EXPOSURE NOTE: The drafting group considered comments submitted based on the 5/14/19 draft of the bulk of the paper and comments submitted on the 7/24/19 draft of Sections VIII “Proposed Changes to the Product Filing Review Handbook” and IX “Proposed State Guidance” (exposed 8/3/19). Please submit comments to Kris DeFrain (kdefrain@naic.org) on this 10/14/19 draft by Friday, Nov. 22, 2019.

---

**Casualty Actuarial and Statistical (C) Task Force**

**Regulatory Review of Predictive Models**

**White Paper**

---

**Table of Contents**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Introduction</td>
<td>2</td>
</tr>
<tr>
<td>II. What is a “Best Practice?”</td>
<td>2</td>
</tr>
<tr>
<td>III. Do Regulators Need Best Practices to Review Predictive Models?</td>
<td>3</td>
</tr>
<tr>
<td>IV. Scope</td>
<td>5</td>
</tr>
<tr>
<td>V. Confidentiality</td>
<td>5</td>
</tr>
<tr>
<td>VI. Guidance for Regulatory Review of Predictive Models (Best Practices)</td>
<td>5</td>
</tr>
<tr>
<td>VII. Predictive Models – Information for Regulatory Review</td>
<td>6</td>
</tr>
<tr>
<td>VIII. Proposed Changes to the <em>Product Filing Review Handbook</em></td>
<td>26</td>
</tr>
<tr>
<td>IX. Proposed State Guidance</td>
<td>35</td>
</tr>
<tr>
<td>X. Other Considerations</td>
<td>35</td>
</tr>
<tr>
<td>Appendix A – Best Practice Development</td>
<td>39</td>
</tr>
<tr>
<td>Appendix B - - Glossary of Terms</td>
<td>40</td>
</tr>
<tr>
<td>Appendix C – Sample Rate-Disruption Template</td>
<td>47</td>
</tr>
<tr>
<td>Appendix D – Information Needed by Regulator Mapped into Best Practices</td>
<td>49</td>
</tr>
<tr>
<td>Appendix E – References</td>
<td>49</td>
</tr>
</tbody>
</table>

---

© 2019 National Association of Insurance Commissioners
I. INTRODUCTION

Insurers’ use of predictive analytics along with big data has significant potential benefits to both consumers and insurers. Predictive analytics can reveal insights into the relationship between consumer behavior and the cost of insurance, lower the cost of insurance for many, and provide incentives for consumers to better control and mitigate loss. However, predictive analytic techniques are evolving rapidly and leaving many regulators without the necessary tools to effectively review insurers’ use of predictive models in insurance applications.

When a rate plan is truly innovative, the insurer must anticipate or imagine the reviewers’ interests because reviewers will respond with unanticipated questions and have unique educational needs. Insurers can learn from the questions, teach the reviewers, and so forth. When that back-and-forth learning is memorialized and retained, filing requirements and insurer presentations can be routinely organized to meet or exceed reviewers’ needs and expectations. Hopefully, this paper helps bring more consistency to the art of reviewing predictive models within a rate filing.

The Casualty Actuarial and Statistical (C) Task Force (CASTF) has been charged with identifying best practices to serve as a guide to state insurance departments in their review of predictive models underlying rating plans. There were two charges given to CASTF by the Property and Casualty Insurance (C) Committee at the request of the Big Data (EX) Working Group:

A. Draft and propose changes to the Product Filing Review Handbook to include best practices for review of predictive models and analytics filed by insurers to justify rates.

B. Draft and propose state guidance (e.g., information, data) for rate filings that are based on complex predictive models.

This paper will identify best practices when reviewing predictive models and analytics filed by insurers to justify rates and provide state guidance for review of rate filings based on predictive models. Upon adoption of this paper by the Executive (EX) Committee and Plenary, the Task Force will evaluate how to incorporate these best practices into the Product Filing Review Handbook and will recommend such changes to the Speed to Market (EX) Working Group.

II. WHAT IS A “BEST PRACTICE”?

A best practice is a form of program evaluation in public policy. At its most basic level, a practice is a “tangible and visible behavior…[based on] an idea about how the actions…will solve a problem or achieve a goal” (Bardach, 2016). Best practices are used to maintain quality as an alternative to mandatory legislated standards and can be based on self-assessment or benchmarking. Therefore, a best practice represents an effective method of problem solving. The “problem” regulators want to solve is probably better posed as seeking an answer to this question: How can regulators determine that predictive models, as used in rate filings, are compliant with state laws and regulations?

Key Regulatory Principles

In this paper, best practices are based on the following principles that promote a comprehensive and coordinated review of predictive models across states:

1. State insurance regulators will maintain their current rate regulatory authority.

2. State insurance regulators will be able to share information to aid companies in getting insurance products to market more quickly.

3. State insurance regulators will share expertise and discuss technical issues regarding predictive models.

4. State insurance regulators will maintain confidentiality, where appropriate, regarding predictive models.

In this paper, best practices are presented in the form of guidance to regulators who review predictive models and to insurance companies filing rating plans that incorporate predictive models. Guidance will identify specific information

---

1 In this paper, references to “model” or “predictive model” are the same as “complex predictive model” unless qualified.

© 2019 National Association of Insurance Commissioners
useful to a regulator in the review of a predictive model, comment on what might be important about that information and, where appropriate, provide insight as to when the information might identify an issue the regulator needs to be aware of or explore further.

III. DO REGULATORS NEED BEST PRACTICES TO REVIEW PREDICTIVE MODELS?

The term “predictive model” refers to a set of models that use statistics to predict outcomes. When applied to insurance, the model is chosen to estimate the probability or expected value of an outcome given a set amount of input data; for example, models can predict the frequency of loss, the severity of loss, or the pure premium. The generalized linear model (GLM) is a commonly used predictive model in insurance applications, particularly in building an insurance product’s rating plan.

Depending on definitional boundaries, predictive modeling can sometimes overlap with the field of machine learning. In this modeling space, predictive modeling is often referred to as predictive analytics.

Before GLMs became vogue, rating plans were built using univariate methods. Univariate methods were considered intuitive and easy to demonstrate the relationship to costs (loss and/or expense). Today, many insurers consider univariate methods too simplistic since they do not take into account the interaction (or dependencies) of the selected input variables.

According to many in the insurance industry, GLMs introduce significant improvements over univariate-based rating plans by automatically adjusting for correlations among input variables. Today, the majority of predictive models used in personal automobile and home rating plans are GLMs. However, GLM results are not always intuitive, and the relationship to costs may be difficult to explain. This is a primary reason regulators can benefit from best practices.

A GLM consists of three elements:

- A target variable, Y, which is a random variable that is independent and follows a probability distribution from the exponential family, defined by a selected variance function and dispersion parameter.
- A linear predictor η = Xβ.
- A link function g such that E(Y) = μ = g⁻¹(η).

As can be seen in the description of the three GLM components above, it may take more than a casual introduction to statistics to comprehend the construction of a GLM. As stated earlier, a downside to GLMs is that it is more challenging to interpret the GLMs output than with univariate models.

If the underlying data is not credible, then no model will improve that credibility, and segmentation methods could make credibility worse. GLM software provides point estimates and allows the modeler to consider standard errors and confidence intervals. GLMs effectively assume that the underlying datasets are 100% credible no matter their size. If some segments have little data, the resulting uncertainty would not be reflected in the GLM parameter estimates themselves (although it might be reflected in the standard errors, confidence intervals, etc.). Even though the process of selecting relativities often includes adjusting the raw GLM output, the resultant selections are not typically then credibility-weighted with any complement of credibility. Nevertheless, selected relativities based on GLM model output may differ from GLM point estimates.

Because of this presumption in credibility, which may or may not be valid in practice, the modeler and the regulator reviewing the model would need to engage in thoughtful consideration when incorporating GLM output into a rating plan to ensure that model predictiveness is not compromised by any lack of actual credibility. Another consideration is the availability of big data, both internal and external, that may result in the selection of predictor variables that have spurious

4 A more thorough exploration of different predictive models will be found in many statistics’ books, including Geisser, Seymour (September 2016). Predictive Inference: An Introduction. New York: Chapman & Hall.
5 The generalized linear model (GLM) is a flexible family of models that are unified under a single method. Types of GLM include logistic regression, Poisson regression, gamma regression and multinomial regression.
6 More information on model elements can be found in most statistics’ books.
7 Sometimes insurers do review complements of credibility and further weight the GLM output with those complements. While this may not be a standard practice today, new techniques could result in this becoming more standard in the future.
8 GLMs provide confidence intervals, credibility methods do not. There are techniques such as penalized regression that blend credibility with a GLM and improve a model’s ability to generalize.

© 2019 National Association of Insurance Commissioners
correlation with the target variable. Therefore, to mitigate the risk that model credibility or predictiveness is lacking, a complete filing for a rating plan that incorporates GLM output should include validation evidence for the rating plan, not just the statistical model.

To further complicate regulatory review of models in the future, modeling methods are evolving rapidly and not limited just to GLMs. As computing power grows exponentially, it is opening up the modeling world to more sophisticated forms of data acquisition and data analysis. Insurance actuaries and data scientists seek increased predictiveness by using even more complex predictive modeling methods. Examples of these are predictive models utilizing random forests, decision trees, neural networks, or combinations of available modeling methods (often referred to as ensembles). These evolving techniques will make the regulators’ understanding and oversight of filed rating plans incorporating predictive models even more challenging.

In addition to the growing complexity of predictive models, many state insurance departments do not have in-house actuarial support or have limited resources to contract out for support when reviewing rate filings that include use of predictive models. The Big Data (EX) Working Group identified the need to provide states with guidance and assistance when reviewing predictive models underlying filed rating plans.9 The Working Group circulated a proposal addressing aid to state insurance regulators in the review of predictive models as used in personal automobile and home insurance rate filings. This proposal was circulated to all of the Working Group members and interested parties on December 19, 2017, for a public comment period ending January 12, 2018.10 The Big Data Working Group effort resulted in the new CASTF charges (see the Introduction section) with identifying best practices that provide guidance to states in the review of predictive models.

So, to get to the question asked by the title of this section: Do regulators need best practices to review predictive models? It might be better to ask this question another way: Are best practices in the review of predictive models of value to regulators and insurance companies? The answer is “yes” to both questions. Regulatory best practices need to be developed that do not unfairly or inordinately create barriers for insurers and ultimately consumers while providing a baseline of analysis for regulators to review the referenced filings. Best practices will aid regulatory reviewers by raising their level of model understanding. With regard to scorecard models and the model algorithm, there is often not sufficient support for relative weight, parameter values, or scores of each variable. Best practices can potentially aid in addressing this problem.

However, best practices are not intended to create standards for filings that include predictive models. Rather, best practices will assist the states in identifying the model elements they should be looking for in a filing that will aid the regulator in understanding why the company believes that the filed predictive model improves the company’s rating plan, making that rating plan fairer to all consumers in the marketplace. To make this work, both regulators and industry need to recognize that:

- Best practices merely provide guidance to regulators in their essential and authoritative role over the rating plans in their state.
- All states may have a need to review predictive models whether that occurs with approval of rating plans or in a market conduct exam. Best practices help the regulator identify elements of a model that may influence the regulatory review as to whether modeled rates are appropriately justified. Each regulator needs to decide if the insurer’s proposed rates are compliant with state laws and regulations and whether to act on that information.
- Best practices will lead to improved quality in predictive model reviews across states, aiding speed to market and competitiveness of the state marketplace.
- Best practices provide a framework for states to share knowledge and resources to facilitate the technical review of predictive models.

---

10 All comments received by the end of January were posted to the NAIC website March 12 for review.

© 2019 National Association of Insurance Commissioners
As adopted by the Casualty Actuarial and Statistical (C) Task Force on XX/XX/XX

• Best practices aid training of new regulators and/or regulators new to reviewing predictive models. (This is especially useful for those regulators who do not actively participate in NAIC discussions related to the subject of predictive models.)

• Each regulator adopting best practices will be better able to identify the resources needed to assist their state in the review of predictive models.

Lastly, from this point on in this paper, best practices will be referred to as “guidance.” This reference is in line with the intent of this paper to support individual state autonomy in the review of predictive models.

IV. SCOPE

The focus of this paper will be on GLMs used to create personal automobile and home insurance rating plans.

The legal and regulatory constraints (including state variations) are likely to be more evolved, and challenging, for personal automobile and home insurance. Through review of these personal lines, the knowledge needed to review predictive models, and guidance in this paper regarding GLMs may be transferrable when the review involves GLMs applied to other lines of business. Modeling depends on context, so the GLM reviewer has to be alert for data challenges and business applications that differ from the more familiar personal lines. For example, compared to personal lines, modeling rates in commercial lines is more likely to encounter low volumes of historical data, dependence on advisory loss costs, unique large accounts with large deductibles and package products that create policies from numerous line-of-business and coverage building blocks. Commercial lines commonly use individual risk modifications following experience, judgment, and/or expense considerations. A regulator may never see how models impact commercial excess and surplus lines filings. Also, the legal and regulatory constraints (including state variations) are likely to be more prevalent, and challenging in personal lines, which is the basis of this paper’s guidance. A GLM rate model for personal lines in 2019 is either an update or a late-adopter's defensive tactic. Adopting a GLM for commercial lines rating plans has a shorter history and thus is less familiar to many regulators.

Guidance offered here might be useful (with deeper adaptations) when starting to review different types of predictive models. If the model is not a GLM, some listed items might not apply. Not all predictive models generate p-values or F tests. Depending on the model type, other considerations might be important. When transferring guidance to other lines of business and other types of model, unique considerations may arise depending on the context in which a predictive model is proposed to be deployed, the uses to which it is proposed to be put, and the potential consequences for the insurer, its customers and its competitors. This paper does not delve into these possible considerations but regulators should be prepared to address them as they arise.

V. CONFIDENTIALITY

Regulatory reviewers are required to protect confidential information in accordance with applicable State law. However, insurers should be aware that a rate filing might become part of the public record. Each state determines the confidentiality of a rate filing, supplemental material to the filing, when filing information might become public, the procedure to request that filing information be held confidentially, and the procedure by which a public records request is made. It is incumbent on an insurer to be familiar with each state’s laws regarding the confidentiality of information submitted with their rate filing.

Though state authority, regulations and rules governing confidentiality always apply, this reliance should be revisited if the NAIC or another third party becomes involved in the review process on behalf of the states.

VI. GUIDANCE FOR REGULATORY REVIEW OF PREDICTIVE MODELS (BEST PRACTICES)

Best practices will help the regulator understand if a predictive model is cost based, if the predictive model is compliant with state law, and how the model improves, the company’s rating plan. Best practices can, also, make the regulator's review more consistent across states and more efficient, and assist companies in getting their products to market faster. With this in mind, the regulator's review of predictive models should:

1. Ensure that the selected rating factors, based on the model or other analysis, produce rates that are not excessive, inadequate, or unfairly discriminatory.
   a. Review the overall rate level impact of the proposed revisions to rate level indications provided by the filer.
b. Review the premium disruption for individual policyholders and how the disruptions can be explained to individual consumers.

c. Review the individual input characteristics to and output factors from the predictive model (and its sub-models), as well as, associated selected relativities to ensure they are not unfairly discriminatory.

2. Obtain a clear understanding of the data used to build and validate the model, and thoroughly review all other aspects of the model, including assumptions, adjustments, variables, submodels used as input, and resulting output.

   a. Determine that individual input characteristics to a predictive model and their resulting rating factors are related to the expected loss or expense differences in risk.

   b. Determine that the data used as input to the predictive model is accurate, including a clear understanding how missing values, erroneous values and outliers are handled.

   c. Determine that any adjustments to the raw data are handled appropriately, including but not limited to, trending, development, capping, removal of catastrophes.

   d. Obtain a clear understanding of how often each risk characteristic, used as input to the model, is updated and whether the model is periodically rerun, so model output reflects changes to non-static risk characteristics.

   e. Determine whether internal and external data used in relation to the model is compatible with practices allowed in the jurisdiction and do not reflect prohibited characteristics.

   f. Obtain a clear understanding of how the selected predictive model was built.

3. Evaluate how the model interacts with and improves the rating plan.

   a. Obtain a clear understanding of the characteristics that are input to a predictive model (and its sub-models), their relationship to each other and their relationship to non-modeled characteristics/variables used to calculate a risk’s premium.

   b. Obtain a clear understanding why the insurer believes this type of model works in an insurance risk application.

   c. Obtain a clear understanding of how model output interacts with non-modeled characteristics/variables used to calculate a risk’s premium.

   d. Obtain a clear understanding of how the predictive model was integrated into the insurer’s state rating plan and how it improves that plan.

   e. For predictive model refreshes, determine whether sufficient validation was performed to ensure the model is still a good fit.

4. Enable competition and innovation to promote the growth, financial stability, and efficiency of the insurance marketplace.

   a. Enable innovation in the pricing of insurance through acceptance of predictive models, provided they are actuarially sound and in compliance with state laws.

   b. Protect the confidentiality of filed predictive models and supporting information in accordance with state law.

   c. Review predictive models in a timely manner to enable reasonable speed to market.

VII. PREDICTIVE MODELS – INFORMATION FOR REGULATORY REVIEW

This section of the paper identifies the information a regulator may need to review a predictive model used by an insurer to support a personal automobile or home insurance rating plan. The list is lengthy but not exhaustive. It is not intended to limit the authority of a regulator to request additional information in support of the model or filed rating plan. Nor is every item on the list intended to be a requirement for every filing. However, the items listed should help guide a regulator to obtain sufficient information to determine if the rating plan meets state specific filing and legal requirements.
Documentation of the design and operational details of the model is required to ensure business continuity and transparency of models used. Granularity of documentation takes into account the level of management or key function at which it is intended to be used. Documentation should be sufficiently detailed and complete to enable a third party to form a sound judgment on the suitability of the model for the intended purpose. The theory, assumptions, methodologies, software and empirical bases should be explained, as well as the data used in developing and implementing the model. Relevant testing and ongoing performance testing need to be documented. Key model limitations and overrides need to be pointed out so that stakeholders understand the circumstances under which the model does not work effectively. End-user documentation should be provided and key reports using the model results described. Major changes to the model need to be shared in a timely manner and documented, and IT controls should be in place, such as a record of versions, change control and access to model.11

Many information elements listed below are probably confidential, proprietary or trade secret and should be treated as such according to state law. Regulators should be aware of their state laws on confidentiality when requesting data from insurers that may be proprietary or trade secret. For example, some proprietary models may have contractual terms (with the insurer) that prevent disclosure to the public. Without clear necessity, exposing this data to additional dissemination may hinder the model's protection.12

Though the list seems long, the insurer should already have internal documentation on the model for more than half of the information listed. The remaining items on the list require either minimal analysis (approximately 25%) or deeper analysis to generate the information for a regulator (approximately 25%).

The “Importance to Regulator’s Review” ranking of information a regulator may need to review is based on the following level criteria:

**Level 1** - This information is necessary to begin the review of a predictive model. These data elements pertain to basic information about the type and structure of the model, the data and variables used, the assumptions made, and the goodness of fit. Ideally, this information would be included in the filing documentation with the initial submission of a filing made based on a predictive model.

**Level 2** - This information is necessary to continue the review of all but the most basic models; such as those based only on the filer’s internal data and only including variables that are in the filed rating plan. These data elements provide more detailed information about the model and address questions arising from review of the information in Level 1. Insurers concerned with speed to market may also want to include this information in the filing documentation.

**Level 3** - This information is necessary to continue the review of a model where concerns have been raised and not resolved based on review of the information in Levels 1 and 2. These data elements address even more detailed aspects of the model including (Drafting Note: Insert List after levels are assigned). This information does not necessarily need to be included with the initial submission, unless specifically requested in a particular jurisdiction, as it is typically requested only if the reviewer has concerns that the model may not comply with state laws.

**Level 4** - This information is necessary to continue the review of a model where concerns have been raised and not resolved based on the information in Levels 1, 2, and 3. This most granular level of detail is addressing the basic building blocks of the model and does not necessarily need to be included by the filer with the initial submission, unless specifically requested in a particular jurisdiction. It is typically requested only if the reviewer has serious concerns that the model may produce rates or rating factors that are excessive, inadequate, or unfairly discriminatory.

---

11 Michele Bourdeau, The Modeling Platform ISSUE 4 • DECEMBER 2016 Model Risk Management: An Overview, Page 6; Published by the Modeling Section of the Society of Actuaries.

12 There are some models that are made public by the vendor and would not result in a hindrance of the model's protection.
### A. Selecting Model Input

<table>
<thead>
<tr>
<th>Section</th>
<th>Information Element</th>
<th>Level of Importance to the Regulator’s Review</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1.a</td>
<td>Review the details of all sources for both insurance and non-insurance data used as input to the model (only need sources for filed input characteristics included in the filed model). For each source, obtain a list all data elements used as input to the model that came from that source.</td>
<td>1</td>
<td>Request details of all data sources, whether internal to the company or from external sources. For insurance experience (policy or claim), determine whether data are aggregated by calendar, accident, fiscal or policy year and when it was last evaluated. For each data source, get a list all data elements used as input to the model that came from that source. For insurance data, get a list all companies whose data is included in the datasets.</td>
</tr>
<tr>
<td>A.1.b</td>
<td>Reconcile aggregated insurance data underlying the model with available external insurance reports.</td>
<td>4</td>
<td>Accuracy of insurance data should be reviewed. Aggregated data is straight from the insurer's data banks without modification (e.g., not scrubbed or transformed). The dataset would not be adjusted for data selection or model building. The company should provide some form of reasonability check that the data makes sense when checked against other audited sources.</td>
</tr>
<tr>
<td>A.1.c</td>
<td>Review the geographic scope and geographic exposure distribution of the raw data for relevance to the state where the model is filed.</td>
<td>2</td>
<td>The company should explain how the data used to build the model makes sense for a specific state. The regulator should inquire which states were included in the data underlying the model build, testing and validation. The company should provide an explanation where the data came from geographically and that it is a good representation for a state, i.e., the distribution by state should not introduce a geographic bias. For example, there could be a bias by peril or wind-resistant building codes. Evaluate whether the data is relevant to the loss potential for which it is being used. For example, verify that hurricane data is only used where hurricanes can occur.</td>
</tr>
</tbody>
</table>
### 2. Sub-Models

<table>
<thead>
<tr>
<th></th>
<th>A.2.a</th>
<th>Consider the relevance of (e.g., is there a bias) of overlapping data or variables used in the model and sub-models.</th>
<th>1</th>
<th>Check if the same variables/datasets were used in both the model, a submodel or as stand-alone rating characteristics. If so, verify there was no double-counting or redundancy.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A.2.b</td>
<td>Determine if the sub-model was previously approved (or accepted) by the regulatory agency.</td>
<td>1</td>
<td>If the sub-model was previously approved, that may change the extent of the sub-model’s review. If approved, verify when and that it was the same model currently under review. However, previous approvals do not necessarily confer a guarantee of ongoing approval, for example when statutes and regulations have changed or if a model's indications have been undermined by subsequent empirical experience. However, knowing whether a model has been previously approved can help focus the regulator's efforts and determine whether or not the prior decision needs to be revisited.</td>
</tr>
<tr>
<td></td>
<td>A.2.c</td>
<td>Determine if sub-model output was used as input to the GLM; obtain the vendor name, and the name and version of the sub-model.</td>
<td>1</td>
<td>To accelerate the review of the filing, get the name and contact information for a representative from the vendor. The company should provide the name of the third-party vendor and a contact in the event the regulator has questions. The &quot;contact&quot; can be an intermediary at the insurer, e.g., a filing specialist, who can place the regulator in direct contact with a Subject Matter Expert (SME) at the vendor. Examples of such sub-models include credit/financial scoring algorithms and household composite score models. Sub-models can be evaluated separately and in the same manner as the primary model under evaluation. A sub-model contact for additional information should be provided. SMEs on sub-model may need to be brought into the conversation with regulators (whether in-house or 3rd-party sub-models are used).</td>
</tr>
<tr>
<td></td>
<td>A.2.d</td>
<td>If using catastrophe model output, identify the vendor and the model settings/assumptions used when the model was run.</td>
<td>1</td>
<td>For example, it is important to know hurricane model settings for storm surge, demand surge, long/short-term views. To accelerate the review of the filing, get contact information for the SME that ran the model and an SME from the vendor. The &quot;SME&quot; can be an intermediary at the insurer, e.g., a filing specialist, who can place the regulator in direct contact with the appropriate SMEs at the insurer or model vendor.</td>
</tr>
</tbody>
</table>
|   | A.2.e | If using catastrophe model output (a sub-model) as input to the GLM under review, verify whether loss associated with the modeled output was removed from the loss experience datasets. | 1 | If a weather-based sub-model is input to the GLM under review, loss data used to develop the model should not include loss experience associated with the weather-based sub-model. Doing so could cause distortions in the modeled results by double counting such losses when determining relativities or loss loads in the filed rating plan. For example, redundant losses in the data may occur when non-hurricane wind losses are included in the data while also using a severe...
convective storm model in the actuarial indication. Such redundancy may also occur with the inclusion of fluvial or pluvial flood losses when using a flood model, inclusion of freeze losses when using a winter storm model or including demand surge caused by any catastrophic event.

Note that, the rating plan or indications underlying the rating plan, may provide special treatment of large losses and non-modeled large loss events. If such treatments exist, the company should provide an explanation how they were handled. These treatments need to be identified and the company/regulator needs to determine whether model data needs to be adjusted. For example, should large BI losses, in the case of personal automobile insurance, be capped or excluded, or should large non-catastrophe wind/hail claims in home insurance be excluded from the model's training, test and validation data?

If using output of any scoring algorithms, obtain a list of the variables used to determine the score and provide the source of the data used to calculate the score.

If any adjustments were made to aggregated data, e.g., transformations, binning and/or categorizations. If any, identify the name of the characteristic/variable and obtain a description of the adjustment.

3. Adjustments to Data

A.3.a  Determine if premium, exposure, loss or expense data were adjusted (e.g., developed, trended, adjusted for catastrophe experience or capped) and, if so, how? Do the adjustments vary for different segments of the data and, if so, identify the segments and how was the data adjusted?

A.3.b  Identify adjustments that were made to aggregated data, e.g., transformations, binning and/or categorizations. If any, identify the name of the characteristic/variable and obtain a description of the adjustment.

The rating plan or indications underlying the rating plan may provide special treatment of large losses and non-modeled large loss events. If such treatments exist, the company should provide an explanation how they were handled. These treatments need to be identified and the company/regulator needs to determine whether model data needs to be adjusted. For example, should large bodily injury (BI) liability losses in the case of personal automobile insurance be excluded, or should large non-catastrophe wind/hail claims in home insurance be excluded from the model's training, test and validation data? Look for anomalies in the data that should be addressed. For example, is there an extreme loss event in the data? If other processes were used to load rates for specific loss events, how is the impact of those losses considered? Examples of losses that can contribute to anomalies in the data are large losses or flood, hurricane or severe convective storm losses for personal auto comprehensive or home insurance.
Ask for aggregated data (one data set of pre-adjusted/scrubbed data and one data set of post-adjusted/scrubbed data) that allows the regulator to focus on the univariate distributions and compare raw data to adjusted/binmed/transformed/etc. data.  

Though most regulators may never ask for aggregated data and do not plan to rebuild any models, a regulator may ask for this aggregated data or subsets of it. It would be useful to the regulator if the percentage of exposures and premium for missing information from the model data by category were provided. This data can be displayed in either graphical or tabular formats.

Determine how missing data was handled. This is most relevant for variables that have been "scrubbed" or adjusted. The regulator should be aware of assumptions the modeler made in handling missing, null or "not available" values in the data. If adjustments or re-coding of values were made, they should be explained. It may be useful to the regulator if the percentage of exposures and premium for missing information from the model data were provided. This data can be displayed in either graphical or tabular formats.

If duplicate records exist, determine how they were handled. Look for a discussion of how outliers were handled. If necessary, the regulator may want to investigate further by getting a list (with description) of the outliers and determine what adjustments were made to each outlier. To understand the filer's response, the regulator should ask for the filer's materiality standard.

Obtain documentation on the methods used to merge data from different sources or filter data based on particular characteristics and a description of any preliminary analyses, data checks, and logical tests performed on the data and the results of those tests.  

This should explain how data from separate sources was merged or how subsets of policies, based on selected characteristics, are filtered to be included in the data underlying the model and the rationale for that filtering.

Obtain documentation on the insurer's process for reviewing the appropriateness, reasonableness, consistency and comprehensiveness of the data, including a discussion of the rational relationship the data has to the predicted variable.  

An example is when by-peril or by-coverage modeling is performed; the documentation should be for each peril/coverage and make rational sense. For example, if “murder” or “theft” data are used to predict the wind peril, provide support and a rational explanation of their use.
### A.4.c
Identify material findings the company had during their data review and obtain an explanation of any potential material limitations, defects, bias or unresolved concerns found or believed to exist in the data. If issues or limitations in the data influenced modeling analysis and/or results, obtain a description of those concerns and an explanation how modeling analysis was adjusted and/or results were impacted.

<table>
<thead>
<tr>
<th>Section</th>
<th>Information Element</th>
<th>Level of Importance to Regulator's Review</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.4.c</td>
<td>Identify material findings the company had during their data review and obtain an explanation of any potential material limitations, defects, bias or unresolved concerns found or believed to exist in the data. If issues or limitations in the data influenced modeling analysis and/or results, obtain a description of those concerns and an explanation how modeling analysis was adjusted and/or results were impacted.</td>
<td>1</td>
<td>A response of &quot;none&quot; or &quot;n/a&quot; may be an appropriate response.</td>
</tr>
</tbody>
</table>

### B. Building the Model

#### 1. High-Level Narrative for Building the Model

<table>
<thead>
<tr>
<th>Section</th>
<th>Information Element</th>
<th>Level of Importance to Regulator's Review</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.1.a</td>
<td>Identify the type of model underlying the rate filing (e.g. Generalized Linear Model – GLM, decision tree, Bayesian Generalized Linear Model, Gradient-Boosting Machine, neural network, etc.). Understand the model's role in the rating system and provide the reasons why that type of model is an appropriate choice for that role.</td>
<td>1</td>
<td>It is important to understand if the model in question is a GLM, and therefore these best practices are applicable or, if it is some other model type, in which case other reasonable review approaches may be considered. There should be an explanation of why the model (using the variables included in it) is appropriate for the line of business. If by-peril or by-coverage modeling is used, the explanation should be by-peril/coverage. Note, if the model is not a GLM, the guidance and information elements in this white paper may not apply in their entirety.</td>
</tr>
<tr>
<td>B.1.b</td>
<td>Identify the software used for model development. Obtain the name of the software vendor/developer, software product and a software version reference used in model development.</td>
<td>3</td>
<td>Changes in software from one model version to the next may explain if such changes, over time, contribute to changes in the modeled results. The company should provide the name of the third-party vendor and a &quot;contact&quot; in the event the regulator has questions. The &quot;contact&quot; can be an intermediary at the insurer who can place the regulator in direct contact with appropriate SMEs. Open-source software/programs used in model development should be identified by name and version the same as if from a vendor. If version is not known, simply state such, e.g., &quot;R is the software source.&quot;</td>
</tr>
<tr>
<td>B.1.c</td>
<td>Obtain a description how the available data was divided between model training, test and validation datasets. The description should include an explanation why the selected approach was deemed most appropriate, and whether the company made any further subdivisions of available data and reasons for the subdivisions (e.g., a portion separated from training data to support testing of components during model building). Determine if the validation data was accessed before model training was completed and, if so, obtain an explanation why that came to occur.</td>
<td>1</td>
<td>It would be unexpected if validation data were used for any purpose other than validation.</td>
</tr>
<tr>
<td>B.1.d</td>
<td>Obtain a brief description of the development process, from initial concept to final model and filed rating plan.</td>
<td>1</td>
<td>The narrative should have the same scope as the filing.</td>
</tr>
<tr>
<td>B.1.e</td>
<td>Obtain a narrative on whether loss ratio, pure premium or frequency/severity analyses were performed and, if separate frequency/severity modeling was performed, how pure premiums were determined.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>B.1.f</td>
<td>Identify the model’s target variable.</td>
<td>1</td>
<td>A clear description of the target variable is key to understanding the purpose of the model. It may also prove useful to obtain a sample calculation of the target variable in Excel format, starting with the “raw” data for a policy, or a small sample of policies, depending on the complexity of the target variable calculation.</td>
</tr>
<tr>
<td>B.1.g</td>
<td>Obtain a description of the variable selection process.</td>
<td>1</td>
<td>The narrative regarding the variable selection process may address matters such as the criteria upon which variables were selected or omitted, identification of the number of preliminary variables considered in developing the model versus the number of variables that remained, and any statutory or regulatory limitations that were taken into account when making the decisions regarding variable selection.</td>
</tr>
<tr>
<td>B.1.h</td>
<td>In conjunction with variable selection, obtain a narrative on how the Company determine the granularity of the rating variables during model development.</td>
<td>2</td>
<td>This discussion should include discussion of how credibility was considered in the process of determining the level of granularity of the variables selected.</td>
</tr>
<tr>
<td>B.1.i</td>
<td>Determine if model input data was segmented in any way. For example, was modeling performed on a by-coverage, by-peril, or by-form basis? If so, obtain a description of data segmentation and the reasons for data segmentation.</td>
<td>1</td>
<td>The regulator would use this to follow the logic of the modeling process.</td>
</tr>
</tbody>
</table>
### B.1.j

If adjustments to the model were made based on credibility considerations, obtain an explanation of the credibility considerations and how the adjustments were applied. 

2

Adjustments may be needed given models do not explicitly consider the credibility of the input data or the model’s resulting output; models take input data at face value and assume 100% credibility when producing modeled output.

### 2. Medium-Level Narrative for Building the Model

#### B.2.a

At crucial points in model development, if selections were made among alternatives regarding model assumptions or techniques, obtain a narrative on the judgment used to make those selections. 

2

#### B.2.b

If post-model adjustments were made to the data and the model was rerun, obtain an explanation on the details and the rationale for those adjustments. 

2

Evaluate the addition or removal of variables and the model fitting. It is not necessary for the company to discuss each iteration of adding and subtracting variables, but the regulator should gain a general understanding how these adjustments were done, including any statistical improvement measures relied upon.

#### B.2.c

Obtain a description of the testing that was performed during the model-building process and a discussion of why interaction terms were included (or not included). 

3

There should be a description of testing that was performed during the model-building process. Examples of tests that may have been performed include univariate testing and review of a correlation matrix.

#### B.2.e

For the GLM, identify the link function used. Identify which distribution was used for the model (e.g., Poisson, Gaussian, log-normal, Tweedie). Obtain an explanation why the link function and distribution were chosen. Obtain the formulas for the distribution and link functions, including specific numerical parameters of the distribution. Obtain a discussion of applicable convergence criterion. 

1

Solving the GLM is iterative and the modeler can check to see if fit is improving. At some point convergence occurs, though when it occurs can be subjective or based on threshold criteria. The convergence criterion should be documented with a brief explanation of why it was selected.

#### B.2.f

Obtain a narrative on the formula relationship between the data and the model outputs, with a definition of each model input and output. The narrative should include all coefficients necessary to evaluate the predicted pure premium, relativity or other value, for any real or hypothetical set of inputs. 

2

B.4.l and B.4.m will show the mathematical functions involved and could be used to reproduce some model predictions.

#### B.2.g

If there were data situations in which GLM weights were used, obtain an explanation of how and why they were used. 

3

Investigate whether identical records were combined to build the model.

### 3. Predictor Variables
### B.3.a
Obtain a complete data dictionary, including the names, types, definitions and uses of each predictor variable, offset variable, control variable, proxy variable, geographic variable, geodemographic variable and all other variables in the model used on their own or as an interaction with other variables (including sub-models and external models).

### Types of variables might be continuous, discrete, Boolean, etc. Definitions should not use programming language or code. For any variable(s) intended to function as a control or offset, obtain an explanation of their rationale and impact. Also, for any use of interaction between variables, obtain an explanation of its rationale and impact.

### B.3.b
Obtain a list of predictor variables considered but not used in the final model, and the rationale for their removal.

The rationale for this requirement is to identify variables that the company finds to be predictive but ultimately may reject for reasons other than loss-cost considerations (e.g., price optimization). Also, look for variables the company tested and then rejected. This item could help address concerns about data dredging. The reasonableness of including a variable with given significance level could depend greatly on the other variables the company evaluated for inclusion in the model and the criteria for inclusion or omission. For instance, if the company tested 1,000 similar variables and selected the one with the lowest p-value of 0.001, this would be a far, far weaker case for statistical significance than if that variable was the only one the company evaluated. Note, context matters.

### B.3.c
Obtain a correlation matrix for all predictor variables included in the model and sub-model(s).

While GLMs accommodate collinearity, the correlation matrix provides more information about the magnitude of correlation between variables. The company should indicate what statistic was used (e.g., Pearson, Cramer's V). The reviewer should understand what statistic was used to produce the matrix; but should not specify the statistic.

