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](#_Toc8388747)

[II. What is a “Best Practice?” 2](#_Toc8388748)

[III. Do Regulators Need Best Practices to Review Predictive Models? 3](#_Toc8388749)

[IV. Scope 4](#_Toc8388750)

[V. Confidentiality 5](#_Toc8388751)

[VI. Guidance for Regulatory Review of Predictive Models (Best Practices) 5](#_Toc8388752)

[VII. Predictive Models – Information for Regulatory Review 6](#_Toc8388753)

[VIII. Proposed Changes to the *Product Filing Review Handbook* 24](#_Toc8388754)

[IX. Proposed State Guidance 24](#_Toc8388755)

[X. Other Considerations 24](#_Toc8388756)

[XI. Recommendations Going Forward 24](#_Toc8388757)

[XII. Appendix A – Best Practice Development 25](#_Toc8388758)

[XIII. Appendix B - - Glossary of Terms 26](#_Toc8388759)

[XIV. Appendix C – Sample Rate-Disruption Template 28](#_Toc8388760)

[XV. Appendix D – Information Needed by Regulator Mapped into Best Practices 30](#_Toc8388761)

[XVI. Appendix E – References 30](#_Toc8388762)

# 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[[1]](#footnote-1) 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:

## 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.

## 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*.*

# 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]](#footnote-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]](#footnote-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:

### 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.

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.

# 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[[4]](#footnote-4). 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)[[5]](#footnote-5) 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[[6]](#footnote-6):

* 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.[[7]](#footnote-7),[[8]](#footnote-8) 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.[[9]](#footnote-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]](#footnote-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.

### 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.

# 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.

# 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.

# 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:

### Ensure that the selected rating factors, based on the model or other analysis, produce rates that are not excessive, inadequate, or unfairly discriminatory.

### Review the overall rate level impact of the proposed revisions to rate level indications provided by the filer.

### Review the premium disruption for individual policyholders and how the disruptions can be explained to individual consumers.

### 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.

### 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.

* 1. 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.
	2. 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.
	3. Determine that any adjustments to the raw data are handled appropriately, including but not limited to, trending, development, capping, removal of catastrophes.
	4. 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.
	5. 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.
	6. Obtain a clear understanding of how the selected predictive model was built.
1. Evaluate how the model interacts with and improves the rating plan.
	1. 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.
	2. Obtain a clear understanding why the insurer believes this type of model works in an insurance risk application.
	3. Obtain a clear understanding of how model output interacts with non-modeled characteristics/variables used to calculate a risk’s premium.
	4. Obtain a clear understanding of how the predictive model was integrated into the insurer’s state rating plan and how it improves that plan.
	5. For predictive model refreshes, determine whether sufficient validation was performed to ensure the model is still a good fit.
2. Enable competition and innovation to promote the growth, financial stability, and efficiency of the insurance marketplace.
3. Enable innovation in the pricing of insurance through acceptance of predictive models, provided they are actuarially sound and in compliance with state laws.
4. Protect the confidentiality of filed predictive models and supporting information in accordance with state law.
5. Review predictive models in a timely manner to enable reasonable speed to market.

# 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]](#footnote-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]](#footnote-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 (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.

## **Selecting Model Input**

|  |  |  |  |
| --- | --- | --- | --- |
| **Section** | **Information Element** | **Level of Importance to the Regulator’s Review** | **Comments** |
| 1. Available Data Sources |
| A.1.a | 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.  | 1 | 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. 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.Note that reviewing source details should not make a difference when the model is new or refreshed; refreshed models would report the prior version list with the incremental changes due to the refresh. |
| A.1.b | Reconcile aggregated insurance data underlying the model with available external insurance reports. | 4 | 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. |
| A.1.c | Review the geographic scope and geographic exposure distribution of the raw data for relevance to the state where the model is filed.  | 2 | 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. |
|  |  |  |  |
| 2. Sub-Models |
| A.2.a | Consider the relevance of (e.g., is there a bias) of overlapping data or variables used in the model and sub-models. | 1 | 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. |
| A.2.b | Determine if the sub-model was previously approved (or accepted) by the regulatory agency.  | 1 | 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. |
| A.2.c | 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, 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 "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.d | 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.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? |
| A.2.f | 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. |
| 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?  | 2 | 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. |
| 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. | 1 |   |
| 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/binned/transformed/etc. data. | 4 | This is most relevant for variables that have been "scrubbed" or adjusted.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. |
| A.3.d | Determine how missing data was handled. | 1 | 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. |
| A.3.e | If duplicate records exist, determine how they were handled. | 1 |   |
| A.3.f | Determine if there were any material outliers identified and subsequently adjusted during the scrubbing process.  | 3 | 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. |
| 4. Data Organization |
| A.4.a | 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. | 2 | 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. |
| A.4.b | 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. | 2 | 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. | 1 | A response of "none" or "n/a" may be an appropriate response. |

