Casualty Actuarial and Statistical (C) Task Force

Regulatory Review of Predictive Models

Table of Contents

I. INTRODUCTION

Insurers' use of predictive analytics along with big data has significant potential benefits to both consumers and insurers by transforming the insurer-consumer experience into a more meaningful relationship. Predictive analytics can reveal insights into consumer behavior, lower the cost of insurance for many, and provide tools for the consumer 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 history, filing requirements and insurer presentations can be routinely organized to meet or exceed reviewers' needs and expectations. Hopefully, this paper helps bring more consistency and even uniformity 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] underlying rating plans. There were two charges given to CASTF by the Property and Casualty Insurance (C) Committee at the request of the Big Data (EX) Working Group:

A. <u>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.</u>

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

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

II. WHAT IS A "BEST PRACTICE?"

A best practice is a form of program evaluation in public policy. At its most basic level, a practice is a "tangible and visible behavior… [based on] an idea about how the actions…will solve a problem or achieve a goal" [2]. Best practices are used to maintain quality as an alternative to mandatory legislated standards and can be based on self-assessment or benchmarking.[3] Therefore, a best practice represents an effective method of problem solving.

A. <u>Key Regulatory Principles</u>

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

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

In this paper, best practices are presented in the form of guidance to regulators who review predictive models and to insurance companies filing rating plans that incorporate predictive models. Guidance will identify specific information useful to a regulator in the review of a predictive model, comment on what might be important about that information

---

[1] In this paper, reference to "model" or "predictive model" are the same as "complex predictive model" unless qualified.
[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.
[3] Bogan, C.E. and English, M.J. (1994). Benchmarking for Best Practices: Winning Through Innovative Adaptation. New York: McGraw-Hill.

and, where appropriate, provide insight as to when the information might identify an issue the regulator needs to be aware of or explore further.

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

The term "predictive model" refers to a set of models that use statistics to predict outcomes[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] 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 and they may imply an assumption of a constant variance across the range of a target variable.

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

A GLM consists of three elements[6]:

- A probability distribution from the exponential family.
- A linear predictor $\eta = X\beta$.
- A link function g such that $E(Y) = \mu = g^{-1}(\eta)$.

As can be seen in the description of the three GLM components above, it may take more than a casual introduction to statistics to comprehend the construction of a GLM. As stated earlier, a downside to GLMs is that it is more challenging to interpret the GLMs output than with univariate models. In addition, GLM output is assumed, as part of the model design, to be 100% credible no matter the size of the underlying data set. 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.

To further complicate regulatory review of models in the future, modeling methods are evolving rapidly and not limited 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 are seeking 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.[7] The Working Group circulated a proposal addressing aid

[4] A more thorough exploration of different predictive models will be found in many statistics' books, including Geisser, Seymour(September 2016). *Predictive Inference: An Introduction*. New York: Chapman & Hall.

[5] The generalized linear model (GLM) is a flexible family of models that are unified under a single method. Types of GLM include logistic regression, Poisson regression, gamma regression and multinomial regression.

[6] More information on model elements can be found in most statistics' books.

[7] Minutes of the Big Data (EX) Working Group, March 9, 2018: https://secure.naic.org/secure/minutes/2018_spring/ex_it_tf.pdf?59

to state insurance regulators in the review of predictive models as used in private passenger automobile and homeowners' 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.[8] 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. Best practices will aid regulatory reviewers by raising their level of model understanding. 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 the company more competitive in the marketplace. To make this work, both regulators and industry need to recognize that:

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

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

IV. SCOPE

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

Guidance and the knowledge needed to review predictive models identified in this paper are, in large part, transferrable to other types of predictive models, to other lines of business, or other insurance endeavors, e.g., commercial automobile, workers' compensation, marketing, underwriting, or claims. When transferring guidance to other lines of business or other insurance endeavors, 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 of an insurer acting on the output of any given predictive model. This paper does not delve into these possible considerations but regulators should be prepared to address them as they arise.

V. CONFIDENTIALITY

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

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

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

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

VII. PREDICTIVE MODELS – INFORMATION FOR REGULATORY REVIEW

This section of the paper identifies the information a regulator may need to review a predictive model used by an insurer to support a filed P/C 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 required 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.