### B.3.d
Obtain an rational explanation for why an increase in each predictor variable should increase or decrease frequency, severity, loss costs, expenses, or any element or characteristic being predicted.

The explanation should go beyond demonstrating correlation. Considering possible causation is relevant, but proving causation is neither practical nor expected. If no rational explanation can be provided, greater scrutiny may be appropriate. For example, the regulator should look for unfamiliar predictor variables and, if found, the regulator should seek to understand the rational connection that variable has to increasing or decreasing the target variable.

### B.3.e
If the modeler made use of one or more dimensionality reduction techniques, such as a Principal Component Analysis (PCA), obtain a narrative about that process, an explanation why that technique was chosen, and a description of the step-by-step process used to transform observations (usually correlated) into a set of linearly uncorrelated variables. In each instance, obtain a list of the pre-transformation and post-transformation variable names, and an explanation how the results of the dimensionality reduction technique was used within the model.

### 4. Adjusting Data, Model Validation and Goodness-of-Fit Measures

© 2019 National Association of Insurance Commissioners
B.4.a Obtain a description of the methods used to assess the statistical significance/goodness of the fit of the model to validation data, such as lift charts and statistical tests. Compare the model's projected results to historical actual results and verify that modeled results are reasonably similar to actual results from validation data.

For models that are built using multi-state data, validation data for some segments of risk is likely to have low credibility in individual states. Nevertheless, some regulators require model validation on state-only data, especially when analysis using state-only data contradicts the countrywide results. State-only data might be more applicable but could also be impacted by low credibility for some segments of risk. Look for geographic stability measures, e.g., across states or territories within state.

B.4.b For all variables (discrete or continuous), review the appropriate parameter values, confidence intervals, chi-square tests, p-values and any other relevant and material tests. Determine if model development data, validation data, test data or other data was used for these tests.

Typical p-values greater than 5% are large and should be questioned. Reasonable business judgment can sometimes provide legitimate support for high p-values. Reasonableness of the p-value threshold could also vary depending on the context of the model. For example, the threshold might be lower when many candidate variables were evaluated for inclusion in the model.

Overall lift charts and/or statistical tests using validation data may not provide enough of the picture. If there is concern about one or more individual variables, the reviewer may obtain, for each discrete variable level, the parameter value, confidence intervals, chi-square tests, p-values and any other relevant and material tests. For variables that are modeled continuously, it may be sufficient to obtain statistics around the modeled parameters; for example, confidence intervals around each level of an AOI curve might be more than what is needed.

B.4.c Identify the threshold for statistical significance and explain why it was selected. Obtain a reasonable an appropriately supported explanation for keeping the variable for each discrete variable level where the p-values were not less than the chosen threshold.

Typical p-values greater than 5% are large and should be questioned. Reasonable business judgment can sometimes provide legitimate support for high p-values. Reasonableness of the p-value threshold could also vary depending on the context of the model. For example, the threshold might be lower when many candidate variables were evaluated for inclusion in the model.

Overall lift charts and/or statistical tests using validation data may not provide enough of the picture. If there is concern about one or more individual variables, the reviewer may obtain, for each discrete variable level, the parameter value, confidence intervals, chi-square tests, p-values and any other relevant and material tests.
For overall discrete variables, review type 3 chi-square tests, p-values, F tests and any other relevant and material test. Determine if model development data, validation data, test data or other data was used for these tests.

Typical p-values greater than 5% are large and should be questioned. Reasonable business judgment can sometimes provide legitimate support for high p-values. Reasonableness of the p-value threshold could also vary depending on the context of the model, e.g., the threshold might be lower when many candidate variables were evaluated for inclusion in the model.

Overall lift charts and/or statistical tests using validation data may not provide enough of the picture. If there is concern about one or more individual variables, the reviewer may obtain, for each discrete variable level, the parameter value, confidence intervals, chi-square tests, p-values and any other relevant and material tests. For variables that are modeled continuously, it may be sufficient to obtain statistics around the modeled parameters; for example, confidence intervals around each level of an AOI curve might be more than what is needed.

Obtain evidence that the model fits the training data well, for individual variables, for any relevant combinations of variables and for, the overall model.

For a GLM, such evidence may be available using chi-square tests, p-values, F tests and/or other means. The steps taken during modeling to achieve goodness-of-fit are likely to be numerous and laborious to describe, but they contribute much of what is generalized about GLM. We should not assume we know what they did and ask "how?" Instead, we should ask what they did and be prepared to ask follow-up questions.

For continuous variables, provide confidence intervals, chi-square tests, p-values and any other relevant and material test. Determine if model development data, validation data, test data or other data was used for these tests.

Typical p-values greater than 5% are large and should be questioned. Reasonable business judgment can sometimes provide legitimate support for high p-values. Reasonableness of the p-value threshold could also vary depending on the context of the model, e.g., the threshold might be lower when many candidate variables were evaluated for inclusion in the model.

Overall lift charts and/or statistical tests using validation data may not provide enough of the picture. If there is concern about one or more individual variables, the reviewer may obtain, for each discrete variable level, the parameter value, confidence intervals, chi-square tests, p-values and any other relevant and material tests. For variables that are modeled continuously, it may be sufficient to obtain statistics around the modeled parameters; for example, confidence intervals around each level of an AOI curve might be more than what is needed.
### Draft: 10/15/2019
As adopted by the Casualty Actuarial and Statistical (C) Task Force on XX/XX/XX

© 2019 National Association of Insurance Commissioners

<table>
<thead>
<tr>
<th>B.4.g</th>
<th>Obtain a description how the model was tested for stability over time.</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Evaluate the build/test/validation datasets for potential time-sensitive model distortions (e.g., a winter storm in year 3 of 5 can distort the model in both the testing and validation datasets). Obsolescence over time is a model risk (e.g., old data for a variable or a variable itself may no longer be relevant). If a model being introduced now is based on losses from years ago, the reviewer should be interested in knowing whether that model would be predictive in the proposed context. Validation using recent data from the proposed context might be requested. Obsolescence is a risk even for a new model based on recent and relevant loss data. The reviewer may want to inquire as to the following: What steps, if any, were taken during modeling to prevent or delay obsolescence? What controls will exist to measure the rate of obsolescence? What is the plan and timeline for updating and ultimately replacing the model? The reviewer should also consider that as newer technologies enter the market (e.g., personal automobile) their impact may change claim activity over time (e.g., lower frequency of loss). So, it is not necessarily a bad thing that the results are not stable over time.</td>
<td></td>
</tr>
<tr>
<td>B.4.h</td>
<td>Obtain a narrative on how potential concerns with overfitting were addressed.</td>
<td>2</td>
</tr>
<tr>
<td>B.4.i</td>
<td>Obtain support demonstrating that the GLM assumptions are appropriate.</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Visual review of plots of actual errors is usually sufficient. The reviewer should look for a conceptual narrative covering these topics: How does this particular GLM work? Why did the rate filer do what it did? Why employ this design instead of alternatives? Why choose this particular distribution function and this particular link function? A company response may be at a fairly high level and reference industry practices. If the reviewer determines that the model makes no assumptions that are considered to be unreasonable, the importance of this item may be reduced.</td>
<td></td>
</tr>
<tr>
<td>B.4.j</td>
<td>Obtain 5-10 sample records with corresponding output from the model for those records.</td>
<td>4</td>
</tr>
</tbody>
</table>

5. “Old Model” Versus “New Model”
**B.5.a**  Obtain an explanation why this model is an improvement to the current rating plan. If it replaces a previous model, find out why it is better than the one it is replacing; determine how the company reached that conclusion and identify metrics relied on in reaching that conclusion. Look for an explanation of any changes in calculations, assumptions, parameters, and data used to build this model from the previous model.

Regulators should expect to see improvement in the new class plan’s predictive ability or other sufficient reason for the change.

**B.5.b**  Determine if two Gini coefficients were compared and obtain a narrative on the conclusion drawn from this comparison.

One example of a comparison might be sufficient. This is relevant when one model is being updated or replaced. Regulators should expect to see improvement in the new class plan’s predictive ability. This information element requests a comparison of Gini coefficient from the prior model to the Gini coefficient of proposed model. It is expected that there should be improvement in the Gini coefficient. A higher Gini coefficient indicates greater differentiation produced by the model and how well the model fits that data. This comparison is not applicable to initial model introduction. Reviewer can look to CAS monograph for information on Gini coefficients.

**B.5.c**  Determine if double lift charts were analyzed and obtain a narrative on the conclusion drawn from this analysis.

One example of a comparison might be sufficient. Note that “not applicable” is an acceptable response.

**B.5.d**  If replacing an existing model, obtain a list of any predictor variables used in the old model that are not used in the new model. Obtain an explanation why these variables were dropped from the new model. Obtain a list of all new predictor variables in the new model that were not in the prior old model.

Useful to differentiate between old and new variables so the regulator can prioritize more time on variables not yet reviewed.

**6. Modeler Software**

**B.6.a**  Request access to SMEs (e.g., modelers) who led the project, compiled the data, built the model, and/or performed peer review.

The filing should contain a contact that can put the regulator in touch with appropriate SMEs and key contributors to the model development to discuss the model.
# 1. General Impact of Model on Rating Algorithm

| C.1.a | In the actuarial memorandum or explanatory memorandum, for each model and sub-model (including external models), look for a narrative that explains each model and its role (how it was used) in the rating system. | 1 | The "role of the model" relates to how the model integrates into the rating plan as a whole and where the effects of the model are manifested within the various components of the rating plan. This is not intended as an overarching statement of the model's goal, but rather a description of how specifically the model is used. This item is particularly important, if the role of the model cannot be immediately discerned by the reviewer from a quick review of the rate and/or rule pages. (Importance is dependent on state requirements and ease of identification by the first layer of review and escalation to the appropriate review staff.) |
| C.1.b | Obtain an explanation of how the model was used to adjust the rating algorithm. | 1 | Models are often used to produce factor-based indications, which are then used as the basis for the selected changes to the rating plan. It is the changes to the rating plan that create impacts. Consider asking for an explanation of how the model was used to adjust the rating algorithm. |
| C.1.c | Obtain a complete list of characteristics/variables used in the proposed rating plan, including those used as input to the model (including sub-models and composite variables) and all other characteristics/variables (not input to the model) used to calculate a premium. For each characteristic/variable, determine if it is only input to the model, whether it is only a separate univariate rating characteristic, or whether it is both input to the model and a separate univariate rating characteristic. The list should include transparent descriptions (in plain language) of each listed characteristic/variable. | 1 | Examples of variables used as inputs to the model and used as separate univariate rating characteristics might be criteria used to determine a rating tier or household composite characteristic. |

## 2. Relevance of Variables and Relationship to Risk of Loss

| C.2.a | Obtain a narrative regarding how the characteristics/rating variables included in the filed rating plan relate to the risk of insurance loss (or expense) for the type of insurance product being priced. | 2 | The narrative should include a discussion of the relevance each characteristic/rating variable has on consumer behavior that would lead to a difference in risk of loss (or expense). The narrative should include a logical and intuitive relationship to cost, and model results should be consistent with the expected direction of the relationship. This explanation would not be needed if the connection between variables and risk of loss (or expense) has already been illustrated. |

## 3. Comparison of Model Outputs to Current and Selected Rating Factors
C.3.a  Compare relativities indicated by the model to both current relativities and the insurer's selected relativities for each risk characteristic/variable in the rating plan.  

1  “Significant difference” may vary based on the risk characteristic/variable and context. However, the movement of a selected relativity should be in the direction of the indicated relativity; if not, an explanation is necessary as to why the movement is logical.

C.3.b  Obtain documentation and support for all calculations, judgments, or adjustments that connect the model's indicated values to the selected values.  

1  The documentation should include explanations for the necessity of any such adjustments and explain each significant difference between the model's indicated values and the selected values. This applies even to models that produce scores, tiers, or ranges of values for which indications can be derived. This information is especially important if differences between model indicated values and selected values are material and/or impact one consumer population more than another.

C.3.c  For each characteristic/variable used as both input to the model (including sub-models and composite variables) and as a separate univariate rating characteristic, obtain a narrative how each characteristic/variable was tempered or adjusted to account for possible overlap or redundancy in what the characteristic/variable measures.  

2  Modeling loss ratio with these characteristics/variables as control variables would account for possible overlap. The insurer should address this possibility or other considerations, e.g., tier placement models often use risk characteristics/variables that are also used elsewhere in the rating plan. One way to do this would be to model the loss ratios resulting from a process that already uses univariate rating variables. Then the model/composite variables would be attempting to explain the residuals.

4. Responses to Data, Credibility and Granularity Issues

C.4.a  Determine what, if any, consideration was given to the credibility of the output data.  

2  At what level of granularity is credibility applied. If modeling was by-coverage, by-form or by-peril, explain how these were handled when there was not enough credible data by coverage, form or peril to model.

C.4.b  If the rating plan is less granular than the model, obtain an explanation why.  

2  This is applicable if the insurer had to combine modeled output in order to reduce the granularity of the rating plan.

C.4.c  If the rating plan is more granular than the model, obtain an explanation why.  

2  A more granular rating plan implies that the insurer had to extrapolate certain rating treatments, especially at the tails of a distribution of attributes, in a manner not specified by the model indications.

5. Definitions of Rating Variables

© 2019 National Association of Insurance Commissioners
<table>
<thead>
<tr>
<th>C.5.a</th>
<th>Obtain a narrative on adjustments made to model output, e.g., transformations, binning and/or categorizations. If adjustments were made, obtain the name of the characteristic/variable and a description of the adjustment.</th>
<th>2</th>
<th>If rating tiers or other intermediate rating categories are created from model output, the rate and/or rule pages should present these rating tiers or categories. The company should provide an explanation how model output was translated into these rating tiers or intermediate rating categories.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.6.a</td>
<td>Obtain aggregated state-specific, book-of-business-specific univariate historical experience data, separately for each year included in the model, consisting of loss ratio or pure premium relativities and the data underlying those calculations for each category of model output(s) proposed to be used within the rating plan. For each data element, obtain an explanation whether it is raw or adjusted and, if the latter, obtain a detailed explanation for the adjustments.</td>
<td>4</td>
<td>For example, were losses developed/undeveloped, trended/untrended, capped/uncapped, etc.? Univariate indications should not necessarily be used to override more sophisticated multivariate indications. However, they do provide additional context and may serve as a useful reference.</td>
</tr>
<tr>
<td>C.6.b</td>
<td>Obtain an explanation of any material (especially directional) differences between model indications and state-specific univariate indications.</td>
<td>4</td>
<td>Multivariate indications may be reasonable as refinements to univariate indications, but possibly not for bringing about significant reversals of those indications. For instance, if the univariate indicated relativity for an attribute is 1.5 and the multivariate indicated relativity is 1.25, this is potentially a plausible application of the multivariate techniques. If, however, the univariate indicated relativity is 0.7 and the multivariate indicated relativity is 1.25, a regulator may question whether the attribute in question is negatively correlated with other determinants of risk. Credibility of state data should be considered when state indications differ from modeled results based on a broader data set. However, the relevance of the broader data set to the risks being priced should also be considered. Borderline reversals are not of as much concern.</td>
</tr>
<tr>
<td>C.7.a</td>
<td>Obtain a listing of the top five rating variables that contribute the most to large swings in premium, both as increases and decreases.</td>
<td>4</td>
<td>These rating variables may represent changes to rating factors, be newly introduced to the rating plan, or have been removed from the rating plan.</td>
</tr>
<tr>
<td>C.7.b</td>
<td>Determine if the insurer performed sensitivity testing to identify significant changes in premium due to small or incremental change in a single risk characteristic. If such testing was performed, obtain a narrative that discusses the testing and provides the results of that testing.</td>
<td>3</td>
<td>One way to see sensitivity is to analyze a graph of each risk characteristic’s/variable’s possible relativities. Look for significant variation between adjacent relativities and evaluate if such variation is reasonable and credible.</td>
</tr>
</tbody>
</table>
For the proposed filing, obtain the impacts on expiring policies and describe the process used by management, if any, to mitigate those impacts.

Some mitigation efforts may substantially weaken the connection between premium and expected loss and expense, and hence may be viewed as unfairly discriminatory by some states.

<table>
<thead>
<tr>
<th>C.7.e</th>
<th>Obtain a rate disruption/dislocation analysis, demonstrating the distribution of percentage and/or dollar impacts on renewal business (created by rerating the current book of business), and sufficient information to explain the disruptions to individual consumers.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The analysis should include the largest dollar and percentage impacts arising from the filing, including the impacts arising specifically from the adoption of the model or changes to the model as they translate into the proposed rating plan.</td>
</tr>
<tr>
<td></td>
<td>While the default request would typically be for the distribution/dislocation of impacts at the overall filing level, the regulator may need to delve into the more granular variable-specific effects of rate changes if there is concern about particular variables having extreme or disproportionate impacts, or significant impacts that have otherwise yet to be substantiated.</td>
</tr>
<tr>
<td></td>
<td>See Appendix C for an example of a disruption analysis.</td>
</tr>
</tbody>
</table>

Obtain exposure distributions for the model's output variables and show the effects of rate changes at granular and summary levels, including the overall impact on the book of business.

See Appendix C for an example of an exposure distribution.
C.7.f  Identify policy characteristics, used as input to a model or sub-model, that remain "static" over a policy's lifetime versus those that will be updated periodically. Obtain a narrative on how the company handles policy characteristics that are listed as "static," yet change over time.

Some examples of "static" policy characteristics are prior carrier tenure, prior carrier type, prior liability limits, claim history over past X years, or lapse of coverage. These are specific policy characteristics usually set at the time new business is written, used to create an insurance score or to place the business in a rating/underwriting tier, and often fixed for the life of the policy. The reviewer should be aware, and possibly concerned, how the company treats an insured over time when the insured’s risk profile based on "static" variables changes over time but the rate charged, based on a new business insurance score or tier assignment, no longer reflect the insured’s true and current risk profile.

A few examples of "non-static" policy characteristics are age of driver, driving record and credit information (FCRA related). These are updated automatically by the company on a periodic basis, usually at renewal, with or without the policyholder explicitly informing the company.

C.7.g  Obtain a means to calculate the rate charged a consumer.

The filed rating plan should contain enough information for a regulator to be able to validate policy premium. However, for a complex model or rating plan, a score or premium calculator via Excel or similar means would be ideal, but this could be elicited on a case-by-case basis. Ability to calculate the rate charged could allow the regulator to perform sensitivity testing when there are small changes to a risk characteristic/variable. Note that this information may be proprietary.

C.7.h  In the filed rating plan, be aware of any non-insurance data used as input to the model (customer-provided or other). In order to respond to consumer inquiries, it may be necessary to inquire as to how consumers can verify their data and correct errors.

If the data is from a third-party source, the company should provide information on the source. Depending on the nature of the data, data should be documented and an overview of who owns it and the topic of consumer verification should be addressed, including how consumers can verify their data and correct errors.

8. Accurate Translation of Model into a Rating Plan

© 2019 National Association of Insurance Commissioners
Obtain sufficient information to understand how the model outputs are used within the rating system and to verify that the rating plan’s manual, in fact, reflects the model output and any adjustments made to the model output.

| C.8.a | | 1 | The regulator can review the rating plan's manual to see that modeled output is properly reflected in the manual's rules, rates, factors, etc. |
The Task Force was charged to propose modifications to the 2016 Product Filing Review Handbook to reflect best practices for the regulatory review of predictive analytics. The following are the titled sections in Chapter Three “The Basics of Property and Casualty Rate Regulation.” Proposed changes are shown as tracked changes.

**Data Adjustments**

...Because the insurance contracts will be written to cover future accident periods, the past data needs to be adjusted to reflect the anticipated future premiums and costs. These adjustments will provide a profit/loss picture if no rate change occurs. Calculations can then be made to determine the overall rate need (or indication). …

**Interaction between Rating Variables (Multivariate Analysis)**

If each rating variable is evaluated separately, statistically significant interactions between rating variables may not be identified and, thus, may not be included in the rating plan. If the pricing of rating variables is evaluated separately for each rating variable, there is potential to miss the interaction between rating variables. Care should be taken to have a multivariate analysis when practical. In some instances, a multivariate analysis is not possible. But, with computing power growing exponentially, insurers believe they have found many ways to improve their operations and competitiveness through use of complex predictive models in all areas of their insurance business.

**Approval of Classification Systems**

With rate changes, companies sometimes propose revisions to their classification system. Because the changes to classification plans can be significant and have large impacts on the consumers’ rates, regulators should focus on these changes.

Some items of proposed classification can sometimes be deemed to be against public policy, such as the use of education or occupation. You should be aware of your state’s laws and regulations regarding which rating factors are allowed, and you should require definitions of all data elements that can affect the charged premium. Finding rating or underwriting characteristics that may violate public policy is becoming more difficult for regulators with the increasing and innovative ways insurers use predictive models.

**Rating Tiers**

Some states allow an insurer to have multiple rate levels, or rating tiers, within a single company. These rating tiers are another way of classifying risks for rating purposes. Typically, there are requirements for rating tiers: the underwriting rules for each tier should be mutually exclusive, clear, and objective; there should be a distinction between the expected losses or expenses for each tier; and the placement process should be auditable. Tiers within a company are mainly seen in personal lines products.
One particular concern with rating tiers would be the analyses of whether a plan produces unfair discrimination. Questions arise around the time-sensitive aspects of the underwriting criteria and any related re-evaluation of the tiers upon renewal. For example, consider two tiers where the insured is placed in the “high” tier because of a lapse of insurance in the prior 12 months. The question is: What happens upon renewal after there has no longer been a lapse of insurance for 12 months? Does the insured get slotted in the “low” tier as he would if he was new business? Some statutes limit the amount of time that violations, loss history, or insurance scores can be used, and some statutes might only allow credit history to be used for re-rating at the policyholder’s request. Regulators should consider the acceptability of differences in rates between existing and new policyholders when they have the same current risk profile.

Insurers also can create different rating levels by having separate companies within a group. While regulators should examine rating tiers within an insurer to a high degree of regulatory scrutiny, there tends to be less scrutiny with differences in rates that exist between affiliated companies. Workers’ compensation insurers are more likely to obtain rating tiers using separate companies.

**Rate Justification: New Products** – (No change is proposed.)

**Predictive Modeling**

The ability of computers to process massive amounts of data has led to the expansion of the use of predictive modeling in insurance ratemaking. Predictive models have enabled insurers to build rating, marketing, underwriting and claim models with significant segmentation predictive power and are increasingly being applied in such areas as claims modeling and used in helping insurers to price risks more effectively.

Key new rating variables that are being incorporated into insurers’ predictive models include homeowners’ home rates by peril, homeowners’home rating by building characteristics, vehicle history, usage-based auto insurance, and credit characteristics.

Data quality within and communication about models are of key importance with predictive modeling. Depending on definitional boundaries, predictive modeling can sometimes overlap with the field of machine learning. In the modeling space, predictive modeling is often referred to as predictive analytics.

Insurers’ use of predictive analytics along with big data has significant potential benefits to both consumers and insurers. Predictive analytics can reveal insights into the relationship between consumer behavior and the cost of insurance, lower the cost of insurance for many, and provide incentives for consumers to better control and mitigate loss. However, predictive analytic techniques are evolving rapidly and leaving many regulators without the necessary tools to effectively review insurers’ use of predictive models in insurance applications. To aid the regulator in the review of predictive models, best practices have been developed for generalized linear models or “GLMs”). GLMs are commonly used in personal automobile and home insurance applications.

The term “predictive model” refers to a set of models that use statistics to predict outcomes. Then applied to insurance, the model is chosen to estimate the probability or expected value of an outcome given a set amount of input data; for example, models can predict the frequency of loss, the severity of loss, or the pure premium.

To further complicate regulatory review of models in the future, modeling technology and methods are evolving rapidly. GLMs are relatively transparent and their output and consequences are much clearer than many other complex models. But as computing power grows exponentially, it is opening up the modeling world to more sophisticated forms of data acquisition and data analysis. Insurance actuaries and data scientists seek increased predictiveness by using even more complex predictive modeling methods. Examples of these are predictive models utilizing logistic regression, K-nearest neighbor classification, random forests, decision trees, neural networks, or combinations of available modeling methods (often referred to as ensembles). These evolving techniques will make the regulators’ understanding and oversight of filed rating plans even more challenging.

**A. Generalized Linear Models**

The generalized linear model (GLM) is a commonly used predictive model in insurance applications, particularly in building an insurance product’s rating plan. Because of this and the fact most Property and Casualty regulators are most concerned...
about personal lines, NAIC has developed a white paper for guidance in reviewing GLMs for personal automobile and home insurance.

Before GLMs became vogue, rating plans were built using univariate methods. Univariate methods were considered easy to understand and easy to demonstrate the relationship to costs (loss and/or expense). However, many consider univariate methods too simplistic since they do not take into account the interaction (or dependencies) of the selected input variables. GLMs introduce significant improvements over univariate-based rating plans by automatically adjusting for correlations among input variables. Today, the majority of predictive models used in personal automobile and home insurance rating plans are GLMs. But, GLM results are not always easy to understand and the relationship to costs may be difficult to explain.

A GLM consists of three elements:

- A target variable, Y, which is a random variable that is independent and follows a probability distribution from the exponential family, defined by a selected variance function and dispersion parameter.
- A linear predictor $\eta = X\beta$.
- A link function $g$ such that $E(Y) = \mu = g^{-1}(\eta)$.

As can be seen in the description of the three GLM components above, it may take more than a casual introduction to statistics to comprehend the construction of a GLM. As stated earlier, a downside to GLMs is that it is more challenging to interpret the GLMs output than with univariate models.

B. Credibility of GLM Output

If the underlying data is not credible no model will improve that credibility, and segmentation methods could make credibility worse. GLM software provides point estimates and allows the modeler to consider standard errors and confidence intervals. GLMs effectively assume that the underlying datasets are 100% credible no matter their size. If some segments have little data, the resulting uncertainty would not be reflected in the GLM parameter estimates themselves (although it might be reflected in the standard errors, confidence intervals, etc.). Even though the process of selecting relativities often includes adjusting the raw GLM output, the resultant selections are not typically credibility-weighted with any complement of credibility. New footnotes: "This is not always true. Sometimes insurers do review complements of credibility and further weight the GLM output with those complements. While this may not be a standard practice today, new techniques could result in this becoming more standard in the future." And "GLMs provide confidence intervals; credibility methods do not. There are techniques such as penalized regression that blend credibility with a GLM and improve a model's ability to generalize." Nevertheless, selected relativities based on GLM model output may differ from GLM point estimates.

Because of this presumption in credibility, which may or may not be valid in practice, the modeler and the regulator reviewing the model would need to engage in thoughtful consideration when incorporating GLM output into a rating plan to ensure that model predictiveness is not compromised by any lack of actual credibility. Therefore, to mitigate the risk that model credibility or predictiveness is lacking, a complete filing for a rating plan that incorporates GLM output should include validation evidence for the rating plan, not just the statistical model.

C. What is a "Best Practice"?

A best practice is a form of program evaluation in public policy. At its most basic level, a practice is a “tangible and visible behavior… [based on] an idea about how the actions…will solve a problem or achieve a goal”14. Best practices can maintain quality as an alternative to mandatory legislated standards and can be based on self-assessment or benchmarking.15

Therefore, a best practice represents an effective method of problem solving. The "problem" regulators want to solve is probably better posed as seeking an answer to this question: How can regulators determine that predictive models, as used in rate filings, are compliant with state laws and regulations? However, best practices are not intended to create standards for filings that include predictive models.

13 Refer to NAIC’s white paper titled Regulatory Review of Predictive Models, found at the NAIC website.
Best practices are based on the following principles that promote a comprehensive and coordinated review of predictive models across states:

- State insurance regulators will maintain their current rate regulatory authority.
- State insurance regulators will be able to share information to aid companies in getting insurance products to market more quickly.
- State insurance regulators will share expertise and discuss technical issues regarding predictive models.
- State insurance regulators will maintain confidentiality, where appropriate, regarding predictive models.

D. Regulatory Review of Predictive Models

The legal and regulatory constraints (including state variations) are likely to be more evolved, and challenging, for personal automobile and home insurance. Through review of these personal lines, the knowledge needed to review predictive models and guidance are more straightforward. When the review involves GLMs applied to other lines of business, modeling depends on context, so the GLM reviewer has to be alert for data challenges and business applications that differ from the more familiar personal lines. For example, compared to personal lines, modeling rates in commercial lines is more likely to encounter low volumes of historical data, dependence on advisory loss costs, unique large accounts with large deductibles, and package products that create policies from numerous line-of-business and coverage building blocks.

Commercial lines commonly use individual risk modifications following experience, judgment, and/or expense considerations. A regulator may never see how models impact commercial excess and surplus lines filings.

Guidance offered here and in the NAIC’s white paper might be useful (with deeper adaptations) when starting to review different types of predictive models. If the model is not a GLM, however, some of the GLM guidance might not apply. For example, not all predictive models generate p-values or F tests. Depending on the model type under review, other considerations might be important that were not as important in the review of a GLM. Also, when transferring GLM guidance to other lines of business, unique considerations may arise depending on the context in which a predictive model is proposed to be deployed, the uses to which it is proposed to be put, and the potential consequences for the insurer, its customers and its competitors. This guidance does not delve into these possible considerations, but regulators should be prepared to address them as they arise.

Best practices will help the regulator understand if a predictive model is cost based, if the predictive model is compliant with state law, and how the model improves the company’s rating plan. Best practices can also increase the consistency among the regulatory review processes used across states and improve the efficiency of each regulator’s review thereby assisting companies in getting their products to market faster. With this in mind, the regulator’s review of predictive models should:

1. Ensure that the selected rating factors based on the model or other analysis produce rates that are not excessive, inadequate, or unfairly discriminatory.
   a. Review the overall rate level impact of the proposed revisions to rate level indications provided by the filer.
   b. Review the premium disruption for individual policyholders and how the disruptions can be explained to individual consumers.
   c. Review the individual input characteristics to and output factors from the predictive model (and its sub-models), as well as, associated selected relativities to ensure they are not unfairly discriminatory.

2. Obtain a clear understanding of how the data used to build and validate the model, and thoroughly review all other aspects of the model, including assumptions, adjustments, variables, submodels used as input, and resulting output.
   a. Determine that individual input characteristics to a predictive model and their resulting rating factors are related to the expected loss or expense differences in risk.
   b. Determine that the data used as input to the predictive model is accurate, including a clear understanding how missing values, erroneous values and outliers are handled.
   c. Determine that any adjustments to the raw data are handled appropriately, including but not limited to, trending, development, capping, removal of catastrophes.
   d. Obtain a clear understanding of how often each risk characteristic, used as input to the model, is updated and whether the model is periodically rerun, so model output reflects changes to non-static risk characteristics.
   e. Obtain a clear understanding of how the selected predictive model was built.
   f. Determine whether internal and external data used in relation to the model is compatible with practices allowed in the jurisdiction and do not reflect characteristics prohibited in the state.

3. Evaluate how the model interacts with and improves the rating plan.
a. Obtain a clear understanding of the characteristics that are input to a predictive model (and its sub-models), their relationship to each other and their relationship to non-modeled characteristics/variables used to calculate a risk’s premium.
b. Obtain a clear understanding why the insurer believes this type of model works in an insurance risk application.
c. Obtain a clear understanding of how model output interacts with non-modeled characteristics/variables used to calculate a risk’s premium.
d. Obtain a clear understanding of how the predictive model was integrated into the insurer’s state rating plan and how it improves that plan.
e. For predictive model refreshes, determine whether sufficient validation was performed to ensure the model is still a good fit.

4. Enable competition and innovation to promote the growth, financial stability, and efficiency of the insurance marketplace.
a. Enable innovation in the pricing of insurance through acceptance of predictive models, provided they are actuarially sound and in compliance with state laws.
b. Review predictive models in a timely manner to enable reasonable speed to market.

e. Information Needed to Follow Best Practices

To assist the regulator in following each best practice, the Casualty Actuarial and Statistical Task Force created a white paper titled Regulatory Review of Predictive Models. The paper contains a list of information elements and considerations that should be useful during the review of a model underlying a rating plan. To further assist the regulator, the information elements were mapped into the best practices listed above in Section XV of the paper.

Note that, in the white paper, CASTF focused on the GLM since it is a commonly used predictive model in insurance applications, particularly in building an insurance product’s rating plan. Combined with the fact most property and casualty regulators are very concerned about personal lines, the white paper is focused on the review of GLMs for personal auto mobile and home insurance rate making applications.

Guidance offered here might be useful (with deeper adaptations) when starting to review different types of predictive models. If the model is not a GLM, some listed items might not apply. For example, not all predictive models generate p-values or F-tests. Depending on the model type, other considerations might be important. When transferring guidance to other lines of business and other types of model, unique considerations may arise depending on the context in which a predictive model is proposed to be deployed, the uses to which it is proposed to be put, and the potential consequences for the insurer, its customers and its competitors. This paper does not delve into these possible considerations, but regulators should be prepared to address them as they arise.

F. Confidentiality

Regulatory reviewers are required to protect confidential information in accordance with applicable State law. However, insurers should be aware that a rate filing might become part of the public record. Each state determines the confidentiality of a rate filing, supplemental material to the filing, when filing information might become public, the procedure to request that filing information be held confidentially, and the procedure by which a public records request is made. It is incumbent on an insurer to be familiar with each state’s laws regarding the confidentiality of information submitted with their rate filing.

Advisory Organizations – (No change is proposed.)

Workers’ Compensation Special Rules – (No change is proposed.)

Premium Selection Decisions

- Indicated Rate Change vs. Selected Rate Change

After applying credibility, the indicated rate change should reflect the company’s best estimate of their premium needs given their current or expected book of business. However, insurance companies also have other business considerations including competition, marketing, legal concerns, impact of the rate change on retention, etc. A company might wish to deviate from their indicated rate change and should justify those decisions, within the constraints of the law.

© 2019 National Association of Insurance Commissioners
Capping and Transition Rules

With advances in technology, it is possible for companies to introduce capping of rates on individual policies with an aim toward gradually increasing policyholders' rates, rather than making large modifications all at one time. Similarly, premiums are often proposed to be modified when an insurer acquires another company’s book of business or decides to move from or to an advisory organization’s plan. These types of proposed capping are sometimes called “renewal premium capping,” “rate capping,” “rate stability program,” or “transition rules.”

Transition rules for individual policyholders can get quite complex and you need to be aware of your state’s positions on premium capping rules. Any premium capping and transition rules require weighing the pros and cons of the potential for unfair discrimination (with some customers not paying the rate commensurate with the risks they have) vs. rate stability for existing policyholders.

If premium capping or transition rules are allowed, additional decisions will need to be made:

- Which rates should get capped?
- Do rate decreases get capped? If so, what is the impact if the policyholder asks to be quoted as new business?
- Do all rate increases get capped or only above a certain percentage?
- How much time will lapse or how many renewal cycles will occur before the new rates are in place or different rating plans are merged?
- Should the insured be told what the final premium will be once no more capping is applied?
- What is the impact on their rate capping?
- How many rate-capping rules can be implemented at any given time?

When premium capping or transition rules have been incorporated, future indicated rate changes and rating factor analyses need to properly reflect the fully approved rate changes. If the overall approved rate change was +10%, yet capping resulted in only 8% being implemented in the first year, the remaining amount to recognize the full 10% should be reflected in the premium on-level adjustment. Otherwise, the indicated rate would be redundant.

Some states encourage more frequent filing of rate changes that can help to avoid the need of premium capping and transition rules. Some states might prefer capping of individual rating variables, rather than capping for individual policyholders.

Installment Plans – (No change is proposed.)

Policy Fees – (No change is proposed.)

Potential Questions to Ask Oneself as a Regulator

Every filing will be different and will result in different regulatory analyses. But the following are some questions the regulator might ask oneself in a rate filing review:

1. Regarding data:
   a. Is the data submitted with the filing enough information for a regulatory review?
   b. Is the number of years of experience appropriate?
   c. Did the company sufficiently analyze and control their quality of data?

2. Regarding the support and justification of rates:
   a. Did they propose rate changes without justification?
   b. Are proposals based on judgment or competitive analysis? If so, are the results reasonable and acceptable? Are there inappropriate marketing practices?
   c. Are the assumptions (loss development, trend, expense load, profit provision, credibility etc.) used to develop the rate indication appropriate? Are they supported with data and are deviations from data results sufficiently explained?
d. Is the weighting of data by year (or credibility) properly justified or does it appear random?
   • Is there more weight being placed on data in one year solely because it produces a higher indicated rate change?
   • If there are two indications being weighted together and one is for a rate increase and one is a rate decrease, is the weighting justified?

e. Is there satisfactory explanation about why a proposed rate change deviates from the indicated rate change?

