methodology

2.1 INITIAL RESEARCH

The researchers conducted an initial literature review on the use of predictive analytics and big data in ratemaking, actuarial and statistical techniques being used or considered by insurance companies, and the regulatory environment surrounding the use and review of big data and complex statistical techniques. The literature used for this purpose is included in the References section of this report.

2.2 QUESTIONNAIRE AND INTERVIEWS

The researchers developed a questionnaire based on their initial findings in applicable literature. The questionnaire covered topics such as the following:

- Current U.S. regulatory atmosphere for insurance ratemaking with respect to emerging actuarial and statistical techniques being used and considered by insurance companies and reviewed by state insurance authorities.
- Applicability of current regulations for rate filing and rate review to advanced modeling techniques and data collection.
- Actuarial and statistical techniques and models used in rate development.
- Sources of data, data review, and model validation.
- Key variables considered in the development or review of an insurance company's ratemaking system.
- Current and emerging issues in ratemaking and analytics techniques in modern rating systems.
- Best practices in the analysis and review of big data, analytics techniques, predictive analytics, and advanced models in setting and reviewing rates.
- Implications from a full implementation of big data and predictive analytics to policyholders, insurance companies, and regulators.
- Review of actuarial experience to pricing assumptions.
- Potential areas of misuse of predictive analytics in the ratemaking process.

Of the 12 panelists who were interviewed, most provided a written response, while a few opted to provide verbal responses. The panelists who were interviewed had one or more of the following characteristics:

- Practical experience in the use of predictive analytics in setting rates for life, annuity, health, or property and casualty products, or for long-term care liabilities.
- Practical experience in reviewing or validating predictive analytics and advanced modeling techniques for multiple firms in the insurance, consulting, and regulatory environments.
- Experience in setting or reviewing rates in multiple states in the U.S.
- Experience with multiple advanced modeling tools and techniques.

The researchers then consolidated all questionnaire responses. Since the sample size was small, and the questions were open-ended, the responses were not conducive to presenting a distribution of results by practice area (life, health, P&C) or type of work (industry, consulting, regulatory).

The panelists interviewed for the purpose of this research have backgrounds as regulators, consultants, and industry insurance practitioners. These panelists have a wide range of experience including life insurance, health insurance, long-term care, and property and casualty insurance. The panelists' experience includes developing models for rate determination, peer review of models, model validation, reviewing assumptions and methodologies used in models, and reviewing filed rates from a regulatory perspective. Several panelists are known speakers at actuarial conferences. The researchers summarized the research and enhanced the results of the research based on the panelists' written and verbal responses to the questionnaire.

Section 3: Summary of Research Results

Our findings from the research performed and input from the panelists are summarized in the sections that follow based on certain key themes identified.

3.1 EVOLUTION OF RATE DETERMINATION PROCESS IN VARIOUS PRACTICE AREAS

Companies determine the appropriate rates to be charged for the risks embedded in the products they offer through various types of models and tools. These models and tools have become more advanced over time, from the use of basic statistical techniques to the use of predictive analytical tools, including more advanced forms of Generalized Linear Models to better understand and segment the risks for pricing. The Generalized Linear Model is a flexible generalization of ordinary linear regression that allows for response variables that have error distribution models other than a normal distribution.

Health insurance and property and casualty insurance companies generally need to file for their rates to be approved by the state they are doing business in, with some states and some products using a "file and use" approach. Life insurance rates are constrained by competition and must comply with all laws regarding policy forms and rates in that state. A common theme across industries is that the insurance rates must not be unfairly discriminatory. The researchers observed that the property and casualty companies are further along in the process of using predictive analytics and advanced modeling techniques for determining their rates than the life insurance and health insurance industries, but the use of predictive analytics and other advanced modeling techniques is still emerging across all the three industries. The life insurance industry is beginning to use more advanced techniques and big data for products that have accelerated underwriting features. The health insurance industry appears to be in the very early stages of using the advanced techniques to price their products.

As insurers consider the use of predictive analytics in the determination of rates, regulators want to ensure that the determined rates are appropriate and not unfairly discriminatory. Because the regulatory space is evolving, there is limited guidance on how the use of predictive analytics in the determination of rates could inadvertently lead to unfair discrimination and how companies can check their models and demonstrate that their models do not produce unfairly discriminatory results.

3.2 REGULATORY REQUIREMENTS

For property and casualty insurance, the rate determination and filing process varies by state. Each state has laws and regulations concerning the pricing of insurance products, and the details vary greatly from state to state and from product to product. Most states have statutes that require insurance rates to be "not excessive, not inadequate, and not unfairly discriminatory." There are a handful of states that are providing guidance for companies that use GLM models to develop rates. Some of the information being requested by states relates to the following:

- Technical information: type of model, model construction, sensitivities, and dependence within the model, model limitations, formulas used in the model
- Description of the dataset: source of data, how data is selected, adjustments to data, sources of non-insurance data
- Variables: definitions, selection process, reconciliation of rate data to model data
- Output: provide statistical measures to indicate predictive value and goodness-of-fit, provide output coefficients

For some health insurance products, insurers have to comply with the federal Patient Protection and Affordable Care Act (ACA) and associated provisions and regulations such as Adjusted Community Rating, Risk Adjustment Transfers, and minimum Medical Loss-Ratio Requirements. Other federal health regulations that health insurers consider include the Mental Health Parity Addiction and Equity Act (MHPAEA) and the Health Insurance Portability and Accountability Act (HIPAA). There are also state-specific laws that govern rate determination and rate filing.

States do not typically regulate ratemaking in life insurance since competition leads to the proper adjustment of premiums.

State regulators protect consumers by evaluating whether insurance policy forms comply with state law, include mandated benefits, and meet readability standards, and they review whether the premium rates are fair, reasonable, and non-discriminatory. The nature of the rate review, rating rules, and forms varies somewhat among the states depending on their laws and regulations. For personal property and casualty lines, about half of the states require insurers to file rates and receive prior approval before they go into effect. With the exception of workers' compensation and medical malpractice, commercial property and casualty lines in many states are subject to a competitive rating approach. Under such a system, regulators typically retain authority to disapprove rates if they find that competition is not working.

Premiums for life insurance and annuity products generally are not subject to regulatory approval, although regulators may seek to ensure that policy benefits are commensurate with the premiums charged. Many states subject health insurance rates to prior approval, with other lines using a "file and use" system or no provisions for review.

Actuaries rely on the American Academy of Actuaries Code of Professional Conduct and all applicable Actuarial Standards of Practice in developing predictive models for ratemaking and rate filing purposes. The panelists mentioned that actuaries should ensure, through internal or external peer reviews, that the predictive analytics tools they use are consistent with the ASOPs, applicability guidelines of the ASOPs, and any state and federal regulations that are applicable to the products being priced.