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%).

A. Selecting Model Input

	Information	Importance to Regulator's Review "Essential" or "May Be Requested"	Comments
1. Available Data Sources			
A.1.a	Provide details of all data sources including the experience period for insurance data and when the data was last recorded or updated.	Essential	This information can be used to evaluate the completeness of the data source, integrity of the data source, relevance of the data to the predictive timeframe, the potential for historical bias, transparency to insured of the data source, and the ability of the insured to make corrections to the data source.
A.1.b	Specify the companies whose data is included in the datasets.	May Be Requested	If the filer is part of a group, do the datasets include data from affiliated companies? If so, which companies? If the filer is an advisory organization, what companies are used? Are the companies included in the data relevant and compatible to the company that filed the rating plan?
A.1.c	Provide the geographical scope and geographic exposure distribution of the data.	Essential	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.
A.1.d	List each data source. For each source, list all data elements used as input to the model that came from that source.	Essential	
A.1.e	Specify the type of data (e.g., accident year or policy year, text, numeric).	Essential	
A.1.f	Explain if internal or external data was used and if external data was used, disclose reliance on data supplied by others.	Essential	

A.1.g	Provide details of any non-insurance data used (customer-provided or other), including who owns this data, how consumers can verify their data and correct errors, whether the data was collected by use of a questionnaire/checklist, whether it was voluntarily reported by the applicant, and whether any of the variables are subject to the Fair Credit Reporting Act. If the data is from an outside source, what steps were taken to verify the data was accurate?	Essential	If the data is from a third-party source, the company should provide information on how the source addresses the questions in this consideration.

	2. *Sub-Models*		
A.2.a	Disclose reliance on sub-model output used as input to this model. If a sub-model was relied upon, provide the vendor name, and the name and version of the sub-model. If the sub-model was built/created in-house, provide contact information for the person responsible for the sub-model.	Essential	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.2.b	If using catastrophe model output, identify the vendor and the model settings/assumptions used when the model was run.	Essential	For example, it is important to know hurricane model settings for storm surge, demand surge, long/short-term views.
A.2.c	If using catastrophe model output (a sub-model) as input to the GLM under review, disclose whether loss associated with the modeled output was removed from the loss experience datasets.	Essential	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.
A.2.d	If using output of any scoring algorithms, provide a list of the variables used to determine the score and provide the source of the data used to calculate the score.	Essential	Any sub-model should be reviewed in the same manner as the primary model that uses the sub-model's output as input.
A.2.e	Was the sub-model previously approved (or accepted) by the regulatory agency?	Essential	If the sub-model was previously approved, that may change the extent of the sub-model's review.

3. Adjustments and Scrubbing			
A.3.a	Provide pre-scrubbed data distributions for each input.	May Be Requested	Compare these distributions to A.3.g
A.3.b	How was missing data handled?	Essential	
A.3.c	If duplicate records exist, how were they handled?	Essential	
A.3.d	Were any data outliers identified and subsequently adjusted? Name the outliers and explain the adjustments made to these outliers.	Essential	
A.3.e	Were premium, exposure, loss or expense data 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, what are the segments and how was the data adjusted?	Essential	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, those losses should be removed from the input data, e.g., large losses, flood, hurricane or severe convective storm models for PPA comprehensive or homeowners' loss.
A.3.f	What adjustments were made to raw data, e.g., transformations, binning and/or categorizations? If so, name the characteristic/variable and describe the adjustment.	Essential	
A.3.g	Provide post-scrubbed data distributions for each input.	May Be Requested	Compare these distributions to A.3.a
4. Data Organization			
A.4.a	Document the method of organization for compiling data, including procedures to merge data from different sources and a description of any preliminary analyses, data checks, and logical tests performed on the data and the results of those tests.	Essential	This should explain how data from separate sources was merged.
A.4.b	Document the process for reviewing the appropriateness, reasonableness, consistency and comprehensiveness of the data, including a justification of why the data makes sense.	Essential	For example, if by-peril modeling is performed, the documentation should be for each peril and make intuitive sense. For example, if "murder" or "theft" rates are used to predict the wind peril, provide support and a logical explanation.
A.4.c	Disclose material findings from the data review and identify any potential material limitations, defects, bias or	Essential	

	unresolved concerns found or believed to exist in the data.		
A.4.d	For any errors or material limitations in the data, explain how they were corrected.	Essential	
5. Final Data Information			
A.5.a	If the raw data selected to build the model is in a format that can be made available to the regulator, provide it.	May Be Requested	