3. Regarding differences in assumptions from previous filings:
   a. Have methodologies changed significantly?
   b. Are assumptions for the weighting of years or credibility significantly different? Or does there appear to be some manipulation to the rate indication?

4. Is there unfair discrimination?
   a. Do classifications comply with state requirements?
   b. Are proposed rates established so that different classes will produce the same underwriting results?
   c. If predictive models are used in the rating plan, are there concerns related to input variables that are prohibited or proxies for prohibited variables?

5. What do you need to communicate?
   a. Can you explain why you are taking a specific action on the filing?
   b. What do you need to tell the Consumer Services Department?
      • Can you explain the impact of the rate change on current business? How big is the company and how much of the market is impacted?
      • What are the biggest changes in the filing (and the ones on which consumer calls might be expected)?
      • What is the maximum rate change impact on any one policyholder?

Questions to Ask a Company

If you remain unsatisfied that the company has satisfactorily justified the rate change, then consider asking additional questions of the company. Questions should be asked of the company when they have not satisfied statutory or regulatory requirements in the state or when any current justification is inadequate and could have an impact on the rate change approval or the amount of the approval.

If there are additional items of concern, the company can be notified so they will make appropriate modifications in future filings.

The CASTF white paper, Regulatory Review of Predictive Models, documents questions that a regular may want to ask when reviewing a model. These questions are listed in the Predictive Model – Information for Regulatory Review section of the white paper. Note that although the white paper focuses on GLMs for personal automobile and home insurance, some of the concepts may be transferable to other types of models and other lines of business.
Additional Ratemaking Information

The Casualty Actuarial Society (CAS) and the Society of Actuaries (SOA) have extensive examination syllabi that contain a significant amount of ratemaking information, on both the basic topics covered in this chapter and on advanced ratemaking topics. The CAS and SOA websites contain links to many of the papers included in the syllabus. Recommended reading is the Foundations of Casualty Actuarial Science, which contains chapters on ratemaking, risk classification, and individual risk rating.

Other Reading

Some additional background reading is recommended:

  - Chapter 1: Introduction
  - Chapter 3: Ratemaking
  - Chapter 6: Risk Classification
  - Chapter 9: Investment Issues in Property-Liability Insurance
  - Chapter 10: Only the section on Regulating an Insurance Company, pp. 777–787
- Casualty Actuarial Society (CAS) Statements of Principles, especially regarding property and casualty ratemaking.
- Association of Insurance Compliance Professionals: “Ratemaking—What the State Filer Needs to Know.”
- Review of filings and approval of insurance company rates.

Summary

Rate regulation for property/casualty lines of business requires significant knowledge of state rating laws, rating standards, actuarial science, statistical modeling and many data concepts.

- Rating laws vary by state, but the rating laws are usually grouped into prior approval, file and use or use and file (competitive), no file (open competition), and flex rating.
- Rate standards typically included in the state rating laws require that “Rates shall not be inadequate, excessive, or unfairly discriminatory.”
- A company will likely determine their indicated rate change by starting with historical years of underwriting data (earned premiums, incurred loss and loss adjustment expenses, general expenses) and adjusting that data to reflect the anticipated ultimate level of costs for the future time period covered by the policies. Numerous adjustments are made to the data. Common premium adjustments are on-level premium, audit, and trend. Common loss adjustments are trend, loss development, Catastrophe/large loss provisions, and an adjusting and other (A&O) loss adjustment expense provision. A profit/contingency provision is also calculated to determine the indicated rate change.
- Once an overall rate level is determined, the rate change gets allocated to the classifications and other rating factors.
- Individual risk rating allows manual rates to be modified by an individual policyholder’s own experience.
- Advisory organizations provide the underlying loss costs for companies to be able to add their own expenses and profit provisions (with loss cost multipliers) to calculate their insurance rates.
- Casualty Actuarial Society’s Statement of Principles Regarding Property and Casualty Insurance Ratemaking provides guidance and guidelines for the numerous actuarial decisions and standards employed during the development of rates.
- NAIC model laws also include special provisions for workers’ compensation business, penalties for not complying with laws, and competitive market analysis to determine whether rates should be subject to prior approval provisions.
- Best practices for reviewing predictive models are provided in the CASTF white paper titled Regulatory Review of Predictive Models. Although the white paper focuses on GLMs for personal automobile and home insurance, some of the concepts may be transferrable to other types of models and other lines of insurance.

© 2019 National Association of Insurance Commissioners
While this chapter provides an overview of the rate determination/actuarial process and regulatory review, state statutory or administrative rule may require the examiner to adopt different standards or guidelines than the ones described.

No additional changes are proposed to the Product Filing Review Handbook.
IX. PROPOSED STATE GUIDANCE

This paper acknowledges that different states will apply the guidance within it differently, based on variations in the legal environment pertaining to insurance regulation in those states, as well as the extent of available resources, including staff members with actuarial and/or statistical expertise, the workloads of those staff members, and the time that can be reasonably allocated to predictive-model reviews. States with prior-approval authority over personal-lines rate filings often already require answers in connection with many of the information elements expressed in this paper. However, states – including those with and without prior-approval authority – may also use the guidance in this paper to choose which model elements to focus on in their reviews and/or to train new reviewers, as well as to gain an enhanced understanding of how predictive models are developed, supported, and deployed in their markets. Ultimately, the insurance regulators within each state will decide how best to tailor the guidance within this paper to achieve the most effective and successful implementation, subject to the framework of statutes, regulations, precedents, and processes that comprise the insurance regulatory framework in that state.

X. OTHER CONSIDERATIONS

During the development of this guidance, topics arose that are not addressed in this paper. These topics may need addressing during the regulator’s review of a predictive model. A few of these topics may be discussed elsewhere by the NAIC as either technical or policy matters. All of these topics should probably be addressed by each state on a case-by-case basis. Below is a listing of topics that CASTF thought might be important for future discussion and consideration but is beyond the scope of this paper, as well as CASTF’s current charges.

- Discuss when rating variables or rating plans become too granular.
  - The granularity of data refers to the size in which data fields are sub-divided. For example, data could be at the state level or could be subdivided into county or further into zip code or even census tracks. Insurers were instituting data warehouse initiatives that greatly improved the granularity and accessibility of data that could be analyzed for ratemaking purposes. So, despite the fact that sophisticated statistical techniques existed much earlier than this, it was the circumstances of enhanced computing power and better data that enabled their usage in classification ratemaking. Perhaps the most important trigger in the widespread adoption of multivariate methods was competitive pressure. When one or more companies implement improved classification ratemaking, they gain a competitive advantage and put the rest of the industry in a position of adverse selection and decreased profitability. [footnote: Basic Ratemaking, Fifth Edition, May 2016; Geoff Werner, FCAS, MAAA and Claudine Modlin, FCAS, MAAA]
  - The science of classification requires balancing two objectives: grouping risks into a sufficient number of levels to ensure the risks within each group are homogeneous while being careful not to create too many granularly defined groups that may lead to instability in the estimated costs. [footnote: Basic Ratemaking, Fifth Edition, May 2016; Geoff Werner, FCAS, MAAA and Claudine Modlin, FCAS, MAAA]
  - Concern has been expressed that when fields are sub-divided too finely, model results may be less reliable. It is commonly assumed that the more data you have, the better. But, the more granular the data, the harder it may be to see the forest for the trees. More granular data used as input to predictive models may make it easier to measure short-term effects, but it can make it harder to measure long-term effects because of more noise in the data. However, more granular data may make anomalies in the data more apparent and make it easier to scrub the data.
  - Therefore, it may be of value to provide guidance around granularity, such as: When are rating variables or rating plans too granular? How is granularity handled during the development of the model or during the selection of rate relativities?
- Discuss the scientific mindset of open inquiry and its relevance to the best practice white paper.
  - This white paper has taken the position that regulatory actuaries, especially when they review predictive models, are in a prime position to be the torchbearers for the scientific approach by maintaining the commitment to open but rigorous, systematic, and principled inquiry.
  - This white paper does not prescribe any specific answers regarding which treatments are to be considered logical or rational. Such answers cannot be arrived at without considering the context of a given jurisdiction’s laws, marketplace, and the specific nature of insurers’ proposals. Therefore, to preempt any arguments by some interested parties that the paper may prescribe specific solutions or restrictions – it clearly is not.
  - As actuaries, if regulators are to practice the discipline called “actuarial science,” it is incumbent upon us to adopt the proper scientific mindset of open inquiry – where no questions are off limits and continued...
systematic exploration and progress are the hallmarks of the scientific approach. Any insistence that certain questions must not be asked, or certain concepts must not be explored, entails a departure from the realm of science into the realm of dogma. If pursued, it would limit the role of regulators and quickly deprive them of broader relevance.

- Discuss correlation vs causality in general and in relation to Actuarial Standard of Practice (ASOP) 12.
  - There were many criticisms during each exposure of this white paper that this paper goes beyond the requirement of Actuarial Standard of Practice #12 and establishes a new standard for the company’s actuaries. There is no question that the issues at hand need to be explored further by states collectively (through NAIC) or on a case-by-case state basis. What a state does with the results of a discussion of rational or logical connections between particular attributes and the risk of insurance loss is subject to the framework of statutes, regulations, precedents, and processes that comprise the insurance regulatory framework in that state.
  - The very act of discussion of the rational, logical, or plausible relationships of individual risk attributes to the risk of insurance loss – and all related implications, such as perception by consumers, legislators, and media; philosophical considerations of fairness; interactions with public policy as determined by the relevant policymaking bodies; and relevance to the evolution of the insurance industry, consumer products, and overall impacts on the incentives and opportunities available to consumers, is crucial to engage in and continue to do so for as long as new predictive models are being developed, new variables are being introduced, and consumer premiums as well as insurer underwriting decisions are being affected. In other words, the discussion needs to continue indefinitely in a variety of venues and evolve along with the industry and the broader society. We, as insurance professionals, cannot insulate ourselves from participation in the conceptual discourse.
  - This white paper, in general, establishes that a rating/model variable should not only be correlated to expected costs but that there should be a rational explanation as to why the correlation exists. While it is difficult to prove causation, and such a proof is not a standard against which rate filings are evaluated in any jurisdiction, there is an immense difference of both degree and kind between proving causation and discussing a rational or logical connection between a particular variable and the risk of insurance loss. It is a non sequitur to assert that the lack of requirement for the former (proof) confers immunity upon insurers in regard to the latter (discussion and expression of plausibility).
  - Discussion of the Actuarial Standards of Practice has been consciously excluded from this paper for a number of reasons. Firstly, only actuaries are obligated to adhere to the ASOPs. Models are created by, supported by, and filed by professionals who are often not actuaries, e.g., data scientists, modelers, and other professions, who are not bound by ASOPs. Secondly, ASOPs do not supersede state laws. Thirdly, ASOPs represent a floor, not a ceiling, for regulatory actuaries who also need to consider state laws and public policy concerns.
  - Finally, ASOP 12 specifically is often misquoted by filers who cite the phrase, in Section 3.2.2 of ASOP 12, that states “it is not necessary for the actuary to establish a cause and effect relationship between the risk characteristic and expected outcome in order to use a specific risk characteristic” while omitting the leading phrase “while the actuary should select risk characteristics that are related to expected outcomes.”

- Discussion of data mining being in conflict with standard scientific model and increase in “false positives.”
  - Throughout this white paper, the regulator asks the modeler to go beyond correlation and document their basic, causal understanding of how variables used in a model or rating plan are related to risk. A correlation alone is not the final arbiter of the validity of findings, but causal understanding can be employed to assess which correlations may be entirely due to chance, what are non-causal relationships, and which are most likely to be enduring causal relationships. Though this white paper does not delve deeply into how these relationships can be identified and documented, the paper does ask the modeler to provide their understanding of these relationships. The future consideration is whether the regulator should take a deeper dive into the causal relationships of variables used in a model or rating plan.
  - The American Statistical Association (ASA) expressed some degree of alarm at approaches similar to data mining (Wasserstein and Lazer, 2016). In a formal statement of the ASA, the association warned against a purely “cookbook” approach to statistics: “… a p-value near .05 taken by itself offers only weak evidence of the null hypothesis” (page 129). Lastly, the ASA warned strongly against an over reliance on data mining: “Cherry-picking promising findings, also known by such terms as data dredging, significance chasing… and "p-hacking," leads to a spurious excess of statistically significant results … and should be vigorously avoided” (page 131).
A problem that will increase significantly with the adoption of data mining techniques and the increasing availability of very large data sets that dwarf anything available even just a decade ago is that data mining will dramatically increase the rate of "false positives" - the technique will inevitably churn up numerous associations between variables that are simply random, non-meaningful correlations resulting purely from chance. The apparent disregard of causality that seems common among practitioners of data mining techniques will significantly magnify the problem. Causality forms the basis of the standard model of all natural and social sciences. Evaluations of models should consider the nature of observed relationships within the context of prior substantive knowledge.

- Regulators are often responding to consumer inquiries regarding how a policy premium is calculated and why the premium, or change in premium, is so high.
  - The white paper identified the following best practices:
    - 1.b. Review the premium disruption for individual policyholders and how the disruptions can be explained to individual consumers. ...and information elements that may assist the regulator's and consumer's understanding of the premium being charged.
    - C.2.a. Provide an explanation how the characteristics/rating variables, included in the filed rating plan, logically and intuitively relate to the risk of insurance loss (or expense) for the type of insurance product being priced. Include a discussion of the relevance each characteristic/rating variable has on consumer behavior that would lead to a difference in risk of loss (or expense).
    - C.7.f. Explain how the insurer will help educate consumers to mitigate their risk.
    - C.7.h. Identify sources to be used at "point of sale" to place individual risks within the matrix of rating system classifications. How can a consumer verify their own "point-of-sale" data and correct any errors?
    - C.7.j. Provide the regulator with a description of how the company will respond to consumers' inquiries about how their premium was calculated.
  - The main challenge to consumers is lack of transparency: trying to understand the data and analytics being used to determine their eligibility for products and the price they are being charged. It may not be clear to the consumer how they are being underwritten or what behaviors they can modify or steps they can take to get a better rate. A potential issue with pricing based on predictive analytics is that it can lead to more granular pricing, which benefits some consumers but not others. This broader distributed range of prices could be perceived as unfair. Privacy issues are also a concern for consumers because of a lack of transparency regarding how data is collected and used. [footnote: Big Data and the Role of the Actuary, American Academy of Actuaries, Big Data Task Force, June 2018.]
  - Though regulators may inquire about the above information elements, they often deal with consumers directly on topics such as the following:
    - Determine the extent the model causes premium disruption for individual policyholders, and how the insurer will explain the disruption to individual consumers that inquire about it.
    - Explain how the consumer can mitigate their insurance risk.
    - Assist a consumer in verifying their "point-of-sale" data.
    - Determine the means available to a consumer to correct or contest individual data input values that may be in error.
    - Assist the consumer in understanding how often each risk characteristics (used as input to the model or in the rating plan) is updated or if the risk characteristic is static.
    - 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.
  - As a future consideration, NAIC or a state may want to explore, with insurers, how to improve communications with the consumer on these topics.

- Discuss guidelines for insurers’ handling of consumer-generated data in insurance transactions.
  - Does a consumer have the right to know what data is being used to determine the consumers' premium, where that data came from, and how the consumer can address errors in the data? To what extent is the insurer accountable for the quality of the data used to calculate a consumer's premium, whether that data is internal or external to the insurer's operations? To what extent should the insurer inform the consumer (transparency) and when should the insurer inform the consumer how their premium is calculated? If the consumer is properly informed, the consumer may make physical and behavioral changes to lower their risk, and
subsequently their premium. "This issue deals with consumers' ownership and control of the data they create through interactions with the insurer or devices provided by or monitored by the insurer as well as the permissible uses of those data by insurers." [Center for Economic Justice, comments to the NAIC Accelerated Underwriting (A) Working Group, September 29, 2019]

- Discuss the development of new tools and techniques for monitoring consumer market outcomes resulting from insurers' use of Big Data analytics in property and casualty rating plans.
  o "While regulators have historically pursued consumer protection by reviewing insurers' forms and rates on the front end, the variety and volume of new data sources and complexity of algorithms require a revision to the historical regulatory approach. Consumer protection in an era of Big Data analytics requires regulators to collect and analyze granular data on actual consumer market outcomes. This is necessary not only because comprehensive review on the front end is likely no longer possible, but also because actual market outcomes may differ dramatically from intended or purported market outcomes. Stated differently, it is no longer sufficient (if it ever was) to rely on a front-end assessment of a data source or algorithm to ensure fair consumer treatment and the absence of unfair discrimination. Routine analysis of actual consumer market outcomes is needed. It is also completely feasible today." [footnote: Center for Economic Justice, comments to the NAIC Accelerated Underwriting (A) Working Group, September 29, 2019]

- Discuss revision to model laws regarding advisory organizations.
  o Organizations not licensed or supervised as advisory organizations are engaging in precisely the same type of activities as licensed advisory organizations – collecting data from insurers, analyzing the data and combining it with other data and information, and producing collective pricing and claim settlement recommendations in the form of algorithms. The vendors of algorithms are providing the same type of guidance as the archetype of advisory organizations, the Insurance Services Office, by producing loss cost recommendations. To ensure that data brokers and vendors of algorithms who are engaged in advisory organization activities are properly licensed and supervised, advisory organization model laws could be revised. [Center for Economic Justice, comments to the NAIC Accelerated Underwriting (A) Working Group, September 29, 2019]

- Discuss the need for NAIC to update and strengthen information-sharing platforms and protocols.
- Discuss paper topic beyond GLMs and personal automobile and home insurance applications.
  o The scope of this white paper was narrowed to GLMs as used in personal automobile and home insurance rating applications. Many commenters expressed concern that the paper's scope is too narrow. NAIC may want to expand these best practices or create new best practices for other lines of business, other insurance applications (other than personal automobile and home insurance filings), and other types of models.
APPENDIX A – BEST PRACTICE DEVELOPMENT

Best-practices development is a method for reviewing public policy processes that have been effective in addressing particular issues and could be applied to a current problem. This process relies on the assumptions that top performance is a result of good practices and these practices may be adapted and emulated by others to improve results.16

The term “best practice” can be a misleading one due to the slippery nature of the word “best”. When proceeding with policy research of this kind, it may be more helpful to frame the project as a way of identifying practices or processes that have worked exceptionally well and the underlying reasons for their success. This allows for a mix-and-match approach for making recommendations that might encompass pieces of many good practices17.

Researchers have found that successful best-practice analysis projects share five common phases:

A. **Scope**
   The focus of an effective analysis is narrow, precise and clearly articulated to stakeholders. A project with a broader focus becomes unwieldy and impractical. Furthermore, Bardach urges the importance of realistic expectations in order to avoid improperly attributing results to a best practice without taking into account internal validity problems.

B. **Identify Top Performers**
   Identify outstanding performers in this area to partner with and learn from. In this phase, it is key to recall that a best practice is a tangible behavior or process designed to solve a problem or achieve a goal (i.e. reviewing predictive models contributes to insurance rates that are not unfairly discriminatory). Therefore, top performers are those who are particularly effective at solving a specific problem or regularly achieve desired results in the area of focus.

C. **Analyze Best Practices**
   Once successful practices are identified, analysts will begin to observe, gather information and identify the distinctive elements that contribute to their superior performance. Bardach suggests it is important at this stage to distill the successful elements of the process down to their most essential idea. This allows for flexibility once the practice is adapted for a new organization or location.

D. **Adapt**
   Analyze and adapt the core elements of the practice for application in a new environment. This may require changing some aspects to account for organizational or environmental differences while retaining the foundational concept or idea. This is also the time to identify potential vulnerabilities of the new practice and build in safeguards to minimize risk.

E. **Implementation and Evaluation**
   The final step is to implement the new process and carefully monitor the results. It may be necessary to make adjustments, so it is likely prudent to allow time and resources for this. Once implementation is complete, continued evaluation is important to ensure the practice remains effective.

---


APPENDIX B - GLOSSARY OF TERMS

Adjusting Data – Adjusting data is when the modeler makes any change to the raw data. For example, capping losses, on-leveling, binning, transformation of the data, etc. This includes scrubbing of the data.

Aggregated Data - Aggregated data is straight from the insurer's data banks without modification (e.g., not scrubbed, transformed). Aggregated datasets are those compiled prior to data selection and model building.

Composite Characteristic - A composite characteristic is an individual risk characteristic used to create a composite variable.

Composite Score - A composite score is a number arrived at through the combination of multiple variables by means of a sequence of mathematical steps - for example, a credit-based insurance scoring model.

Composite Variable - A composite variable is a variable created by combining two or more individual risk characteristics of the insured into a single variable.

Continuous Variable - A continuous variable is a numeric variable that represents a measurement on a continuous scale. Examples include age, amount of insurance (in dollars), and population density. https://www.casact.org/pubs/monographs/papers/05-Goldburd-Khare-Tevet.pdf

Control Variable - Control variables are variables whose relativities are not used in the final rating algorithm but are included when building the model. They are included in the model so that other correlated variables do not pick up their signal. For example, state and year are frequently included in models as control variables so that the different experiences and distributions between states and across time do not influence the rating factors used in the final rating algorithm. [11]

Correlation Matrix - A correlation matrix is a table showing correlation coefficients between sets of variables. Each random variable (X) in the table is correlated with each of the other values in the table (X). This allows you to see which pairs have the highest correlation. Below is a correlation matrix showing correlation coefficients for combinations of 5 variables B1:B5. The diagonal of the table is always a set of ones, because the correlation between a variable and itself is always 1. You could fill in the upper-right triangle, but these would be a repeat of the lower-left triangle (because B1:B2 is the same as B2:B1). In other words, a correlation matrix is also a symmetric matrix. [xx]

<table>
<thead>
<tr>
<th></th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>B5</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>1</td>
<td>0.53</td>
<td>0.73</td>
<td>0.67</td>
<td>0.43</td>
</tr>
<tr>
<td>B2</td>
<td>0.53</td>
<td>1</td>
<td>0.44</td>
<td>0.36</td>
<td>0.71</td>
</tr>
<tr>
<td>B3</td>
<td>0.73</td>
<td>0.44</td>
<td>1</td>
<td>0.72</td>
<td>0.56</td>
</tr>
<tr>
<td>B4</td>
<td>0.67</td>
<td>0.36</td>
<td>0.72</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>B5</td>
<td>0.43</td>
<td>0.71</td>
<td>0.56</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
Data Dredging - Data dredging is also referred to as data fishing, data snooping, data butchery, and p-hacking. It is the misuse of data analysis to find patterns in data that can be presented as statistically significant when, in fact, there is no real underlying effect. This is done by performing many statistical tests on the data and only paying attention to those that come back with significant results, instead of stating a single hypothesis about an underlying effect before the analysis and then conducting a single test for it.

The process of data dredging involves automatically testing huge numbers of hypotheses about a single data set by exhaustively searching—perhaps for combinations of variables that might show a correlation, and perhaps for groups of cases or observations that show differences in their mean or in their breakdown by some other variable.

Conventional tests of statistical significance are based on the probability that a particular result would arise if chance alone were at work, and necessarily accept some risk of mistaken conclusions of a certain type (mistaken rejections of the null hypothesis). This level of risk is called the significance. When large numbers of tests are performed, some produce false results of this type, hence 5% of randomly chosen hypotheses turn out to be significant at the 5% level, 1% turn out to be significant at the 1% significance level, and so on, by chance alone. When enough hypotheses are tested, it is virtually certain that some will be statistically significant but misleading, since almost every data set with any degree of randomness is likely to contain (for example) some spurious correlations. If they are not cautious, researchers using data mining techniques can be easily misled by these results.

The multiple comparisons hazard is common in data dredging. Moreover, subgroups are sometimes explored without alerting the reader to the number of questions at issue, which can lead to misinformed conclusions.[fnzz]

Data Source - A data source is the original repository of the information used to build the model. For example, information from internal insurance data, an application, a vendor, credit bureaus, government websites, a sub-model, verbal information provided to agents, external sources, consumer information databases, etc.

Discrete Variable - A discrete variable is a variable that can only take on a countable number of values. Examples include number of claims, marital status, and gender.

Discrete Variable Level - Discrete variables are generally referred to as "factors" (not to be confused with rating factors), with values that each factor can take being referred to as "levels".

Double-Lift Chart - Double lift charts are similar to simple quantile plots, but rather than sorting based on the predicted loss cost of each model, the double lift chart sorts based on the ratio of the two models' predicted loss costs. Double lift charts directly compare the results of two models.[12]

Exponential Family - The exponential family is a class of distributions that have certain properties that are used in fitting GLMs. It includes many well-known distributions, such as the Normal, Poisson, Gamma, Tweedie, and Binomial distributions. [13]

Fair Credit Reporting Act – The Fair Credit Reporting Act (FCRA), 15 U.S.C. § 1681 (FCRA) is U.S. Federal Government legislation enacted to promote the accuracy, fairness and privacy of consumer information contained in the files of consumer reporting agencies. It was intended to protect consumers from the willful and/or negligent inclusion of inaccurate information in their credit reports. To that end, the FCRA regulates the collection, dissemination and use of consumer information, including consumer credit information.[14] Together with the Fair Debt Collection Practices Act (FDCPA), the FCRA forms the foundation of consumer rights law in the United States. It was originally passed in 1970 and is enforced by the US Federal Trade Commission, the Consumer Financial Protection Bureau and private litigants.

© 2019 National Association of Insurance Commissioners
Generalized Linear Model - Generalized linear models (GLMs) are a means of modeling the relationship between a variable whose outcome we wish to predict and one or more explanatory variables. The predicted variable is called the target variable and is denoted \( y \). In property/casualty insurance ratemaking applications, the target variable is typically one of the following:

- Claim count (or claims per exposure)
- Claim severity (i.e., dollars of loss per claim or occurrence)
- Pure premium (i.e., dollars of loss per exposure)
- Loss ratio (i.e., dollars of loss per dollar of premium)

For quantitative target variables such as those above, the GLM will produce an estimate of the expected value of the outcome. For other applications, the target variable may be the occurrence or non-occurrence of a certain event. Examples include:

- Whether or not a policyholder will renew his/her policy.
- Whether a submitted claim contains fraud.

For such variables, a GLM can be applied to estimate the probability that the event will occur.

The explanatory variables, or predictors, are denoted \( x_1, \ldots, x_p \), where \( p \) is the number of predictors in the model. Potential predictors are typically any policy term or policyholder characteristic that an insurer may wish to include in a rating plan. Some examples are:

- Type of vehicle, age, or marital status for personal auto insurance.
- Construction type, building age, or amount of insurance (AOI) for home insurance. [15]

Geodemographic - Geodemographic segmentation (or analysis) is a multivariate statistical classification technique for discovering whether the individuals of a population fall into different groups by making quantitative comparisons of multiple characteristics with the assumption that the differences within any group should be less than the differences between groups. Geodemographic segmentation is based on two principles:

1. People who live in the same neighborhood are more likely to have similar characteristics than are two people chosen at random.
2. Neighborhoods can be categorized in terms of the characteristics of the population that they contain. Any two neighborhoods can be placed in the same category, i.e., they contain similar types of people, even though they are widely separated.
Granularity of Data - The granularity of data refers to the size in which data fields are sub-divided.

For example, a postal address can be recorded, with coarse granularity, as a single field:
- address = 200 2nd Ave. South #358, St. Petersburg, FL 33701-4313 USA

Or, with fine granularity, as multiple fields:
- street address = 200 2nd Ave. South #358
- city = St. Petersburg
- state = FL
- postal code = 33701-4313
- country = USA

Or, even finer granularity:
- street = 2nd Ave. South
- address number = 200
- suite/apartment number = #358
- city = St. Petersburg
- state = FL
- postal-code = 33701
- postal-code-add-on = 4313
- country = USA

Home Insurance – Home insurance covers damage to the property, contents, and outstanding structures (if applicable), as well as loss of use, liability and medical coverage. The perils covered, and amount of insurance provided are detailed in the policy contract.

Insurance Data - Data collected by the insurance company.

Interaction Term - Two predictor variables are said to interact if the effect of one of the predictors on the target variable depends on the level of the other. Suppose that predictor variables $X_1$ and $X_2$ interact. A GLM modeler could account for this interaction by including an interaction term of the form $X_1X_2$ in the formula for the linear predictor. For instance, rather than defining the linear predictor as $\eta = \beta_0 + \beta_1X_1 + \beta_2X_2$, they could set $\eta = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_1X_2$.

The following two plots of modeled personal auto bodily injury pure premium by age and gender illustrate this effect. The plots are based on two otherwise identical log-link GLMs, built using the same fictional dataset, with the only difference between the two being that the second model includes the Age-Gender interaction term while the first does not. Notice that the male curve in the first plot is a constant multiple of the female curve, while in the second plot the ratios of the male to female values differ from age to age.

Lift Chart - See definition of quantile plot.
Linear Predictor - A linear predictor is the linear combination of explanatory variables \( (X_1, X_2, \ldots, X_k) \) in the model, e.g., \( \beta_0 + \beta_1X_1 + \beta_2X_2. \) [18]

Link Function - The link function, \( \eta \) or \( g(\mu) \), specifies the link between random and systematic components. It describes how the expected value of the response relates to the linear predictor of explanatory variables; e.g., \( \eta = g(E(Y_i)) = E(Y_i) \) for linear regression, or \( \eta = \logit(\pi) \) for logistic regression. [19]

Missing data - Missing data occurs when some records contain blanks or "Not Available" or "Null" where variable values should be.

Non-Insurance Data - Non-insurance data is data provided by another party other than the insurance company.

Offset Variable – Offset variables (or factors) are model variables with a known or pre-specified coefficient. Their relativities are included in the model and the final rating algorithm, but they are generated from other studies outside the multivariate analysis and fixed (not allowed to change) in the model when it is run. Examples of offset variables include limit and deductible relativities that are more appropriately derived via loss elimination analysis. The resulting relativities are then included in the multivariate model as offsets. Another example is using an offset factor to account for the exposure in the records; this does not get included in the final rating algorithm. [20]

Overfitting – Overfitting is the production of an analysis that corresponds too closely or exactly to a particular set of data and may, therefore, fail to fit additional data or predict future observation reliably. [21]

PCA Approach (Principal Component Analysis) – The PCA method creates multiple new variables from correlated groups of predictors. Those new variables exhibit little or no correlation between them—thereby making them potentially more useful in a GLM. A PCA in a filing can be described as “a GLM within a GLM.” One of the more common applications of PCA is geodemographic analysis, where many attributes are used to modify territorial differentials on, for example, a census block level.

Personal Automobile Insurance – Personal automobile insurance is insurance for privately owned motor vehicles and trailers for use on public roads not owned or used for commercial purposes. This includes personal auto combinations of private passenger auto, motorcycle, financial responsibility bonds, recreational vehicles and/or other personal auto. Policies include any combination of coverage such as the following: auto liability, personal injury protection (PIP), medical payments (MP), uninsured/underinsured motorist (UM/UIM); specified causes of loss, comprehensive, and collision. [22]

Post-model Adjustment - Post-model adjustment is any adjustment made to the output of the model including but not limited to adjusting rating factors or removal of variables.

Probability Distribution – A probability distribution is a statistical function that describes all the possible values and likelihoods that a random variable can take within a given range. The chosen probability distribution is supposed to best represent the likely outcomes.

Proxy Variable - A proxy variable is any characteristic that indirectly captures the effect of another characteristic whether or not that characteristic is used in the insurer’s rating plan.
Quantile Plot - A quantile plot is a visual representation of a model's ability to accurately differentiate between the best and the worst risks. Data is sorted by predicted value from smallest to largest, the data is then bucketed into quantiles with the same volume of exposures, within each bucket calculate the average predicted value and the average actual value. Plot for each quantile the actual and the predicted value. The first quantile contains the risks that the model predicts have the best experience and the last quantile contains the risks predicted to have the worst experience. The plot shows three things: how well the model predicts actual values by quantile, the predicted value should be increasing as the quantile increases, and the lift of the model, the difference between the first and last quantile, the larger it indicates the model’s ability to distinguish between the best and worst risk.[23] An example follows:

Rating Algorithm – A rating algorithm is the mathematical or computational component of the rating plan used to calculate an insured’s premiums.

Rating Category - A rating category is the same as a rating characteristic can be quantitative or qualitative.

Rating Characteristic - A rating characteristic is a specific risk criterion of the insured used to define the level of the rating variable that applies to the insured. Ex. Rating variable- Driver age, Rating characteristic- Age 42

Rating Factor – A rating factor is the numerical component included in the rate pages of the rating plan’s manual. Rating factors are used together with the rating algorithm to calculate the insured’s premiums.

Rating Plan – The rating plan describes in detail how to combine the various components in the rules and rate pages to calculate the overall premium charged for any risk that is not specifically pre-printed in a rate table. The rating plan is very specific and includes explicit instructions, such as:

- the order in which rating variables should be considered;
- how the effect of rating variables is applied in the calculation of premium (e.g., multiplicative, additive, or some unique mathematical expression);
- the existence of maximum and minimum premiums (or in some cases the maximum discount or surcharge that can be applied);
- specifics associated with any rounding that takes place.

If the insurance product contains multiple coverages, then separate rating plans by coverage may apply.[24]

Rating System - The rating system is the insurance company’s IT infrastructure that produces the rates derived from the rating algorithm.
Rating Tier - A rating tier is rating based on a combination of rating characteristics rather than a single rating characteristic resulting in a separation of groups of insureds into different rate levels within the same or separate companies. Often, rating tiers are used to differentiate quality of risk, e.g., substandard, standard, or preferred.

Rating Treatment - Rating treatment is the manner in which an aspect of the rating affects an insured's premium.

Rating Variable - A rating variable is a risk criterion of the insured used to modify the base rate in a rating algorithm.


Raw Data - Raw data is data before scrubbing, transformation etc. takes place when received from a source.

Sample Record - A sample record is one line of data from a data source including all variables. For example:

<table>
<thead>
<tr>
<th>Rating</th>
<th>ZIP</th>
<th>Gender</th>
<th>Type</th>
<th>Age</th>
<th>ZipZipCode</th>
<th>FootLeak</th>
<th>Roof</th>
<th>RoomLeak</th>
<th>Exterior</th>
<th>NeighBor</th>
<th>Cells</th>
<th>Living</th>
<th>NeighBorLeak</th>
<th>Shelf</th>
<th>Elev</th>
<th>FireAssist</th>
<th>Term</th>
<th>Term</th>
<th>ScrubbedData</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>04234</td>
<td>garage</td>
<td>basement</td>
<td>25700</td>
<td>25100</td>
<td>25100</td>
<td>25300</td>
<td>25300</td>
<td>25300</td>
<td>25300</td>
<td>25300</td>
<td>25300</td>
<td>25300</td>
<td>25300</td>
<td>25300</td>
<td>25300</td>
<td>25300</td>
<td>25300</td>
<td></td>
</tr>
</tbody>
</table>

Scrubbed Data - Scrubbed data is data reviewed for errors, where "N/A" has been replaced with a value, and where most transformations have been performed. Data that has been "scrubbed" is now in a useable format to begin building the model.

Scrubbing Data - Scrubbing is the process of editing, amending, or removing data in a dataset that is incorrect, incomplete, improperly formatted, or duplicated.

SME - Subject Matter Expert.

Sub-Model - A sub-model is any model that provides input into another model.

Transformation - A transformation is a change to a variable by taking a function of that variable, for example, when age's value is replaced by the value (age)^2. The result is called a transformation variable.

Voluntarily Reported Data - Voluntarily reported data is data directly obtained by a company from a consumer. Examples would be data taken directly of an application for insurance or obtained verbally by a company representative.

Univariate Model – A univariate model is a model that only has one independent variable.
APPENDIX C – SAMPLE RATE-DISRUPTION TEMPLATE

<table>
<thead>
<tr>
<th>State Division of Insurance - EXAMPLE for Rate Disruption</th>
<th>Template Updated October 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>● First, fill in the boxes for minimum and maximum individual impacts, shaded in light blue. Default values in the cells are examples only.</td>
<td></td>
</tr>
<tr>
<td>● The appropriate percent-change ranges will then be generated based on the maximum/minimum changes.</td>
<td></td>
</tr>
<tr>
<td>● For every box shaded in light green, replace “ENTER VALUE” with the number of affected insureds within the corresponding change range.</td>
<td></td>
</tr>
<tr>
<td>● Once all values are filled in, use the “Charts” feature in Excel to generate a histogram to visually display the spread of impacts.</td>
<td></td>
</tr>
</tbody>
</table>

**NOTE:** Values of Minimum % Change, Maximum % Change, and Total Number of Insureds must reconcile to the Rate/Rule Schedule in SERFF.