Certain emerging guidance was identified in the course of the research:

- The National Association of Insurance Commissioners (NAIC) is in the process of developing centralized resources for regulators evaluating the use of Big Data in the property and casualty line of business.
- The Actuarial Standards Board recently issued Actuarial Standard of Practice (ASOP) No. 54, which provides guidance for the pricing of life insurance and annuity products.
- The New York Department of Financial Services (NYDFS) recently issued a circular letter titled "Use of External Consumer Data and Information Sources in Underwriting for Life Insurance." The NYDFS wants to ensure that there is no potential negative impact on the availability and affordability of life insurance for protected classes of customers, as well as to ensure transparency for consumers when external data sources or predictive models are used to make decisions that limit or decline insurance access or affect the rates for the consumers.
- The American Academy of Actuaries (AAA) issued a long-term care credibility monograph that serves as a guide for the use of credibility in long-term care pricing.

Several industry and consulting panelists suggested that, because there is limited regulatory guidance on the use of predictive models, they often rely on ASOPs and any available state or federal laws in developing and reviewing the predictive models used for ratemaking. Several panelists often rely on the Applicability Guidelines for ASOPs, which are intended to suggest which ASOPs might provide guidance to them on more common assignments.

Some panelists thought that regulators appeared to set broad guidance without specific background on the current processes that each insurer currently uses and the varying implications of the guidance for each company. A common theme provided by the panelists is that regulators are not currently equipped with all the necessary tools to regulate advanced predictive models. The current set of guidance from regulators is broad, leading to often conflicting instructions by states and sometimes between analysts within a state. Consulting panelists suggested that the predictive analytics models are tools and the purpose for which they are used should be understood by both the industry panelists and regulators. As advanced data and modeling techniques evolve, it was recommended by some respondents that there should be clear and streamlined regulations that are consistent across states.

Some regulatory panelists stated that, while innovation in insurance companies is encouraged, companies should focus on demonstrating that the rates developed using predictive analytics are not unfairly discriminatory to consumers. In addition, regulators and consultants recommended high standards and controls on data quality and model validation.

3.3 CURRENT TECHNIQUES AND CONSIDERATIONS USED IN INSURANCE RATEMAKING

The researchers observed various techniques used in insurance ratemaking for property and casualty companies, including univariate methods, multivariate methods, and machine-learning techniques. Machine learning is an application of artificial intelligence that provides systems with the ability to automatically learn and improve from experience without being explicitly programmed. It focuses on the development of computer programs that can access data and use it to learn for themselves.

There are two basic univariate approaches for determining an overall rate level: the pure-premium method and the loss-ratio method.

The pure-premium method determines an indicated average rate and involves projecting the average loss and loss adjustment expenses per exposure and the average fixed expenses per exposure to the period that the rates will be in effect. The sum of those two is then adjusted for variable expenses and the target profit percentage by dividing by one minus the sum of the variable expense provision and target profit percentage to get the indicated average rate.

The loss-ratio method compares the estimated percentage of each premium dollar needed to cover future losses, loss adjustment expenses, and other fixed expenses to the amount of each premium dollar that is available to pay for such costs. The sum of the projected loss and the loss adjustment expense ratio, and the fixed expense ratio is divided by one minus the sum of the variable expense provision and the target profit percentage to get the indicated change factor. The major difference between the pure-premium and loss-ratio approaches is that the loss-ratio approach uses premium as opposed to exposure.

The loss-ratio approach requires obtaining premiums at the current rate level for each level of the variable being analyzed, but if that level of detail is not available, it is possible to make an adjustment to the pure-premium approach such that the pure-premium approach is performed using exposures adjusted by the exposure-weighted average relativity of all other variables. This is referred to as the Adjusted Pure-Premium approach.

Univariate methods are limited because they do not account for the effect of other rating variables. There are multivariate classification ratemaking techniques that consider all rating variables simultaneously and automatically adjust for exposure correlations between rating variables and allow for the interaction and interdependency between two or more rating variables.

Multivariate methods also attempt to remove unsystematic effects in the data (noise) and capture only the systematic effects in the data (signal) as much as possible. Multivariate techniques also produce model diagnostics and information about the appropriateness of the model fitted. The most commonly used multivariate technique in recent years is generalized linear models (GLMs) because of their relative transparency.

Another multivariate technique that may be gaining in popularity is neural networks, though this approach tends to be less transparent. GLMs and other linear models focus on expressing a relationship between an observed response variable and a number of predictor or explanatory variables. Classification and regression trees, random forests, and neural networks are often used as well as GLMs.

Machine-learning techniques, particularly Artificial Neural Networks (ANNs), are increasingly popular in several disciplines, but not necessarily insurance due to a lack of interpretability. The explosion in the variety and volume of available data, coupled with cheap data storage and fast computing power, have made ANNs a key tool of data scientists. ANNs allow modeling of nonlinear processes and can be a useful tool for solving problems such as retention modeling, fraud detection, claims triage, and traditional pricing models. ANNs attempt to digitally replicate the biological mechanism of the human brain and its neurons. Although ANNs have been on the upswing in a variety of fields, the insurance sector has not utilized these "brain-like" techniques on a large scale. The insurance field is still heavily skewed in favor of the more familiar and traditional data-mining techniques, such as GLMs and rule-based algorithms.

ANNs are like a collection of nested regression models residing in its layers. The "neurons" in the first layer represent the input (predictor) variables, which are used to build the first set of regression models. The predictions from these regressions are, in turn, used as inputs into the next series of regressions in the following (hidden) layers of "neurons" and so on. When the "signal" reaches the final (output) layer of the ANN, another regression is applied to produce the final predictions. In this sense, ANNs are a more advanced form of GLMs. GLMs are based on a single regression equation whose predictions are easier to explain through regression coefficients. Techniques to "unpack" the output from ANNs to make the output easier to compare to the output of other models include sensitivity analysis, Neural Interpretation Diagrams, Fuzzy Logic theory, and modeling ANNs' predictions to derive a simple mathematical expression.

According to survey results, life insurance companies are using or exploring the use of multivariate techniques. Health insurance companies tend to use more basic techniques such as trend analysis and linear regression tools, though some use GLMs.

Various panelists provided further information on the techniques currently being used. Several companies are using a combination of internal and external data for their ratemaking models, and internal and external (vendor) models. Several companies use multiple techniques, with the most

popular one being GLM. Basic tools such as trending and basic forecasting are still being used for certain benefit features. Some panelists indicated that, while machine-learning and neural network techniques are being used on a smaller scale, they expect these advanced techniques to be used more as companies gain a better understanding of them and regulations on the use of predictive models become more developed.