## **Building the Model**

|  |  |  |  |
| --- | --- | --- | --- |
| **Section** | **Information Element** | **Level of Importance to Regulator’s Review** | **Comments** |
| 1. High-Level Narrative for Building the Model |
| B.1.a | 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. | 1 | 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. |
| B.1.b | Identify the software used for model development. Obtain the name of the software vender/developer, software product and a software version reference used in model development. | 3 | 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." |
| B.1.c | 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. | 1 | It would be unexpected if validation data were used for any purpose other than validation. |
| B.1.d | Obtain a brief description of the development process, from initial concept to final model and filed rating plan. | 1 | The narrative should have the same scope as the filing. |
| B.1.e | 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. | 1 |   |
| B.1.f | Identify the model’s target variable. | 1 | 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. |
| B.1.g | Obtain a description of the variable selection process. | 1 | 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. |
| B.1.h | In conjunction with variable selection, obtain a narrative on how the Company determine the granularity of the rating variables during model development. | 2 | This discussion should include discussion of how credibility was considered in the process of determining the level of granularity of the variables selected. |
| B.1.i | 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. | 1 | The regulator would use this to follow the logic of the modeling process. |
| 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).  | 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 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). | 3 | 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.  | 3 | 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. | 2 |   |
| 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. | 1 | 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. | 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, 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. | 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, confidence intervals, chi-square tests, p-values and any other relevant and material tests.  |
| B.4.d | 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. | 2 | 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.e | Obtain evidence that the model fits the training data well, for individual variables, for any relevant combinations of variables and for, the overall model. | 2 | 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.  |
| B.4.f | 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. | 2 | 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 a description how the model was tested for stability over time. | 2 | 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. |
| B.4.h | Obtain a narrative on how potential concerns with overfitting were addressed. | 2 |   |
| B.4.i | Obtain support demonstrating that the GLM assumptions are appropriate. | 3 | 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.j | Obtain 5-10 sample records with corresponding output from the model for those records. | 4 |   |
| 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.  | 2 | 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. | 3 | 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. | 2 | 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.  | 2 | 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. | 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. |

## **The Filed Rating Plan**

|  |  |  |  |
| --- | --- | --- | --- |
| **Section** | **Information Element** | **Level of Importance to Regulator’s Review** | **Comments** |
| 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 |
| C.5.a | 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. | 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. |
|  |  |  |  |
| 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 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. | 4 | 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.  | 4 | 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.  | 4 | These rating variables may represent changes to rating factors, 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 that testing. | 3 | 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. | 2 | 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. | 3 | 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.  | 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. |

# 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.

**Product Filing Review Handbook, August 2016**

**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; 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 may 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.. 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 predictive power

homehome

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.

## 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[[13]](#footnote-13) 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.

## 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.

## 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]](#footnote-14). Best practices can maintain quality as an alternative to mandatory legislated standards and can be based on self-assessment or benchmarking.[[15]](#footnote-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.

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.*

## 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 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.

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.
2. Review the overall rate level impact of the proposed revisions to rate level indications provided by the filer.
3. Review the premium disruption for individual policyholders and how the disruptions can be explained to individual consumers.
4. 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.
5. 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.
6. 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.
7. 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.
8. Determine that any adjustments to the raw data are handled appropriately, including but not limited to, trending, development, capping, removal of catastrophes.
9. 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.
10. Obtain a clear understanding of how the selected predictive model was built.
11. 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.
12. Evaluate how the model interacts with and improves the rating plan.
13. 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.
14. Obtain a clear understanding why the insurer believes this type of model works in an insurance risk application.
15. Obtain a clear understanding of how model output interacts with non-modeled characteristics/variables used to calculate a risk’s premium.
16. Obtain a clear understanding of how the predictive model was integrated into the insurer’s state rating plan and how it improves that plan.
17. For predictive model refreshes, determine whether sufficient validation was performed to ensure the model is still a good fit.
18. Enable competition and innovation to promote the growth, financial stability, and efficiency of the insurance marketplace.
19. Enable innovation in the pricing of insurance through acceptance of predictive models, provided they are actuarially sound and in compliance with state laws.
20. Protect the confidentiality of filed predictive models and supporting information in accordance with state law.
21. Review predictive models in a timely manner to enable reasonable speed to market.