<table>
<thead>
<tr>
<th>Uncapped</th>
<th>Capped (If Applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Minimum % Change</strong></td>
<td><strong>Minimum % Change</strong></td>
</tr>
<tr>
<td>-30.000%</td>
<td>-15.000%</td>
</tr>
<tr>
<td><strong>Maximum % Change</strong></td>
<td><strong>Maximum % Change</strong></td>
</tr>
<tr>
<td>30.000%</td>
<td>15.000%</td>
</tr>
<tr>
<td><strong>Total Number of Insureds</strong></td>
<td><strong>Total Number of Insureds</strong></td>
</tr>
<tr>
<td>(Auto-Calculated)</td>
<td>(Auto-Calculated)</td>
</tr>
<tr>
<td>1994</td>
<td>1994</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Uncapped Rate Disruption</th>
<th>Capped Rate Disruption (If Applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Percent-Change Range</strong></td>
<td><strong>Number of Insureds in Range</strong></td>
</tr>
<tr>
<td>-30% to &lt;-25%</td>
<td>2</td>
</tr>
<tr>
<td>-25% to &lt;-20%</td>
<td>90</td>
</tr>
<tr>
<td>-20% to &lt;-15%</td>
<td>130</td>
</tr>
<tr>
<td>-15% to &lt;-10%</td>
<td>340</td>
</tr>
<tr>
<td>-10% to &lt;-5%</td>
<td>245</td>
</tr>
<tr>
<td>-5% to &lt;0%</td>
<td>12</td>
</tr>
<tr>
<td>Exactly 0%</td>
<td>12</td>
</tr>
<tr>
<td>&gt;0% to &lt;5%</td>
<td>150</td>
</tr>
<tr>
<td>5% to &lt;10%</td>
<td>160</td>
</tr>
<tr>
<td>10% to &lt;15%</td>
<td>401</td>
</tr>
<tr>
<td>15% to &lt;20%</td>
<td>201</td>
</tr>
</tbody>
</table>

**EXAMPLE Uncapped Rate Disruption**

![Graph showing number of insureds in different percent-change ranges](https://via.placeholder.com/150)

© 2019 National Association of Insurance Commissioners
State Division of Insurance - EXAMPLE for Largest Percentage Increase

- Fill in fields highlighted in light green. Fields highlighted in red are imported from the Template for Rate Disruption.

### Largest Percentage Increase

<table>
<thead>
<tr>
<th>Uncapped Change</th>
<th>30.00%</th>
<th>Uncapped Dollar Change</th>
<th>$165.00</th>
<th>Current Premium</th>
<th>$500.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capped Change (If Applicable)</td>
<td>15.00%</td>
<td>Capped $ Change (If Applicable)</td>
<td>$82.50</td>
<td>Proposed Premium</td>
<td>$632.50</td>
</tr>
</tbody>
</table>

### Characteristics of Policy (Fill in Below)

- For Auto Insurance: At minimum, identify the age and gender of each named insured, limits by coverage, territory, make / model of vehicle(s), prior accident / violation history, and any other key attributes whose treatments are affected by this filing.
- For Home Insurance: At minimum, identify age and gender of each named insured, amount of insurance, territory, construction type, protection class, any prior loss history, and any other key attributes whose treatments are affected by this filing.

### Automobile Policy: Three Insureds - Male (Age 54), Female (Age 49), and Male (Age 25). Territory: Las Vegas, Zip Code 89105.

- No prior accidents, 1 prior speeding conviction for 25-year-old male. Policy receives EFT discount and loyalty discount.

### Most Significant Impacts to This Policy (Identify attributes - e.g., base-rate change or changes to individual rating variables)

**NOTE:** If capping is proposed to apply for this policy, include the impact of capping at the end, after displaying uncapped impacts by attribute. Add rows as needed. Total percent and dollar impacts should reconcile to the values presented above in this exhibit.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>% Impact (Uncapped)</th>
<th>Dollar Impact (Uncapped)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insured Age (M/25)</td>
<td>12.00%</td>
<td>$66.00</td>
</tr>
<tr>
<td>COMP Deductible</td>
<td>10.00%</td>
<td>$61.60</td>
</tr>
<tr>
<td>Territory (89105)</td>
<td>4.00%</td>
<td>$27.10</td>
</tr>
<tr>
<td>Vehicle Symbol (2003 Honda Accord)</td>
<td>1.46%</td>
<td>$10.29</td>
</tr>
<tr>
<td>Effect of Capping</td>
<td>-1.54%</td>
<td>-$82.50</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>15.00%</strong></td>
<td><strong>$82.50</strong></td>
</tr>
</tbody>
</table>

What lengths of policy terms does the insurer offer in this book of business?

- [ ] 12-Month Policies
- [ ] 6-Month Policies
- [ ] 3-Month Policies
- [ ] Other (Specify)

© 2019 National Association of Insurance Commissioners
### APPENDIX D – INFORMATION NEEDED BY REGULATOR MAPPED INTO BEST PRACTICES

**TBD**

### APPENDIX E – REFERENCES


**DRAFTING NOTE 10/15/19: WILL NEED TO CORRECT ALL FOOTNOTES AND INSERT REFERENCES. THE FOLLOWING IS SAVED HERE UNTIL MOVED TO THE CORRECT LOCATION.**

To see that this second definition accounts for the interaction, note that it is equivalent to $\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2$ and to $\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2$, with $\beta_1 = \beta_1 + \beta_3 X_2$ and $\beta_2 = \beta_2 + \beta_3 X_1$. Since $\beta_1$ is a function of $X_2$ and $\beta_2$ is a function of $X_1$, these two equivalences say that the effect of $X_1$ depends on the level of $X_2$ and vice versa.

REFERENCES:
November 21, 2019

Kris DeFrain, FCAS, MAAA, CPCU
Director of Research and Actuarial Services
National Association of Insurance Commissioners (NAIC) Central Office

Via Email

Re: CASTF Regulatory Review of Predictive Models White Paper

Dear Kris,

As the American Academy of Actuaries\textsuperscript{1} senior property/casualty fellow, I appreciate this opportunity to comment further on the Casualty Actuarial and Statistical Task Force (CASTF) draft white paper discussing best practices for the Regulatory Review of Predictive Models (RRPM). My comments herein relate to the discussion draft released on October 15, 2019.

Throughout the RRPM paper, the philosophical benefits of predictive analytics and big data are well documented. Additionally, the paper describes well the challenges inherent in reviewing the models and regulating rates resulting from these models. The American Academy of Actuaries remains committed to effective actuarial practice in this area. It was our great pleasure to once again host a day-long session on predictive modeling at the NAIC Insurance Summit this past June. I will also point out anew that in 2018, the Academy produced a monograph, \textit{Big Data and the Role of the Actuary}, which includes extensive sections on regulatory and professionalism considerations. My comments here will be brief, and it is my hope that they will be helpful to you and the CASTF members.

First, I will focus on a technical point. In Section VI, we see that the wording has been modified to emphasize the concept that input characteristics and rating factors are related to the expected loss or expense differences in risk. I strongly agree with this shift from the earlier version. However, I note that in A.4.b and B.3.d, the wording of “rational relationship” and “rational explanation” are less precise. I was not certain if this was intentional or an oversight. The current Section VI wording would seem to lead to better actuarial practice.

\textsuperscript{1} The American Academy of Actuaries is a 19,500+ member professional association whose mission is to serve the public and the U.S. actuarial profession. For more than 50 years, the Academy has assisted public policymakers on all levels by providing leadership, objective expertise, and actuarial advice on risk and financial security issues. The Academy also sets qualification, practice, and professionalism standards for actuaries in the United States.
My second point is more philosophical than technical. It is one that I have raised before, and I suspect that you have received similar feedback from others. In short, the new requirements of the RRPM process have the potential to become unwieldly. Rate filers and regulatory reviewers will have to perform considerably more detailed work as a result. I certainly note throughout the document that you reference the fact that the RRPM structure is meant as guidance to regulators and is not binding. I also note considerable conciliatory language throughout. Knowing many of the CASTF members personally, I have no doubt that this is genuine.

Finally, I note that you have somewhat scaled back the list of potential request items and in some cases reduced the Level of Importance. All of this notwithstanding, one can hope that the near-term learning curve that will result will be brief and that delays will not be onerous.

In closing, I wish to reiterate that the American Academy of Actuaries remains committed to working with CASTF on this matter.

If you have any questions about these comments, contact me (gibson@actuary.org) or Marc Rosenberg, senior casualty policy analyst, at 202-785-7865 or rosenberg@actuary.org.

Sincerely,

Richard Gibson, MAAA, FCAS
Senior Casualty Fellow
American Academy of Actuaries
November 22, 2019

Kris DeFrain, FCAS, MAAA, CPCU
Director, Research and Actuarial Services
National Association of Insurance Commissioners (NAIC)
NAIC Central Office
1100 Walnut Street, Suite 1500 Kansas City, MO
64106-2197

Sent via e-mail at kdefrain@naic.org


The American Property Casualty Insurance Association (APCIA)\(^1\) appreciates the opportunity to provide comment on the NAIC Casualty Actuarial and Statistical Task Force (CASTF) exposure draft, dated October 15, 2019, regarding the *Regulatory Review of Predictive Models*.

The APCIA remains committed to working collaboratively with the Task Force in support of innovation and the effort to leverage the advancements in technology and data analytics to effectively respond to the changing risks and needs of our insurance consumers. The APCIA believes that development of best practices regarding the regulatory review of predictive models can foster beneficial upfront dialogue between the filing company and regulator that supports an efficient and effective review appropriately focused on ensuring compliance with applicable regulatory rating standards. However, the APCIA cautions against developing best practices that could create new standards or establish information elements that extend the statutory scope of the rate review process.

The following outlines our priority items of interest in the October 15, version of the white paper:

1. Section VI. *Guidance for Regulatory Review of Predictive Models (Best Practices)*
   a. Best Practices #1: Remove the newly added language "or other analysis". These Best Practices relate to the review of a predictive model used to inform how an insured's premium is determined. It is unnecessary to broaden the scope beyond predictive models to include "other analysis".
   b. Best Practices #3: The regulator should "Consider whether the model is an update to or resolves a model submitted within a previously approved filing or, is completely new to the rating plan." This can help the reviewer save time by using information from the previous review of the

\(^1\) Representing nearly 60 percent of the U.S. property casualty insurance market, the American Property Casualty Insurance Association (APCIA) promotes and protects the viability of private competition for the benefit of consumers and insurers. APCIA represents the broadest cross-section of home, auto, and business insurers of any national trade association. APCIA members represent all sizes, structures, and regions, protecting families, communities, and businesses in the U.S. and across the globe.
predictive model.

c. Opening paragraph and Best Practices #3: Remove "and improves." "Improve" applies subjectivity and may result in different interpretations from different stakeholders (regulators, consumers, agents, insurers). The Best Practices should be agnostic to how the new or resolved predictive model impacts the rating plan. After the predictive model is reviewed, then the state's DOI may determine if the predictive model "improved" the rating plan relative to their regulatory responsibilities.

d. Best Practice 4a: revise "actuarially sound" to "consistent with actuarial standards of practice, other applicable professional standards." Otherwise, we suggest that CASTF define the term "actuarially sound" in the glossary. As stressed in previous APCIA feedback, reference to applicable actuarial standards and principles should give the regulator greater comfort and knowledge of the general professional practices that guide the elements of the actuary's filed rating plan, including the relevant aspects and use of any predictive model.

2. Section VII - Comments on specific Information Elements

a. C.1.c - Revise this information element to include a list of the relevant characteristics/variables from the associated predictive model. A predictive model may only impact one part of the rating plan, for example solving for expenses as opposed to indemnity costs. It will be extremely burdensome and unnecessarily exhaustive to both the company producing the filing and the DOI reviewing the filing to discuss all variables used in any capacity to produce a rating plan. A relevant list of variables will aid in reviewing the filed predictive model for prohibited variables and appropriate modeling procedures, which is consistent with the goals of these Best Practices. A review of the information element as stated is a review of the entire rating plan, consistent with the responsibilities of a DOI filing reviewer, but beyond the scope of this White Paper.

b. B.4.b, B.4.c, B.4.d - P-values, Statistical Significance, and all listed statistics should not be required. For example, P-values are not always used to develop a model. The APCIA suggests rewording these information elements to align with the Best Practice #2a, specifically determining how the resulting rating factors (or coefficients) are related to the differences in risk. The insurer can then provide a narrative and results for their process of following this Best Practice. Otherwise, this is an exercise where an insurer will provide p-values that are not used in developing a model. A reviewer at the DOI will question high p-values, but the modelers will not have an adequate response since p-values were not reviewed. The information elements as written are unduly prescribing how to develop a predictive model.

3. Section X Other Considerations

a. Discuss Correlation vs. Causality. This discussion is establishing a newly defined standard in reviewing a predictive model. However, the discussion of correlation vs. causality is not limited to the variables/characteristics used in a predictive model. Consistency should be applied to all variables/characteristics used in any capacity in a rating plan. The review of a predictive model should have similar Best Practices with those for reviewing a rating plan when the requested information (causation) is not unique to a predictive model. Rational explanations for the predictiveness of a variable are subjective and each regulator, actuary, data scientist, etc. has his or her own opinion. It is understood that regulators have the authority to impose standards of a rating plan to ensure that state laws are followed, e.g. avoiding unfair discrimination and specifically unaccepted variables. However, the APCIA encourages each DOI that may
implement or revise this subjective standard to determine and communicate their position relative to the specific statutes and regulations of their respective state. Clear communication on how companies should discuss causation vs. correlation and how a DOI will evaluate the discussion will provide companies with the understanding of how to build a rating plan for that state.

b. C.2.a. The APCIA reiterates our objection to guidance that suggests all characteristics and rating variables can be isolated and related to the risk of insurance loss in a manner that is logical and intuitive to any regulator or consumer regardless of their background or expertise. We agree that an insurer should be able to show that there is statistically significant correlation between a predictive variable and loss. However, an intuitive explanation is not proof of causation. An intuitive explanation may be illusory. Instead, the focus should be on identifying variables that are unfairly discriminatory. So long as a variable is not shown to be unfairly discriminatory, as that standard is currently and in past applied, its use should be permitted without requirement of an intuitive explanation. Insurers could be required to attest to the fact its variables are not unfairly discriminatory. A regulator could then object to use of a variable that the regulator can demonstrate as unfairly discriminatory.

The APCIA believes that for this white paper to meet its purpose of providing effective and practical regulatory guidance to improve the quality of predictive model reviews across states and aid speed to market and the competitiveness of the state marketplace, measures should be taken prior to adoption to demonstrate its efficacy for regulators and the industry. To that end, the APCIA strongly suggests “field testing” the Best Practices in Section VI using the Information Elements for Regulatory Review in Section VII. This idea of field testing was raised by a regulator member of the Task Force during its October 15, conference call.

Why is Field Testing needed? The CASTF White Paper drafting group stated, “We believe that there is a misunderstanding between the terms "best practices" and "information elements" that have been identified in this paper. Many comments appear to interpret "information elements" to mean "best practices" and as such have concerns. We believe the concerns raised in this and other similar comments is with the "information elements" that regulators may find helpful when applying the "best practices."” However, the White Paper does not clarify which Best Practices are supported by each information element. Therefore, DOI’s are required to understand all information elements to determine which to use when implementing these Best Practices in reviewing a predictive model filing. The White Paper would benefit from field testing that could provide more practical guidance in applying these valuable Best Practices.

What does it achieve? The CASTF members can review true examples from companies in providing information elements to satisfy the Best Practices. The examples can be discussed by those with knowledge of predictive modeling to educate and give guidance on adequate responses provided by a company or responses that require additional regulator questions.

Why does it need to occur before the final adoption of the White Paper? Field testing may identify needed revisions to the White Paper before it is finalized and implemented as information elements may provide too much or too little information to evaluate the Best Practices. If the CASTF does not leverage field testing, these practical learnings will occur individually in each state as the Best Practices are applied. Clarity and consistency of implementation across many states will aid
regulators because it increases the ability to discuss findings with each other or leverage potential NAIC assistance in reviewing predictive models. Clarity in how states will implement these Best Practices will aid companies in preparing documentation of a predictive model. The documentation often occurs during or immediately after solving the model. Companies are better able to provide filing requirements if known ahead of time because model documentation may be months before the model is implemented in a state’s rating plan and provided to the DOI in a filing.

Thank you again for the opportunity to comment. We look forward to working with the Task Force to achieve a solution that benefits regulators, insurers and ultimately our consumers.

****

Respectfully Submitted,

David Kodama, Jr.
Assistant Vice President, Research & Policy Analysis
From California DOI

Draft: 10/15/2019
As adopted by the Casualty Actuarial and Statistical (C) Task Force on XX/XX/XX

EXPOSURE NOTE: The drafting group considered comments submitted based on the 5/14/19 draft of the bulk of the paper and comments submitted on the 7/24/19 draft of Sections VIII “Proposed Changes to the Product Filing Review Handbook” and IX “Proposed State Guidance” (exposed 8/3/19). Please submit comments to Kris DeFrain (kdefrain@naic.org) on this 10/14/19 draft by Nov. ____, 2019.

Casualty Actuarial and Statistical (C) Task Force
Regulatory Review of Predictive Models

Table of Contents
I. Introduction ....................................................................................................................................................................... 2
II. What is a “Best Practice?” ................................................................................................................................................ 2
III. Do Regulators Need Best Practices to Review Predictive Models? ........................................................................ 3
IV. Scope ................................................................................................................................................................................. 5
V. Confidentiality .................................................................................................................................................................. 5
VI. Guidance for Regulatory Review of Predictive Models (Best Practices) ............................................................... 6
VII. Predictive Models – Information for Regulatory Review .......................................................................................... 7
VIII. Proposed Changes to the Product Filing Review Handbook .................................................................................. 27
IX. Proposed State Guidance ............................................................................................................................................... 35
X. Other Considerations ....................................................................................................................................................... 35
XI. Recommendations Going Forward ............................................................................................................................ 39
XII. Appendix A – Best Practice Development .............................................................................................................. 40
XIII. Appendix B – Glossary of Terms .............................................................................................................................. 41
XIV. Appendix C – Sample Rate-Disruption Template ................................................................................................... 50
XV. Appendix D – Information Needed by Regulator Mapped into Best Practices .......................................................... 52
XVI. Appendix E – References ............................................................................................................................................ 52

© 2019 National Association of Insurance Commissioners
I. INTRODUCTION

Insurers’ use of predictive analytics along with big data has significant potential benefits to both consumers and insurers. Predictive analytics can reveal insights into the relationship between consumer behavior and the cost of insurance, lower the cost of insurance for many, and provide incentives for consumers to better control and mitigate loss. However, predictive analytic techniques are evolving rapidly and leaving many regulators without the necessary tools to effectively review insurers’ use of predictive models in insurance applications.

When a rate plan is truly innovative, the insurer must anticipate or imagine the reviewers’ interests because reviewers will respond with unanticipated questions and have unique educational needs. Insurers can learn from the questions, teach the reviewers, and so forth. When that back-and-forth learning is memorialized and retained, filing requirements and insurer presentations can be routinely organized to meet or exceed reviewers’ needs and expectations. Hopefully, this paper helps bring more consistency and to the art of reviewing predictive models within a rate filing.

The Casualty Actuarial and Statistical (C) Task Force (CASTF) has been charged with identifying best practices to serve as a guide to state insurance departments in their review of predictive models underlying rating plans. There were two charges given to CASTF by the Property and Casualty Insurance (C) Committee at the request of the Big Data (EX) Working Group:

A. Draft and propose changes to the Product Filing Review Handbook to include best practices for review of predictive models and analytics filed by insurers to justify rates.
B. Draft and propose state guidance (e.g., information, data) for rate filings that are based on complex predictive models.

This paper will identify best practices when reviewing predictive models and analytics filed by insurers with regulators to justify rates and provide state guidance for review of rate filings based on predictive models. Upon adoption of this paper by the Executive (EX) Committee and Plenary, the Task Force will evaluate how to incorporate these best practices into the Product Filing Review Handbook and will recommend such changes to the Speed to Market (EX) Working Group.

II. WHAT IS A “BEST PRACTICE?”

A best practice is a form of program evaluation in public policy. At its most basic level, a practice is a “tangible and visible behavior…[based on] an idea about how the actions…will solve a problem or achieve a goal”2. Best practices are used to maintain quality as an alternative to mandatory legislated standards and can be based on self-assessment or benchmarking.3 Therefore, a best practice represents an effective method of problem solving. The "problem" regulators want to solve is probably better posed as seeking an answer to this question: How can regulators determine that predictive models, as used in rate filings, are compliant with state laws and regulations?

Key Regulatory Principles

In this paper, best practices are based on the following principles that promote a comprehensive and coordinated review of predictive models across states:

1. State insurance regulators will maintain their current rate regulatory authority.
2. State insurance regulators will be able to share information to aid companies in getting insurance products to market more quickly.
3. State insurance regulators will share expertise and discuss technical issues regarding predictive models.
4. State insurance regulators will maintain confidentiality, where appropriate, regarding predictive models.

In this paper, best practices are presented in the form of guidance to regulators who review predictive models and to insurance companies filing rating plans that incorporate predictive models. Guidance will identify specific information
useful to a regulator in the review of a predictive model, comment on what might be important about that information and, where appropriate, provide insight as to when the information might identify an issue the regulator needs to be aware of or explore further.

III. DO REGULATORS NEED BEST PRACTICES TO REVIEW PREDICTIVE MODELS?

The term “predictive model” refers to a set of models that use statistics to predict outcomes. When applied to insurance, the model is chosen to estimate the probability or expected value of an outcome given a set amount of input data; for example, models can predict the frequency of loss, the severity of loss, or the pure premium. The generalized linear model (GLM) is a commonly used predictive model in insurance applications, particularly in building an insurance product’s rating plan.

Depending on definitional boundaries, predictive modeling can sometimes overlap with the field of machine learning. In this modeling space, predictive modeling is often referred to as predictive analytics.

Before GLMs became vogue, rating plans were built using univariate methods. Univariate methods were considered intuitive and easy to demonstrate the relationship to costs (loss and/or expense). Today, many insurers consider univariate methods too simplistic since they do not take into account the interaction (or dependencies) of the selected input variables.

According to many in the insurance industry, GLMs introduce significant improvements over univariate-based rating plans by automatically adjusting for correlations among input variables. Today, the majority of predictive models used in private passenger personal automobile and homeowners home rating plans are GLMs. However, GLM results are not always intuitive, and the relationship to costs may be difficult to explain. This is a primary reason regulators can benefit from best practices.

A GLM consists of three elements:

- A target variable, each component of Y, which is a random variable that is independent and follows an independent and a probability distribution from the exponential family, defined by or more generally, a selected variance function and dispersion parameter.
- A linear predictor \( \eta = X\beta \).
- A link function \( g \) such that \( E(Y) = \mu = g^{-1}(\eta) \).

As can be seen in the description of the three GLM components above, it may take more than a casual introduction to statistics to comprehend the construction of a GLM. As stated earlier, a downside to GLMs is that it is more challenging to interpret the GLMs output than with univariate models.

If the underlying data is not credible, then no model will improve that credibility, and segmentation methods could make credibility worse. GLM software provides point estimates and allows the modeler to consider standard errors and confidence intervals. GLM output is typically assumed to be 100% credible no matter the size of the underlying data set. GLMs effectively assume that the underlying datasets are 100% credible no matter their size. If some segments have little data, the resulting uncertainty would not be reflected in the GLM parameter estimates themselves (although it might be reflected in the standard errors, confidence intervals, etc.). Even though the process of selecting relativities often includes adjusting the raw GLM output, the resultant selections are not typically then credibility-weighted with any complement of credibility. Nevertheless, selected relativities based on GLM model output may differ from GLM point estimates.

Because of this presumption in credibility, which may or may not be valid in practice, the modeler and the regulator reviewing the model would need to engage in thoughtful consideration when incorporating GLM output into a rating plan.
to ensure that model predictiveness is not compromised by any lack of actual credibility. Another consideration is the availability of big data, both internal and external, that may result in the selection of predictor variables that have spurious correlation with the target variable. Therefore, to mitigate the risk that model credibility or predictiveness is lacking, a complete filing for a rating plan that incorporates GLM output should include validation evidence for the rating plan, not just the statistical model.

To further complicate regulatory review of models in the future, modeling methods are evolving rapidly and not limited just to GLMs. As computing power grows exponentially, it is opening up the modeling world to more sophisticated forms of data acquisition and data analysis. Insurance actuaries and data scientists seek increased predictiveness by using even more complex predictive modeling methods. Examples of these are predictive models utilizing random forests, decision trees, neural networks, or combinations of available modeling methods (often referred to as ensembles). These evolving techniques will make the regulators’ understanding and oversight of filed rating plans incorporating predictive models even more challenging.

In addition to the growing complexity of predictive models, many state insurance departments do not have in-house actuarial support or have limited resources to contract out for support when reviewing rate filings that include use of predictive models. The Big Data (EX) Working Group identified the need to provide states with guidance and assistance when reviewing predictive models underlying filed rating plans. The Working Group circulated a proposal addressing aid to state insurance regulators in the review of predictive models as used in private passenger personal automobile and homeowners homeowners-attached insurance rate filings. This proposal was circulated to all of the Working Group members and interested parties on December 19, 2017, for a public comment period ending January 12, 2018. The Big Data Working Group effort resulted in the new CASTF charges (see the Introduction section) with identifying best practices that provide guidance to states in the review of predictive models.

So, to get to the question asked by the title of this section: Do regulators need best practices to review predictive models? It might be better to ask this question another way: Are best practices in the review of predictive models of value to regulators and insurance companies? The answer is “yes” to both questions. Regulatory best practices need to be developed that do not unfairly or inordinately create barriers for insurers and ultimately consumers while providing a baseline of analysis for regulators to review the referenced filings. Best practices will aid regulatory reviewers by raising their level of model understanding. With regard to scorecard models and the model algorithm, there is often not sufficient support for relative weight, parameter values, or scores of each variable. Best practices can potentially aid in addressing this problem.

However, best practices are not intended to create standards for filings that include predictive models. Rather, best practices will assist the states in identifying the model elements they should be looking for in a filing that will aid the regulator in understanding why the company believes that the filed predictive model improves the company’s rating plan, making that rating plan fairer to all consumers in the marketplace. To make this work, both regulators and industry need to recognize that:

---


9 All comments received by the end of January were posted to the NAIC website March 12 for review.
• Best practices merely provide guidance to regulators in their essential and authoritative role over the rating plans in their state.
• All states may have a need to review predictive models whether that occurs with approval of rating plans or in a market conduct exam. Best practices help the regulator identify elements of a model that may influence the regulatory review as to whether modeled rates are appropriately justified. Each regulator needs to decide if the insurer’s proposed rates are compliant with state laws and regulations and whether to act on that information.
• Best practices will lead to improved quality in predictive model reviews across states, aiding speed to market and competitiveness of the state marketplace.
• Best practices provide a framework for states to share knowledge and resources to facilitate the technical review of predictive models.
• Best practices aid training of new regulators and/or regulators new to reviewing predictive models. (This is especially useful for those regulators who do not actively participate in NAIC discussions related to the subject of predictive models.)
• Each regulator adopting best practices will be better able to identify the resources needed to assist their state in the review of predictive models.

Lastly, from this point on in this paper, best practices will be referred to as “guidance.” This reference is in line with the intent of this paper to support individual state autonomy in the review of predictive models.

IV. SCOPE

The focus of this paper will be on GLMs used to create private passenger personal automobile and home insurance rating plans.

The legal and regulatory constraints (including state variations) are likely to be more evolved, and challenging, for personal automobile and home insurance. Through review of these personal lines, the knowledge needed to review predictive models, and guidance in this paper regarding GLMs for personal automobile and home insurance may be transferrable when the review involves GLMs applied to other lines of business. Modeling depends on context, so the GLM reviewer has to be alert for data challenges and business applications that differ from the more familiar personal lines. For example, compared to personal lines, modeling for rates in commercial lines is more likely to encounter low volumes of historical data, dependence on advisory loss costs, unique large accounts with some large deductibles and package products that build create policies from numerous line-of-business and coverage building blocks. Commercial lines commonly use individual risk modifications following experience, judgment, and/or expense considerations. A regulator may never see how models impact commercial excess and surplus lines filings. The legal and regulatory constraints (including state variations) are likely to be more evolved, and challenging, in personal lines. A GLM rate model for personal lines in 2019 is either an update or a late-adopter's defensive tactic. Adopting GLM for commercial lines has a shorter history. Also, the legal and regulatory constraints (including state variations) are likely to be more prevalent, and challenging in personal lines, which is the basis of this paper’s guidance. A GLM rate model for personal lines in 2019 is either an update or a late-adopter's defensive tactic. Adopting a GLM for commercial lines rating plans has a shorter history and thus is less familiar to many regulators.

Guidance offered here might be useful (with deeper adaptations) when starting to review different types of predictive models. If the model is not a GLM, some listed items might not apply. Not all predictive models generate p-values or F tests. Depending on the model type, other considerations might be important. When transferring guidance to other lines of business and other types of model, unique considerations may arise depending on the context in which a predictive model is proposed to be deployed, the uses to which it is proposed to be put, and the potential consequences for the insurer, its customers and its competitors. This paper does not delve into these possible considerations but regulators should be prepared to address them as they arise.

V. CONFIDENTIALITY

Regulatory reviewers are required to protect confidential information in accordance with applicable State law. However, insurers should be aware that a rate filing might become part of the public record. Each state determines the confidentiality of a rate filing, supplemental material to the filing, when filing information might become public, the procedure to request
that filing information be held confidentially, and the procedure by which a public records request is made. It is incumbent on an insurer to be familiar with each state’s laws regarding the confidentiality of information submitted with their rate filing.

Though state authority, regulations and rules governing confidentiality always apply, this reliance should be revisited if the NAIC or another third party becomes involved in the review process on behalf of the states.

VI. GUIDANCE FOR REGULATORY REVIEW OF PREDICTIVE MODELS (BEST PRACTICES)

Best practices will help the regulator understand if a predictive model is cost based, if the predictive model is compliant with state law, and how the model improves, the company’s rating plan. Best practices can, also, make the regulator's review more consistent across states and more efficient, and assist companies in getting their products to market faster.

With this in mind, the regulator's review of predictive models should:

1. Ensure that the selected rating factors, developed based on the model or other analysis, produce rates that are not excessive, inadequate, or unfairly discriminatory.
   a. Review the overall rate level impact of the proposed revisions proposed based on the predictive model output in comparison to rate level indications provided by the filer.
   b. Review the premium disruption for individual policyholders and how the disruptions can be explained to individual consumers.
   c. Review the individual input characteristics to and output factors from the predictive model (and its submodels), as well as, associated selected relativities to ensure they are not unfairly discriminatory.

2. Thoroughly review all aspects of Obtain a clear understanding of the data used to build and validate the model, and thoroughly review all other aspects of the model, including the source data, assumptions, adjustments, variables, submodels used as input, and resulting output.
   a. Determine that individual input characteristics to a predictive model and their resulting rating factors 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.
   b. Determine that the data used as input to the predictive model is accurate, including a clear understanding how missing values, erroneous values and outliers are handled.
   c. Determine that any adjustments to the raw data are handled appropriately, including but not limited to, trending, development, capping, removal of catastrophes.
   d. Determine that rating factors from a predictive model are related to expected loss or expense differences in risk. Each rating factor should have a demonstrable actual relationship to expected loss or expense.
   e. Obtain a clear understanding of how often each risk characteristic, used as input to the model, is updated and whether the model is periodically rerun, so model output reflects changes to non-static risk characteristics.
   f. Determine whether internal and external data used in relation to the model is compatible with practices allowed in the jurisdiction and do not reflect prohibited characteristics.
   g. Obtain a clear understanding of how the selected predictive model was built.

3. Evaluate how the model interacts with and improves the rating plan.
   a. Obtain a clear understanding of the characteristics that are input to a predictive model (and its sub-models), their relationship to each other and their relationship to non-modeled characteristics/variables used to calculate a risk’s premium.
   b. Obtain a clear understanding of how the selected predictive model was built and why the insurer believes this type of model works in an private passenger automobile or homeowner’s insurance risk application.
   c. Obtain a clear understanding of how model output interacts with non-modeled characteristics/variables used to calculate a risk’s premium.
4. Enable competition and innovation to promote the growth, financial stability, and efficiency of the insurance marketplace.
   a. Enable innovation in the pricing of insurance through acceptance of predictive models, provided they are actuarially sound and in compliance with state laws.
   b. Protect the confidentiality of filed predictive models and supporting information in accordance with state law.
   c. Review predictive models in a timely manner to enable reasonable speed to market.

VII. PREDICTIVE MODELS – INFORMATION FOR REGULATORY REVIEW

This section of the paper identifies the information a regulator may need to review a predictive model used by an insurer to support a filed personal automobile or home insurance rating plan. The list is lengthy but not exhaustive. It is not intended to limit the authority of a regulator to request additional information in support of the model or filed rating plan. Nor is every item on the list intended to be a requirement for every filing. However, the items listed should help guide a regulator to obtain sufficient information to determine if the rating plan meets state specific filing and legal requirements.

Documentation of the design and operational details of the model is required to ensure business continuity and transparency of models used. granularity of documentation takes into account the level of management or key function at which it is intended to be used. Documentation should be sufficiently detailed and complete to enable a third party to form a sound judgment on the suitability of the model for the intended purpose. The theory, assumptions, methodologies, software and empirical bases should be explained, as well as the data used in developing and implementing the model. Relevant testing and ongoing performance testing need to be documented. Key model limitations and overrides need to be pointed out so that stakeholders understand the circumstances under which the model does not work effectively. End-user documentation should be provided and key reports using the model results described. Major changes to the model need to be shared in a timely manner and documented, and IT controls should be in place, such as a record of versions, change control and access to model.

Many information elements listed below are probably confidential, proprietary or trade secret and should be treated as such according to state law. Regulators should be aware of their state laws on confidentiality when requesting data from insurers that may be proprietary or trade secret. For example, some proprietary models may have contractual terms (with the insurer) that prevent disclosure to the public. Without clear necessity, exposing this data to additional dissemination may hinder the model's protection.

Though the list seems long, the insurer should already have internal documentation on the model for more than half of the information listed. The remaining items on the list require either minimal analysis (approximately 25%) or deeper analysis to generate the information for a regulator (approximately 25%).

The “Importance to Regulator’s Review” ranking of information a regulator may need to review is based on the following level criteria:

**Level 1** - This information is necessary to begin the review of a predictive model. These data elements pertain to basic information about the type and structure of the model, the data and variables used, the assumptions made, and the goodness of fit. Ideally, this information would be included in the filing documentation with the initial submission of a filing made based on a predictive model.

---

11 Michele Bourdeau, The Modeling Platform ISSUE 4 • DECEMBER 2016 Model Risk Management: An Overview, Page 6; Published by the Modeling Section of the Society of Actuaries.
12 There are some models that are made public by the vendor and would not result in a hindrance of the model’s protection.
Level 2 - This information is necessary to continue the review of all but the most basic models; such as those based only on the filer’s internal data and only including variables that are in the filed rating plan. These data elements provide more detailed information about the model and address questions arising from review of the information in Level 1. Insurers concerned with speed to market may also want to include this information in the filing documentation.

Level 3 - This information is necessary to continue the review of a model where concerns have been raised and not resolved based on review of the information in Levels 1 and 2. These data elements address even more detailed aspects of the model including (to be listed after we assign levels). This information does not necessarily need to be included with the initial submission, unless specifically requested in a particular jurisdiction, as it is typically requested only if the reviewer has concerns that the model may not comply with state laws.

Level 4 - This information is necessary to continue the review of a model where concerns have been raised and not resolved based on the information in Levels 1, 2, and 3. This most granular level of detail is addressing the basic building blocks of the model and does not necessarily need to be included by the filer with the initial submission, unless specifically requested in a particular jurisdiction. It is typically requested only if the reviewer has serious concerns that the model may produce rates or rating factors that are excessive, inadequate, or unfairly discriminatory.