3.4 DATA SOURCES AND TYPES OF VARIABLES ALLOWED

The types of variables allowed to be used in rating could vary depending on the line of business. The rating variables are used to segment the insured population into different groups of similar risks for the purposes of rating. The criteria for selecting variables may be summarized into the following criteria categories: actuarial or statistical, operational, social, and legal.

Actuarial or Statistical Criteria: Most panelists believe that companies should consider the following actuarial or statistical criteria to help ensure the accuracy and reliability of the potential rating variable:

- Statistical Significance: The rating variable should be a statistically significant risk differentiator. The expected cost estimates should carry for the different levels of the rating variable, and the estimated differences should be within an acceptable level of statistical confidence and be relatively stable from one year to the next.
- Homogeneity: The levels of a rating variable should represent distinct groups of risks with similar expected costs. The groups should be defined such that the risk potential is homogeneous within groups and heterogeneous between groups. If a group of insureds contains materially different risks, then the risks should be subdivided further by creating more levels of an existing variable or by introducing additional rating variables.
- Credibility: The number of risks in each group should be large enough, stable enough, or both for the actuary to be able to accurately estimate the cost.

Operational Criteria: Most panelists believe that companies should consider practical and operational constraints after identifying the statistical criteria of their variables. For rating variables to be considered practical, they should have the following qualities:

- Objective: The levels within a rating variable should have objective definitions that are not ambiguous. This limits the opportunities for administrative errors. Rating variables such as age, gender, and marital status are objective, while maturity is subjective.
- Inexpensive to administer: The operational cost to obtain the information necessary to properly classify and rate a given risk should not be too high. The cost of obtaining and verifying information should not exceed the value of the additional accuracy.
- Verifiable: The levels of a rating variable should not be easily manipulated by the insured or distribution channel and should be easy for the insurer to verify. For example, it is easier

to verify the total number of miles driven in a recent period than the total number of miles estimated to be driven in an upcoming period.

Social Criteria: Most panelists believe that companies may want to consider public perception as they identify the rating variables to be used. The following items affect social acceptability of using a particular risk characteristic as a rating variable:

- Affordability: It is desirable for insurance to be affordable for all risks, and companies may combine classes and introduce subsidies for risk classes that imply additional costs.
- Causality: A rating variable having an intuitive and direct relationship to insurance costs increases the social acceptability of that variable.
- Controllability: A controllable variable is one that is under the control of the insured. It is preferable for the insured to have some ability to control the class to which they belong and, as a result, be able to affect the premium charged.
- Privacy Concerns: Policyholders may be reluctant to disclose personal information and may be willing to pay more to disclose less information. Insurers need to consider what information requires the express consent of the insured before it is used.

Legal Criteria: The risk classification may be affected by state and federal statutes and regulations. Generally, constitutions govern statutes and statutes govern regulations. The rate classification system must be in compliance with the applicable laws and regulations of each jurisdiction in which a company is writing business. Actuaries need to be familiar with the laws and regulations of each jurisdiction in which their company writes insurance and assure that the classification rating is in compliance with that jurisdiction's laws and regulations. This usually requires working with other professionals, such as lawyers or regulatory compliance experts, in determining what is acceptable and what is not.

The panelists indicated that the data used in predictive analytics should include internal and external data. The internal data is generally from the information that insurers currently have from policyholders. The panelists shared some of the third-party data that they have considered or relied on for the purpose of predictive analytics. Some companies considered the use of some of these data as applicable for their lines of business and concluded that, although some of the third-party data was beneficial for ratemaking purposes, the potential costs of using other external data sources outweighed the projected additional benefits from using them. Examples of third-party data sources include credit-based models, electronic medical records, wearable devices and telematics, motor vehicle records, health engagement data, Carfax vendors, and social media. For health, Centers for Disease Control and Prevention data on outbreaks of epidemics, such as influenza and wellness program participation, were identified as specific external data sources. Of this list, some panelists noted that insurers are reluctant to use social media data.

3.5 APPROACH TO REVIEW AND VALIDATE THE MODEL RESULTS

Effective insurance ratemaking heavily relies on models working effectively to capture risk and appropriately account for the key elements of rate determination. Model validation is the process of performing an independent challenge and thorough assessment of the reasonableness and adequacy of a model based on peer review and testing across multiple dimensions, including design, data, assumptions, results, and governance. Proper model validation and governance are necessary to mitigate financial model risk due to the potential negative impact of models failing to appropriately price products as designed. Through model review and validation, companies can confirm that their models are working as designed, identify limitations of their models, and manage the associated risks in the models.

The panelists indicated that independent peer review is a common approach to model review. The independent peer reviewers should understand the purpose of the model, data used to develop and validate the model, assumptions and methodology used for the model, and expected results from the model.

The panelists also indicated that model review and validation occur both internally through internal audit, data science, and risk management teams, and externally through consultants, reinsurers, and third-party vendors. Some companies review their model through a "Train/Test/Validate" approach where three different datasets are used to train the model, test the model, and then validate the model.

A general consensus among the panelists was the recommendation that companies need to understand their ratemaking model broadly (in other words, business users and not just technical experts understand the model), have effective controls for the model, and be able to communicate the workings and the results of the model efficiently, even if the companies are relying on an external consultant or vendor. The use of newer visual techniques, including animation, to explain the model was recommended.

In addition, the panelists recommended that results from advanced tools should be frequently compared to results from traditional tools as a way of assessing the value of the advanced tools. They also suggested that adding new variables to the model or increasing the complexity of the model should only be done when there is a corresponding material improvement in the quality of the model. In addition, companies should periodically review their actual results compared to what was expected and understand the differences.

One panelist indicated that an effective model validation process should be periodic and include a review of the following: data, application code, assumptions and methodologies, plan code mapping, product features, controls, and model performance and outcome analysis. Each of these items is detailed further below, including additional comments provided by other respondents.

Data: The model validator should verify that internal and/or external data used for the model are demonstrably appropriate, accurate, and complete. The limitations of the data or implications that the data has on the estimates produced from the model should also be understood. The data input to the model and the data output from the model should be verified.

In Machine Learning, the dataset should be split into train and test subsets. The training dataset should be used to prepare or train the model, using both the input from the training data and the output from the training data to develop the algorithm in the model. The test dataset is treated like new data and the output values are withheld from the algorithm of the machine learning data. The predictions from the trained model using the test dataset are compared to the known output values of the test data to understand how well the model is working.

Application Code: The model validator should ensure that the model is appropriately coded and stress tested. The results of the models should be verified outside the model using Excel or some other software package. The coding in the model should meet requirements set by applicable regulatory authorities and follow best practices such as the ASOPs.

Assumptions and Methodologies: The assumptions used in the model should be recent, reasonable, applicable, and entered into the model correctly. Assumptions should be documented, reviewed, and formally approved. The rationale and credibility for the assumptions should also be documented. The model validators should understand the sensitivity of various key assumptions in the model, their impact on the estimates from the model, and whether those impacts are appropriate. The assumptions used and the rationale for the assumptions should be documented.