## 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.

## 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.

* 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?

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 syllabi. 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:

* *Foundations of Casualty Actuarial Science*, Fourth Edition (2001):
	+ 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.
* Casualty Actuarial Society (www.casact.org): “Basic Ratemaking.”
* American Institute for Chartered Property Casualty Underwriters: “Insurance Operations, Regulation, and Statutory Accounting,” Chapter Eight.
* Association of Insurance Compliance Professionals: “Ratemaking—What the State Filer Needs to Know.”
* Review of filings and approval of insurance company rates.
* NAIC’s Casualty Actuarial and Statistical Task Force white paper: “Regulatory Review of Predictive Models.”

**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.

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.*

# 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.

# 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. 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. 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 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.
	+ 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 is 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.
	+ "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.
	+ 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 paper topic 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 fiings), 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]](#footnote-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 practices[[17]](#footnote-17).

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

## 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.

## 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.

## 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.

## 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.

## 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**Correlation Matrix** - A correlation matrix is a table showing correlation coefficients between sets of variables. Each random variable (Xi) in the table is correlated with each of the other values in the table (Xj). 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] |
| **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". https://www.casact.org/pubs/dpp/dpp04/04dpp1.pdf |
| **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. |
| **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 x1 . . . xp, 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.[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 USAOr, with fine granularity, as multiple fields:· street address = 200 2nd Ave. South #358· city = St. Petersburg· state = FL· postal code = 33701-4313· country = USAOr, 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. [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. Suppose that predictor variables X1 and X2 interact. A GLM modeler could account for this interaction by including an interaction term of the form X1X2 in the formula for the linear predictor. For instance, rather than defining the linear predictor as η = β0 + β1X1 + β2X2, they could set η = β0 + β1X1 + β2X2 + β3X1X2.[ww]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 (X1, X2, ... Xk) in the model... e.g., β0 + β1x1 + β2x2.[18] |
| **Link Function** - The link function, η or g(μ), 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., η = g(E(Yi)) = E(Yi) for linear regression, or η = logit(π) 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 distguish betwee 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. https://www.casact.org/library/studynotes/werner\_modlin\_ratemaking.pdf |
| **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: |
| **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. |

 DRAFTING NOTE 10/15/19: WILL NEED TO CORRECT ALL FOOTNOTES. THE FOLLOWING IS ADDED FOR DRAFTING PURPOSES:

|  |
| --- |
| [11] www.casact.org/cms/pdf/Practical\_Guide\_for\_Evaluating\_Predictive\_Models\_Closter\_Carmean.pdf |
| [12] https://www.casact.org/newsletter/index.cfm?fa=viewart&id=6540 |
| [13] www.casact.org/pubs/monographs/papers/05-Goldburd-Khare-Tevet.pdf |
| [14] Dlabay, Les R.; Burrow, James L.; Brad, Brad (2009). Intro to Business. Mason, Ohio: South-Western Cengage Learning. p. 471. ISBN 978-0-538-44561-0. |
| [15] www.casact.org/pubs/monographs/papers/05-Goldburd-Khare-Tevet.pdf |
| [16] www.casact.org/library/studynotes/Werner\_Modlin\_Ratemaking.pdf |
| [17] https://docs.microsoft.com/en-us/analysis-services/data-mining/lift-chart-analysis-services-data-mining |
| [18] www.casact.org/pubs/monographs/papers/05-Goldburd-Khare-Tevet.pdf |
| [19] newonlinecourses.science.psu.edu/stat504/node/216 |
| [20] www.casact.org/cms/pdf/Practical\_Guide\_for\_Evaluating\_Predictive\_Models\_Closter\_Carmean.pdf |
| [21] https://en.wikipedia.org/wiki/OxfordDictionaries.com |
| [22] https://www.naic.org/documents/industry\_pcm\_p\_c\_2019.pdf |
| [23] https://www.casact.org/pubs/monographs/papers/05-Goldburd-Khare-Tevet.pdf |
| [24] https://www.casact.org/library/studynotes/Werner\_Modlin\_Ratemaking.pdf |
| [zz] https://en.wikipedia.org/wiki/Data\_dredging |
| [yy] https://en.wikipedia.org/wiki/Granularity#Data\_granularity |
| [xx] https://www.statisticshowto.datasciencecentral.com/correlation-matrix |
| [ww] To see that this second definition accounts for the interaction, note that it is equivalent to η = β0 + β1’X1 + β2X2 and to η = β0 + β1X1 + β2’X2, with β1’ = β1+ β3X2 and β2’ = β2 + β3X1. Since β1’ is a function of X2 and β2’ is a function of X1, these two equivalences say that the effect of X1 depends on the level of X2 and vice versa.REFERENCES:• “Generalized Linear Models for Insurance Rating.” CAS Monograph Series Number 5, by Mark Goldburd et al., Casualty Actuarial Society, 2016, pp. 52-58. Accessed at: https://www.casact.org/pubs/monographs/papers/05-Goldburd-Khare-Tevet.pdf• An Introduction to Statistical Learning: with Applications in R, by Gareth James et al., Springer, 2017, pp. 87–90. Accessed at: http://faculty.marshall.usc.edu/gareth-james/ISL/ISLR%20Seventh%20Printing.pdf |

# Appendix C – Sample Rate-Disruption Template







# Appendix D – Information Needed by Regulator Mapped into Best Practices

TBD

# Appendix E – References

https://www.casact.org/pubs/monographs/papers/05-Goldburd-Khare-Tevet.pdf

1. In this paper, references to “model” or “predictive model” are the same as “complex predictive model” unless qualified. [↑](#footnote-ref-1)
2. Bardach, E. and Patashnik, E. (2016.) *A Practical Guide for Policy Analysis, The Eightfold Path to More Effective Problem Solving.* Thousand Oaks, CA: CQ Press. See Appendix A for an overview of Bardach’s best-practice analysis. [↑](#footnote-ref-2)
3. Bogan, C.E. and English, M.J. (1994). Benchmarking for Best Practices: Winning Through Innovative Adaptation. New York: McGraw-Hill. [↑](#footnote-ref-3)
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. [↑](#footnote-ref-4)
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. [↑](#footnote-ref-5)
6. More information on model elements can be found in most statistics’ books. [↑](#footnote-ref-6)
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. [↑](#footnote-ref-7)
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." [↑](#footnote-ref-8)
9. Minutes of the Big Data (EX) Working Group, March 9, 2018: https://secure.naic.org/secure/minutes/2018\_spring/ex\_it\_tf.pdf?59 [↑](#footnote-ref-9)
10. All comments received by the end of January were posted to the NAIC website March 12 for review. [↑](#footnote-ref-10)
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. [↑](#footnote-ref-11)
12. There are some models that are made public by the vendor and would not result in a hindrance of the model's protection. [↑](#footnote-ref-12)
13. Refer to NAIC’s white paper titled *Regulatory Review of Predictive Models, found at the NAIC website*. [↑](#footnote-ref-13)
14. Bardach, E. and Patashnik, E. (2016.) *A Practical Guide for Policy Analysis, The Eightfold Path to More Effective Problem Solving.* Thousand Oaks, CA: CQ Press. See Appendix A for an overview of Bardach’s best-practice analysis. [↑](#footnote-ref-14)
15. Bogan, C.E. and English, M.J. (1994). Benchmarking for Best Practices: Winning Through Innovative Adaptation. New York: McGraw-Hill. [↑](#footnote-ref-15)
16. Ammons, D. N. and Roenigk, D. J. 2014. Benchmarking and Interorganizational Learning in Local Government. *Journal of Public Administration Research and Theory,* Volume 25, Issue 1. P 309-335. https://doi.org/10.1093/jopart/muu014 [↑](#footnote-ref-16)
17. Bardach, E. and Patashnik, E. 2016. *A Practical Guide for Policy Analysis: The Eightfold Path to More Effective Problem Solving.* Thousand Oaks, CA. CQ Press. [↑](#footnote-ref-17)