A. Selecting Model Input

<table>
<thead>
<tr>
<th>Section</th>
<th>Information Element</th>
<th>Level of Importance to the Regulator’s Review</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Available Data Sources</td>
<td>Request details of all data sources, whether internal to the company or from external sources. For insurance experience (policy or claim), determine whether data are aggregated by calendar, accident, fiscal or policy year data and when it was last evaluated. For each data source, get a list all data elements used as input to the model that came from that source. For insurance data, get a list all companies whose data is included in the datasets.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>A.1.a</td>
<td>Reconcile raw aggregated insurance data with available external insurance reports.</td>
<td>2</td>
<td>Accuracy of insurance data should be reviewed as well. Aggregated data is straight from the insurer’s data banks without modification (e.g., not scrubbed or transformed). The dataset would not be adjusted for</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>A.1.c</td>
<td>Review the geographic scope and geographic exposure distribution of the raw data for relevance to the state where the model is filed.</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The company should explain how the data used to build the model makes sense for a specific state. The regulator should inquire which states were included in the data underlying the model build, testing and validation. The company should provide an explanation where the data came from geographically and that it is a good representation for a state, i.e., the distribution by state should not introduce a geographic bias. For example, there could be a bias by peril or wind-resistant building codes. Evaluate whether the data is relevant to the loss potential for which it is being used. For example, verify that hurricane data is only used where hurricanes can occur.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A.1.d</td>
<td>Be aware of any non-insurance data used (customer-provided or other), including who owns this data, how consumers can verify their data and correct errors, whether the data was collected by use of a questionnaire/checklist, whether it was voluntarily reported by the applicant, and whether any of the variables are subject to the Fair Credit Reporting Act. If the data is from an outside source, determine the steps that were taken by the company to verify the data was accurate.</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>If the data is from a third-party source, the company should provide information on the source. Depending on the nature of the data, data should be documented and an overview of who owns it and the topic of consumer verification should be addressed.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 2. Sub-Models

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A.2.a</td>
<td>Consider the relevance of (e.g., is there a bias) of overlapping data or variables used in the model and sub-models.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Check if the same variables/datasets were used in both the model, a submodel or as stand-alone rating characteristics. If so, verify there was no double-counting or redundancy.</td>
<td></td>
</tr>
<tr>
<td>A.2.b</td>
<td>Determine if the sub-model was previously approved (or accepted) by the regulatory agency.</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>If the sub-model was previously approved, that may change the extent of the sub-model’s review. If approved, verify when and that it was the same model currently under review. However, previous approvals do not necessarily confer a guarantee of ongoing approval, for example when statutes and regulations have changed or if a model’s indications have been undermined by subsequent empirical experience. However, knowing whether a model has been previously approved can help focus the regulator’s efforts and determine whether or not the prior decision needs to be revisited.</td>
<td></td>
</tr>
</tbody>
</table>
| A.2.bc | Determine if sub-model output was used as input to the GLM; obtain the vendor name, and the name and version of the sub-model. | 1 | To accelerate the review of the filing, the regulator needs to know the name of 3rd party vendor and contact information for a representative from the vendor, whether model or sub-model. The company should provide the name of the third-party vendor and a contact in the event the regulator has questions. The "contact" can be an intermediary at the insurer, e.g., a filing specialist, who can place the regulator in direct contact with a Subject Matter Expert (SME) at the vendor.

Examples of such sub-models include credit/financial scoring algorithms and household composite score models. Sub-models can be evaluated separately and in the same manner as the primary model under evaluation. A sub-model contact for additional information should be provided. SMEs on sub-model may need to be brought into the conversation with regulators (whether in-house or 3rd-party sub-models are used). |
| A.2.cd | If using catastrophe model output, identify the vendor and the model settings/assumptions used when the model was run. | 1 | For example, it is important to know hurricane model settings for storm surge, demand surge, long/short-term views. To accelerate the review of the filing, get contact information for the SME that ran the model and an SME from the vendor. The "SME" can be an intermediary at the insurer, e.g., a filing specialist, who can place the regulator in direct contact with the appropriate SMEs at the insurer or model vendor. |
| A.2.de | If using catastrophe model output (a sub-model) as input to the GLM under review, verify whether loss associated with the modeled output was removed from the loss experience datasets. | 1 | If a weather-based sub-model is input to the GLM under review, loss data used to develop the model should not include loss experience associated with the weather-based sub-model. Doing so could cause distortions in the modeled results by double counting such losses when determining relativities or loss loads in the filed rating plan. For example, redundant losses in the data may occur when non-hurricane wind losses are included in the data while also using a severe convective storm model in the actuarial indication. Such redundancy may also occur with the inclusion of fluvial or pluvial flood losses when using a flood model, inclusion of freeze losses when using a winter storm model or including demand surge caused by any catastrophic event. Note that, the rating plan or indications underlying the rating plan, may provide special treatment of large losses and non-modeled large loss events. If such treatments exist, the company should provide an explanation how they were handled. These treatments need to be identified and the company/regulator needs to determine whether model data needs to be adjusted. For example, should large BI losses, in the case of personal automobile insurance, be capped or excluded, or should large non-catastrophe wind/hail claims in...
### Adjustments to Data

| A.2.a | If using output of any scoring algorithms, obtain a list of the variables used to determine the score and provide the source of the data used to calculate the score. | 1 | Any sub-model should be reviewed in the same manner as the primary model that uses the sub-model’s output as input. |

#### A.3.a
Determine if premium, exposure, loss or expense data were adjusted (e.g., developed, trended, adjusted for catastrophe experience or capped) and, if so, how? Do the adjustments vary for different segments of the data and, if so, identify the segments and how was the data adjusted?

#### A.3.b
Identify adjustments that were made to raw aggregated data, e.g., transformations, binning and/or categorizations. If any, identify the name of the characteristic/variable and obtain a description of the adjustment.

#### A.3.c
Ask for aggregated data (one data set of pre-adjusted/scrubbed data and one data set of post-adjusted/scrubbed data) that allows the regulator to focus on the univariate distributions and compare raw data to adjusted/binmed/transformed/etc. data.

#### A.3.d
Determine how missing data was handled.

---

© 2019 National Association of Insurance Commissioners
null or “not available” values in the data. If adjustments or re-coding of values were made, they should be explained. It may be useful to the regulator if the percentage of exposures and premium for missing information from the model data were provided. This data can be displayed in either graphical or tabular formats.

<table>
<thead>
<tr>
<th>A.3.e</th>
<th>If duplicate records exist, determine how they were handled.</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.3.f</td>
<td>Determine if there were any material data outliers identified and subsequently adjusted during the scrubbing process. Get a list (with description) of the outliers and determine what adjustments were made to those outliers.</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Look for a discussion of how outliers were handled. If necessary, the regulator may want to investigate further by getting a list (with description) of the outliers and determine what adjustments were made to each outlier. To understand the filer's response, the regulator should ask for the filer's materiality standard.</td>
<td></td>
</tr>
</tbody>
</table>

4. Data Organization

<table>
<thead>
<tr>
<th>A.4.a</th>
<th>Obtain documentation on the methods used to compile and organize data, including procedures to merge data from different sources or filter data based on particular characteristics and a description of any preliminary analyses, data checks, and logical tests performed on the data and the results of those tests.</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>This should explain how data from separate sources was merged or how subsets of policies, based on selected characteristics, are filtered to be included in the data underlying the model and the rationale for that filtering.</td>
<td></td>
</tr>
<tr>
<td>A.4.b</td>
<td>Obtain documentation on the insurer's process for reviewing the appropriateness, reasonableness, consistency and comprehensiveness of the data, including a discussion of the rational relationship the data has to the predicted variable.</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>An example is when by-peril or by-coverage modeling is performed; the documentation should be for each peril/coverage and make intuitive-rational sense. For example, if “murder” or “theft” data are used to predict the wind peril, provide support and an intuitive-rational explanation of their use.</td>
<td></td>
</tr>
<tr>
<td>A.4.c</td>
<td>Identify material findings the company had during their data review and obtain an explanation of any potential material limitations, defects, bias or unresolved concerns found or believed to exist in the data. If issues or limitations in the data influenced modeling analysis and/or results, obtain a description of those concerns and an explanation how modeling analysis was adjusted and/or results were impacted.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>A response of “none” or “n/a” may be an appropriate response.</td>
<td></td>
</tr>
</tbody>
</table>
### B. Building the Model

<table>
<thead>
<tr>
<th>Section</th>
<th>Information Element</th>
<th>Level of Importance to Regulator’s Review</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.1.a</td>
<td>Identify the type of model underlying the rate filing (e.g. Generalized Linear Model – GLM, decision tree, Bayesian Generalized Linear Model, Gradient-Boosting Machine, neural network, etc.). Understand the model’s role in the rating system and provide the reasons why that type of model is an appropriate choice for that role.</td>
<td>1</td>
<td>It is important to understand if the model in question is a GLM, and therefore these best practices are applicable or, if it is some other model type, in which case other reasonable review approaches may be considered. There should be an explanation of why the model (using the variables included in it) is appropriate for the line of business. If by-peril or by-coverage modeling is used, the explanation should be by-peril/coverage. Note, if the model is not a GLM, the guidance and information elements in this white paper may not apply in their entirety.</td>
</tr>
<tr>
<td>B.1.b</td>
<td>Identify the software used for model development. Obtain the name of the software vendor/developer, software product and a software version reference used in model development.</td>
<td>2</td>
<td>Changes in software from one model version to the next may explain if such changes, over time, contribute to changes in the modeled results. The company should provide the name of the third-party vendor and a “contact” in the event the regulator has questions. The “contact” can be an intermediary at the insurer who can place the regulator in direct contact with appropriate SMEs. Open-source software/programs used in model development should be identified by name and version the same as if from a vendor. If version is not known, simply state such, e.g., “R is the software source.”</td>
</tr>
<tr>
<td>B.1.c</td>
<td>Obtain a description how the available data was divided between model training, test and validation datasets. The description should include an explanation why the selected approach was deemed most appropriate, and whether the company made any further subdivisions of available data and reasons for the subdivisions (e.g., a portion separated from training data to support testing of components during model building). Determine if the validation data was accessed before model training was completed and, if so, obtain an explanation why and why that came to occur.</td>
<td>1</td>
<td>It would be unexpected if validation data were used for any purpose other than validation.</td>
</tr>
<tr>
<td>B.1.d</td>
<td>Obtain a brief description of the development process, from initial concept to final model and filed rating plan (in less than three pages of narrative).</td>
<td>1</td>
<td>The narrative should have the same scope as the filing.</td>
</tr>
</tbody>
</table>

© 2019 National Association of Insurance Commissioners
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.1.e</td>
<td>Obtain a narrative on whether loss ratio, pure premium or frequency/severity analyses were performed and, if separate frequency/severity modeling was performed, how pure premiums were determined.</td>
</tr>
<tr>
<td>B.1.f</td>
<td>Identify the model’s target variable.</td>
</tr>
<tr>
<td>B.1.g</td>
<td>Obtain a detailed description of the variable selection process.</td>
</tr>
<tr>
<td>B.1.h</td>
<td>In conjunction with variable selection, obtain a narrative on how the Company determine the granularity of the rating variables during model development.</td>
</tr>
<tr>
<td>B.1.i</td>
<td>Determine if model input data was segmented in any way. For example, was modeling performed on a by-coverage, by-peril, or by-form basis? If so, obtain a description of data segmentation and the reasons for data segmentation.</td>
</tr>
<tr>
<td>B.1.j</td>
<td>If adjustments to the model were made based on credibility considerations, obtain an explanation of the credibility considerations and how the adjustments were applied.</td>
</tr>
</tbody>
</table>

2. Medium-Level Narrative for Building the Model

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.2.a</td>
<td>At crucial points in model development, if selections were made among alternatives regarding model assumptions or techniques, obtain a narrative on the judgment used to make those selections.</td>
</tr>
<tr>
<td>B.2.b</td>
<td>If post-model adjustments were made to the data and the model was rerun, obtain an explanation on the details and the rationale for those adjustments.</td>
</tr>
</tbody>
</table>

Adjustments may be needed given models do not explicitly consider the credibility of the input data or the model’s resulting output; models take input data at face value and assume 100% credibility when producing modeled output.

Evaluating the addition or removal of variables and the model fitting. It is not necessary for the company to discuss each iteration of adding and subtracting variables, but the regulator should gain a general understanding of how these adjustments were done, including any statistical improvement measures relied upon.
B.2.c Obtain a description of univariate balancing and the **testing that was performed during the model-building process**, including an explanation of the thought processes involved and a discussion of why interaction terms were included (or not included).

22 Further elaboration from B.2.b. There should be a description of testing that was performed during the model-building process. Examples of tests that may have been performed include univariate testing and review of a correlation matrix.

B.2.d Obtain a description of the 2-way balancing and testing that was performed during the model-building process, including an explanation of the thought processes of including (or not including) interaction terms.

2 Further elaboration from B.2.a and B.2.b.

B.2.e For the GLM, identify the link function used. Identify which distribution was used for the model (e.g., Poisson, Gaussian, log-normal, Tweedie). Obtain an explanation why the link function and distribution were chosen. Obtain the formulas for the distribution and link functions, including specific numerical parameters of the distribution. Obtain a discussion of applicable convergence criterion.

1 Solving the GLM is iterative and the modeler can check to see if fit is improving. At some point convergence occurs, though when it occurs can be subjective or based on threshold criteria. The convergence criterion should be documented with a brief explanation of why it was selected.

B.2.f Obtain a narrative on the formula relationship between the data and the model outputs, with a definition of each model input and output. The narrative should include all coefficients necessary to evaluate the predicted pure premium, relativity or other value, for any real or hypothetical set of inputs.

2 B.4.l and B.4.m will show the mathematical functions involved and could be used to reproduce some model predictions.

B.2.g If there were data situations in which GLM weights were used, obtain an explanation of how and why they were used.

3 Investigate whether identical records were combined to build the model.

### 3. Predictor Variables

| B.3.a | Obtain a complete data dictionary, including the names, types, definitions and uses of each predictor variable, offset variable, control variable, proxy variable, geographic variable, geodemographic variable and all other variables in the model used on their own or as an interaction with other variables (including sub-models and external models). |
| 1 | Types of variables might be continuous, discrete, Boolean, etc. Definitions should not use programming language or code. For any variable(s) intended to function as a control or offset, obtain an explanation of their rationale and impact. Also, for any use of interaction between variables, obtain an explanation of its rationale and impact. |

| B.3.b | Obtain a list of predictor variables considered but not used in the final model, and the rationale for their removal. |
| 4 | The rationale for this requirement is to identify variables that the company finds to be predictive but ultimately may reject for reasons other than loss-cost considerations (e.g., price optimization). Also, look for variables the company tested and then rejected. This item could help address concerns about data dredging. The reasonableness of including a variable with given significance level could depend greatly on the other variables the company evaluated for inclusion in the model and the criteria for inclusion or omission. For instance, if the company tested 1,000 similar variables and selected the one with the lowest p-value of 0.001, this would be a far, far weaker case for statistical |

© 2019 National Association of Insurance Commissioners
### 3. Model Development

**B.3.c** Obtain a correlation matrix for all predictor variables included in the model and sub-model(s).

- **Example:** While GLMs accommodate collinearity, the correlation matrix provides more information about the magnitude of correlation between variables. The company should indicate what statistic was used (e.g., Pearson, Cramer's V). The reviewer should understand what statistic was used to produce the matrix, but should not specify the statistic.

**B.3.d** Obtain an *intuitive rational* explanation for why an increase in each predictor variable should increase or decrease frequency, severity, loss costs, expenses, or any element or characteristic being predicted.

- **Example:** The explanation should go beyond demonstrating correlation. Considering possible causation is relevant, but proving causation is neither practical nor expected. If no *intuitive rational* explanation can be provided, greater scrutiny may be appropriate. For example, the regulator should look for unfamiliar predictor variables and, if found, the regulator should seek to understand the rational connection that variable has to increasing or decreasing the target variable.

**B.3.e** If the modeler made use of one or more dimensionality reduction techniques, such as a Principal Component Analysis (PCA), obtain a narrative about that process, an explanation why that technique was chosen, and a description of the step-by-step process used to transform observations (usually correlated) into a set of linearly uncorrelated variables. In each instance, obtain a list of the pre-transformation and post-transformation variable names, and an explanation how the results of the dimensionality reduction technique was used within the model.

### 4. Adjusting Data, Model Validation and Goodness-of-Fit Measures

**B.4.a** Obtain a description of the methods used to assess the statistical significance/goodness of the fit of the model to validation data, such as lift charts and statistical tests. Compare the model's projected results to historical actual results and verify that modeled results are reasonably similar to actual results from validation data.

- **Example:** For models that are built using multi-state data, validation data for some segments of risk is likely to have low credibility in individual states. Nevertheless, some regulators require model validation on State-only data, especially when analysis using state-only data contradicts the countrywide results. State-only data might be more applicable but could also be impacted by low credibility for some segments of risk. Look for geographic stability measures, e.g., across states or territories within state.

**B.4.b** Obtain a description of any adjustments that were made in the data with respect to scaling for discrete variables or binning the data.

- **Example:** A 3.3 addresses pre-modeling adjustments to data. In the mid-level narrative context, B.2.a addresses judgments of any kind made during modeling. Only choices made at “crucial points in model development” need be discussed.
<table>
<thead>
<tr>
<th>B.4.c</th>
<th>Obtain a description of any transformations made for continuous variables.</th>
</tr>
</thead>
</table>

2. A.3.f addresses pre-modeling transformations to data. In the mid-level narrative context, B.2.a addresses transformations of any kind made during modeling. Only choices made at "crucial points in model development" need be discussed.

To build a unique model with acceptable goodness of fit to the training data, important steps have been taken. Such steps may have been numerous, and at least some of the judgments involved may be difficult to describe and explain. Nevertheless, neither the model filer nor the reviewer can assume these steps are immaterial, generally understood, or implied by the model's generic form. The model filer should anticipate regulatory concerns in its initial submission by identifying and explaining the model fitting steps it considers most important. If a reviewer has regulatory concerns not resolved by the initial submission, appropriate follow-up inquiries are likely to depend on the particular circumstance.

<table>
<thead>
<tr>
<th>B.4.db</th>
<th>For each discrete all variables (discrete or continuous) level, review the appropriate parameter values, confidence intervals, chi-square tests, p-values and any other relevant and material tests. Determine if model development data, validation data, test data or other data was used for these tests.</th>
</tr>
</thead>
</table>

1. Typical p-values greater than 5% are large and should be questioned. Reasonable business judgment can sometimes provide legitimate support for high p-values. Reasonableness of the p-value threshold could also vary depending on the context of the model. For example, the threshold might be lower when many candidate variables were evaluated for inclusion in the model.

Overall lift charts and/or statistical tests using validation data may not provide enough of the picture. If there is concern about one or more individual variables, the reviewer may obtain, for each discrete variable level, the parameter value, the parameter value, confidence intervals, chi-square tests, p-values and any other relevant and material tests. For variables that are modeled continuously, it may be sufficient to obtain statistics around the modeled parameters; for example, confidence intervals around each level of an AOI curve might be more than what is needed.
**B.4.e** Identify the threshold for statistical significance and explain why it was selected. Obtain a reasonable and appropriately supported explanation for keeping the variable for each discrete variable level where the p-values were not less than the chosen threshold.

Typical p-values greater than 5% are large and should be questioned. Reasonable business judgment can sometimes provide legitimate support for high p-values. Reasonableness of the p-value threshold could also vary depending on the context of the model. For example, the threshold might be lower when many candidate variables were evaluated for inclusion in the model.

Overall lift charts and/or statistical tests using validation data may not provide enough of the picture. If there is concern about one or more individual variables, the reviewer may obtain, for each discrete variable level, the parameter value, confidence intervals, chi-square tests, p-values and any other relevant and material tests. For variables that are modeled continuously, it may be sufficient to obtain statistics around the modeled parameters; for example, confidence intervals around each level of an AOI curve might be more than what is needed.

**B.4.f** For overall discrete variables, review type 3 chi-square tests, p-values, F tests and any other relevant and material test. Determine if model development data, validation data, test data or other data was used for these tests.

Typical p-values greater than 5% are large and should be questioned. Reasonable business judgment can sometimes provide legitimate support for high p-values. Reasonableness of the p-value threshold could also vary depending on the context of the model, e.g., the threshold might be lower when many candidate variables were evaluated for inclusion in the model.

Overall lift charts and/or statistical tests using validation data may not provide enough of the picture. If there is concern about one or more individual variables, the reviewer may obtain, for each discrete variable level, the parameter value, confidence intervals, chi-square tests, p-values and any other relevant and material tests. For variables that are modeled continuously, it may be sufficient to obtain statistics around the modeled parameters; for example, confidence intervals around each level of an AOI curve might be more than what is needed.

**B.4.g** Obtain evidence that the model fits the training data well, for individual variables, for any relevant combinations of variables and for, the overall model.

For a GLM, such evidence may be available using chi-square tests, p-values, F tests and/or other means.

The steps taken during modeling to achieve goodness-of-fit are likely to be numerous and laborious to describe, but they contribute much of what is generalized about GLM. We should not assume we know what they did and ask "how?" Instead, we should ask what they did and be prepared to ask follow up questions.
For continuous variables, provide confidence intervals, chi-square tests, p-values and any other relevant and material test. Determine if model development data, validation data, test data or other data was used for these tests.

Typical p-values greater than 5% are large and should be questioned. Reasonable business judgment can sometimes provide legitimate support for high p-values. Reasonableness of the p-value threshold could also vary depending on the context of the model, e.g., the threshold might be lower when many candidate variables were evaluated for inclusion in the model.

Overall lift charts and/or statistical tests using validation data may not provide enough of the picture. If there is concern about one or more individual variables, the reviewer may obtain, for each discrete variable level, the parameter value, confidence intervals, chi-square tests, p-values and any other relevant and material tests. For variables that are modeled continuously, it may be sufficient to obtain statistics around the modeled parameters; for example, confidence intervals around each level of an AOI curve might be more than what is needed.

Evaluate the build/test/validation datasets for potential time-sensitive model distortions (e.g., a winter storm in year 3 of 5 can distort the model in both the testing and validation datasets).

Obsolescence over time is a model risk (e.g., old data for a variable or a variable itself may no longer be relevant). If a model being introduced now is based on losses from years ago, the reviewer should be interested in knowing whether that model would be predictive in the proposed context. Validation using recent data from the proposed context might be requested. Obsolescence is a risk even for a new model based on recent and relevant loss data. The reviewer may want to inquire as to the following: What steps, if any, were taken during modeling to prevent or delay obsolescence? What controls will exist to measure the rate of obsolescence? What is the plan and timeline for updating and ultimately replacing the model?

The reviewer should also consider that as newer technologies enter the market (e.g., personal automobile) their impact may change claim activity over time (e.g., lower frequency of loss). So, it is not necessarily a bad thing that the results are not stable over time.

Obtain a description how the model was tested for stability over time.

Obtain a narrative on how potential concerns with overfitting were addressed.
**B.4.k**
Obtain support demonstrating that the GLM assumptions are appropriate.

Obtain support demonstrating that the GLM assumptions are appropriate. Visual review of plots of actual errors is usually sufficient.

The reviewer should look for a conceptual narrative covering these topics: How does this particular GLM work? Why did the rate filer do what it did? Why employ this design instead of alternatives? Why choose this particular distribution function and this particular link function? A company response may be at a fairly high level and reference industry practices. If the reviewer determines that the model makes no assumptions that are considered to be unreasonable, the importance of this item may be reduced.

**B.4.l**
Obtain 5-10 sample records with corresponding output from the model for those records.

Obtain 5-10 sample records with corresponding output from the model for those records.

---

**5. “Old Model” Versus “New Model”**

**B.5.a**
Obtain an explanation why this model is an improvement to the current rating plan.

If it replaces a previous model, find out why it is better than the one it is replacing; determine how the company reached that conclusion and identify metrics relied on in reaching that conclusion. Look for an explanation of any changes in calculations, assumptions, parameters, and data used to build this model from the previous model.

Regulators should expect to see improvement in the new class plan’s predictive ability or other sufficient reason for the change.

**B.5.b**
Determine if two Gini coefficients were compared and obtain a narrative on the conclusion drawn from this comparison.

Determine if two Gini coefficients were compared and obtain a narrative on the conclusion drawn from this comparison.

One example of a comparison might be sufficient. This is relevant when one model is being updated or replaced. Regulators should expect to see improvement in the new class plan’s predictive ability. This information element requests a comparison of Gini coefficient from the prior model to the Gini coefficient of proposed model. It is expected that there should be improvement in the Gini coefficient. A higher Gini coefficient indicates greater differentiation produced by the model and how well the model fits that data. This comparison is not applicable to initial model introduction. Reviewer can look to CAS monograph for information on Gini coefficients.

**B.5.c**
Determine if double lift charts were analyzed and obtain a narrative on the conclusion drawn from this analysis.

Determine if double lift charts were analyzed and obtain a narrative on the conclusion drawn from this analysis.

One example of a comparison might be sufficient. Note that "not applicable" is an acceptable response.

**B.5.d**
If replacing an existing model, obtain a list of any predictor variables used in the old model that are not used in the new model. Obtain an explanation why these variables were dropped from the new model.

Obtain a list of all new predictor variables in the new model that were not in the prior old model.

Useful to differentiate between old and new variables so the regulator can prioritize more time on new variables not yet reviewed.
6. Modeler Software

| B.6.a | Request access to SMEs (e.g., modelers) who led the project, compiled the data, built the model, and/or performed peer review. | 3 | The filing should contain a contact that can put the regulator in touch with appropriate SMEs and key contributors to the model development to discuss the model. |

C. The Filed Rating Plan

<table>
<thead>
<tr>
<th>Section</th>
<th>Information Element</th>
<th>Level of Importance to Regulator’s Review</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. General Impact of Model on Rating Algorithm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.1.a</td>
<td>In the actuarial memorandum or explanatory memorandum, for each model and sub-model (including external models), look for a narrative that explains each model and its role (how it was used) in the rating system.</td>
<td>1</td>
<td>The &quot;role of the model&quot; relates to how the model integrates into the rating plan as a whole and where the effects of the model are manifested within the various components of the rating plan. This is not intended as an overarching statement of the model's goal, but rather a description of how specifically the model is used. This item is particularly important, if the role of the model cannot be immediately discerned by the reviewer from a quick review of the rate and/or rule pages. (Importance is dependent on state requirements and ease of identification by the first layer of review and escalation to the appropriate review staff.)</td>
</tr>
<tr>
<td>C.1.b</td>
<td>Obtain an explanation of how the model was used to adjust the rating algorithm.</td>
<td>1</td>
<td>Models are often used to produce factor-based indications, which are then used as the basis for the selected changes to the rating plan. It is the changes to the rating plan that create impacts. Consider asking for an explanation of how the model was used to adjust the rating algorithm.</td>
</tr>
<tr>
<td>C.1.c</td>
<td>Obtain a complete list of characteristics/variables used in the proposed rating plan, including those used as input to the model (including sub-models and composite variables) and all other characteristics/variables (not input to the model) used to calculate a premium. For each characteristic/variable, determine if it is only input to the model, whether it is only a separate univariate rating characteristic, or whether it is both input to the model and a separate univariate rating characteristic. The list should include transparent descriptions (in plain language) of each listed characteristic/variable.</td>
<td>1</td>
<td>Examples of variables used as inputs to the model and used as separate univariate rating characteristics might be criteria used to determine a rating tier or household composite characteristic.</td>
</tr>
</tbody>
</table>

2. Relevance of Variables and Relationship to Risk of Loss
<table>
<thead>
<tr>
<th>C.2.a</th>
<th>Obtain a narrative regarding how the characteristics/rating variables, included in the filed rating plan, logically and intuitively relate to the risk of insurance loss (or expense) for the type of insurance product being priced.</th>
<th>2</th>
<th>The narrative should include a discussion of the relevance each characteristic/rating variable has on consumer behavior that would lead to a difference in risk of loss (or expense). The narrative should include a logical and intuitive relationship to cost, and model results should be consistent with the expected direction of the relationship. This explanation would not be needed if the connection between variables and risk of loss (or expense) has already been illustrated.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.3.a</td>
<td>Compare relativities indicated by the model to both current relativities and the insurer's selected relativities for each risk characteristic/variable in the rating plan.</td>
<td>1</td>
<td>“Significant difference” may vary based on the risk characteristic/variable and context. However, the movement of a selected relativity should be in the direction of the indicated relativity; if not, an explanation is necessary as to why the movement is logical.</td>
</tr>
<tr>
<td>C.3.b</td>
<td>Obtain documentation and support for all calculations, judgments, or adjustments that connect the model's indicated values to the selected values.</td>
<td>1</td>
<td>The documentation should include explanations for the necessity of any such adjustments and explain each significant difference between the model's indicated values and the selected values. This applies even to models that produce scores, tiers, or ranges of values for which indications can be derived. This information is especially important if differences between model indicated values and selected values are material and/or impact one consumer population more than another.</td>
</tr>
<tr>
<td>C.3.c</td>
<td>For each characteristic/variable used as both input to the model (including sub-models and composite variables) and as a separate univariate rating characteristic, obtain a narrative how each characteristic/variable was tempered or adjusted to account for possible overlap or redundancy in what the characteristic/variable measures.</td>
<td>2</td>
<td>Modeling loss ratio with these characteristics/variables as control variables would account for possible overlap. The insurer should address this possibility or other considerations, e.g., tier placement models often use risk characteristics/variables that are also used elsewhere in the rating plan. One way to do this would be to model the loss ratios resulting from a process that already uses univariate rating variables. Then the model/composite variables would be attempting to explain the residuals.</td>
</tr>
</tbody>
</table>

### 3. Comparison of Model Outputs to Current and Selected Rating Factors

- **C.3.a**
  - Compare relativities indicated by the model to both current relativities and the insurer's selected relativities for each risk characteristic/variable in the rating plan.
  - “Significant difference” may vary based on the risk characteristic/variable and context. However, the movement of a selected relativity should be in the direction of the indicated relativity; if not, an explanation is necessary as to why the movement is logical.

- **C.3.b**
  - Obtain documentation and support for all calculations, judgments, or adjustments that connect the model's indicated values to the selected values.
  - The documentation should include explanations for the necessity of any such adjustments and explain each significant difference between the model's indicated values and the selected values. This applies even to models that produce scores, tiers, or ranges of values for which indications can be derived. This information is especially important if differences between model indicated values and selected values are material and/or impact one consumer population more than another.

- **C.3.c**
  - For each characteristic/variable used as both input to the model (including sub-models and composite variables) and as a separate univariate rating characteristic, obtain a narrative how each characteristic/variable was tempered or adjusted to account for possible overlap or redundancy in what the characteristic/variable measures.
  - Modeling loss ratio with these characteristics/variables as control variables would account for possible overlap. The insurer should address this possibility or other considerations, e.g., tier placement models often use risk characteristics/variables that are also used elsewhere in the rating plan. One way to do this would be to model the loss ratios resulting from a process that already uses univariate rating variables. Then the model/composite variables would be attempting to explain the residuals.
| C.4.a | Determine what, if any, consideration was given to the credibility of the output data. | 2 | At what level of granularity is credibility applied. If modeling was by-coverage, by-form or by-peril, explain how these were handled when there was not enough credible data by coverage, form or peril to model. |
| C.4.b | If the rating plan is less granular than the model, obtain an explanation why. | 2 | This is applicable if the insurer had to combine modeled output in order to reduce the granularity of the rating plan. |
| C.4.c | If the rating plan is more granular than the model, obtain an explanation why. | 2 | A more granular rating plan implies that the insurer had to extrapolate certain rating treatments, especially at the tails of a distribution of attributes, in a manner not specified by the model indications. |

5. Definitions of Rating Variables

| C.5.a | Obtain a narrative on adjustments made to raw model output, e.g., transformations, binning and/or categorizations. If adjustments were made, obtain the name of the characteristic/variable and a description of the adjustment. | 2 | If rating tiers or other intermediate rating categories are created from model output, the rate and/or rule pages should present these rating tiers or categories. The company should provide an explanation how model output was translated into these rating tiers or intermediate rating categories. |
| C.5.b | Obtain a complete list and description of any rating tiers or other intermediate rating categories that translate the model output into some other structure that is then presented within the rate and/or rule pages. | 1 | - |

6. Supporting Data

| C.6.a | Obtain aggregated state-specific, book-of-business-specific univariate historical experience data, separately for each year included in the model, consisting of loss ratio or pure premium relativities and the data underlying those calculations, at minimum, earned exposures, earned premiums, incurred losses, loss ratios and loss ratio relativities for each category of model output(s) proposed to be used within the rating plan. For each data element, obtain an explanation whether it is raw or adjusted and, if the latter, obtain a detailed explanation for the adjustments. | 34 | For example, were losses developed/undeveloped, trended/untrended, capped/uncapped, etc? Univariate indications should not necessarily be used to override more sophisticated multivariate indications. However, they do provide additional context and may serve as a useful reference. |
C.6.b Obtain an explanation of any material (especially directional) differences between model indications and state-specific univariate indications.

Multivariate indications may be reasonable as refinements to univariate indications, but possibly not for bringing about significant reversals of those indications. For instance, if the univariate indicated relativity for an attribute is 1.5 and the multivariate indicated relativity is 1.25, this is potentially a plausible application of the multivariate techniques. If, however, the univariate indicated relativity is 0.7 and the multivariate indicated relativity is 1.25, a regulator may question whether the attribute in question is negatively correlated with other determinants of risk. Credibility of state data should be considered when state indications differ from modeled results based on a broader data set. However, the relevance of the broader data set to the risks being priced should also be considered. Borderline reversals are not of as much concern.

7. Consumer Impacts

C.7.a Obtain a listing of the top five rating variables that contribute the most to large swings in premium, both as increases and decreases.

These rating variables may represent changes to ratios of factor relativities, be newly introduced to the rating plan, or have been removed from the rating plan.

C.7.b Determine if the insurer performed sensitivity testing to identify significant changes in premium due to small or incremental change in a single risk characteristic. If such testing was performed, obtain a narrative that discusses the testing and provides the results of the testing.

One way to see sensitivity is to analyze a graph of each risk characteristic’s variable’s possible relativities. Look for significant variation between adjacent relativities and evaluate if such variation is reasonable and credible.

C.7.c For the proposed filing, obtain the impacts on expiring policies and describe the process used by management, if any, to mitigate those impacts.

Some mitigation efforts may substantially weaken the connection between premium and expected loss and expense, and hence may be viewed as unfairly discriminatory by some states.
| C.7.d | Obtain a rate disruption/dislocation analysis, demonstrating the distribution of percentage and/or dollar impacts on renewal business (created by rerating the current book of business), and sufficient information to explain the disruptions to individual consumers. | 2 | The analysis should include the largest dollar and percentage impacts arising from the filing, including the impacts arising specifically from the adoption of the model or changes to the model as they translate into the proposed rating plan. While the default request would typically be for the distribution/dislocation of impacts at the overall filing level, the regulator may need to delve into the more granular variable-specific effects of rate changes if there is concern about particular variables having extreme or disproportionate impacts, or significant impacts that have otherwise yet to be substantiated. See Appendix C for an example of a disruption analysis. |
| C.7.e | Obtain exposure distributions for the model's output variables and show the effects of rate changes at granular and summary levels, including the overall impact on the book of business. | 3 | See Appendix C for an example of an exposure distribution. |
| C.7.f | Identify policy characteristics, used as input to a model or sub-model, that remain "static" over a policy's lifetime versus those that will be updated periodically. Obtain a narrative on how the company handles policy characteristics that are listed as "static," yet change over time. | 3 | Some examples of "static" policy characteristics are prior carrier tenure, prior carrier type, prior liability limits, claim history over past X years, or lapse of coverage. These are specific policy characteristics usually set at the time new business is written, used to create an insurance score or to place the business in a rating/underwriting tier, and often fixed for the life of the policy. The reviewer should be aware, and possibly concerned, how the company treats an insured over time when the insured’s risk profile based on "static" variables changes over time but the rate charged, based on a new business insurance score or tier assignment, no longer reflect the insured’s true and current risk profile. A few examples of "non-static" policy characteristics are age of driver, driving record and credit information (FCRA related). These are updated automatically by the company on a periodic basis, usually at renewal, with or without the policyholder explicitly informing the company. |
### C.7.g
Obtain a means to calculate the rate charged a consumer.

The filed rating plan should contain enough information for a regulator to be able to validate policy premium. However, for a complex model or rating plan, a score or premium calculator via Excel or similar means would be ideal, but this could be elicited on a case-by-case basis. Ability to calculate the rate charged could allow the regulator to perform sensitivity testing when there are small changes to a risk characteristic/variable. Note that this information may be proprietary.

### C.7.h
In the filed rating plan, be aware of any non-insurance data used as input to the model (customer-provided or other). In order to respond to consumer inquiries, it may be necessary to inquire as to how consumers can verify their data and correct errors.

If the data is from a third-party source, the company should provide information on the source. Depending on the nature of the data, data should be documented and an overview of who owns it and the topic of consumer verification should be addressed, including how consumers can verify their data and correct errors.

### 8. Accurate Translation of Model into a Rating Plan

C.8.a
Obtain sufficient information to understand how the model outputs are used within the rating system and to verify that the rating plan's manual, in fact, reflects the model output and any adjustments made to the model output.

The regulator can review the rating plan's manual to see that modeled output is properly reflected in the manual's rules, rates, factors, etc.
VIII. PROPOSED CHANGES TO THE PRODUCT FILING REVIEW HANDBOOK

The Task Force was charged to propose modifications to the 2016 Product Filing Review Handbook to reflect best practices for the regulatory review of predictive analytics. The following are the titled sections in Chapter Three “The Basics of Property and Casualty Rate Regulation.” Proposed changes are shown as tracked changes.