The model validation team should understand the methods used in the model and why those methods are appropriate for the intended purpose of the model. To understand the methods to be used in the models, the starting point should be the identification of the risks that need to be modeled and ensure that there is a process in place to periodically assess the modeled risks. Once this is done, the method for modeling these risks to determine rates should be reviewed to understand how it is both applicable and appropriate. The methods used and the rationale for the methods should be documented.

Plan Code Mapping: To speed up run time, the data is often grouped into subsets using similar characteristics. The model validator then evaluates whether the product features are preserved by the plan code mapping. In addition, the plan code mapping needs to be periodically reviewed to ensure its continuous applicability to the dataset.

Product Features: The model validator should understand the product features and confirm that the model is correctly representing how the products actually work, as well as related policyholder

behavior. Model documentation with respect to product features and benefits should include simplifications in the model, product guarantees, and product features not modeled.

Controls: Some models fall under Sarbanes-Oxley (SOX) or Model Audit Rule (MAR), which include some of the needed controls in place. The controls in place for rate determination should include those around accuracy of the data, access to the models, independent and multiple layers of review, and model change management. In addition, the models should be reviewed by at least the internal audit and/or a risk team before deployment.

Model Performance and Outcome Analysis: To appropriately review a model, the intended use of the model must be fully understood, as this will allow the model validation process to focus on all key details of the model. The model validation should identify the use of the model, confirm that the model is consistent and applicable for the intended use, and ensure that the model is not being used for purposes outside the capabilities of the model. In addition to specifying the intended purpose of the model, documentation should specify the purposes for which its use is not appropriate.

The actual estimates produced from the model should be compared against historical data to determine how well the model makes predictions. Some tools used by actuaries include:

- Benchmarking and replication, including static validation, parallel testing against other models, and spreadsheet replicas.
- Outcome analysis, including dynamic validation, back-testing, and out-of-sample testing.

The model should be sound and stable within the defined range of use and under a range of scenarios. The validation process should include scenario testing, stress testing, sensitivity testing, and extreme value testing to understand the limitations of the model. Results across different levels of input and granularity can be reviewed. In addition, aggregation benefits across different levels of result aggregation can be reviewed.

The NAIC has tasked its Casualty Actuarial and Statistical Task Force (CASTF) with identifying best practices to serve as guidance to state insurance departments and insurers in their review of complex models underlying rating plans. Although these best practices were designed for the review of property and casualty insurance predictive models to support a filed property and casualty insurance rating plan, the identified best practices would benefit other industries. The identified rate review best practices in the October 2018 draft include the following:

- Encourage competition among insurers.
- Protect the confidentiality of filed predictive models and supporting information (according to state law).
- Review a predictive model efficiently.

- Obtain a clear understanding of the characteristics that are input into a predictive model (and its sub-models), their relationship to each other, and their relationship to nonmodeled characteristics/variables used to calculate a risk's premium.
- Determine whether individual input characteristics to a predictive model are related to the expected loss or expense differences in risk. Each input characteristic should have an intuitive or demonstrable actual relationship to expected loss or expense.
- Determine that the data used as input into the predictive model is accurate, including a clear understanding how missing values, erroneous values, and outliers are handled.
- Ensure that individual input characteristics to a predictive model (and its sub-models) are not unfairly discriminatory.
- Obtain a clear understanding of how the selected predictive model was built and why the insurer believes this type of model works in a private passenger automobile or homeowner's insurance risk application.
- Determine whether individual output characteristics from a predictive model are related to expected loss or expense differences in risk. Each output characteristic should have a demonstrable actual relationship to expected loss or expense.
- Obtain a clear understanding of how model output interacts with non-modeled characteristics/variables used to calculate a risk's premium.
- Ensure that individual outputs from a predictive model and their associated selected relativities are not unfairly discriminatory.
- Obtain a clear understanding of how the predictive model was integrated into the insurer's state rating plan and how it improves the state rating plan (this latter element is only applicable when a new or revised model is introduced into an existing rating plan).
- Determine the extent to which the model causes premium disruption for individual policyholders, and how the insurer will explain the disruption to individual consumers who inquire about it.
- Determine the means available to a consumer to correct or contest individual data input values that may be in error.
- Obtain a clear understanding how often each risk characteristic used as input to the model is updated and whether the model is periodically rerun to reflect changes to non-static characteristics.
- Given an insurer's rating plan relies on a predictive model, and knowing all characteristics of a risk, a regulator should be able to audit/calculate the risk's premium without consultation with the insurer.

The panelists identified the following best practices in the use of predictive analytics:

- Modeling teams should follow sound predictive modeling processes. This includes gaining consensus on the business problem and testing whether there is available data to solve the problem. Once the modeling team is in a position to build, they should create a holdout dataset to use to validate the model and ensure the model is not over-fit to the data. In

other words, they should consider the goal of the model being to provide a reasonable projection using various datasets, and not only one dataset. Modeling teams should allow time to obtain quality data and for stability in the data and model outputs.

- Actuaries should understand all aspects of the predictive analytics, including the data used in the model, the assumptions and methods used in the model, and the biases in the model, even if they are developed by external vendors or consultants. Actuaries also need to verify that techniques and data used by vendors are not unfairly discriminatory.
- Data, assumptions used, and methods used in the predictive analytics and the rationale for their use should be fully documented and the documentation should follow ASOP 41. Actuaries should also understand and follow ASOPs that apply to the models and the products underlying the models. There should be high standards and controls on data sufficiency and cleanliness.
- There should be clear justification of factors used and their impact on rating. Approaches to demonstrate that factors used are not masking those that would constitute unlawful discrimination are needed. In addition, the accuracy of those factors' impact on pricing and reserving need to be determined by actuaries with multiple aspects of testing involved.
- To the extent that the data used to determine rates is not considered credible, actuaries should include corresponding provisions for adverse claims deviation.
- Modeling teams should be comfortable explaining the model to any stakeholder, including senior leadership.
- Model review and validation is necessary. Multiple layers of testing and review by independent third parties are highly recommended.
- Actuaries should continue to increase their knowledge and skill set around the use of big data and advanced modeling techniques. Knowledge of data science tools such as R, SAS, Python, SQL, and Tableau is encouraged.
- There should be more collaboration among actuaries of different practices, and between the Casualty Actuarial Society (CAS) and SOA. This will allow for effective transfer of skills across practices and cross-function modeling.