B. Building the Model

	Information	**Importance to Regulator's Review** "Essential" or "May Be Requested"	Comments
1. High-Level Narrative for Building the Model			
B.1.a	Identify the type of model (e.g. Generalized Linear Model – GLM, decision tree, Bayesian Generalized Linear Model, Gradient-Boosting Machine, neural network, etc.), describe its role in the rating system and provide the reasons why that type of model is an appropriate choice for that role.	Essential	If by-peril or by-coverage modeling is used, the explanation should be by-peril/coverage.
B.1.b	A description of why the model (using the variables included in it) is appropriate for the line of business.	Essential	If by-peril, by-form or by-coverage modeling is used, the explanation should be by-peril/coverage/form.
B.1.c	Describe the model review process, from initial concept to final model. Keep this in overview narrative mode, less than 3 pages.	Essential	
B.1.d	Describe whether loss ratio, pure premium or frequency/severity analyses was performed and, if separate frequency/severity modeling was performed, how pure premiums were determined.	Essential	
B.1.e	What is the model's target variable?	Essential	A clear description of the target variable is key to understanding the purpose of the model.
B.1.f	Provide a detailed description of the variable selection process.	Essential	
B.1.g	Was input data segmented in any way, e.g., was modeling performed on a by-coverage or by-peril basis or by-form? Explain the form of data segmentation and the reasons for data segmentation.	Essential	The regulator would use this to follow the logic of the modeling process.
B.1.h	Describe any limitations or concerns in the analysis resulting from data issues and discuss the resulting impact on the modeling results.	Essential	

B.1.i	How data credibility (or lack thereof) was accounted for in the model building?	Essential	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	Describe any judgment used throughout the modeling process. Disclose assumptions used in constructing the model and provide support for these assumptions.	Essential	
B.2.b	If post-model adjustments were made to the data and the model was rerun, explain the details and the rationale. It is not necessary to discuss each iteration of adding and subtracting variables, but the regulator should be provided with a general description of how that was done, including any measures relied upon.	Essential	Evaluate the addition or removal of variables and the model fitting.
B.2.c	Describe the univariate testing and balancing that was performed during the model-building process, including a verbal summary of the thought processes involved.	Essential	Further elaboration from B.2.b.
B.2.d	Describe the 2-way testing and balancing that was performed during the model-building process, including a verbal summary of the thought processes of including (or not including) interaction terms.	Essential	Further elaboration from B.2.a and B.2.b.
B.2.e	For the GLM, what was the link function used? What distribution was used for the model (e.g., Poisson, Gaussian, log-normal, Tweedie)? Explain why the link function distribution was chosen. Provide the formulas for the distribution and link functions, including specific numerical parameters of the distribution.	Essential	
B.2.f	Were there data situations GLM weights were used? Describe these.	May Be Requested	Investigate whether identical records were combined to build the model.

	3. Predictor Variables		
B.3.a	Provide the names, descriptions and uses of each predictor variable, offset variable, control variable, proxy variable, geographic variable, geodemographic variable and all other variables in the model; explanations should not use programming language or code.	Essential	
B.3.b	For each predictor variable, state whether the variable is continuous, discrete or Boolean.	Essential	
B.3.c	Provide an intuitive argument for why an increase in each predictor variable should increase or decrease frequency, severity, loss costs, expenses, or whatever is being predicted.	Essential	
B.3.d	If the modeler used a Principal Component Analysis (PCA) approach, provide a narrative about that process, explain why PCA was used, and describe the step-by-step process used to transform observations (usually correlated) into a set of linearly uncorrelated variables. Include a listing of the PCA variable and its principal components.	Essential	
	4. Massaging Data, Model Validation and Goodness-of-Fit Measures		
B.4.a	Provide a description of how the available raw data was divided between model development, test and validation datasets. Describe all circumstances under which the testing and validation datasets were accessed.	Essential	