CHAPTER THREE
The Basics of Property and Casualty Rate Regulation

No changes are proposed to the following sections at the beginning of Chapter Three: Introduction; Rating Laws; Rate Standards; Rate Justification and Supporting Data; Number of Years of Historical Data; Segregation of Data; Data Adjustments; Premium Adjustments; Losses and LAE (perhaps just DCC) Adjustments; Catastrophe or Large Loss Provisions; Loss Adjustment Expenses; Data Quality; Rate Justification: Overall Rate Level; Contingency Provision; Credibility; Calculation of Overall Rate Level Need: Methods (Pure Premium and Loss Ratio Methods); Rate Justification: Rating Factors; Calculation of Deductible Rating Factors; Calculation of Increased Limit Factors; and Credibility for Rating Factors.

Data Adjustments

...Because the insurance contracts will be written to cover future accident periods, the past data needs to be adjusted to reflect the anticipated future premiums and costs. These adjustments will provide a profit/loss picture if no rate change occurs. Calculations can then be made to determine the overall rate need (or indication). …

Interaction between Rating Variables (Multivariate Analysis)

If each rating variable is evaluated separately, statistically significant interactions between rating variables may not be identified and, thus, may not be included in the rating plan. If the pricing of rating variables is evaluated separately for each rating variable, there is potential to miss the interaction between rating variables. Care should be taken to have a multivariate analysis when practical. In some instances, a multivariate analysis is not possible. But, with computing power growing exponentially, insurers believe they have found many ways to improve their operations and competitiveness through use of complex predictive models in all areas of their insurance business.

Approval of Classification Systems

With rate changes, companies sometimes propose revisions to their classification system. Because the changes to classification plans can be significant and have large impacts on the consumers’ rates, regulators should focus on these changes.

Some items of proposed classification can sometimes be deemed to be against public policy, such as the use of education or occupation. You should be aware of your state’s laws and regulations regarding which rating factors are allowed, and you should require definitions of all data elements that can affect the charged premium. Finding rating or underwriting characteristics that may violate public policy is becoming more difficult for regulators with the increasing and innovative ways insurers use predictive models.

Rating Tiers

Some states allow an insurer to have multiple rate levels, or rating tiers, within a single company. These rating tiers are another way of classifying risks for rating purposes. Typically, there are requirements for rating tiers: the underwriting rules for each tier should be mutually exclusive, clear, and objective; there should be a distinction between the expected losses or expenses for each tier; and the placement process should be auditable. Tiers within a company are mainly seen in personal lines products.

© 2019 National Association of Insurance Commissioners
One particular concern with rating tiers would be the analyses of whether a plan produces unfair discrimination. Questions arise around the time-sensitive aspects of the underwriting criteria and any related re-evaluation of the tiers upon renewal. For example, consider two tiers where the insured is placed in the “high” tier because of a lapse of insurance in the prior 12 months. The question is: What happens upon renewal after there has no longer been a lapse of insurance for 12 months? Does the insured get slotted in the “low” tier as he would if he was new business? Some statutes limit the amount of time that violations, loss history, or insurance scores can be used, and some statutes might only allow credit history to be used for re-rating at the policyholder’s request. Regulators should consider the acceptability of differences in rates between existing and new policyholders when they have the same current risk profile.

Insurers also can create different rating levels by having separate companies within a group. While regulators should examine rating tiers within an insurer to a high degree of regulatory scrutiny, there tends to be less scrutiny with differences in rates that exist between affiliated companies. Workers’ compensation insurers are more likely to obtain rating tiers using separate companies.

**Rate Justification: New Products** – (No change is proposed.)

### Predictive Modeling

The ability of computers to process massive amounts of data has led to the expansion of the use of predictive modeling in insurance ratemaking. Predictive models have enabled insurers to build rating, marketing, underwriting and claim models with significant segmentation predictive power and are increasingly being applied in such areas as claims modeling and used in helping insurers to price risks more effectively.

Key new rating variables that are being incorporated into insurers' predictive models include homeowners' home rates by peril, homeowners'home rating by building characteristics, vehicle history, usage-based auto insurance, and credit characteristics.

Data quality within and communication about models are of key importance with predictive modeling. Depending on definitional boundaries, predictive modeling can sometimes overlap with the field of machine learning. In the modeling space, predictive modeling is often referred to as predictive analytics.

Insurers' use of predictive analytics along with big data has significant potential benefits to both consumers and insurers. Predictive analytics can reveal insights into the relationship between consumer behavior and the cost of insurance, lower the cost of insurance for many, and provide incentives for consumers to better control and mitigate loss. However, predictive analytic techniques are evolving rapidly and leaving many regulators without the necessary tools to effectively review insurers' use of predictive models in insurance applications. To aid the regulator in the review of predictive models, best practices have been developed for generalized linear models or “GLMs”). GLMs are commonly used in personal automobile and home insurance applications.

The term “predictive model” refers to a set of models that use statistics to predict outcomes. When applied to insurance, the model is chosen to estimate the probability or expected value of an outcome given a set amount of input data; for example, models can predict the frequency of loss, the severity of loss, or the pure premium.

To further complicate regulatory review of models in the future, modeling technology and methods are evolving rapidly. GLMs are relatively transparent and their output and consequences are much clearer than many other complex models. But as computing power grows exponentially, it is opening up the modeling world to more sophisticated forms of data acquisition and data analysis. Insurance actuaries and data scientists seek increased predictiveness by using even more complex predictive modeling methods. Examples of these are predictive models utilizing logistic regression, K-nearest neighbor classification, random forests, decision trees, neural networks, or combinations of available modeling methods (often referred to as ensembles). These evolving techniques will make the regulators’ understanding and oversight of filed rating plans even more challenging.

#### A. Generalized Linear Models

The generalized linear model (GLM) is a commonly used predictive model in insurance applications, particularly in building an insurance product’s rating plan. Because of this and the fact most Property and Casualty regulators are most concerned
About personal lines, NAIC has developed a white paper for guidance in reviewing GLMs for personal automobile and home insurance.

Before GLMs became vogue, rating plans were built using univariate methods. Univariate methods were considered easy to understand and easy to demonstrate the relationship to costs (loss and/or expense). However, many consider univariate methods too simplistic since they do not take into account the interaction (or dependencies) of the selected input variables. GLMs introduce significant improvements over univariate-based rating plans by automatically adjusting for correlations among input variables. Today, the majority of predictive models used in personal automobile and home insurance rating plans are GLMs. But, GLM results are not always easy to understand and the relationship to costs may be difficult to explain.

A GLM consists of three elements:

- A target variable, Y, which is a random variable that is independent and follows a probability distribution from the exponential family, defined by a selected variance function and dispersion parameter.
- A linear predictor η = Xβ.
- A link function g such that E(Y) = μ = g−1(η).

As can be seen in the description of the three GLM components above, it may take more than a casual introduction to statistics to comprehend the construction of a GLM. As stated earlier, a downside to GLMs is that it is more challenging to interpret the GLMs output than with univariate models.

B. Credibility of GLM Output

If the underlying data is not credible no model will improve that credibility, and segmentation methods could make credibility worse. GLM software provides point estimates and allows the modeler to consider standard errors and confidence intervals. GLMs effectively assume that the underlying datasets are 100% credible no matter their size. If some segments have little data, the resulting uncertainty would not be reflected in the GLM parameter estimates themselves (although it might be reflected in the standard errors, confidence intervals, etc.). Even though the process of selecting relativities often includes adjusting the raw GLM output, the resultant selections are not typically credibility-weighted with any complement of credibility. [New footnotes: "This is not always true. Sometimes insurers do review complements of credibility and further weight the GLM output with those complements. While this may not be a standard practice today, new techniques could result in this becoming more standard in the future." And "GLMs provide confidence intervals; credibility methods do not. There are techniques such as penalized regression that blend credibility with a GLM and improve a model's ability to generalize."] Nevertheless, selected relativities based on GLM model output may differ from GLM point estimates.

Because of this presumption in credibility, which may or may not be valid in practice, the modeler and the regulator reviewing the model would need to engage in thoughtful consideration when incorporating GLM output into a rating plan to ensure that model predictiveness is not compromised by any lack of actual credibility. Therefore, to mitigate the risk that model credibility or predictiveness is lacking, a complete filing for a rating plan that incorporates GLM output should include validation evidence for the rating plan, not just the statistical model.

C. What is a “Best Practice”??

A best practice is a form of program evaluation in public policy. At its most basic level, a practice is a “tangible and visible behavior...[based on] an idea about how the actions...will solve a problem or achieve a goal.”14. Best practices can maintain quality as an alternative to mandatory legislated standards and can be based on self-assessment or benchmarking.15 Therefore, a best practice represents an effective method of problem solving. The "problem" regulators want to solve is probably better posed as seeking an answer to this question: How can regulators determine that predictive models, as used in rate filings, are compliant with state laws and regulations? However, best practices are not intended to create standards for filings that include predictive models.

13 Refer to NAIC’s white paper titled Regulatory Review of Predictive Models, found at the NAIC website.
Best practices are based on the following principles that promote a comprehensive and coordinated review of predictive models across states:

- State insurance regulators will maintain their current rate regulatory authority.
- State insurance regulators will be able to share information to aid companies in getting insurance products to market more quickly.
- State insurance regulators will share expertise and discuss technical issues regarding predictive models.
- State insurance regulators will maintain confidentiality, where appropriate, regarding predictive models.

D. Regulatory Review of Predictive Models

The legal and regulatory constraints (including state variations) are likely to be more evolved, and challenging, for personal automobile and home insurance. Through review of these personal lines, the knowledge needed to review predictive models and guidance may be transferrable when the review involves GLMs applied to other lines of business. Modeling depends on context, so the GLM reviewer has to be alert for data challenges and business applications that differ from the more familiar personal lines. For example, compared to personal lines, modeling rates in commercial lines is more likely to involve low volumes of historical data, dependence on advisory loss costs, unique large accounts with large deductibles, and package products that create policies from numerous line-of-business and coverage building blocks.

Commercial lines commonly use individual risk modifications following experience, judgment, and/or expense considerations. A regulator may never see how models impact commercial excess and surplus lines filings.

Best practices will help the regulator understand if a predictive model is cost based, if the predictive model is compliant with state law, and how the model improves the company’s rating plan. Best practices can also increase the consistency among the regulatory review processes used across states and improve the efficiency of each regulator’s review thereby assisting companies in getting their products to market faster. With this in mind, the regulator’s review of predictive models should:

1. Ensure that the selected rating factors based on the model or other analysis produce rates that are not excessive, inadequate, or unfairly discriminatory.
   a. Review the overall rate level impact of the proposed revisions to rate level indications provided by the filer.
   b. Review the premium disruption for individual policyholders and how the disruptions can be explained to individual consumers.
   c. Review the individual input characteristics to and output factors from the predictive model (and its sub-models), as well as, associated selected relativities to ensure they are not unfairly discriminatory.

2. Obtain a clear understanding of how the data used to build and validate the model, and thoroughly review all other aspects of the model, including assumptions, adjustments, variables, submodels used as input, and resulting output.
   a. Determine that individual input characteristics to a predictive model and their resulting rating factors are related to the expected loss or expense differences in risk.
   b. Determine that the data used as input to the predictive model is accurate, including a clear understanding how missing values, erroneous values and outliers are handled.
   c. Determine that any adjustments to the raw data are handled appropriately, including but not limited to, trending, development, capping, removal of catastrophes.
   d. Obtain a clear understanding of how often each risk characteristic, used as input to the model, is updated and whether the model is periodically rerun so model output reflects changes to non-static risk characteristics.
   e. Obtain a clear understanding of how the selected predictive model was built.
   f. Determine whether internal and external data used in relation to the model is compatible with practices allowed in the jurisdiction and do not reflect characteristics prohibited in the state.

3. Evaluate how the model interacts with and improves the rating plan.

© 2019 National Association of Insurance Commissioners
a. Obtain a clear understanding of the characteristics that are input to a predictive model (and its sub-models), their relationship to each other and their relationship to non-modeled characteristics/variables used to calculate a risk’s premium.
b. Obtain a clear understanding why the insurer believes this type of model works in an insurance risk application.
c. Obtain a clear understanding of how model output interacts with non-modeled characteristics/variables used to calculate a risk’s premium.
d. Obtain a clear understanding of how the predictive model was integrated into the insurer’s state rating plan and how it improves that plan.
e. For predictive model refreshes, determine whether sufficient validation was performed to ensure the model is still a good fit.

4. Enable competition and innovation to promote the growth, financial stability, and efficiency of the insurance marketplace.
   a. Enable innovation in the pricing of insurance through acceptance of predictive models, provided they are actuarially sound and in compliance with state laws.
   b. Protect the confidentiality of filed predictive models and supporting information in accordance with state law.
   c. Review predictive models in a timely manner to enable reasonable speed to market.

E. Information Needed to Follow Best Practices

To assist the regulator in following each best practice, the Casualty Actuarial and Statistical Task Force created a white paper titled *Regulatory Review of Predictive Models*. The paper contains a list of information elements and considerations that should be useful during the review of a model underlying a rating plan. To further assist the regulator, the information elements were mapped into the best practices listed above in Section XV of the paper.

Note that, in the white paper, CASTF focused on the GLM since it is a commonly used predictive model in insurance applications, particularly in building an insurance product’s rating plan. Combined with the fact most property and casualty regulators are very concerned about personal lines, the white paper is focused on the review of GLMs for personal auto mobile and home insurance rate making applications.

Guidance offered here might be useful (with deeper adaptations) when starting to review different types of predictive models. If the model is not a GLM, some listed items might not apply. For example, not all predictive models generate p-values or F-tests. Depending on the model type, other considerations might be important. When transferring guidance to other lines of business and other types of model, unique considerations may arise depending on the context in which a predictive model is proposed to be deployed, the uses to which it is proposed to be put, and the potential consequences for the insurer, its customers and its competitors. This paper does not delve into these possible considerations, but regulators should be prepared to address them as they arise.

F. Confidentiality

Regulatory reviewers are required to protect confidential information in accordance with applicable State law. However, insurers should be aware that a rate filing might become part of the public record. Each state determines the confidentiality of a rate filing, supplemental material to the filing, when filing information might become public, the procedure to request that filing information be held confidentially, and the procedure by which a public records request is made. It is incumbent on an insurer to be familiar with each state’s laws regarding the confidentiality of information submitted with their rate filing.

Advisory Organizations – (No change is proposed.)

Workers’ Compensation Special Rules – (No change is proposed.)

Premium Selection Decisions

- Indicated Rate Change vs. Selected Rate Change

After applying credibility, the indicated rate change should reflect the company’s best estimate of their premium needs given their current or expected book of business. However, insurance companies also have other business considerations including competition, marketing, legal concerns, impact of the rate change on retention, etc. A company might wish to deviate from their indicated rate change and should justify those decisions, within the constraints of the law.

© 2019 National Association of Insurance Commissioners 31
Capping and Transition Rules

With advances in technology, it is possible for companies to introduce capping of rates on individual policies with an aim toward gradually increasing policyholders’ rates, rather than making large modifications all at one time. Similarly, premiums are often proposed to be modified when an insurer acquires another company’s book of business or decides to move from or to an advisory organization’s plan. These types of proposed capping are sometimes called “renewal premium capping,” “rate capping,” “a rate stability program,” or “transition rules.”

Transition rules for individual policyholders can get quite complex and you need to be aware of your state’s positions on premium capping rules. Any premium capping and transition rules require weighing the pros and cons of the potential for unfair discrimination (with some customers not paying the rate commensurate with the risks they have) vs. rate stability for existing policyholders.

If premium capping or transition rules are allowed, additional decisions will need to be made:

- Which rates should get capped?
- Do rate decreases get capped? If so, what is the impact if the policyholder asks to be quoted as new business?
- Do all rate increases get capped or only above a certain percentage?
- How much time will lapse or how many renewal cycles will occur before the new rates are in place or different rating plans are merged?
- Should the insured be told what the final premium will be once no more capping is applied?
- How would exposure change be addressed? If the policyholder buys a new car or changes their liability limits, what is the impact on their rate capping?
- How many rate-capping rules can be implemented at any given time?

When premium capping or transition rules have been incorporated, future indicated rate changes and rating factor analyses need to properly reflect the fully approved rate changes. If the overall approved rate change was +10%, yet capping resulted in only 8% being implemented in the first year, the remaining amount to recognize the full 10% should be reflected in the premium on-level adjustment. Otherwise, the indicated rate would be redundant.

Some states encourage more frequent filing of rate changes that can help to avoid the need of premium capping and transition rules. Some states might prefer capping of individual rating variables, rather than capping for individual policyholders.

Installment Plans – (No change is proposed.)

Policy Fees – (No change is proposed.)

Potential Questions to Ask Oneself as a Regulator

Every filing will be different and will result in different regulatory analyses. But the following are some questions the regulator might ask oneself in a rate filing review:

1. Regarding data:
   a. Is the data submitted with the filing enough information for a regulatory review?
   b. Is the number of years of experience appropriate?
   c. Did the company sufficiently analyze and control their quality of data?

2. Regarding the support and justification of rates:
   a. Did they propose rate changes without justification?
   b. Are proposals based on judgment or competitive analysis? If so, are the results reasonable and acceptable? Are there inappropriate marketing practices?
   c. Are the assumptions (loss development, trend, expense load, profit provision, credibility etc.) used to develop the rate indication appropriate? Are they supported with data and are deviations from data results sufficiently explained?

© 2019 National Association of Insurance Commissioners
d. Is the weighting of data by year (or credibility) properly justified or does it appear random?
   • Is there more weight being placed on data in one year solely because it produces a higher indicated rate change?
   • If there are two indications being weighted together and one is for a rate increase and one is a rate decrease, is the weighting justified?

e. Is there a satisfactory explanation about why a proposed rate change deviates from the indicated rate change?

3. Regarding differences in assumptions from previous filings:
   a. Have methodologies changed significantly?
   b. Are assumptions for the weighting of years or credibility significantly different? Or does there appear to be some manipulation to the rate indication?

4. Is there unfair discrimination?
   a. Do classifications comply with state requirements?
   b. Are proposed rates established so that different classes will produce the same underwriting results?
   c. If predictive models are used in the rating plan, are there concerns related to input variables that are prohibited or proxies for prohibited variables?

5. What do you need to communicate?
   a. Can you explain why you are taking a specific action on the filing?
   b. What do you need to tell the Consumer Services Department?
      • Can you explain the impact of the rate change on current business? How big is the company and how much of the market is impacted?
      • What are the biggest changes in the filing (and the ones on which consumer calls might be expected)?
      • What is the maximum rate change impact on any one policyholder?

Questions to Ask a Company

If you remain unsatisfied that the company has satisfactorily justified the rate change, then consider asking additional questions of the company. Questions should be asked of the company when they have not satisfied statutory or regulatory requirements in the state or when any current justification is inadequate and could have an impact on the rate change approval or the amount of the approval.

If there are additional items of concern, the company can be notified so they will make appropriate modifications in future filings.

The CASTF white paper, *Regulatory Review of Predictive Models*, documents questions that a regulator may want to ask when reviewing a model. These questions are listed in the Predictive Model – Information for Regulatory Review section of the white paper. Note that although the white paper focuses on GLMs for personal automobile and home insurance, some of the concepts may be transferable to other types of models and other lines of business.

Additional Ratemaking Information
The Casualty Actuarial Society (CAS) and the Society of Actuaries (SOA) have extensive examination syllabi that contain a significant amount of ratemaking information, on both the basic topics covered in this chapter and on advanced ratemaking topics. The CAS and SOA websites contain links to many of the papers included in the syllabus. Recommended reading is the *Foundations of Casualty Actuarial Science*, which contains chapters on ratemaking, risk classification, and individual risk rating.

**Other Reading**

Some additional background reading is recommended:

  - Chapter 1: Introduction
  - Chapter 3: Ratemaking
  - Chapter 6: Risk Classification
  - Chapter 9: Investment Issues in Property-Liability Insurance
  - Chapter 10: Only the section on Regulating an Insurance Company, pp. 777–787
- Casualty Actuarial Society (CAS) Statements of Principles, especially regarding property and casualty ratemaking.
- Association of Insurance Compliance Professionals: “Ratemaking—What the State Filer Needs to Know.”
- Review of filings and approval of insurance company rates.

**Summary**

Rate regulation for property/casualty lines of business requires significant knowledge of state rating laws, rating standards, actuarial science, statistical modeling and many data concepts.

- Rating laws vary by state, but the rating laws are usually grouped into prior approval, file and use or use and file (competitive), no file (open competition), and flex rating.
- Rate standards typically included in the state rating laws require that “Rates shall not be inadequate, excessive, or unfairly discriminatory.”
- A company will likely determine their indicated rate change by starting with historical years of underwriting data (earned premiums, incurred loss and loss adjustment expenses, general expenses) and adjusting that data to reflect the anticipated ultimate level of costs for the future time period covered by the policies. Numerous adjustments are made to the data. Common premium adjustments are on-level premium, audit, and trend. Common loss adjustments are trend, loss development, Catastrophe/large loss provisions, and an adjusting and other (A&O) loss adjustment expense provision. A profit/contingency provision is also calculated to determine the indicated rate change.
- Once an overall rate level is determined, the rate change gets allocated to the classifications and other rating factors.
- Individual risk rating allows manual rates to be modified by an individual policyholder’s own experience.
- Advisory organizations provide the underlying loss costs for companies to be able to add their own expenses and profit provisions (with loss cost multipliers) to calculate their insurance rates.
- CAS provides guidance and guidelines for the numerous actuarial decisions and standards employed during the development of rates.
- NAIC model laws also include special provisions for workers’ compensation business, penalties for not complying with laws, and competitive market analysis to determine whether rates should be subject to prior approval.

Best practices for reviewing predictive models are provided in the CASTF white paper titled *Regulatory Review of Predictive Models*. Although the white paper focuses on GLMs for personal automobile and home insurance, some of the concepts may be transferrable to other types of models and other lines of insurance.

While this chapter provides an overview of the rate determination/actuarial process and regulatory review, state statutory or administrative rule may require the examiner to adopt different standards or guidelines than the ones described.
No additional changes are proposed to the Product Filing Review Handbook.

IX. PROPOSED STATE GUIDANCE

TBD – placeholder for guidance for rate filings that are based on predictive models

This paper acknowledges that different states will apply the guidance within it differently, based on variations in the legal environment pertaining to insurance regulation in those states, as well as the extent of available resources, including staff members with actuarial and/or statistical expertise, the workloads of those staff members, and the time that can be reasonably allocated to reviews of predictive models. States with prior approval authority over personal lines rate filings often already require information that is expressed in this paper. However, states including those with and without prior approval authority – may also use the guidance in this paper to choose which model elements to focus on in their reviews and/or to train new reviewers, as well as to gain an enhanced understanding of how predictive models are developed, supported, and deployed in their markets. Ultimately, the insurance regulators within each state will decide how best to tailor the guidance within this paper to achieve the most effective and successful implementation, subject to the framework of statutes, regulations, precedents, and processes that comprise the insurance regulatory framework in that state.

X. OTHER CONSIDERATIONS

During the development of this guidance, topics arose that are not addressed in this paper. These topics may need addressing during the regulator’s review of a predictive model. A few of these “topics” may be discussed elsewhere within the NAIC as either technical or policy matters. All of these topics should probably be addressed, if at all possible, by each state on a case-by-case basis. A sampling of topics for consideration in this section include: Below is a listing of topics that CASTF thought might be important for future discussion and consideration but are beyond the scope of this paper, as well as CASTF’s current charges:

• TBD: Discuss when rating variables or rating plans become too granular? How is granularity handled during the development of the model and during the selection of rate relativities? File in a rating plan supported by a model?
  o The granularity of data refers to the size in which data fields are sub-divided. For example, data could be at the state level or could be subdivided into county or further into zip/ZIP code or even census tracks. Insurers who have been instituting data warehouse initiatives that greatly improved the granularity and accessibility of data that could be analyzed for ratemaking purposes. So, despite the fact that sophisticated statistical techniques existed much earlier than these data warehouses, it was the circumstances of enhanced computing power and better data that enabled its usage in classification ratemaking. Perhaps the most important trigger in the widespread adoption of multivariate methods was competitive pressure. When one or more companies implement improved classification ratemaking, they gain a competitive advantage and put the rest of the industry in a position of adverse selection and decreased profitability. [footnote: Basic Ratemaking, Fifth Edition, May 2016; Geoff Werner, FCAS, MAAA and Claudine Modlin, FCAS, MAAA]
  o The science of classification requires balancing two objectives: grouping risks into a sufficient number of levels to ensure the risks within each group are homogeneous while being careful not to create too many granularity defined groups that may lead to instability in the estimated costs. [footnote: Basic Ratemaking, Fifth Edition, May 2016; Geoff Werner, FCAS, MAAA and Claudine Modlin, FCAS, MAAA]
  o Concern has been expressed that when fields are subdivided too finely, model results may be less reliable.
    • It is commonly assumed that having a larger volume of data is preferable. However, even with a larger volume of data, if the model is overly granular the more data you have, the better. But, the more granular the data, it may be the harder/difficult it may be to see the forest for the trees. More granular data used as input to predictive models may make it easier to measure short-term effects, but it can also make it harder to measure long-term effects, due to-- because of more relatively greater noise in the data. However, more granular data may make anomalies in the data more apparent and make it easier to scrub the data.
    • Therefore, it may be of value to provide guidance around granularity, such as: When are rating variables or rating plans too granular? How is granularity handled during the development of the model or during the selection of rate relativities?
As adopted by the Casualty Actuarial and Statistical (C) Task Force on XX/XX/XX

Draft: 10/15/2019

1. **TBD:** Discuss the regulator’s scientific mindset of unbiased and open inquiry and its relevance to the best practice white paper.
   - This white paper has taken the position that regulatory reviewers of models, both actuaries and non-actuaries, especially when they review predictive models, are in a prime position to be the torchbearers for the scientific approach for an unbiased view, by maintaining the commitment to open but rigorous, systematic, and principled inquiry and exploration.
   - This white paper does not prescribe any specific answers regarding which treatments are to be considered rational or logical. Such answers cannot be assumed-determined without considering the context of a given jurisdiction’s laws, marketplace, and the specific nature of insurers’ proposals. Therefore, to preempt any arguments by some interested parties that the paper may prescribe specific solutions or restrictions – it clearly does not.
   - As actuaries, if regulators adopt the discipline called “actuarial science,” it is incumbent upon us regulators to adopt the proper scientific mindset of open inquiry – where no questions are off limits and continued systematic exploration and progress are the hallmarks of the scientific approach. Any insistence that certain questions must not be asked, or certain concepts must not be explored, entails a departure from the realm of science into the realm of dogma. If pursued, it would limit the role of regulators and quickly deprive them of broader relevance.

2. **TBD:** Discuss correlation vs causality in general and in relation to Actuarial Standard of Practice (ASOP) 12.
   - There were many criticisms during each exposure of this white paper that this paper goes beyond the requirement of Actuarial Standard of Practice #12 and establishes a new standard for the company’s actuaries. This topic may need to be explored further by states collectively through NAIC or on a case-by-case state basis. What a state does with the results of a discussion of rational or logical connections between particular attributes and the risk of insurance loss is subject to the framework of statutes, regulations, precedents, and processes that comprise the insurance regulatory framework in that state.
   - The very act of discussion of the rational, logical, or plausible relationships of individual risk attributes to the risk of insurance loss—and all related implications, such as perception by consumers, legislators, and media; philosophical considerations of fairness; interactions with public policy as determined by the relevant policymaking bodies; and relevance to the evolution of the insurance industry, consumer products, and overall impacts on the incentives and opportunities available to consumers—is crucial to engage in and continue to do so for as long as new predictive models are being developed, new variables are being introduced, and consumer premiums as well as insurer underwriting decisions are being affected. In other words, the discussion needs to continue indefinitely in a variety of venues and evolve along with the industry and the broader society. We, as insurance professionals, cannot insulate ourselves from participation in the conceptual discourse.
   - This white paper, in general, establishes that a rating modeled variable should not only be correlated to expected costs but that there should be a rational explanation as to why the correlation exists. While it is difficult to prove causation, and such a proof is not a standard against which rate filings are evaluated in any jurisdiction, there is an immense difference of both degree and kind between proving causation and discussing a rational or logical connection between a particular variable and the risk of insurance loss. It is a non sequitur to assert that the lack of requirement for the former (proof) confers immunity upon insurers in regard to the latter (discussion and expression of plausibility).
   - Discussion of the Actuarial Standards of Practice has been consciously excluded from this paper for a number of reasons. First, legally obligated to adhere to the ASOPs, Models are created by, supported by, and filed by professionals who, if they are not actuaries, e.g., data scientists, modelers, and other professions, who are not bound by ASOPs. Second, ASOPs do not supersede state laws. Third, ASOPs represent a floor, not a ceiling, for regulatory actuaries who also need to consider state laws and public policy concerns.
   - Finally, ASOP 12 specifically is often misquoted by filers who cite the phrase, in Section 3.2.2 of ASOP 12, that states “it is not necessary for the actuary to establish a cause and effect relationship between the risk characteristic and expected outcome in order to use a specific risk characteristic” while ommitting the leading phrase “while the actuary should select risk characteristics that are related to expected outcomes.”

3. **TBD:** Will the following guidance provided in this white paper increase or pressure state regulators to take adverse actions?
   - Discussion of data mining as it being in conflict with the standard scientific model and the increase in “false positives.”

© 2019 National Association of Insurance Commissioners
• Throughout this white paper, the regulator asks the modeler to go beyond correlation and document their basic, causal understanding of how variables used in a model or rating plan are related to risk. A correlation alone is not the final arbiter of the validity of findings, but causal understanding can be employed to assess which correlations may be entirely due to chance, what are non-causal relationships, and which are most likely to be enduring causal relationships. Though this white paper does not delve deeply into how these relationships can be identified and documented, the paper does ask the modeler to provide their understanding of these relationships. The future consideration is whether the regulator should take a deeper dive into the causal relationships of variables used in a model or rating plan.

• The American Statistical Association (ASA) expressed some degree of alarm at approaches similar to data mining (Wasserstein and Lazer, 2016). In a formal statement of the ASA, the association warned against a purely "cookbook" approach to statistics: "... a p-value near .05 taken by itself offers only weak evidence of the null hypothesis" (page 129). Lastly, the ASA warned strongly against an over reliance on data mining: "Cherry-picking promising findings, also known by such terms as data dredging, significance chasing, and "p-hacking," leads to a spurious excess of statistically significant results, and should be vigorously avoided" (page 131).

• A problem that will increase significantly with the increased adoption of data mining techniques and the increasing growing availability of very large data sets that dwarf anything available even just a decade ago is that data mining will dramatically increase the rate of "false positives" - the technique Data mining will inevitably churn up numerous associations between variables that are simply random, non-meaningful correlations resulting purely from chance. The apparent disregard of causality that seems common among practitioners of data mining techniques will significantly magnify the problem. Causality forms the basis of the standard model of all natural and social sciences. Evaluations of models should consider the nature of observed relationships within the context of prior substantive knowledge.

• Because of these issues regarding data-mining and false positives stated in the prior paragraphs, throughout this white paper, the regulator asks the modeler to go beyond correlation and document their basic, causal understanding of how variables used in a model or rating plan are related to risk. A correlation alone is not the final arbiter of the validity of findings, but causal understanding can be employed to assess which correlations may be entirely due to chance, what are non-causal relationships, and which are most likely to be enduring causal relationships. Though this white paper does not delve deeply into how these relationships can be identified and documented, the paper does ask the modeler to provide their understanding of these relationships. The future consideration is whether the regulator should take a deeper dive into the causal relationships of variables used in a model or rating plan.

• TBD: Explain how the insurer will help educate consumers to mitigate their risk. Discuss the multitude of Regulators are often responding to consumer inquiries to which regulators respond regarding how a policy premium is calculated.

• 1.b. Review the premium disruption for individual policyholders and how the disruptions can be explained to individual consumers. …and information elements that may assist the regulator's and consumer's understanding of the premium being charged.

• C.2.a Provide an explanation how the characteristics/rating variables, included in the filed rating plan, logically and intuitively relate to the risk of insurance loss (or expense) for the type of insurance product being priced. Include a discussion of the relevance each characteristic/rating variable has on consumer behavior that would lead to a difference in risk of loss (or expense).

• C.7.f Explain how the insurer will help educate consumers to mitigate their risk.

• C.7.h Identify sources to be used at "point of sale" to place individual risks within the matrix of rating system classifications. How can a consumer verify their own "point-of-sale" data and correct any errors?

• C.7.i Provide the regulator with a description of how the company will respond to consumers' inquiries about how their premium was calculated.

• The main challenge to consumers is lack of transparency trying to understand the data and analytics being used to determine their eligibility for products and the price they are being charged. It may not be clear to the consumer how they are being underwritten or what behaviors they can modify or steps they can take to get a better rate. A potential issue with pricing based on predictive analytics is that it can lead to more granular
Discuss guidelines for insurers' handling of consumer-generated data in insurance transactions.

- TBD: Discuss guidelines for insurers' handling of data from external sources.

- Discuss guidelines for insurers' handling of consumer-generated data in insurance transactions.

  - Does a consumer have the right to know what data is being used to determine the consumers' premium, where that data came from, and how the consumer can address errors in the data? To what extent is the insurer accountable for the quality of the data used to calculate a consumer's premium, whether that data is internal or external to the insurer's operations? To what extent should the insurer inform the consumer about the types of data used in the rating process? How can a consumer verify the accuracy of the data used in the rating process? How can a consumer correct any errors in the data?

- Discuss the development of new tools and techniques for monitoring consumer market outcomes resulting from insurers' use of Big Data analytics in property and casualty rating plans.

  - What tools and techniques are insurers using to monitor consumer market outcomes? How are these tools and techniques improving the accuracy and reliability of insurer predictive models? How are insurers using these tools and techniques to improve the fairness and transparency of insurer rating plans?

- Discuss the costs associated with the use of Big Data analytics in property and casualty rating plans.

  - What are the costs associated with the use of Big Data analytics in property and casualty rating plans? How are insurers justifying these costs to consumers? How are insurers ensuring that these costs are reasonable and transparent?

- Discuss the use of Big Data analytics in other insurance lines.

  - What role is Big Data analytics playing in other insurance lines, such as life, health, and accident and health insurance? How are insurers using Big Data analytics to improve the accuracy and fairness of their rating plans in these lines of business?
recommendations. To ensure that data brokers and vendors of algorithms who are engaged in advisory organization activities are properly licensed and supervised, advisory organization model laws could be revised. [Center for Economic Justice, comments to the NAIC Accelerated Underwriting (A) Working Group, September 29, 2019]

- TBD: Discuss the need for NAIC to update and strengthen information-sharing platforms and protocols.
- Discuss paper topics beyond GLMs and personal automobile and home insurance applications.
  - The scope of this white paper was narrowed to GLMs as used in personal automobile and home insurance rating applications. Many commenters expressed concern that the paper's scope is too narrow. NAIC may want to expand these best practices or create new best practices for other lines of business, other insurance applications (other than personal automobile and home filings), and other types of models.

XI. RECOMMENDATIONS GOING FORWARD

The following are examples of topics that may be included in the recommendations:

- TBD: Discuss confidentiality as it relates to filings submitted via SERFF.
- TBD: Discuss confidentiality as it relates to state statutes and regulations.
- TBD: Discuss policyholder disclosure when complex predictive model underlies a rating plan.
- TBD: Discuss the need for NAIC to update and strengthen information-sharing platforms and protocols.
- TBD: Determine the means available to a consumer to correct or contest individual data input values that may be in error.
- TBD: Given an insurer’s rating plan relies on a predictive model and knowing all characteristics of a risk, discuss a regulator’s ability/need to audit/calculate the risk’s premium without consultation with the insurer.
- Other TBDs
APPENDIX A – BEST PRACTICE DEVELOPMENT

Best practices development is a method for reviewing public policy processes that have been effective in addressing particular issues and could be applied to a current problem. This process relies on the assumptions that top performance is a result of good practices and these practices may be adapted and emulated by others to improve results. The term “best practice” can be a misleading one due to the slippery nature of the word “best”. When proceeding with policy research of this kind, it may be more helpful to frame the project as a way of identifying practices or processes that have worked exceptionally well and the underlying reasons for their success. This allows for a mix-and-match approach for making recommendations that might encompass pieces of many good practices.

Researchers have found that successful best-practice analysis projects share five common phases:

A. Scope
   The focus of an effective analysis is narrow, precise and clearly articulated to stakeholders. A project with a broader focus becomes unwieldy and impractical. Furthermore, Bardach urges the importance of realistic expectations in order to avoid improperly attributing results to a best practice without taking into account internal validity problems.

B. Identify Top Performers
   Identify outstanding performers in this area to partner with and learn from. In this phase, it is key to recall that a best practice is a tangible behavior or process designed to solve a problem or achieve a goal (i.e. reviewing predictive models contributes to insurance rates that are not unfairly discriminatory). Therefore, top performers are those who are particularly effective at solving a specific problem or regularly achieve desired results in the area of focus.