3.6 CONSULTANT INTERACTION

Consultants provide various services to assist insurers in the process of using predictive analytics for setting and developing rates. Insurance companies often rely on consultants for various aspects of the predictive analytics tool development and testing in the life insurance, health insurance, and property and casualty insurance spaces. Some of the services offered include forecasting and projecting, regulatory filing, understanding data, independent reviews of aspects of the model, and model validation. Regulators sometimes rely on consultants to perform their review of insurers' predictive models and underlying data.

In the course of discussion with the panelists, consultants are assisting insurers with education on advanced modeling techniques, model development, peer review of developed models, model review and validation, data analysis, and more. Although some companies have explored various

advanced techniques such as machine learning, several are still in the process of deciding whether to use them due to their auditability. Some companies use several aspects of machine learning techniques in rate setting, while others are willing to trade off some statistical precision for simpler to explain methods such as GLM. Some companies have reputational risk concerns about using advanced statistical techniques for rate setting. Even when companies are comfortable with their advanced models, there is a reputational concern about life insurance rates being determined by an advanced statistical model. Some companies are further along in their use of advanced techniques like machine learning for their sales processes, such as cross selling or up selling, than they are in their use of advanced techniques for ratemaking.

3.7 CURRENT AND EMERGING ISSUES DURING RATE FILING REVIEW

Three panelists identified several emerging issues, particularly from a regulator's perspective, around the use of predictive analytics and advanced modeling techniques in ratemaking and rate filing, as well as the review of these rates.

1. An insurer's use of big data has significant implications for fairness, access, and affordability of insurance, and the regulator's ability to be up to date with the changes and protect consumers from unfairly discriminatory practices.
2. Market forces alone, or free-market competition, are not sufficient to protect consumers from unfair practices.
3. The current insurance regulatory framework does not provide regulators with the tools to effectively review and assess an insurer's use of predictive analytics.
4. The current insurance regulatory framework does not provide insurance companies with sufficient constraints around their use of predictive analytics for ratemaking.

The panelists identified several issues that they are currently working through when using predictive analytics for rate determination, or when reviewing the rates determined using advanced modeling techniques.

- Advanced tools such as machine learning are more efficient and accurate, but are generally lacking in explainable results. The data science industry is building software to help interpret machine learning results.
- Companies need to understand the quality of the data and any associated limitations.
- When big data for ratemaking is obtained, companies need to sort through the information to determine what data sources and data elements are useful.
- Companies should ensure that they have appropriate resources and document their data review, model validation, and any controls they have in place. Regulators should also increase their skill set to be able to review this information.
- As companies work on innovation and continued use of predictive analytics, it is recommended that regulators work on increasing their understanding of these tools and

provide consistent guidance on how companies can use big data and advanced modeling techniques to set rates. Regulatory ambiguity and uncertainty with regard to what regulations will apply to advanced modeling techniques may limit innovation.
- Companies need to understand what data points could be unfairly discriminatory and should be able to demonstrate that their data and models are not unfairly discriminatory.
- Regulators can add value to the insurance industry by setting guidance around what is unfairly discriminatory.
- Companies and regulators need to consider data privacy and transparency in their use or review of big data.

The panelists also identified several emerging issues that they are preparing for in their continued use of predictive analytics for rate determination or in their review of the rates determined using advanced modeling techniques. Because the sophistication in this area varies from company to company, some of the current issues above were identified by other panelists as emerging ones and are, therefore, also included below.

- Considering the large volume of big data available, companies will need to decide what sources of data are appropriate and would not lead to unfair discrimination. Companies will also decide their risk and reward threshold from using certain data and if the potential benefits outweigh the potential costs.
- Companies who do not understand their models fully or do not have evidence showing that their data or models do not have biases that are unfairly discriminatory will likely receive increased scrutiny from regulators.
- Although there are benefits to using wearables and telematics either through incentives to consumers who opt in or through additional information for the insurer, insurers and regulators need to consider what could make this unfairly discriminatory for those who decide to opt out.
- The impact of new technology, such as self-driving cars and vehicle telematics on current ratemaking processes, is unclear and it is also not yet clear what information would or should be made available to insurers.
- With the increase in availability and use of genetic testing, consumers may have some information that insurers do not, which could result in anti-selection and increase the cost of insurance. Insurers are seeking guidance on whether they can set rates based on genetic information that is available or request this information of consumers. Insurers are seeking regulatory guidance around the use of advanced medical techniques in underwriting and ratemaking, and guidance on what would constitute unfair discrimination.
- Although access to electronic health records at the time of an application could solve the issue of misrepresentation by applicants, it is unclear whether regulators will be comfortable with accessing and using that information.

- Limited understanding of the use of big data and predictive analytics by key stakeholders, such as management and policyholders, can lead to negative press, which could affect insurers' innovation.
- The ability of regulators to keep pace with advanced modeling techniques, considering the various regulatory tasks they are involved with, could slow down the innovation process by insurers.

3.8 MONITORING ACTUAL EXPERIENCE VERSUS PRICING

Once the model is developed, panelists indicated that the sufficiency of the rates charged should be investigated to help the insurer understand how well the model is working. To do this, insurers could perform outcome analysis, such as back-testing, out-of-sample testing, and actual-to-expected comparisons on an ongoing basis. Outcome analysis should be performed prior to implementing the model and done at least on an annual basis after implementation to ensure that it is performing as expected. Error limits should be developed for the outcome analysis results and the actual errors from the model exceeding those limits identified. When limits are exceeded, the company should take certain actions that would have been predetermined. If the model is recalibrated or updated on an annual basis, limits should be developed that monitor the size and frequency of the re-estimates. If updating the model repeatedly results in large changes in the estimates of the model, then certain actions should be required, including external model validation, a recalibration of the model, or possibly the development of an entirely new methodology or model. The action type and triggers should be set in advance of, and in accordance with, the use and risk of the model and should be documented.

3.9 IMPACTS ON FINANCIAL PERFORMANCE

Predictive analytic techniques enable insurers to better understand their data. Proper implementation of predictive analytics techniques can improve an insurer's consistency and efficiency in product pricing and product development. The use of predictive analytics in ratemaking could affect the financial performance of insurers in several ways, including the following:

- Improve pricing by increasing the number of rate segments and price points,
- Efficient underwriting, potentially reducing the employee hours underwriters spend on researching, and analyzing applicants,
- Improve competitive advantage, and
- Potential increase in legal and information technology cost.

Several panelists identified some benefits of big data and advanced modeling techniques in insurance, which have the potential of positively influencing financial performance, but noted that some of them could be drawbacks depending on the stakeholder's point of view:

- Big data and advanced modeling techniques provide new insights and ways to assess premium rating segments and parameters. The improved segmentation and accuracy help match premiums to risks and avoid adverse risk selection.
- The use of big data provides insurers with a better understanding of the risks being priced.
- Big data and advanced modeling and analytics tools could provide insurers with a competitive advantage, at least initially.
- Policyholders potentially have more input in their rates by making positive lifestyle choices, such as exercise that can be monitored, and insurers could potentially reward policyholders for making positive choices.
- Other benefits of big data in insurance, outside of ratemaking, include making insurance desirable for millennials by making the application and underwriting process simpler, faster, and less invasive.