B.4.b	Describe the methods used to assess the statistical significance/goodness of the fit of the model, such as lift charts and statistical tests. Disclose whether the results are based on testing data, validation data and holdout samples. Ensure that the assessment includes model projection results compared to historical actual results to verify that modeled results bear a reasonable relationship to actual results. Discuss the results.	Essential	Some states require state-only data to test the plan, especially for analysis where using the 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.
B.4.c	Describe any adjustments that were made in the data with respect to scaling for discrete variables or binning the data.	Essential	
B.4.d	Describe any transformations made for continuous variables.	Essential	
B.4.e	For each discrete variable level, provide the parameter value, confidence intervals, chi-square tests, p-values and any other relevant and material tests. Were model development data, validation data, test data or other data used for these tests?	Essential	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.
B.4.f	Identify the threshold for statistical significance and explain why it was selected. Provide a verbal defense for keeping the variable for each discrete variable level where the p-values were not less than the chosen threshold.	Essential	See Comment for B.4.e.
B.4.g	For overall discrete variables, provide type 3 chi-square tests, p-values, F tests and any other relevant and material test. Were model development data, validation data, test data or other data used for these tests?	Essential	See Comment for B.4.e.
B.4.h	For continuous variables, provide confidence intervals, chi-square tests, p-values and any other relevant and material test. Were model development data, validation data, test data or other data used for these tests?	Essential	See Comment for B.4.e.
B.4.i	Describe how the model was tested for stability over time.	Essential	Evaluate the build/test/validation datasets for potential model distortions (e.g., a winter storm in year 3 of 5 can distort the model in both the testing and validation datasets).

B.4.j	Describe how the model was tested for geographic stability, e.g., across states or territories within state.	Essential	Evaluate the geographic splits for potential model distortions.
B.4.k	Describe how overfitting was addressed and the results of correlation tests.	Essential	
B.4.l	Provide support demonstrating that the GLM assumptions are appropriate (for example, the choice of error distribution).	Essential	Visual review of plots of actual errors is usually sufficient.
B.4.m	Provide the formula relationship between the data and the model outputs, with a definition of each model input and output. Provide all necessary coefficients to evaluate the predicted value for any real or hypothetical set of inputs.	Essential	B.4.l and B.4.m will show the mathematical functions involved and could be used to reproduce some model predictions.
B.4.n	Provide 5-10 sample records and the output of the model for those records.	Essential	
5. "Old Model" Versus "New Model"			
B.5.a	An explanation of why this model is better than the one it is replacing. How was that conclusion formed? What metrics were relied on for measurement?	Essential	Regulators should expect to see improvement in the new class plan's predictive ability or other sufficient reason for the change.
B.5.b	Were 2 Gini coefficients compared? What was the conclusion drawn from this comparison?	May Be Requested	One example of a comparison might be sufficient.
B.5.c	Were double lift charts analyzed? What was the conclusion drawn from this analysis?	Essential	One example of a comparison might be sufficient.
B.5.d	Provide a list of all **new** predictor variables in the model that were not in the prior model.	Essential	Useful to differentiate between old and new variables so the regulator can prioritize more time on factors not yet reviewed.
B.5.e	Provide a list of predictor variables used in the old model that are not used in the new model. Why were they dropped from the new model?	Essential	

6. Modeler/Software			
B.6.a	Provide the names, contact emails, phone numbers and qualifications of the key persons who: a. Led the project b. Compiled the data c. Built the model d. Performed peer review	Essential	
B.6.b	What software was used? Provide the name of the software vender/developer, software product and a software version reference.	Essential	
B.6.c	When did work to build the model begin and when was the model build finalized?	Essential	

C. The Filed Rating Plan

	Information	Importance to Regulator's Review "Essential" or "May Be Requested"	Comments
1. General Impact of Model on Rating Algorithm			
C.1.a	In the Actuarial Memorandum section on the SERFF Supporting Documentation tab, for each model relied upon, include a document that explains the model and its role in the rating system.	Essential	This item becomes "Essential" 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	Provide an explanation of how the model was used to adjust the rating algorithm.	Essential	
C.1.c	Provide a complete list of all 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 used to calculate a premium. For each characteristic/variable, indicate 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 provide transparent descriptions of each listed characteristic/variable.	Essential	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.
C.1.d	For each characteristic/variable used as both input to the model (including sub-models and composite variables) and as a separate univariate rating characteristic, explain how these are tempered or adjusted to account for possible overlap or redundancy in what the characteristic/variable measures.	Essential	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.