C. Analyze Best Practices
   Once successful practices are identified, analysts will begin to observe, gather information and identify the distinctive elements that contribute to their superior performance. Bardach suggests it is important at this stage to distill the successful elements of the process down to their most essential idea. This allows for flexibility once the practice is adapted for a new organization or location.

D. Adapt
   Analyze and adapt the core elements of the practice for application in a new environment. This may require changing some aspects to account for organizational or environmental differences while retaining the foundational concept or idea. This is also the time to identify potential vulnerabilities of the new practice and build in safeguards to minimize risk.

E. Implementation and Evaluation
   The final step is to implement the new process and carefully monitor the results. It may be necessary to make adjustments, so it is likely prudent to allow time and resources for this. Once implementation is complete, continued evaluation is important to ensure the practice remains effective.


© 2019 National Association of Insurance Commissioners
**APPENDIX B - GLOSSARY OF TERMS**

<table>
<thead>
<tr>
<th>Adjusting Data</th>
<th>Adjusting data refers to any changes made to the raw data. For example, capping losses, on-leveling, binning, transformation of the data, etc. This term includes scrubbing of the data.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregated Data</td>
<td>Aggregated data is from the insurer’s data banks without modification (e.g., not scrubbed or transformed). Aggregated datasets are those compiled prior to data selection and model building.</td>
</tr>
<tr>
<td>Composite Characteristic</td>
<td>A composite characteristic is a combination of two or more individual risk characteristics. Composite characteristics are used to create composite variables.</td>
</tr>
<tr>
<td>Composite Score</td>
<td>A composite score is a number arrived at through the combination of multiple variables by means of a sequence of mathematical steps - for example, a credit-based insurance scoring model.</td>
</tr>
<tr>
<td>Composite Variable</td>
<td>A composite variable is a variable created by combining two or more individual risk characteristics of the insured into a single variable.</td>
</tr>
<tr>
<td>Continuous Variable</td>
<td>A continuous variable is a numeric variable that represents a measurement on a continuous scale. Examples include age, amount of insurance (in dollars), and population density.</td>
</tr>
<tr>
<td>Control Variable</td>
<td>Control variables are variables whose relativities are not used in the final rating algorithm but are included when building the model. They are included so that other correlated variables do not pick up their signal. For example, state and year are frequently included in countrywide models as control variables so that the different experiences and distributions between states and across time do not influence the rating factors used in the final rating algorithm.</td>
</tr>
<tr>
<td>Correlation Matrix</td>
<td>A correlation matrix is a table showing correlation coefficients between sets of variables. Each random variable (X) in the table is correlated with each of the other variables in the table (Y). This allows you to see which pairs of variables have the highest correlation. Below is a sample correlation matrix showing correlation coefficients for combinations of 5 variables B1:B5. The table shows that variables B2 and B4 have the highest correlation coefficient (0.96) in this example. The diagonal of the table is always a set of ones, because the correlation coefficient between a variable and itself is always 1. You could fill in the upper-right triangle, but these would be a mirror-image of the lower-left triangle (because correlation between B1 and B2 is the same as between B2 and B1). In other words, a correlation matrix is also a symmetric matrix.</td>
</tr>
</tbody>
</table>

**Commented [WL29]:** “Aggregation” implies that data is summarized or compiled in some way, whether or not it comes straight from the insurer’s data banks, whether or not it has been modified. Perhaps the author confused between aggregation and raw??? Raw data is defined below.  

**Commented [WL30]:** Is there another way to say this? Seems unclear.
**Data Dredging** - Data dredging is also referred to as data fishing, data snooping, data butchery, and p-hacking. It is the misuse of data analysis to find patterns in data that can be presented as statistically significant when, in fact, there is no real underlying effect. This data dredging is done by performing many statistical tests on the data and only paying attention to those that come back with significant results. Data dredging is in conflict with hypothesis testing, which entails performing at the most a handful of tests to determine the validity of the hypothesis about an underlying effect, instead of stating a single hypothesis about an underlying effect before the analysis and then conducting a single test for it. The process of data dredging involves automatically testing huge numbers of hypotheses about a single data set by exhaustive variables that might show a correlation, and perhaps for groups of cases or observations that show differences in their mean or in their breakdown by some other variable.

Conventional tests of statistical significance are based on the probability that a particular result would arise if chance alone were at work, and necessarily accept some risk of mistaken conclusions of a certain type (mistaken rejections of the null hypothesis). This level of risk is called the significance. When large numbers of tests are performed, some produce false results of this type, hence 5% of randomly chosen hypotheses turn out to be significant at the 5% level, 1% turn out to be significant at the 1% significance level, and so on, by chance alone. When enough hypotheses are tested, it is virtually certain that some will be statistically significant but misleading, since almost every data set with any degree of randomness is likely to contain (for example) some spurious correlations. If they are not cautious, researchers using data mining techniques can be easily misled by these results.

The multiple comparisons hazard is common in data dredging. Moreover, subgroups are sometimes explored without alerting the reader to the number of questions at issue, which can lead to misinformed conclusions.

**Data Source** - A data source is the original repository of the information used to build the model. For example, information from an vendor, credit bureaus, government websites, a sub-model, verbal information provided to agents, external sources, consumer information databases, and so on. The data source is the foundation for the model and is used to build a sub-model.

**Discrete Variable** - A discrete variable is a variable that can only take on a countable number of values/categories. Examples include number of claims, marital status, and gender.

**Discrete Variable Level** - Discrete variables are generally referred to as "factors" (not to be confused with rating factors), with values that each factor can take being referred to as "levels." For example, gender can be a discrete variable with levels male and female.

**Double-Lift Chart** - Double-lift charts are similar to simple quantile plots, but rather than sorting based on the predicted loss cost of each model, the double lift chart sorts based on the ratio of the two models' predicted loss costs. Double lift charts directly compare the results of two models.

**Exponential Family** - The exponential family is a class of distributions that share the same density form and have certain properties that are used in fitting GLMs. It includes many well-known distributions, such as the Normal, Poisson, Gamma, Tweedie, and Binomial distributions, to name a few.

**Fair Credit Reporting Act** - The Fair Credit Reporting Act (FCRA), 15 U.S.C. § 1681 (FCRA) is U.S. Federal Government legislation enacted to promote the accuracy, fairness, and privacy of consumer information contained in the files of consumer reporting agencies. It was intended to protect consumers from the willful and/or negligent inclusion of inaccurate information in their credit reports. To that end, the FCRA regulates the collection, dissemination and use of consumer information, including consumer credit information.[4] Together with the Fair Debt Collection Practices Act (FDCPA), the FCRA forms the foundation of consumer rights law in the United States. It was originally passed in 1970 and is enforced by the US Federal Trade Commission, the Consumer Financial Protection Bureau and private litigants.
Generalized Linear Model - Generalized linear models (GLMs) are a means of modeling the relationship between a variable whose outcome we wish to predict and one or more explanatory variables. The predicted variable is called the target variable and is denoted \( y \). In property/casualty insurance ratemaking applications, the target variable is typically one of the following:

- Claim count (or claims per exposure)
- Claim severity (i.e., dollars of loss per claim or occurrence)
- Pure premium (i.e., dollars of loss per exposure)
- Loss ratio (i.e., dollars of loss per dollar of premium)

For quantitative target variables such as those above, the GLM will produce an estimate of the expected value of the outcome. For other applications, the target variable may be the occurrence or non-occurrence of a certain event. Examples include:

- Whether or not a policyholder will renew his/her policy.
- Whether a submitted claim contains fraud.

For such variables, a GLM can be applied to estimate the probability that the event will occur.

The explanatory variables, or predictors, are denoted \( x_1, \ldots, x_p \), where \( p \) is the number of predictors in the model. Potential predictors are typically any policy term or policyholder characteristic that an insurer may wish to include in a rating plan. Some examples are:

- Type of vehicle, age, or marital status for personal auto insurance.
- Construction type, building age, or amount of insurance (AOI) for home insurance. [15]

Geodemographic - Geodemographics is the study of the population and its characteristics, divided according to regions on a geographical basis. This involves application of clustering techniques to group statistically similar neighbourhoods and areas with the assumption that the differences within any group should be less than the difference between groups. While the main source of data for a geodemographic study is the census data, the use of other sources of relevant data is also prevalent. Geodemographic segmentation (or analysis) is a multivariate statistical classification technique for discovering whether the individuals of a population fall into different groups by making quantitative comparisons of multiple characteristics with the assumption that the differences within any group should be less than the differences between groups.

Geodemographic segmentation is based on two principles:

1. People who live in the same neighborhood are more likely to have similar characteristics than are two people chosen at random.
2. Neighborhoods can be categorized in terms of the characteristics of the population that they contain. Any two neighborhoods can be placed in the same category, i.e., they contain similar types of people, even though they are widely separated.
Granularity of Data - divided

Granularity of data is the level of segmentation at which the data is grouped or summarized. It reflects the level of detail used to slice and dice the data. The granularity of data refers to the size in which data fields are sub-divided.\[yy\]

For example, a postal address can be recorded, with coarse granularity, as a single field:
- address = 200 2nd Ave. South #358, St. Petersburg, FL 33701-4313 USA

Or, with fine granularity, as multiple fields:
- street address = 200 2nd Ave. South
- city = St. Petersburg
- state = FL
- postal code = 33701-4313
- country = USA

Or, even finer granularity:
- street = 2nd Ave. South
- address number = 200
- suite/apartment number = #358
- city = St. Petersburg
- state = FL
- postal-code = 33701
- postal-code-add-on = 4313
- country = USA

Home Insurance - Home insurance covers damage to the property, contents, and outstanding structures (if applicable), as well as loss of use, liability and medical coverage. The perils covered, and the amount of insurance provided and other policy characteristics are detailed in the policy contract. [16]

Insurance Data - Data collected by the insurance company.

Interaction Term - Two predictor variables are said to interact if the effect of one of the predictors on the target variable depends on the level of the other predictor. For instance, rather than defining the linear predictor as $\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2$, they could set $\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$.

The following two plots of modeled personal auto bodily injury pure premium by age and gender illustrate this effect. The plots are GLMs, built using the same fictional dataset, with the only difference between the two being that the second model includes the Age *Gender interaction term while the first does not. Notice that the male curve in the first plot is a constant multiple of the female curve, while in the second plot the ratios of the male to female values differ from age to age.

Lift Chart - See definition of quantile plot.
**Linear Predictor** - A linear predictor is the linear combination of explanatory variables \( (X_1, X_2, \ldots, X_k) \) in the model, e.g., \( \beta_0 + \beta_1x_1 + \beta_2x_2 \). \[18\]

**Link Function** - The link function, \( \eta \) or \( g(\mu) \), specifies the link between random and systematic components. It describes how the expected value of the response relates to the linear predictor of explanatory variables; e.g., \( \eta = g(E(Y_i)) = E(Y_i) \) for linear regression, or \( \eta = \logit(\pi) \) for logistic regression. \[19\]

**Missing data** - Missing data occurs when some records contain blanks or "Not Available" or "Null" where variable values should be available.

**Non-Insurance Data** - Non-insurance data is data provided by another party other than the insurance company.

**Offset Variable** - Offset variables (or factors) are model variables with a known or pre-specified coefficient. Their relativities are included in the model and the final rating algorithm, but they are generated from other studies outside the multivariate analysis, and are fixed (not allowed to change) in the model when it is run. The model does not estimate any coefficients for the offset variables, and they are included in the model, so that the estimated coefficients for other variables in the model would be optimal in their presence. Examples of offset variables include limit and deductible relativities that are more appropriately derived via loss elimination analysis. The resulting relativities are then included in the multivariate model as offsets. Another example is using an offset factor to account for the exposure in the records; this does not get included in the final rating algorithm. \[20\]

**Overfitting** - Overfitting is the production of an analysis that corresponds too closely or exactly to a particular set of data and may, therefore, fail to fit additional data or predict future observations reliably. \[21\]

**PCA Approach (Principal Component Analysis)** - The PCA method creates multiple new variables from correlated groups of predictors. Those new variables exhibit little or no correlation between them—thereby making them potentially more useful in a GLM. A PCA in a filing can be described as "a GLM within a GLM." One of the more common applications of PCA is geodemographic analysis, where many attributes are used to modify territorial differentials on, for example, a census block level.

**Personal Automobile Insurance** - Personal automobile insurance is insurance for privately owned motor vehicles and trailers for use on public roads not owned or used for commercial purposes. This includes private auto combinations of passenger auto, motorcycle, financial responsibility bonds, recreational vehicles and/or other personal auto. Policies include any combination of coverage such as the following: auto liability, personal injury protection (PIP), medical payments (MP), uninsured/underinsured motorist (UM/UIM); specified causes of loss, comprehensive, and collision. \[22\]

**Post-model Adjustment** - Post-model adjustment is any adjustment made to the output of the model including but not limited to adjusting rating factors or removal of variables.

**Probability Distribution** - A probability distribution is a statistical function that describes all the possible values and likelihoods that a random variable can take within a given range. The chosen probability distribution is supposed to best represent the likely outcomes.

**Proxy Variable** - A proxy variable is any characteristic variable that is used instead of a variable of interest (when that variable of interest cannot be measured or used directly), to indirectly capture the effect of another characteristic represented by the variable of interest whether or not that characteristic is used in the insurer’s rating plan. In order for a variable to be a good proxy, it must have a close correlation, not necessarily linear, with the variable of interest.
Quantile Plot - A quantile plot is a visual representation of a model’s ability to accurately differentiate between the best and the worst risks. Data is sorted by predicted value from smallest to largest, and the data is then bucketed into quantiles with the same volume of exposures. Within each bucket, calculate the average predicted value and the average actual value are calculated, and for each quantile the actual and the predicted values are plotted. The first quantile contains the risks that the model predicts have the best experience and the last quantile contains the risks predicted to have the worst experience. The plot shows two things: how well the model predicts actual values by quantile, the predicted value should be increasing as the quantile increases, and the lift of the model, the difference between the first and last quantile, which is a reflection of how large it indicates the model’s ability to distinguish between the best and worst risks. By definition, the average predicted values would be monotonically increasing, but the average actual values may show reversals. An example follows:

Rating Algorithm – A rating algorithm is the mathematical or computational component of the rating plan used to calculate an insured’s premiums.

Rating Category – A rating category is the same as a rating characteristic, and can be quantitative or qualitative.

Rating Characteristic – A rating characteristic is a specific risk criterion of the insured used to define the level of the rating variable that applies to the insured. Ex. Rating variable- Driver age, Rating characteristic- Age 42

Rating Factor – A rating factor is the numerical component included in the rate pages of the rating plan’s manual. Rating factors are used together with the rating algorithm to calculate the insured’s premium.

Rating Plan – The rating plan describes in detail how to combine the various components in the rules and rate pages to calculate the overall premium charged for any risk that is not specifically pre-printed in a rate table. The rating plan is very specific and includes explicit instructions, such as:
- the order in which rating variables should be considered;
- how the effect of rating variables is applied in the calculation of premium (e.g., multiplicative, additive, or some unique mathematical expression);
- the existence of maximum and minimum premiums (or in some cases the maximum discount or surcharge that can be applied);
- specifics associated with any rounding that takes place.

If the insurance product contains multiple coverages, then separate rating plans by coverage may apply.

Rating System - The rating system is the insurance company’s IT infrastructure that produces the rates derived from the rating algorithm.

Commented [WL36]: The graph looks like a regression line. A graph showing pure premium or loss ratio by decile based on both the average predicted and the average actual values may be a more appropriate example for the purposes of this white paper.

Commented [WL37]: What is meant by this phrase? Is this referring to pre-printed base rates by territory (for instance), to which the relativities of other rating variables are applied? Aren’t all aspects of a rating plan pre-printed in a rate manual, comprising the entirety of the rating plan? Isn’t the rating plan the whole population of rates, rating variables and associated relativities and all rules and algorithms detailing how those rates and rating relativities combine together to determine an insured’s premium, whether or not pre-printed in a rate table?
Rating Tier - A rating tier is rating based on a combination of rating characteristics rather than a single rating characteristic resulting in a separation of groups of insureds into different rate levels within the same or separate companies. Often, rating tiers are used to differentiate quality of risk, e.g., substandard, standard, or preferred.

Rating Treatment - Rating treatment is the manner in which an aspect of the rating affects an insured's premium.

Rating Variable - A rating variable is a risk criterion of the insured used to modify the base rate in a rating algorithm.

Raw Data - Raw data is data before scrubbing, transformation etc. takes place ... “as is” when received from a source.

Sample Record - A sample record is one line of data from a data source including all variables. For example:

```
<table>
<thead>
<tr>
<th>Record</th>
<th>Age</th>
<th>Dependent</th>
<th>Homeowner</th>
<th>Married</th>
<th>Age</th>
<th>Zip</th>
<th>Property</th>
<th>Type</th>
<th>Risk</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>42</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>500</td>
<td>Garage</td>
<td>Base</td>
<td>1500</td>
<td>2</td>
</tr>
</tbody>
</table>
```

Scrubbed Data - Scrubbed data is data reviewed for errors, where "N/A" has been replaced with a value, and where most transformations have been performed. Data that has been "scrubbed" is now in a useable format to begin building the model.

Scrubbing Data - Scrubbing is the process of editing, amending, or removing data in a dataset that is incorrect, incomplete, improperly formatted, or duplicated.

SME - Subject Matter Expert.

Sub-Model - A sub-model is any model that provides input into another model.

Variable Transformation - A variable transformation is a change to a variable by taking a function of that variable, for example, when age's value is replaced by the value \(\text{age}^2\). The result is called a transformation variable.

Voluntarily Reported Data - Voluntarily reported data is data directly obtained by a company from a consumer. Examples would be data taken directly from an application for insurance or obtained verbally by a company representative.

Univariate Model – A univariate model is a model that only has one independent variable.

Adjusting Data - TBD

Control Factor - TBD

Data source - TBD

Double lift chart - TBD

Exponential Family - TBD

Fair Credit Reporting Act – The Fair Credit Reporting Act (FCRA), 15 U.S.C. § 1681 (FCRA) is U.S. Federal Government legislation enacted to promote the accuracy, fairness, and privacy of consumer information contained in the files of consumer reporting agencies. It was intended to protect consumers from the willful and/or negligent inclusion of inaccurate information in their credit reports. To that end, the FCRA regulates the collection, dissemination and use of consumer information, including...
consumer credit information. Together with the Fair Debt Collection Practices Act (FDCPA), the FCRA forms the foundation of consumer rights law in the United States. It was originally passed in 1970 and is enforced by the US Federal Trade Commission, the Consumer Financial Protection Bureau and private litigants.

To see that this second definition accounts for the interaction, note that it is equivalent to \( \eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \) and to \( \eta = \beta_0 + \beta_1 X_1 + \beta_3 X_2 \), with \( \beta_3 = \beta_1 + \beta_2 \beta_2 \) and \( \beta_2 = \beta_2 + \beta_1 \beta_3 \). Since \( \beta_2 \) is a function of \( X_1 \) and \( \beta_3 \) is a function of \( X_2 \), these two equivalences say that the effect of \( X_1 \) depends on the level of \( X_2 \) and vice versa.

REFERENCES:

XIV.XIII. APPENDIX C – SAMPLE RATE-DISRUPTION TEMPLATE

● First, fill in the boxes for minimum and maximum individual impacts, shaded in light blue. Default values in the cells are examples only.
● The appropriate percent-change ranges will then be generated based on the maximum/minimum changes.
● For every box shaded in light green, replace “ENTER VALUE” with the number of affected insureds within the corresponding change range.

NOTE: Values of Minimum % Change, Maximum % Change, and Total Number of Insureds must reconcile to the Rate/Rule Schedule in SERFF.

<table>
<thead>
<tr>
<th>Uncapped</th>
<th>Capped (If Applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum % Change</td>
<td>-30.000%</td>
</tr>
<tr>
<td>Maximum % Change</td>
<td>30.000%</td>
</tr>
<tr>
<td>Total Number of Insureds (Auto-Calculated)</td>
<td>1994</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percent-Change Range</th>
<th>Number of Insureds in Range</th>
<th>Percent-Change Range</th>
<th>Number of Insureds in Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>-30% to &lt;-25%</td>
<td>2</td>
<td>-15% to &lt;-10%</td>
<td>452</td>
</tr>
<tr>
<td>-25% to &lt;-20%</td>
<td>90</td>
<td>-10% to &lt;-5%</td>
<td>340</td>
</tr>
<tr>
<td>-20% to &lt;-15%</td>
<td>130</td>
<td>-5% to &lt;0%</td>
<td>245</td>
</tr>
<tr>
<td>-15% to &lt;-10%</td>
<td>230</td>
<td>Exactly 0%</td>
<td>12</td>
</tr>
<tr>
<td>-10% to &lt;-5%</td>
<td>340</td>
<td>0% to &lt;5%</td>
<td>150</td>
</tr>
<tr>
<td>-5% to &lt;0%</td>
<td>245</td>
<td>5% to &lt;10%</td>
<td>160</td>
</tr>
<tr>
<td>Exactly 0%</td>
<td>12</td>
<td>10% to &lt;15%</td>
<td>401</td>
</tr>
<tr>
<td>&gt;0% to &lt;5%</td>
<td>150</td>
<td>15% to &lt;20%</td>
<td>234</td>
</tr>
<tr>
<td>5% to &lt;10%</td>
<td>160</td>
<td>20% to &lt;25%</td>
<td>19</td>
</tr>
<tr>
<td>10% to &lt;15%</td>
<td>401</td>
<td>25% to &lt;30%</td>
<td>12</td>
</tr>
<tr>
<td>15% to &lt;20%</td>
<td>201</td>
<td>30% to &lt;35%</td>
<td>2</td>
</tr>
</tbody>
</table>

EXAMPLE Uncapped Rate Disruption

© 2019 National Association of Insurance Commissioners
EXAMPLE Capped Rate Disruption

State Division of Insurance - EXAMPLE for Largest Percentage Increase

<table>
<thead>
<tr>
<th>Largest Percentage Increase</th>
<th>Corresponding Dollar Increase (for insured receiving largest percentage increase)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncapped Change</td>
<td>30.00% Uncapped Dollar Change $165.00 Current Premium $550.00</td>
</tr>
<tr>
<td>Capped Change (If Applicable)</td>
<td>15.00% Capped $ Change (If Applicable) $82.50 Proposed Premium $632.50</td>
</tr>
</tbody>
</table>

Characteristics of Policy (Fill in Below)

- **For Auto Insurance**: At minimum, identify the age and gender of each named insured, limits by coverage, territory, make / model of vehicle(s), prior accident / violation history, and any other key attributes whose treatments are affected by this filing.
- **For Home Insurance**: At minimum, identify age and gender of each named insured, amount of insurance, territory, construction type, protection class, any prior loss history, and any other key attributes whose treatments are affected by this filing.

Automobile policy: Three insureds - Male (Age 54), Female (Age 49), and Male (Age 25). Territory: Las Vegas, ZIP Code 89105.

<table>
<thead>
<tr>
<th>Vehicle</th>
<th>BI Limits</th>
<th>PD Limits</th>
<th>UM/UIM Limits</th>
<th>MED Limits</th>
<th>COMP Deductible</th>
<th>COLL Deductible</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009 Ford Focus</td>
<td>$50,000 / $100,000</td>
<td>$25,000</td>
<td>$50,000 / $100,000</td>
<td>$5,000</td>
<td>$500</td>
<td>$1,000</td>
</tr>
<tr>
<td>2003 Honda Accord</td>
<td>$25,000 / $50,000</td>
<td>$20,000</td>
<td>$25,000 / $50,000</td>
<td>$1,000</td>
<td>$500</td>
<td>$1,000</td>
</tr>
</tbody>
</table>

No prior accidents, 1 prior speeding conviction for 25-year-old male. Policy receives EFT discount and loyalty discount.

Primary impacts are the increases to the relativities for the age of insured, ZIP Code 89105, COLL Deductible of $1,000, and symbol for 2003 Honda Accord.

Most Significant Impacts to This Policy (Identify attributes - e.g., base-rate change or changes to individual rating variables)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>% Impact (Uncapped)</th>
<th>Dollar Impact (Uncapped)</th>
<th>What lengths of policy terms does the insurer offer in this book of business?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insured Age (M/25)</td>
<td>12.00%</td>
<td>$66.00</td>
<td>Check all options that apply below.</td>
</tr>
<tr>
<td>COLL Deductible ($1,000)</td>
<td>10.00%</td>
<td>$61.60</td>
<td>12-Month Policies</td>
</tr>
<tr>
<td>Territory (89105)</td>
<td>4.00%</td>
<td>$27.10</td>
<td>6-Month Policies</td>
</tr>
<tr>
<td>Vehicle Symbol (2003 Honda Accord)</td>
<td>1.46%</td>
<td>$10.29</td>
<td>3-Month Policies</td>
</tr>
<tr>
<td>Effect of Capping</td>
<td>-11.54%</td>
<td>-$82.50</td>
<td>Other (Specify)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>15.00%</td>
<td>$82.50</td>
<td></td>
</tr>
</tbody>
</table>

© 2019 National Association of Insurance Commissioners

© 2019 National Association of Insurance Commissioners
Draft: 10/15/2019
As adopted by the Casualty Actuarial and Statistical (C) Task Force on XX/XX/XX

APPENDIX D – INFORMATION NEEDED BY REGULATOR MAPPED INTO BEST PRACTICES

TBD

APPENDIX E – REFERENCES


© 2019 National Association of Insurance Commissioners  52
Re: CASTF Regulatory Review of Predictive Models White Paper

Ms. DeFrain,

Several members of the CAS Ratemaking Research Committee have discussed the most recently exposed revisions to the draft white paper on “Regulatory Review of Predictive Models”. This document should not be construed as representing an official response from the CAS membership as a whole. The views are representative of only those members whose names appear at the conclusion of this document.

Having reviewed and discussed the changes to the document, we make the following observations:

- Given the emphasis on GLMs in the guide, should the title be reworded to indicate that the paper refers specifically to such models?
- On page 3, with respect to the issue of credibility in GLMs, we note the following
  - We disagree with the statement credibility weighting GLM outputs is not standard practice.
  - If the sample data are highly variable, then so too are the coefficient estimates. An experienced modeler will recognize this fact and so opine.
  - Estimates of standard error for coefficients and residuals permit one to develop confidence intervals around estimated quantities. Non-parametric sample estimates of the first moment of a distribution – however credible – do not support this.
  - Segmented data may be analyzed using GLMMs – generalized linear mixed models – which have substantially similar assumptions of classic Buhlman-Straub credibility estimators.
- Page 4:
  - Spurious relationships between predictors and targets may occur in smaller samples as well as in “big data” samples.
  - When controlling for non-permitted variables, validation of a rating plan isn't straightforward. Model performance may suffer from the exclusion of variables which are available to the insurer.
- Page 5:
  - The focus on personal auto and property appears to be largely driven by the ratemaking behavior of market participants. The adoption of GLMs for commercial auto writers is noted. However, there is no comment on why the best practices advocated in the document would be inappropriate for commercial auto coverage. One might conjecture that the higher limits on offer in commercial lines would necessitate robust statistical techniques like extreme value theory. Absent clarifying language from the NAIC, this remains a supposition.
• On page 7, we feel that the first paragraph oversteps the bounds of a rate filing review. We suggest that the paragraph be worded as follows:
  - Documentation of the design and operational details of the model is required to ensure business continuity and transparency of models used. Granularity of documentation takes into account the level of management or key function at which it is intended to be used. Documentation should be sufficiently detailed and complete to enable a qualified third party to form a sound judgment on the suitability of the model for the intended purpose. The theory, assumptions, methodologies, software and empirical bases should be explained, as well as the data used in developing and implementing the model. Relevant testing and ongoing performance testing needs to be documented. Key model limitations and overrides need to be pointed out so that stakeholders understand the circumstances under which the model does not work effectively. End-user documentation should be provided and key reports using the model results described. Major changes to the model need to be shared in a timely manner and documented, and IT controls should be in place, such as a record of versions, change control and access to model.

• Page 8:
  - Point A.1.b – We have struggled to understand this point. The notion that aggregated data has not been “scrubbed” suggests that no quality audit has been performed at all. This remark would benefit from further wording which would clarify what is meant here.

• Page 11:
  - Point A.3.d – The modeler should make a statement as to whether there is any systemic reason for missing data.

• Page 13:
  - Point B.1.a – The practice guidelines are highly specific to GLMs. Although reference is made to other model types, the steps involved in compliance for other models is vague. In a sense, this penalizes GLMs, which have existed for far longer and whose development and implementation enjoys more widespread understanding. The lesser detail for non-GLM models could persuade some market participants that the compliance burden may be reduced by gravitating towards techniques which have been less tested. We not sure that’s a desirable outcome.
  - Point B.1.b – Pedantic note, but the R version, as well as the version of all packages used in calculation is incredibly simple to obtain. Simply call `sessionInfo()` at an appropriate stage of the calculation (likely at the end). If I were a regulator who was told that a filer didn’t know what version of R they were using, I’d be highly suspect.
  - Point B.1.c – Although the term “training” is used with little ambiguity, “test” and “validation” are terms that are sometimes interchanged, or the word “validation” may not be used at all. The practice guidelines should be clear on this point.

• Page 15:
  - Point B.2.e – Most software has default convergence criteria. In practice, the modeler would only need to adjust this if the model fails to converge. I would recommend changing this statement to one which requires the modeler to make a statement if they deviated from defaults.
  - Point B.3.a – As worded, this seems to imply that interaction terms warrant more explanation than any other rating variable. In practice, interaction terms are used for the same reason as any other explanatory variable: they improve the model’s predictive power. The classic example is the interaction between age and smoking when predicting mortality. That variable’s use in life insurance rating has been well established and non-controversial.
  - Point B.3.b – We think it is reasonable to require diagnostics for various candidate models.

• Page 15, B.3.b—We have commented on this before. To reiterate, while this is listed as a Level 4 item, assembly of a list of all predictor variables is onerous. Further, can lead to companies having to disclose
intellectual property for types of variables they have experimented with. This will hamper speed to market and hinder innovation.

- Page 16
  - Point B.3.c – The wording in the comment would benefit from more context. When constructing a GLM, the modeler may measure correlation as an aid to decide the set of candidate predictors, or to interpret significance of coefficient estimates. The utility of a statement about Pearson v. Cramer, or a statement about how the matrix was produced is not immediately clear.
  - Point B.3.e – Use of PCA will obviate much of the preceding guidance about a rational relationship between a predictor and a target response.

We make the following comments with respect to terms used in the document:

- The current definition for “insurance data” is “data collected by the insurance company”. We feel that the definition would benefit by citing some examples of data sources which are not insurance data. We presume this would include items such as census or credit data.
- We recommend avoiding use of the term “predictive power” without reference to how this is measured. There are numerous statistical diagnostics – root-mean squared error, mean absolute error, to name two – which have crisp, clear definitions.
- Within the data science literature, although the term “training” is used with little ambiguity, “test” and “validation” are terms that are sometimes interchanged. The word “validation” may not be used at all. The practice guidelines should be clear on the definition of these terms.

Once again, we thank you for your consideration of these points and welcome the opportunity to discuss with you or any members of the CASTF.

Regards,

Sandra Callanan
Greg Frankowiak
Brian Fannin
Joshua Newkirk
David Terné
Honorable Steve Kelley  
Commissioner, Minnesota Department of Commerce  
Chairman, NAIC Casualty Actuarial and Statistical Task Force  
Minnesota Department of Commerce  
85 7th Place East, Suite 280  
Saint Paul, MN 55101  

Honorable James J. Donelon  
Commissioner, Louisiana Department of Insurance  
Vice-Chairman, NAIC Casualty Actuarial and Statistical Task Force  
1702 N. Third Street; P.O. Box 94214;  
Baton Rouge, LA 70802  

Submitted Electronically to kdefrain@naic.org  


Dear Chairman Kelley and Vice Chair Donelon:

I write on behalf of the Consumer Data Industry Association (CDIA) to comment on the exposure draft concerning best practices when reviewing predictive models and analytics. This draft was released by your Casualty Actuarial and Statistical Task Force (“Task Force”) on October 15, 2019. Thank you for allowing CDIA another chance to offer comments on behalf of our consumer reporting agency (“CRA”) members. We offer comments on section VI in the body of the whitepaper and sections A, B and C in the modeling guide.

The Consumer Data Industry Association is the voice of the consumer reporting industry, representing consumer reporting agencies including the nationwide credit bureaus, regional and specialized credit bureaus, background check and residential screening companies, and others. Founded in 1906, CDIA promotes the responsible use of consumer data to help consumers achieve their financial goals, and to help businesses, governments and volunteer organizations avoid fraud and manage risk. Through data and analytics, CDIA members empower economic opportunity all over the world, helping ensure fair and safe transactions for consumers, facilitating competition and expanding consumers’ access to financial and other products suited to their unique needs.

Section VI, 1. c (p. 5) addresses a “Review [of] the individual input characteristics to and output factors from the predictive model (and its sub-models), as well as, associated selected relativities to ensure they are not unfairly discriminatory”. We appreciate your feedback on our initial comments expressing concerns related to including “sub-models” like Credit-Based Insurance Scores (“CBIS”) into the regulatory
review process. However, we do respectfully believe this will increase the burden of regulatory compliance for CRAs, slowdown the speed to market and impede the relationship between insurers and consumers. These new burdens can inject unnecessary friction into consumers who seek quick decisions and competitive prices from their insurers.

We respectfully believe these are "new, proposed obligations". The review of CBIS models has been established and ongoing in many States for close to two decades like you highlight, but those occur in other forms of insurance and not under the forms the Casualty Actuarial and Statistical (C) Task Force is seeking to add to its handbook and make an industry wide practice. The current reviews may include the same CBIS models, but if they are not currently being reviewed then we would argue these are in fact new obligations on CRAs.

Many States have provided certain confidentiality protections from the general public for CBIS models in accordance with their State law, but many is not all states. CDIA members spend significant amounts of time and resources developing their models and complying with current regulations. only takes one employee in one state to make one mistake and decades of hard work, investment and research is available for anyone to view, replicate, deceive or use to commit fraud. We are encouraged by the inclusion of new confidentiality language in Section VII of the Whitepaper, pertaining state confidential, proprietary, or trade secret state laws and relevant contractual provisions, and request inclusion of the language as a proposed change to the Product Filing Review Handbook. Even with the new language, the lack of a national exemption from public records remains a concern because information that has never previously been requested could be subject to the myriad of public disclosure laws around the country. There is no surety to how all states will respond to public records requests.

New language in Section V of the Whitepaper suggests that reliance on state confidentiality authority, regulations, and rules may not govern if the NAIC or another third party becomes involved in the review process on behalf of the states. NAIC or third party participation in the review process causes significant trade secret and proprietary information protection concerns. It is not clear from the new language what protections, law, or authority would apply in such a case. We request clarifying language be added that, as a floor, the confidential, proprietary, and trade secret protections of the state on behalf of which a review is being performed apply.

We understand no information should be confidential from the regulators themselves. However, if the CBIS models are reviewed and accepted elsewhere, it would seem that a repetitive and costly process is occurring for not much if any added value to the final product for the consumers. The credit reporting system is a consistent nationwide process. Exposing individual characteristics of scoring models to public record requests allows competitors access to information that they can use to gain an unfair advantage over another company. It also reduces the incentive to continue to
create new solutions, reducing a competitive environment, which ultimately hurts consumers. Regulators should be able to know whether scoring models are in compliance with the law, but this information should not be accessible as a public record.

The potential for confidentiality concerns is not only with the CRAs, but the companies they work with (date furnishers and lenders) in the credit reporting system and their consumers. We are not convinced that including CBIS in this type of review is mission critical. Yet, if this review needs to be in the process, CDIA recommends the establishment of highly specific rules to protect confidentiality and proprietary information. Additionally, a separate review process of sub-models as an optional request with defined valid concerns would help in addressing concerns.

Credit-based insurance scores do not unfairly discriminatory towards any race, religion, gender, ethnicity, or other established suspect classes and there are studies that show the lack of illegal discrimination. A myth of illegal discrimination pervades many media accounts and public policy debates, but in truth, credit-based insurance scores do not promote redlining or other illegal insurance practices.

Section VI 3.a. (p. 6) addresses how to “[e]valuate how the model interacts with and improves the rating plan” and how to “[o]btain a clear understanding of the characteristics that are input to a predictive model (and its sub-models), their relationship to each other and their relationship to non-modeled characteristics/variables used to calculate a risk’s premium.” We recognize the goal of the regulator in seeking to understand how the individual components of the rating plan interrelate to produce a consumer’s premium, but we feel your comment adds further confusion to our members. The white paper only mentions “characteristics”, but your comment refers to “information that the ‘CRAs use to create CBIS’ is essential to understanding the structure of the CBIS models, the variables used, and their justification.”. CRAs could provide general characteristics of the model without having confidentiality concerns, but the “information they use to create CBIS” appears to be far more specific.