3.10 RISKS TO THE INSURER AND THE POLICYHOLDER

The use of advanced modeling tools and techniques has many benefits but poses several risks to both the insurer and policyholder. The insurer runs the risk of unintentionally charging rates that are unfairly discriminatory and needs to be able to demonstrate fairness of rating to consumer groups and regulators. Before deploying a model, the insurer needs to thoroughly review the model and should be able to defend key assumptions and methods used to management and consumer representatives. The risk factors used for estimating rates for products could change over time, potentially leading to mispricing of the modeled products. As a result, the developed models need to be consistently monitored and improved upon as additional information becomes available.

Some panelists also identified some drawbacks of big data and advanced modeling techniques, and noted that some of these drawbacks could be benefits depending on the stakeholder's point of view:

- Companies could have unrealistic expectations that machine learning and artificial intelligence can deliver solutions to every business problem.
- Internal models could be built by people who do not understand sound actuarial practices. Insufficient understanding of internal or vendor models could lead to setting rates that are unfairly discriminatory. Companies need to be able to explain their models such that the ability of consumers and regulators to understand the rating is not compromised.
- Some companies may rush to be first, while utilizing poor due diligence in the process. This can lead to unintended consequences of volatility, backlash from a regulator, or bad

headlines. Making too big of a mistake too early in the process affects the stability of the model and lengthens time to adoption.

- Certain benefits of risk pooling could be diminished if there is over-segmentation of risks.
- Regulators are at risk of falling behind in their ability to evaluate these models, which could either allow companies to take on too much risk without adequately pricing or reserving for it or lead to over-regulation of a process that is not fully understood. This could lead to requiring companies to hold too much reserves due to the regulator's uncertainty or could prevent companies from using these innovations to make insurance more accessible to the middle markets.
- There are potential data privacy and transparency concerns. Companies need to understand how policyholders' personal data is being used and the regulations involved. Policyholders may have constraints on how their personal data is being used.
- It is possible that consumers will be unjustifiably penalized in the form of higher rates for irrelevant factors.

Some panelists identified several potential implications from implementation of advanced modeling techniques as detailed below:

- Companies need to understand the legal requirements and demonstrate that rates and policies are not unlawful, unfair, inequitable, or misleading. They need to confirm that rates do not misrepresent the benefits of the policy. Actuarial justification is more difficult to review if a near-black-box model is involved.
- Data is now more available, so there is exponentially more information to rely on. For the Company, risk selection is more streamlined, and there is a finer risk classification system, which allows for better pricing.
- It is possible that some policyholders will be unlawfully discriminated against or pay a higher rate than actuarially justified.
- Although there is more available data from the use of telematics, some policyholders prefer to maintain their privacy and do not want to be monitored. However, with the Internet of Things (internet connectivity devices embedded in everyday objects), there is going to be more data available to the industry that the customer is not aware of. Regulations will emerge and will likely be inconsistent across regulators.
- Reputation and Headline risk: The Company needs to stay ahead of public relations by informing investors about what they are doing (such as using big data and predictive analytics to better serve customers and applying ethical best practices while doing so). Companies should be prepared to answer questions about their ratemaking model to public and consumer representatives.
- Regulators need to understand what the risk selection process is at insurance companies in order to be able to regulate the process. Regulators will need to offer consistent guidance, and companies will also need to be prepared to deal with inconsistent guidance from regulators.

- For policyholders, there is potentially easier and faster access to policies and limited invasive procedures such as blood draw. However, more information is collected about them so data governance, privacy issues, and transparency issues need to be understood by the company.

The following potential areas of misuse were identified by some panelists:

- People who are not trained in actuarial professionalism could be doing something that is unfairly discriminatory, either intentionally or unintentionally.
- When using massive datasets, almost any variable can look like it is statistically significant. There is a danger of using models with spurious correlation if the focus is to build a model without a clear understanding of the model or underlying data.
- Another potential issue is using data to reinforce bias in your target variable. Models are only as good as the underlying data they are built on. If the advanced models used for ratemaking contain any biases, either human or from the data itself, the model output will reflect that and end up reinforcing that bias.
- Overtuning of models could lead to inappropriate levels of precision that one might infer could be achieved.
- Models may use rating variables that are not predictive variables.

Section 4: Concluding Remarks

Companies use basic and complex modeling tools and techniques to set rates for their products within the constraints of applicable regulations, statutes and laws. The complex modeling tools rely on big data and advanced modeling techniques. Although these complex techniques have been embraced by other industries and other sections of insurance industries, insurers appear to be cautious in the use of predictive analytics for setting rates. Reasons for this may be due to limited comprehensive regulation on the use of predictive analytics for the purposes of ratemaking and how the insurer can demonstrate that their rates are not unfairly discriminatory, as well as the resource constraints of budget, time, and staff.

The researchers observed that the property and casualty insurance industry appears to be further along in the use of advanced analytics tools for ratemaking than the life and health insurance industries. It also appears that there is more P&C regulatory guidance emerging than in life and health. The life and health insurance industries should leverage some of the tools and knowledge from the property and casualty industry for the purposes of building and validating models and developing regulatory guidance.

Some of the complex models being used by insurance companies include GLMs and machine learning, with the machine learning being more precise than GLM. GLM is generally understood in the industry, but the machine-learning models would typically require more explanation. Of the companies that have explored the more advanced analytical techniques such as machine learning, some companies have decided to use the advanced models for ratemaking, while others have decided that the benefit from added precision pales in comparison to the added effort and cost required to develop, validate, and explain the model.

The increased complexity of the models implies that companies need to be able to demonstrate that their rates are not unfairly discriminatory, and regulators need to be prepared to understand and review multiple variations of these advanced modeling techniques. Consultants can provide assistance to insurers with model development, model validation, data analysis, and education. Consultants can also assist regulators to view companies' complex data and models.

This paper documented several benefits, drawbacks, and implications of implementing advanced modeling techniques. The use of big data provides insurers with a better understanding of the risks being priced and advanced modeling techniques provide new insights and ways to assess premium rating segments and parameters. However, certain benefits of risk pooling could be diminished if there is over-segmentation of risks; there are potential data privacy and transparency concerns, and the possibility exists that consumers will be unjustifiably penalized in the form of higher rates for irrelevant factors.

During their review of literature and the panelists' responses to the questionnaire, the researchers noted that best practices include following the ASOPs and emerging guidance from CASTF. While CASTF supports the property and casualty industry, much of their guidance, such as having a clear

understanding of how the predictive model was built and how the data inputs where chosen and defined, are applicable across all practices areas.