C.1.e	If the filing support includes an update or replacement of an existing model, identify and explain the changes in calculations, assumptions, parameters and data used to build the models. Provide an explanation of why the updated/replacement model is better than the one it is replacing, including, how that conclusion was reached, and the metrics relied upon to reach that conclusion.	Essential	
2. Relevance of Variables / Relationship to Risk of Loss			
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).	Essential	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	Provide a comparison between relativities indicated by the model to both current relativities and the insurer's selected relativities for each risk characteristic/variable in the rating plan. Each significant difference should be highlighted and explained.	Essential	“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	What calculations, judgments and adjustments, if any, were made before using the model output in the rating system? Identify any adjustments that were made to the indicated model to derive the selected model.	Essential	
C.3.c	If the model results in scores, tiers, or ranges of values for which indications are then derived for each such resulting category, provide explanations for filed rating values that deviate from these indications and supporting information/ analyses. For example, identify any adjustments that were made to the factors indicated for each category of the model outputs to derive the factors selected for the rating plan.	Essential	This is especially important if deviations are material and/or impact one consumer population more than another.

4. Responses to Data, Credibility and Granularity Issues			
C.4.a	What consideration was given to the credibility of the output data?	Essential	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 applicable, discuss the rationale for using a model that is more granular than the rating plan.	Essential	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 applicable, discuss the rationale for using a rating plan that is more granular than modeled output.	Essential	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	Provide a transparent presentation and explanation of binning decisions that assign ranges of model outputs to particular rating categories.	Essential	
C.5.b	Provide complete definitions of any rating tiers or other intermediate rating categories that translate the model outputs into some other structure that is then presented within the rate and/or rule pages.	Essential	
6. Supporting Data			
C.6.a	Provide state-specific, book-of-business-specific univariate historical experience data consisting of, at minimum, earned premiums, incurred losses, loss ratios and loss ratio relativities for each category of model output(s) proposed to be used within the rating plan.	Essential	

C.6.b	Provide an explanation of any material (especially directional) differences between model indications and state-specific univariate indications.	Essential	Multivariate indications may be reasonable as refinements to univariate indications, but likely not for bringing about 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.
7. Consumer Impacts			
C.7.a	Identify model changes and rating variables that will cause large premium disruptions.	Essential	
C.7.b	Did the insurer perform sensitivity testing to identify significant changes in premium due to small or incremental change in a single risk characteristic? If so, discuss and provide the results of that testing.	May Be Requested	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.
C.7.c	Measure and describe the impacts on expiring policies and describe the process used by management to mitigate or get comfortable with those impacts.	Essential	
C.7.d	Provide a rate disruption analysis, demonstrating the distribution of percentage impacts on renewal business (create by rerating the current book of business). Include the largest dollar and percentage impacts arising from the filing, including (desirably) the impacts arising specifically from the adoption of the model or changes to the model as they translate into the proposed rating plan.	Essential	While the default request would typically be for the distribution 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	Provide exposure distributions for output variables and show the effects of rate changes at granular and summary levels.	Essential	See Appendix C for an example of an exposure distribution.
C.7.f	Explain how the insurer will help educate consumers to mitigate their risk.	Essential	

C.7.g	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?	Essential	Could be "Essential" if the variables/ characteristics used could 1) have public-policy implications, 2) result in erroneous information being used, or 3) result in many large, disruptive premium changes at renewal. Another consideration to judge “importance” is whether consumers are proactively involved (e.g., use of consumer credit information and credit-report accuracy issues).
C.7.h	Identify rating variables that remain static over a consumer’s lifetime versus those that will be updated periodically. Document guidelines for variables that are listed as static yet for which the underlying consumer attributes may change over time.	Essential	
C.7.i	Provide the regulator with a description of how the company will respond to consumers’ inquiries about how their premium was calculated.	Essential	
C.7.j	Provide the regulator with a means to calculate the rate charged a consumer.	May Be Requested	Especially 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 can allow the regulator to perform sensitivity testing when there are small changes to a risk characteristic/variable.