If these provisions are meant to include information relating to the scoring models that CRAs use to create CBIS, there would be a significant new regulatory burden on CRAs and this would impede the relationship between insurers and consumers. These new burdensome requirements can inject unnecessary friction on to consumers who seek quick decisions and competitive prices from their insurers. Along with heightening the risk of disclosing proprietary information that is currently kept confidential because of its importance.

In “Selecting Model Input” under subsections A.1.a “Available Data Sources”, the original wording caused concern that FCRA requirements would be extended to all external data sources. The edit to this section is appreciated, but we believe application
to contractual disclosure restriction concerns remain. For CBIS models, we feel that review should be restricted to credit variables used in the model, not all credit variables.

Regarding A.2.b of the third exposure draft, former subsection A.2.f, “Determine if the sub-model was previously approved (or accepted) by the regulatory agency,” the review level change is appreciated as it will eliminate unnecessary and duplicative reviews of third-party and vendor models that have been previously approved. To be consistent with the A.2.b review level change, a change from a review level 1 to a 3 or 4 is requested for current A.2.f, former A.2.e, “If using output of any scoring algorithms, obtain a list of the variables used to determine the score and provide the source of the data used to calculate the score”.

Section A.4.c addresses “Identif[ing] material findings the company had during their data review and obtain an explanation of any potential material limitations, defects, bias or unresolved concerns found or believed to exist in the data. If issues or limitations in the data influenced modeling analysis and/or results, obtain a description of those concerns and an explanation how modeling analysis was adjusted and/or results were impacted”. This provision should be recategorized from its current score of 1 to a 3 or 4 score. Existing regulations around actuarial rate making standards and state regulations should prevent these items from entering a “final/proposed” model. This should be categorized as three of four (i.e. if model review uncovers issues).

We have several comments regarding Section B, “building the model”:

- Sec. B.2.c, “Obtain a description of univariate balancing and the testing that was performed during the model-building process, including an explanation of the thought processes involved and a discussion of why interaction terms were included (or not included ).” Only included interactions should be discussed. Interactions not be included, but default are not in a model, and therefore should not need to be justified.
- Secs. B.3.a and B.3.c, Both subsections pose trade secret protection and confidentiality issues.
- Sec. B.3.b, “Obtain[ing] a list of predictor variables considered but not used in the final model, and the rationale for their removal”. The best practices and guidelines should be limited to only the variables that were in the final and proposed models.
- Sec. B.3.d, “Obtain[ing] an rational explanation for why an increase in each predictor variable should increase or decrease frequency, severity, loss costs, expenses, or any element or characteristic being predicted.” CDIA agrees with the current and actuarially accepted practice of rate making guidelines not requiring intuitive or rational explanations of predictive values. We support use of variables that are statistically and actuarially predictive of insurance losses.
Additionally, this subsection poses a risk exposing trade secret and confidential information.

- Secs. B.4.b, through B.4.b  CDIA recommends recategorizing these scores from their current scores of two to a three or four score, along with only making this a requirement if deemed necessary.

- Sec. B.4.c “Identif[i]ing] the threshold for statistical significance and explain why it was selected. Obtain a reasonable and appropriately supported explanation for keeping the variable for each discrete variable level where the p-values were not less than the chosen threshold”. This is a fairly subjective standard. We recommend that it includes more objective and actuarially sound information and decisions. We recommend adding “threshold for statistical significance” into the list of required elements and changing this score from its current one to a three or four.

We have several comments regarding “Section C, “The Filed Rating Plan”:

- Sec. C.1.c, like many other areas, this provision creates potential trade secret and confidentiality issues.

- Sec. C.2.a, “Obtain a narrative regarding how the characteristics/rating variables, included in the filed rating plan, logically and intuitively relate to the risk of insurance loss (or expense) for the type of insurance product being priced.” CDIA appreciates the edits made to the Information Element. “Logical and intuitive” was removed from the “Information Element” box, but not the “Comment” box. We recommend removal of “logical and intuitive” from the “Comment” box for consistency.

- Sec. C.7.h, this new section will impact CBIS and it appears to extend FCRA requirements on all external data. To ease FCRA requirement extension, we request changing the language in the Comment box from “…data should be documented and an overview of who….” and “…consumer verification should be addressed,….” to “…data may need to be documented and an overview…” and “consumer verification may need to be addressed…”.

The “Supporting Data” section, specifically Secs. C.6.a and C.6.b, on “Obtain[ing] an explanation of any material (especially directional) differences between model indications and state-specific univariate indications” pose some concerns for CRAs and could interfere with the insurance process for consumers.

Section VIII of the Whitepaper proposes several changes to the Handbook. Section X, “Other Considerations” of the Handbook suggest advisory organization regulation of model and algorithm vendors. As explained further in this comment, CIBS modelers are already heavily regulated.

Credit Based Insurance Scores are constructed using nationwide data sets. Scoring or grading their performance out at a state level may not be supported or
accurate with this approach. It is also a common occurrence for certain contracts to prevent model providers from sharing distinct or customer specific data with third parties. There are several factors besides credit information and CBIS that go into the rate setting process. Credit Information and CBIS may be the only ones that are consistent and transferrable across the country, while some of the other factors used can and do differ greatly on a state by state basis.

The insurance industry has been using CBIS models for decades and they have been approved by nearly every state’s insurance department for auto and home insurers. Adding the work CASTF proposes will be burdensome and repetitive. The lack of trade secret and proprietary information protection will always remain a source of concern. In the long run we see this as something only large insurers will be able to absorb and the small to medium sized insurers that rely on third parties help will get squeezed out. We strongly feel that this will give large insurers a competitive edge in the marketplace. This will come at great cost to the consumers when their options decrease because of the eventual lack of competition.

There is already a large regulatory review presence on the industry. It is already over seen at the federal level by the Consumer Financial Protection Bureau (CFPB) and Federal Trade Commission (FTC), along with several states implementing their own regulations and the Conference of State Banking Commissioners looking into the industry as well. This increased regulation not only hurts the industry, but the consumers it serves. It will significantly hamper speed to market for the products consumers need and does not appear to add much, if any, benefit to the outcome for the industry and its consumer.

In conclusion, we believe that these potential new best practices will create burdensome regulatory difficulties for our members, speed to market issues for insurance companies, their product and the consumers that need them. CDIA members provide quality products that are already regulated and accepted by the insurance industry. CDIA and its members respectfully request consideration and inclusion of its comments in the task force’s whitepaper. Thank you for the opportunity to comment and please feel free to contact us with any questions you may have.

Sincerely,

Eric J. Ellman
Senior Vice President, Public Policy & Legal Affairs

cc: Members of the Casualty Actuarial and Statistical Task Force (CASTF) of the Property and Casualty Insurance (C) Committee
    Kris DeFrain, NAIC Staff
    Jennifer Gardner, NAIC Staff
Comments for the Center for Economic Justice  

To the Casualty Actuarial Task Force  

Regulatory Review of Predictive Models White Paper  

November 22, 2019  

The Center for Economic Justice offers the following comments on the October 2019 draft of the Regulatory Review of Predictive Models White Paper.  

Section VI, number 1 states: “Ensure that the selected rating factors, based on the model or other analysis, produce rates that are not excessive, inadequate, or unfairly discriminatory.” CEJ suggests the following addition to part c.:  

   c. Review the individual input characteristics to and output factors from the predictive model (and its sub-models), as well as, associated selected relativities to ensure they are not unfairly discriminatory in terms of both a cost-based relationship of the risk classification and an absence of intentionally or unintentional discrimination against protected classes.  

The suggested addition identifies the two prongs of unfair discrimination – the absence of cost differentials necessary to justify different treatment of consumers and the direct or indirect (proxy) use of prohibited classes for different treatment of consumers.  

Section VI, number 2 states: “Obtain a clear understanding of the data used to build and validate the model, and thoroughly review all other aspects of the model, including assumptions, adjustments, variables, submodels used as input, and resulting output.” CEJ suggests the addition of another item under number 2:  

   x. Determine if data used for model development and testing are biased against protected classes of consumers, if insurers have tested the data for such bias and if any action has been taken to eliminate or reduce bias in data.  

While this type of information and testing is implied in other parts of the white paper,  

CEJ suggests explicit identification of this type of data and model testing.
Regarding Section VII, A.1.a:

Request details of any non-insurance data used (customer-provided or other), whether the data was collected by use of a questionnaire/checklist, whether data was voluntarily reported by the applicant, and whether any of the data is subject to the Fair Credit Reporting Act. If the data is from an outside source, find out what steps were taken to verify the data was accurate, complete and unbiased in terms of relevant and representative time frame, representative of potential exposures and uncorrelated with protected classes.

While there are important consumer protection issues associated with insurers’ use of non-FCRA compliant data, it is unclear what a filing reviewer should or might do with information sought in the first (italicized) sentence. CEJ suggests that, whatever the source of the data or the means of obtaining the data from or about consumers, the same regulatory issues and questions apply – those set out in CEJ’s proposed revisions to the second sentence of the section.

CEJ suggests a new section: **Testing for and Minimizing Disparate Impact Unfair Discrimination:**

While regulators must review models for the direct use of prohibited risk classifications (intentional discrimination or disparate treatment), such violations are relatively easy to identify. Insurers’ use of many new databases of non-insurance personal consumer information as well as more intensive and granular databases of insurance personal consumer information (including consumer-generated data through telematics) increases the risk of proxy discrimination against protected classes. A “protected class” of consumers is one associated with prohibited risk classifications, such as race, religion or national origin.

GLMs – like any predictive model -- are developed using historical data. If the historical data incorporates or reflects biased or atypical outcomes, the algorithm will reflect and perpetuate those biases. The scholars Barocas and Selbst note in *Big Data’s Disparate Impact*¹

Advocates of algorithmic techniques like data mining argue that they eliminate human biases from the decision-making process. But an algorithm is only as good as the data it works with. Data mining can inherit the prejudices of prior decision-makers or reflect the widespread biases that persist in society at large. Often, the “patterns” it discovers are simply preexisting societal patterns of inequality and exclusion. Unthinking reliance on data mining can deny members of vulnerable groups full participation in society.

---

Disparate impact unfair discrimination refers to practices which have the effect of discrimination against protected classes and is sometimes referred to disparate effect. Regulatory review of complex predictive models should include a requirement that insurers demonstrate:

1. Testing of bias against protected classes in data used to develop and test the predictive model;
2. Testing of disparate impact against protected classes in the development of the model;
3. Employing statistical tools to minimize disparate impact in the development of the model; and
4. Testing of model output for disparate impact.

One common approach to identifying and minimizing disparate impact unfair discrimination is to utilize a control variable for the prohibited class in the development of the model. A control variable is an independent variable used to control or neutralize effects that might otherwise distort model specifications and output. For example, an insurer developing a national personal auto or homeowners insurance pricing model might use a control variable for state to control for / remove effects of differences among the states in minimum limits requirements, tort frameworks or other state-specific issues that might impact the statistical contribution of other, national, factors to explaining the dependent variable. While a control variable is used in the development of the model, the control variable is not included in the model deployed for use.

Similarly, by using the prohibited class characteristics as independent (control) variables in the development of the model, the remaining independent variables’ contribution (to explaining the dependent variable) is shorn of that part of their contribution that is a function of correlation with the prohibited characteristics. For the independent variables other than race, religion and national origin, what remains is a more accurate picture of the remaining independent variables’ contribution to the target outcome. Consequently, using prohibited class characteristics as control variables simultaneously tests for and minimizes disparate impact.

Testing for disparate impact is consistent with the statistical and actuarial nature of unfair discrimination based on cost-based analysis. One form of insurance unfair discrimination is different treatment of consumers without any demonstrated differences in the cost of the transfer of risk of those consumers. Stated differently, a rate is unfairly discriminatory if consumers of the same risk and hazard are treated differently. The traditional test for unfair discrimination is whether an insurer can demonstrate a difference in expected claims or expenses on the basis of the risk classification.

Testing for and measuring disparate impact is completely consistent with the cost-based tests for unfair discrimination. Actuarial justification is a statistical test – that a particular characteristic of the consumer, vehicle, property or environment is correlated with a particular outcome, like pure premium (average claim cost). The same statistical test can be used to evaluate and minimize disparate impact. Stated differently – if a particular correlation and statistical significance is used to justify, say, insurance credit scoring, those same standards of correlation and statistical significance are reasonable evidence of disparate impact and unfair discrimination on the basis of prohibited factors.
Testing for and minimizing disparate impact improves cost-based insurance pricing models. To the extent that historical data reflects bias and unfair discrimination against protected classes, testing for and minimizing disparate impact can stop the cycle of algorithms reflecting and perpetuating that historic discrimination.
I would suggest the following to “fix” the issue raised (note that the first bullet is moved up a level):

Other Considerations
- Regulators are often responding to consumer inquiries regarding how a policy premium is calculated and why the premium, or change is premium, is so high. The ability of the regulator to respond to these inquiries is included in best practice 1.b, “Review the premium disruption for individual policyholders and how the disruptions can be explained to individual consumers.”
  - The white paper identified the following information elements that may assist in addressing this best practice and a response to a consumer:
    - C.2.a Provide an explanation how the characteristics/rating variables, included in the filed rating plan, logically and intuitively relate to the risk of insurance loss (or expense) for the type of insurance product being priced. Include a discussion of the relevance each characteristic/rating variable has on consumer behavior that would lead to a difference in risk of loss (or expense).
    - C.7.f Explain how the insurer will help educate consumers to mitigate their risk.
    - C.7.h Identify sources to be used at "point of sale" to place individual risks within the matrix of rating system classifications. How can a consumer verify their own "point-of-sale" data and correct any errors?
    - C.7.j Provide the regulator with a description of how the company will respond to consumers’ inquiries about how their premium was calculated.

The white paper has the following note about “best practices” but then lists one best practice and 4 information items. How should I modify the introductory sentence?

Other Considerations
- Regulators are often responding to consumer inquiries regarding how a policy premium is calculated and why the premium, or change is premium, is so high.
  - The white paper identified the following best practices:
    - 1.b. Review the premium disruption for individual policyholders and how the disruptions can be explained to individual consumers. ...and information elements that may assist the regulator's and consumer's understanding of the premium being charged.
    - C.2.a Provide an explanation how the characteristics/rating variables, included in the filed rating plan, logically and intuitively relate to the risk of insurance loss (or expense) for the type of insurance product being priced. Include a discussion of the relevance each characteristic/rating variable has on consumer behavior that would lead to a difference in risk of loss (or expense).
    - C.7.f Explain how the insurer will help educate consumers to mitigate their risk.
    - C.7.h Identify sources to be used at "point of sale" to place individual risks within the matrix of rating system classifications. How can a consumer verify their own "point-of-sale" data and correct any errors?
    - C.7.j Provide the regulator with a description of how the company will respond to consumers’ inquiries about how their premium was calculated.
November 22, 2019

NAIC Casualty Actuarial and Statistical Task Force
Attn: Kris DeFrain, FCAS, MAAA, CPCU

Via Email – kdefrain@naic.org

Re: Comments on the October 15, 2019 Exposed CAS Task Force Draft White Paper –
Best Practices – Regulatory Review of Predictive Models

Dear Ms. DeFrain:

Fair Isaac Corporation (FICO) is pleased to provide its comments on the most recently released draft white paper, Best Practices – Regulatory Review of Predictive Models.

FICO is an independent analytics provider (not a data company) that is dependent on other firms (e.g., consumer reporting agencies, insurance companies, lending institutions) to provide the appropriate and necessary data for FICO analysis and for its development of predictive models. With a focus on innovation that effectively rewards all parties – insurers, lenders, and consumers alike – FICO is recognized as the pioneer in developing the algorithms and underlying analytics used to produce credit scores, credit-based insurance scores, and other risk management scores. FICO fully understands and respects the value of regulatory scrutiny and the need for regulatory flexibility to help ensure that consumers continue to benefit from these scores by enjoying quick, fair access to credit and greater access to more affordable insurance. In previous years, access to affordable insurance involved a lengthy decision process based, in some cases, on subjective and inconsistent underwriting and pricing factors.

In 1993, FICO introduced the first commercially available credit-based insurance scores to US insurers as an additional risk segmentation factor that could be used in their private passenger auto and home insurance underwriting and pricing programs. On behalf of several hundred FICO® Insurance Score clients, over these past 25 years, FICO has met with state departments of insurance and has testified before dozens of state legislative committees. Our goal in each of these interactions was to provide support for our clients’ continued use of FICO Insurance Scores by answering all appropriate regulatory questions to the best of our ability and by offering as much insight into FICO’s proprietary modeling analytics and technologies as possible, while still protecting our intellectual property.

For nearly two decades, in support of successful rate filings throughout the nation by our FICO® Insurance Score clients, FICO has provided model documentation—specific consumer credit characteristics, attributes and weights for the filed model—as well as reason code/factor definitions, and a general discussion of our model development process to all requesting departments of insurance.
able to provide the necessary protections. In addition, FICO has modified its insurance score models as required by those states with specific statutory or regulatory mandates. FICO also offers an insurance score educational website (insurancescores.fico.com) that has been accessed by consumers, regulators, legislators, insurers, agents and other interested parties throughout the nation for a more thorough understanding of FICO’s credit-based insurance scores, the insurance industry’s use of our insurance scores, and general credit management tips.

It remains our hope that home and auto insurers using credit-based insurance scores as one factor in their underwriting and pricing programs will continue to receive filing approvals, just as they have in virtually every state in the nation for the past two decades. The use of credit-based insurance scores should be allowed to continue unabated – grandfathered, as it were - under current regulatory review processes such that such long-standing use of credit by virtually every insurer would not be negatively impacted by the undue burdens proposed by the October 15, 2019 CAS TF draft white paper. State regulatory authorities want to ensure a fair playing field that allows all insurers, not just a few at the top, the opportunity to compete effectively and appropriately for their constituents’ business. The industry’s use of credit-based insurance scores that have been approved for over two decades should not be overshadowed by a newly proposed predictive model review approach that may force market participants such as FICO to withdraw their scores from use and may lead to premium increases for the majority of consumers.

Quite concerning is the fact that in recent months and weeks both Maine and Washington have apparently chosen to “jump the gun” on the traditional NAIC decision-making process. The following blog highlights the problem this is creating for effective industry competition and consumer pricing benefits - https://www.insurancejournal.com/blogs/right-street/2019/10/27/546717.htm.

As was stated in the conclusion of the FTC’s 2007 report – “Credit-Based Insurance Scores: Impacts on Consumers of Automobile Insurance” –

“…..credit-based insurance scores are effective predictors of risk under automobile insurance policies. Using scores is likely to make the price of insurance conform more closely to the risk of loss that consumers pose…..”

Industry studies have shown that the same conclusion could be drawn with respect to homeowner insurance, as well.

Having shared a bit of FICO’s background and our FICO® Insurance Score client support strategies that we certainly hope to continue, the remainder of our comments will focus on our Scores business model and the negative implications the recommendations within the draft white paper will have on FICO’s Scores business. More importantly, if the proposed predictive model review positions remain in place with respect to time-tested, regularly reviewed and approved credit-based insurance scores, there will be significant negative impact seen by virtually all auto and home insurance companies and the vast
majority of consumers – your constituents – across the nation as this key risk segmentation tool is restricted from use in rate filings.

The intellectual property underlying much of our predictive modeling and analytics technology has been developed by FICO data scientists over the past six decades. This development work has taken an enormous amount of time, money, research, know-how, and testing. Given that, however, FICO has been very transparent – sharing our models with state insurance regulators for the past two decades where appropriate protections were in place to avoid exposure of critical intellectual property.

Our goal is to continue to offer an insurance risk management tool to the benefit of the industry and to consumers alike, while still protecting the interest of our shareholders. FICO’s scoring-related trade secrets have substantial independent economic value to the company precisely because they are not generally known by others, including any potential competitors, that could unfairly obtain economic value from their disclosure or use. Forcing disclosure of these intellectual property assets would put them at risk and dissipate their value.

Given the necessary protection of FICO’s intellectual property, including its trade secrets, our belief is that the depth and breadth of the regulatory review of predictive models proposed by the draft white paper presents serious market-restriction issues for FICO, and for the hundreds of FICO® Insurance Score clients doing business in all states that allow for the industry’s significant use of credit-based insurance scores within their well-considered and comprehensive rating programs.

As mentioned previously, we believe the state regulatory practices under which FICO has supported its clients for the past two decades are appropriate and quite sufficiently protect all interests – consumers, regulators, and insurers. These scores have proven time and again over 25+ years to be highly accurate and effective in enabling insurers to more objectively and accurately price risk, while lowering premiums for the majority of consumers. As such, we believe previously approved FICO® Insurance Scores should be excluded from the draft white paper to the benefit of the industry as a whole and to the benefit of your constituents specifically.

The draft white paper’s only references to protection for the intellectual property of an independent analytics provider like FICO are too vague to offer any real protection. The proposal, as highlighted here, leaves the decision about confidentiality of a company’s intellectual property and trade secrets entirely within the discretion of each state regulator.

1. The fourth Key Regulatory Principle: State insurance regulators will maintain confidentiality, where appropriate, regarding predictive models.
2. Section V. CONFIDENTIALITY warns rate filers:

   Insurers and regulators should be aware that a rate filing might become part of the public record. Each state determines the confidentiality of a rate filing, supplemental material to the filing, when filing information might become public, the procedure to request that filing information be held confidentially, and the procedure by which a public records request is made. It is incumbent on an insurer to be familiar with each
state’s laws regarding the confidentiality of information submitted with their rate filing.

FICO strongly supports the following comments from the National Association of Mutual Insurance Companies (NAMIC) stated in its June 28, 2019 letter to the NAIC Casualty Actuarial and Statistical (C) Task Force – “NAMIC wishes to continue to reiterate that exposing confidential and proprietary trade secrets, confidential information, and other business practices simply for accumulation of data in a rate filing, when otherwise unnecessary, is problematic for all involved. The data provided for these requirements subjects the regulator to increased Freedom of Information Act requests, subpoenas, and other types of litigation when there has been no demonstrated harm to consumers or trigger for the inquiry.”

Since FICO cannot be left in a precarious position with respect to the protection of its intellectual property, if the drafted white paper is adopted, as written, by any state without necessary trade secrets and other intellectual property protections in place, FICO may be forced to remove its FICO Insurance Score models from use by our insurance clients in that state, just as is now occurring in some “early adopter” states, creating wholly unnecessary market disruption.

As always, we look to the NAIC to do the right thing for consumers and insurers throughout the nation. We also look forward to working with the NAIC Casualty Actuarial and Statistical Task Force toward a regulatory review approach that protects the interests of all stakeholders, including the vast numbers of US consumers who benefit from the insurance industry’s continued use of credit-based insurance scores to enhance their underwriting and pricing policies based on proven risk characteristics.

Sincerely,

Lamont D. Boyd, CPCU, AIM
Insurance Industry Director, Scores

LamontBoyd@FICO.com
602-317-6143 (mobile)
Insurance Services Office, Inc.
545 Washington Boulevard
Jersey City, NJ 07310-1686
www.iso.com

Stephen C. Clarke, CPCU
Vice President
Government Relations
t 201.469.2656
f 201.748.1760
sclarke@iso.com

Kris DeFrain, FCAS, MAAA, CPCU
Director of Research and Actuarial Science
National Association of Insurance Commissioners (NAIC) Central Office
1100 Walnut Street
Suite 1500
Kansas City, MO 64106-2197

re: 10/15/19 Draft White Paper on Best Practices

Dear Ms. DeFrain,

Insurance Services Office, Inc. (ISO) is a countrywide licensed rating/advisory organization serving the property/casualty market. We have extensive experience and expertise in the development of advisory insurance pricing tools including prospective loss costs, rating plans and predictive analytics, including related regulatory issues.

ISO appreciates the opportunity to provide comments on the Draft White Paper on Best Practices for Regulatory Review of Predictive Models as published by the CASTF in October 2019. The CASTF has addressed the bulk of ISO’s previous comments but we have a few comments.

- B.1.c addresses how validation (hold out) data is used. The GLM paper (Generalized Linear Models for Insurance Rating) that is on the Exam 8 syllabus addresses the use of hold out data. On page 39 it says “Once a final model is chosen, however, we would then go back and rebuild it using all of the data, so that the parameter estimates would be at their most credible.”

- B.3.b asks for a list of predictor variables considered but not used in the final model and the rationale for their removal. While we appreciate that this is a level 4 item we don’t see how the variables not used in a model are relevant to reviewing the filed model. This would be analogous to asking for policy wording considered but not used in a filed policy form.

- Item C.7.h does not have a level ranking.
Respectfully Submitted,

Stephen C. Clarke, CPCU
November 22, 2019

NAIC Casualty Actuarial and Statistical (C) Task Force
c/o Kris DeFrain - kdefrain@naic.org
1100 Walnut Street, Suite 1500
Kansas City, MO 64106-2197

Re: NAMIC Comments on CASTF’s Predictive Model White Paper – October 14, 2019 Exposure

Dear Task Force Chair, Vice Chair, Task Force Members, and Other Interested Regulators,

Please accept the following remarks of the National Association of Mutual Insurance Companies (hereinafter “NAMIC”) on behalf of its member companies regarding the task force request for comments regarding the October 14, 2019 exposure of the latest draft of the Predictive Modeling White Paper. NAMIC wishes to thank the task force for the ability to provide additional comments on the white paper and the continuing transparency of the process as a whole.

While the task force has exhaustively attempted to review and examine all comments and submitted input concerning this endeavor, NAMIC still believes there are concerns that might override any completion of this project in the near term. NAMIC respectfully suggests there remain substantial principles that must be clearly defined and/or established before moving to any type of final product as previously mentioned in our comment letters. Further, to avoid repetitive comments from continually being reiterated we would refer the task force to NAMIC’s comment letters of January 15, 2019; June 28, 2019; and September 9, 2019. While NAMIC may refer to some of the content of those letters, please do not interpret a failure to discuss as relinquishment of concerns already posited to the task force to the extent they were not formally adopted.

As for the topic of predictive modeling broadly, NAMIC does not believe as a whitepaper the document has much description of the positive aspects of analysis of large data sets for policyholders and instead moves to anecdotal or cursory concerns. Consequently, the paper almost assumes that there are “issues” that currently exist despite any real demonstration of the same. It is genuinely believed by NAMIC and its members that large data sets provide a level of detail that promotes healthy and robust insurance products and concomitant marketplaces that benefit all stakeholders.

---

1 NAMIC membership includes more than 1,400 member companies. The association supports regional and local mutual insurance companies on main streets across America and many of the country’s largest national insurers. NAMIC member companies write $268 billion in annual premiums. Our members account for 59 percent of homeowners, 46 percent of automobile, and 29 percent of the business insurance markets. Through our advocacy programs we promote public policy solutions that benefit NAMIC member companies and the policyholders they serve and foster greater understanding and recognition of the unique alignment of interests between management and policyholders of mutual companies.
NAMIC encourages the authors to integrate the positive aspects of large data set analytics more fully throughout the paper. Issues need not necessarily be presupposed, and they should not be presented without including the demonstrable benefits data yields. The ability, for instance, to provide outstanding customer service and products which is demanded by consumers should be further elucidated as there is existing support for these outcomes presently. Future regulators that review this document might be overly persuaded that imminent issues exist that in reality have not been manifested.

Using larger data sets is a natural evolution of the insurance paradigm. Its positive effects outweigh any perceived downside. Although prudent stakeholders plan for potentialities, many positive results from large data analytics are already present and being utilized for the benefit of the public and consumers. Because the insurance industry is so justifiably based on data – to understand the risks involved in order to better underwrite and price – the paper has a responsibility to prominently discuss the benefits to all stakeholders gained through improved efficiency, accuracy, and fairness.

As for the submitted comments and current draft of the white paper, NAMIC wants to thank the task force for accepting a number of comments NAMIC has made concerning this matter including but not limited to revisiting the level of importance definitions and assignments, discussion of a core set of information elements that should be in every filing that includes a model; adding additional sentences/paragraphs on confidentiality and regulatory duties in this regard; attempting to limit the scope of review of data sources; addressing consumer responsibility in data reporting; merging or deleting duplicative matter; clarifying in many instances definitions; adding terminology to the glossary, removal of vague terms such as “intuitive,” and “thought processes;” and revision of the wording in many paragraphs of the elements including adding NAMIC phraseology. This acknowledgment of existing concerns is greatly appreciated and applauded.

Nevertheless, and in continuation of concerns regarding more broad overarching principles concerning the implementation of this process, NAMIC must reiterate existing concerns and posit additional thought regarding the same.

Unpromulgated Model Regulation

NAMIC believes that due to the level of granularity and detail that is being requested in this paper which is, in some instances, unnecessary, excessive, overburdensome, and overly prescriptive without considering any demonstrated need or trigger for the regulator to request is essentially a regulation masked as a white paper. The upfront loading of data being sought not only subjects the same to unwarranted exposure, it is not necessary for the regulator to perform their respective duties in a timely and efficient, yet, legal manner. Inevitably, despite protestations to the contrary, this document as currently written will tie up regulatory discretion, encourage a slow-down in speed to market of products, stifle innovation and be utilized as a manual of necessity.

However, if it is the intention of NAIC to move forward in this regard, NAMIC would suggest this is on the wrong track and should be subjected to model regulation scrutiny and the accompanying process. It appears NAIC is attempting to adopt a national standard requiring uniformity amongst all states. Further, any state who intends to implement such standards
should go through their own rule-making process for the same to be effective. There are many items within regulatory functions that are less onerous or detailed that have gone through this process. Throughout the paper the term guidance is utilized. Additionally, there are too many ramifications concerning this paper including the confidential and proprietary nature of the data subject to exposure that the appropriate process should be followed to ensure maximum input and protection of the scope and concerns that are being and have been previously discussed.

**Drafting Notes**

In relation to the rule-making process concern described above, there are concerns that too many drafting notes are being lost in this process. When the intentioned paper is ultimately released as currently drafted, it will be devoid of many of the drafting statements made to essentially reject alterations, edits, or comment suggestions. When the paper ends up in the regulators’ purview, they will not necessarily have the benefit of this “guidance” concerning the thought processes that ultimately led to the document’s finality. These notes are a part of this process and should therefore be inclusive such as with a model regulation. Intentions, however well-meaning, can be misinterpreted when not fully explained.

**Unresolved Work Streams – “Other Considerations”**

It appears to be somewhat unfair to mention other topics, provide a brief discussion without any resolution, and then state that they are not covered in the paper. There could be a host of suggestions to include in this area from an industry standpoint as well such as exploring the positive aspects of predictive modeling and how it improves consumer experience in many aspects. Providing such a brief level of concern without more discussion leaves the potential for misimpressions and may cause readers/regulators to delve more closely into supposed aspirational conclusions without proper review/discussion or further input on such topics from all stakeholders. These unresolved issues include granularity of rating variables and plans, scientific mindset of open inquiry, correlation vs. causality, and data mining conflicts. We believe these topics should be left for further elucidation if and when such matters are decided to be thoroughly explored from all aspects in a transparent manner but not touched upon in such a cursory manner.

**Continuing Concerns**

While NAMIC provides comments as requested by the task force, we would like to reiterate that discussion should be revisited as to the necessity of this document as currently drafted. NAMIC would implore the task force to consider meaningful baseline analysis requirements that regulators need to review filings and fulfill their legal and regulatory obligations. Each regulator may always request further documentation to quell concerns. However, as currently drafted, we do not feel the same has been sufficiently curtailed. In fact, the paper has continued from its existence to require an extraordinary amount of data and responses that may not be necessary at all. There is a difference between an amount of information needed to perform the required duties of approving such filings and merely satisfying the inquisitive nature of a reviewer. We believe the document encourages delay and continual demand for explanatory conferences that while
important to address legitimate concerns should not be the rule but the exception. This paper is requiring each regulator to be a data modeler and then opine on the actual resultant product which in many instances deviates from traditional legal duties of a regulator. Additionally, NAMIC believes there is terminology that is open to interpretation such as “rational,” that will create confusion and delay.

In closing, NAMIC again wants to thank the task force for the ability to respond to its well-intended and open process. We look forward to providing continued input in regard to this endeavor. However, due to the existing concerns, NAMIC would suggest that there are too many unresolved matters to move forward with finalization in Austin, Texas.

Sincerely,

Andrew Pauley, CPCU
Government Affairs Counsel
National Association of Mutual Insurance Companies (NAMIC)
Dear Ms. DeFrain,

I write you as director of finance, insurance and trade policy at the R Street Institute, a nonprofit, nonpartisan public policy research organization (“think tank”). Our mission is to engage in policy research and outreach to promote free markets and limited, effective government. Since our founding in 2012, R Street has had a successful history of research into public policy regarding the business of insurance. Our most notable contribution in this space has been producing our annual Insurance Regulation Report Card, whose eighth edition will be published in early December.

R Street is not engaged as an insurance underwriter, data firm or model provider. Our thoughts on the Task Force’s Predictive Model White Paper are solely our own, grounded in the principles of limited, effective and efficient government. With that said, we would like to share some of our concerns with this exposed white paper.

While framed as a simple “best practices” document focused on generalized linear models (GLM) deployed in the personal auto and homeowners lines of business, the white paper’s scope clearly goes far beyond that. It proposes new rating standards that do not reflect any existing state actuarial review process for rating plans filed with a GLM. In some places, it even suggests that such standards be applied to other, and possibly even all, model types. Of particular concern is that the white paper in several places recommends states require the collection of information that could breach confidentiality, trade secrets and long-established Actuarial Standards of Practice (ASOPs).

Among the problematic recommendations are the paper’s various instructions concerning raw data, which it defines as “data before scrubbing, transformation etc. takes place.” Regulators should, the paper notes, “review the geographic scope and geographic exposure distribution of the raw data for relevance to the state where the model is filed” as well as “ask for aggregated data... that allows the regulator to focus on the univariate distributions and compare raw data to adjusted/binned/transformed/etc. data.”

The goal appears to be to allow regulators to reproduce a model’s outputs, rather than simply review those outputs. But submitting raw data could pose security risks and may violate contractual obligations.
with third parties. Our view is that, as a general heuristic, regulators should temper their desire to know everything that goes on inside the proverbial “black box” and instead put the appropriate focus on what comes out of it; that is, focus on rates, not the models that produced them.

The confidentiality concerns extend beyond raw data and to the models themselves. The paper has been amended to better recognize the need for confidentiality protections, but it continues to entrust regulators to determine where it is “appropriate” to guard intellectual property and trade secrets. Third-party vendors can only fiscally justify their significant investment in proprietary algorithms if they are granted certainty that such work product will be protected. Without that certainty, few new vendors will come to the market and existing models could be withdrawn. In either case, the effect would be to stifle innovation.

Needless to say, a withdrawal of existing GLMs would cause significant market disruption. State departments of insurance have been deploying well-established review processes for GLMs for years. The results have been vibrant and competitive insurance markets. As R Street has demonstrated in its annual Insurance Regulation Report Card, no state currently has either a personal auto or homeowners insurance market with a Herfindahl-Hirschman Index (HHI) score that would indicate it is highly concentrated. In our forthcoming edition, we find only Alaska, Louisiana and New York had moderately concentrated auto insurance markets in 2018 and only Alaska had a moderately concentrated homeowners market.

This is a marked difference from the situation that prevailed through the 1980s, when some states saw as much as half of all auto insurance consumers shunted into residual market mechanisms. By contrast, according to the Automobile Insurance Plans Service Office (AIPSO), as of 2018, residual markets accounted for less than 0.1% of the market in 34 of the 50 states. Just four states—Maryland, Massachusetts, Rhode Island and North Carolina—have residual markets that account for more than 1% of auto insurance policies. This greatly improved ability of insurers to segment, classify and price risk effectively can be traced directly to the emergence of dynamic models like credit-based insurance scores.

Regulators should be very cautious before adopting any changes that could reverse those victories. To its credit, this updated version of the exposed draft does acknowledge a central weakness at the heart of the project, which is the degree to which regulators are expected to ask “the modeler to go beyond correlation and document their basic, causal understanding of how variables used in a model or rating plan are related to risk.” As the white paper notes, this approach significantly exceeds the requirements established in ASOP No. 12. It is, of course, reasonable to require model predictions to bear some resemblance to the subject being modeled, but causality is notorious difficult to prove, and the standards raised here could make the practice of modeling itself untenable.

---

What also should be noted is the irony that the white paper would recommend that regulators insist upon filings that prove not only the credibility, but the causal nature of modeling assumptions, when the Task Force itself has not begun to credibly demonstrate that such radical shifts in the rate-filing approval process would better serve markets or consumers.

Indeed, given the decades-long record for competitive insurance markets, there simply is no good reason to risk widespread market disruption via unprecedented information requests. Any best practices around regulation of predictive modeling should begin by determining what information is truly “essential” to ensure that rates are sufficient and neither excessive nor unfairly discriminatory.

R.J. Lehmann
Director of Finance, Insurance and Trade Policy
R Street Institute