The insurance industry wants to develop innovative products and, as such, is looking to use big data and advanced statistical models to help them capture the appropriate risks and relationships. The survey results indicated a general consensus that existing state regulations are not keeping pace with setting rates through the use of big data and advanced statistical models, and that there is a need for more regulatory guidance in these areas. The industry needs to work with the regulators to provide guidance to legislative bodies to modernize existing rules. The industry and regulators also need to work together to develop guidance that is consistent both within an insurance department and across state insurance departments. The state regulators need the resources and education to appropriately review rate filings based on new techniques, and the industry data and model documentation needs to be clear and transparent to the regulators. Both the industry and regulators agree on the need to maintain high standards and controls for data quality and model validation.

Section 5: Panelists

Special thanks go to those who took the considerable time and effort required to provide thoughtful and detail-oriented responses to the questionnaire.

Frederick James Andersen, FSA, MAAA
Michael Conover
Ralph Donato
Robert T. Eaton, FSA, MAAA
Gordon K. Hay, FCAS, MAAA, CPCU
Joseph S. Highbarger, FCAS, MAAA
Jeffrey Robert Huddleston, ASA, MAAA, CERA
Linda M. Lankowski, FSA, MAAA
David A. Moore, FSA, MAAA
Michael N. Muldoon, ASA, MAAA
Richard N. Piazza, ACAS, MAAA
Paul E. Stordahl, FSA, MAAA

Section 6: Acknowledgements

The researchers' deepest gratitude goes to those without whose efforts this project could not have come to fruition: the Project Oversight Group and others for their diligent work overseeing questionnaire development, analyzing and discussing respondent answers, and reviewing and editing this report for accuracy and relevance.

Project Oversight Group Members:

Rhonda Kay Ahrens, FSA, MAAA
Patricia L. Allison, FSA, MAAA
Mary J. Bahna-Nolan, FSA, MAAA, CERA
Michael W. Boerner, ASA, MAAA
Cabe W. Chadick, FSA, MAAA
Theresa M. Dziedzic, FSA, MAAA
Laura Alden Hanson, FSA, MAAA
Rachel C. Hemphill, FSA, FCAS, MAAA
Jeffrey L. Johnson, FSA, MAAA
Karen T. Jiang, FSA, MAAA, CERA
Derek Michael Kueker, FSA, MAAA
Raja Malkani, ASA, MAAA
John W. Robinson, FSA, MAAA, FCA
Jeff Wu, FSA

At the Society of Actuaries:

R. Dale Hall, FSA, MAAA, CERA, Managing Director of Research
Cynthia MacDonald, FSA, MAAA, Senior Experience Study Actuary
Mervyn Kopinsky, FSA, EA, MAAA, Experience Study Actuary
Korrel E. Crawford, Senior Research Administrator

References

The following documents have been reviewed to prepare portions of the information provided in this report (along with links, if available).

- Tata Consultancy Services Limited. 2011. White Paper on Predictive Analytics in Property and Casualty Insurance: Real Value or More Empty Promises. https://www.the-digitalinsurer.com/wp-content/uploads/2013/12/80-Consulting_Whitepaper_PredictiveAnalytics-Property-Casualty-Insurance_05_2011.pdf

- Werner, Modlin for Willis Towers Watson. 2016. Basic Ratemaking. https://www.casact.org/library/studynotes/werner_modlin_ratemaking.pdf

- Glowacki. 2012. Effective Model Validation. http://www.milliman.com/insight/insurance/Effective-model-validation/

- Society of Actuaries. 2014. Article from "The Actuary Magazine." August/September 2014 – Volume 11, Issue 4. What is the Scope of Model Validation? https://theactuarymagazine.org/wp-content/uploads/2016/12/act-2014-vol11-iss4.pdf

- A Public Policy Practice Note. 2017. Model Governance: Some Consideration for Practicing Life Actuaries. https://www.actuary.org/files/publications/Model_Governance_PN_042017.pdf

- Brownlee. 2017. How to Train a Final Machine Learning Model. https://machinelearningmastery.com/train-final-machine-learning-model/

- Hogan Lovells Publications. 2018. NAIC Developing Best Practices for Regulatory Review of Predictive Models and Analytics in Rate Filings. https://www.hoganlovells.com/en/publications/naic-developing-best-practices-forregulatory-review-of-predictive-models-and-analytics-in-rate-filings

- Casualty Actuarial and Statistical (C) Task Force. 2018. Draft White Paper on Regulatory Review of Predictive Models. https://www.naic.org/documents/cmte_c_catf_exposure_predictive_model_white_paper.pdf

- Casualty Actuarial and Statistical (C) Task Force. State Rate Filings – GLM and Other Model Checklists https://www.naic.org/cmte_c_catf.htm

- National Association of Insurance Commissioners. 2011. State Insurance Regulation. https://www.naic.org/documents/topics_white_paper_hist_ins_reg.pdf

- National Association of Insurance Commissioners. State Insurance Regulation: History, Purpose and Structure. https://www.naic.org/documents/consumer_state_reg_brief.pdf
- About the Affordable Care Act. https://www.hhs.gov/healthcare/about-theaca/index.html
- American Academy of Actuaries. 2009. Code of Professional Conduct. https://www.actuary.org/pdf/prof/code_of_conduct.pdf
- American Academy of Actuaries. 2019. Actuarial Standards of Practice. http://www.actuarialstandardsboard.org/standards-of-practice/
- American Academy of Actuaries. Applicability Guidelines for Actuarial Standards of Practice. http://www.actuary.org/content/applicability-guidelines-actuarial-standardspractice-0
- Birnbaum. Center for Economic Justice. NAIC Summit. 2018. Ethical Algorithms in Insurance. https://www.naic.org/insurance_summit/documents/insurance_summit_2018_MR_04-3.pdf
- Wenman. 2017. May I have Some Neural Networks with My Insurance Data, Please? https://www.pinnacleactuaries.com/blog/may-i-have-some-neural-networks-with-myinsurance-data-please
- Willis Towers Watson. 2017. Infographic: How U.S. P&C insurers use predictive models and big data. 2016 Predictive Modeling and Big Data Survey (U.S.). https://www.willistowerswatson.com/en-US/insights/2017/03/infographic-how-us-pand-c-insurers-use-predictive-models-and-big-data
- Deloitte. Predictive Modeling for Life Insurance. Ways Life Insurers Can Participate in the Business Analytics Revolution. https://www.soa.org/globalassets/assets/files/research/projects/research-pred-mod-life-batty.pdf
- New York Department of Financial Services. 2019. Use of External Consumer Data and Information Sources in Underwriting for Life Insurance. https://dfs.ny.gov/industry_guidance/circular_letters/cl2019_01

Appendix A: Questionnaire

A.1 PROJECT SCOPE

Review and highlight the current U.S. regulatory atmosphere for insurance ratemaking, and the actuarial and statistical techniques being used and considered by insurance companies and reviewed by state insurance authorities. This project will highlight the intersection of U.S. insurance ratemaking and analytics techniques in advanced modern rating systems. The evolution of the product filing process and evolving regulation with the emergence of predictive analytics and Generalized Linear Models for rating will be considered. In addition, the report will highlight the current environment and emerging issues, the types of variables allowed to be used in rating, and provide the concepts considered in order to be used in a rating system, from a statistical perspective, and is not a proxy for disallowed rating variables.