8. Accurate Translation of Model into a Rating Plan			
C.8.a	Provide sufficient information for the reviewer to be able to understand how the model outputs are used within the rating system and to verify that the rating plan, in fact, reflects the model output and any adjustments made to the model output.	Essential	

VIII. PROPOSED CHANGES TO THE *PRODUCT FILING REVIEW HANDBOOK*

TBD – placeholder to include best practices for review of predictive models and analytics filed by insurers to justify rates

IX. PROPOSED STATE GUIDANCE

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

X. OTHER CONSIDERATIONS

During the development of this paper a 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 issues may be discussed elsewhere within NAIC. All of these issues, if addressed, will be addressed by each state on a case-by-case basis. The topics for consideration include:

- TBD

XI. RECOMMENDATIONS GOING FORWARD

- TBD

XII. 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[9].

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[10].

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

A. Scope

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

B. Identify Top Performers

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

C. Analyze Best Practices

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

D. Adapt

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

E. Implementation and evaluation

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

---

[9] 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

[10] 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.

XIII. APPENDIX B - - GLOSSARY OF TERMS

Offset vs. control factors - TBD

Probability Distribution - TBD

Exponential Family - TBD

Linear Predictor - TBD

Link Function - TBD

Univariate Model - TBD

Generalized Linear Model - TBD

Private Passenger Automobile Insurance – TBD

Homeowners Insurance – TBD

Rating algorithm – TBD

Rating plan – TBD

Rating system – TBD

PCA approach (Principal Component Analysis) – The PCA approach is also known as factor analysis. These methods create 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.

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.[11] 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.

Overfitting - TBD

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:

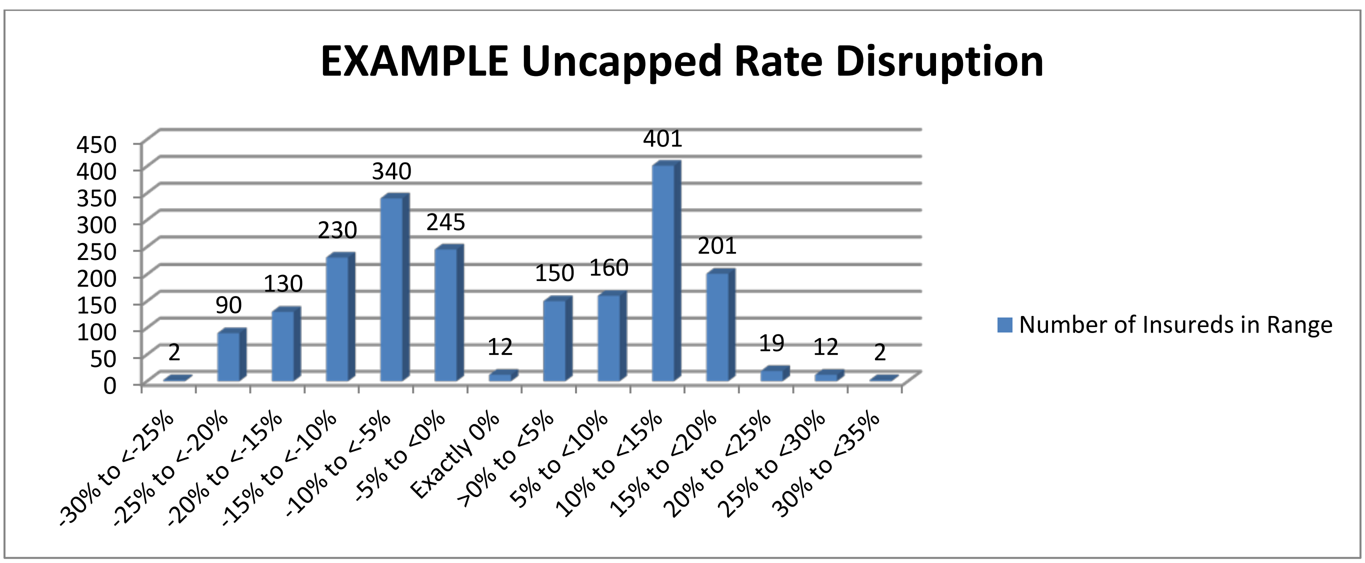
1. People who live in the same neighborhood are more likely to have similar characteristics than are two people chosen at random.
2. Neighborhoods can be categorized in terms of the characteristics of the population that they contain. Any two neighborhoods can be placed in the same category, i.e., they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