This will be accomplished using research approaches that include reviewing technical specifics of appropriate literature along with interviews of individual experts regarding current practices. The results of these approaches will be analyzed and summarized in a final report.

A.2 GENERAL NOTES ON QUESTIONNAIRE

Thank you for participating in this questionnaire. Please note the following:

1. The questions below are intended for you to respond to from your particular perspective and focus - regulatory, industry and consulting services. Please respond from your personal point of view. You are encouraged to network with others from your company to enhance your response. If you have comments as a response to a particular question from a different perspective also, please provide them and indicate as such. Additionally, please respond to the questions as it pertains to your area of practice (Life, Health, P&C, other), but indicate if your responses are applicable to multiple practice areas. For example, if the majority of your work is in health insurance, then please respond from this perspective. You are free to comment on other practice areas and should indicate when that is the case. Please do not include any specific information regarding current or planned actual pricing.

2. If you are a consultant please indicate the nature of the work your responses reflect (regulatory, industry or other)

3. If a particular question within your category does not apply to you, please state “Not Applicable”.

4. Your responses are confidential. The final research report will be based on anonymized responses and would not involve specific information on any particular company or client.
5. To the extent possible, for each answer you provide, please include the applicable market(s), industry(ies), and product(s).

6. This survey is intended to be updated electronically within this MS Word document. The "Response" section for each question will expand as responses are typed in.

Questionnaire:

1. Please provide a brief description of your organization and role.

    a. Your organization, industry category (regulatory, industry, or consulting), and area of practice (property & casualty, health, or life insurance):

    b. Your role within your organization:

2. Please share your thoughts on the current U.S. regulatory atmosphere for insurance ratemaking with respect to emerging actuarial and statistical techniques being used and considered by insurance companies and reviewed by state insurance authorities. Please be as specific as possible and reference applicable statutes, regulations, ASOPs, etc.

3. What regulations are you focused on during your rate development or rate review process? Please be specific and reference applicable statutes, regulations, ASOPs, etc.

4. How well do current regulations for rate filing / reviewing of rates apply to modern techniques and data collection? What regulations work well and what, if anything, needs to change?

5. What actuarial and statistical techniques or models do you use, have considered using, or have observed being used in rate development? Are these models developed by a vendor or developed in-house? Are Generalized Linear Models commonly used or are newer or other modeling techniques now being used? Please share advantages and disadvantages of the models from your perspective.

6. To what extent do you review the models used, and if so what techniques do you use to validate the models (describe your methodology)? Alternatively, does your company/client typically rely on consultants for all or part of this task?

7. What are key variables for the development / review of an insurance company's ratemaking system? Please specify by product type.

8. What are current issues you observe in ratemaking and analytics techniques in advanced modern rating systems, i.e. use of big data, predictive models, etc.?
9. What are emerging issues you are expecting or gearing up for in ratemaking and analytics techniques in advanced modern rating systems?

10. What best practices would you recommend to look out for in the analysis and review of big data, analytics techniques, predictive analytics, and advanced models in setting and reviewing rates?

11. What would you consider best practices that could be used across multiple practices (life insurance, health insurance, and/or property and casualty insurance)?

12. What would you consider benefits and drawbacks of using advanced models and big data and the risks to insurance companies, regulators and policyholders?

13. When setting and reviewing rates that are developed using big data and advanced models, how do you determine the accuracy of the data and the predictive power of the model?

14. What implications do you project from a full implementation of big data and predictive analytics? Please consider implications to policyholders as well as any financial, legal and personnel implication for insurance companies and regulators.

15. In order of importance, what are the sources of and data elements from non-traditional external data (for example, motor vehicle records and credit scores for life insurance, usage-based insurance devises and education for P&C insurance, and credit card spending habits and wearable devices for health insurance) you consider applicable for rate setting? Please specify the practice area (life, health, P&C). How do you ensure applicability and relevance of the data for its intended use?

16. What processes do you have in place to review actual experience versus pricing assumptions? How frequently is each process performed?

17. Please share your thoughts on how statistical and predictive models could potentially be misused in the ratemaking process. How should this be taken into consideration from a rate development, review, and regulatory perspective? (e.g. being used to circumvent allowable rating factors).

About The Society of Actuaries

The Society of Actuaries (SOA), formed in 1949, is one of the largest actuarial professional organizations in the world dedicated to serving more than 32,000 actuarial members and the public in the United States, Canada and worldwide. In line with the SOA Vision Statement, actuaries act as business leaders who develop and use mathematical models to measure and manage risk in support of financial security for individuals, organizations and the public.

The SOA supports actuaries and advances knowledge through research and education. As part of its work, the SOA seeks to inform public policy development and public understanding through research. The SOA aspires to be a trusted source of objective, data-driven research and analysis with an actuarial perspective for its members, industry, policymakers and the public. This distinct perspective comes from the SOA as an association of actuaries, who have a rigorous formal education and direct experience as practitioners as they perform applied research. The SOA also welcomes the opportunity to partner with other organizations in our work where appropriate.

The SOA has a history of working with public policymakers and regulators in developing historical experience studies and projection techniques as well as individual reports on health care, retirement and other topics. The SOA's research is intended to aid the work of policymakers and regulators and follow certain core principles:

Objectivity: The SOA's research informs and provides analysis that can be relied upon by other individuals or organizations involved in public policy discussions. The SOA does not take advocacy positions or lobby specific policy proposals.

Quality: The SOA aspires to the highest ethical and quality standards in all of its research and analysis. Our research process is overseen by experienced actuaries and nonactuaries from a range of industry sectors and organizations. A rigorous peer-review process ensures the quality and integrity of our work.

Relevance: The SOA provides timely research on public policy issues. Our research advances actuarial knowledge while providing critical insights on key policy issues, and thereby provides value to stakeholders and decision makers.

Quantification: The SOA leverages the diverse skill sets of actuaries to provide research and findings that are driven by the best available data and methods. Actuaries use detailed modeling to analyze financial risk and provide distinct insight and quantification. Further, actuarial standards require transparency and the disclosure of the assumptions and analytic approach underlying the work.

Society of Actuaries
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org