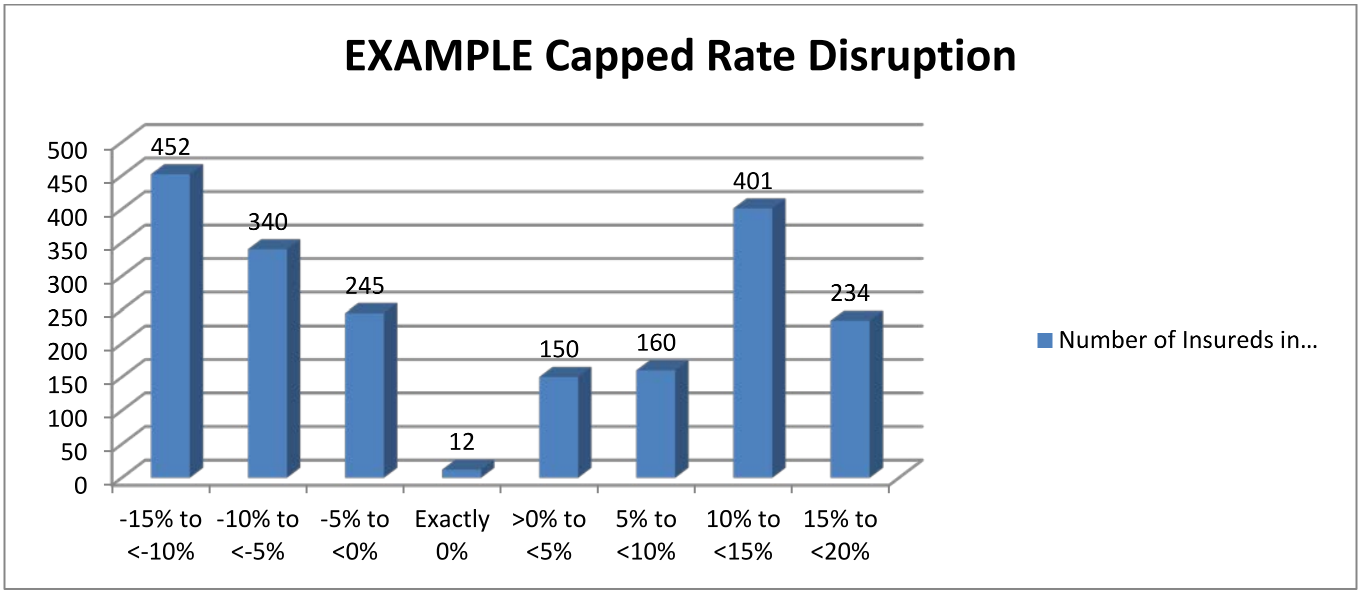
Etc.

[11] 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.

XIV. APPENDIX C – SAMPLE RATE-DISRUPTION TEMPLATE

State Division of Insurance - EXAMPLE for Rate Disruption		*Template Updated October 2018*	
● First, fill in the boxes for minimum and maximum individual impacts, shaded in light blue. Default values in the cells are examples only.			
● The appropriate percent-change ranges will then be generated based on the maximum/minimum changes.			
● For every box shaded in light green, replace "ENTER VALUE" with the number of affected insureds within the corresponding change range.			
● Once all values are filled in, use the "Charts" feature in Excel to generate a histogram to visually display the spread of impacts.			
NOTE: Values of Minimum % Change, Maximum % Change, and Total Number of Insureds must reconcile to the Rate/Rule Schedule in SERFF.			
	Uncapped		***Capped (If Applicable)***
Minimum % Change	-30.000%	**Minimum % Change**	-15.000%
Maxmium % Change	30.000%	**Maxmium % Change**	15.000%
Total Number of Insureds (Auto-Calculated)	1994	**Total Number of Insureds (Auto-Calculated)**	1994
	Uncapped Rate Disruption	***Capped Rate Disruption (If Applicable)***	
Percent-Change Range	**Number of Insureds in Range**	**Percent-Change Range**	**Number of Insureds in Range**
-30% to <-25%	2	-15% to <-10%	452
-25% to <-20%	90	-10% to <-5%	340
-20% to <-15%	130	-5% to <0%	245
-15% to <-10%	230	Exactly 0%	12
-10% to <-5%	340	>0% to <5%	150
-5% to <0%	245	5% to <10%	160
Exactly 0%	12	10% to <15%	401
>0% to <5%	150	15% to <20%	234
5% to <10%	160		
10% to <15%	401		
15% to <20%	201		
20% to <25%	19		
25% to <30%	12		
30% to <35%	2		





State Division of Insurance - EXAMPLE for Largest Percentage Increase	*Template Updated October 2018*

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

Largest Percentage Increase		***Corresponding Dollar Increase (for Insured Receiving Largest Percentage Increase)***			
Uncapped Change	30.00%	**Uncapped Dollar Change**	$165.00	**Current Premium**	$550.00
Capped Change (If Applicable)	15.00%	**Capped $ Change (If Applicable)**	$82.50	**Proposed Premium**	$632.50

Characteristics of Policy (Fill in Below)

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

● **For Home Insurance:** At minimum, identify age and gender of each named insured, amount of insurance, territory, construction type, protection class, any prior loss history, and any other key attributes whose treatments are affected by this filing.

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

Vehicle:	**BI Limits:**	**PD Limits:**	**UM/UIM Limits:**	**MED Limits:**	**COMP Deductible:**	**COLL Deductible:**
2009 Ford Focus	$50,000 / $100,000	$25,000	$50,000 / $100,000	$5,000	$500	$1,000
2003 Honda Accord	$25,000 / $50,000	$10,000	$25,000 / $50,000	$1,000	$500	$1,000

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

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

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

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

Attribute	**% Impact (Uncapped)**	**Dollar Impact (Uncapped)**
Insured Age (M/25)	12.00%	$66.00
COLL Deductible ($1,000)	10.00%	$61.60
Territory (89105)	4.00%	$27.10
Vehicle Symbol (2003 Honda Accord)	1.46%	$10.29
Effect of Capping	-11.54%	-$82.50
TOTAL	**15.00%**	**$82.50**

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

Check all options that apply below.

☐ 12-Month Policies
☑ 6-Month Policies
☐ 3-Month Policies
☐ Other (SPECIFY)

State Division of Insurance - EXAMPLE for Largest Dollar Increase				*Template Updated October 2018*		
● Fill in fields highlighted in light green.						
Largest Dollar Increase		***Corresponding Percentage Increase (for Insured Receiving Largest Dollar Increase)***				
Uncapped Change	$306.60	**Current Premium**	$2,555.00	**Uncapped Percent Change**		12.00%
Capped Change (If Applicable)	$306.60	**Proposed Premium**	$2,861.60	**Capped % Change (If Applicable)**		12.00%
Characteristics of Policy (Fill in Below)						

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

● **For Home Insurance:** At minimum, identify age and gender of each named insured, amount of insurance, territory, construction type, protection class, any prior loss history, and any other key attributes whose treatments are affected by this filing.

Automobile policy: Two insureds - Male (Age 33), Female (Age 32). **Territory:** Reno, ZIP Code 89504.

Vehicle:	**BI Limits:**	**PD Limits:**	**UM/UIM Limits:**	**MED Limits:**	**COMP Deductible:**	**COLL Deductible:**
2016 Tesla Model S	$200,000 / $600,000	$50,000	$200,000 / $600,000	$10,000	$2,500	$2,500
2015 Mercedes-Benz C-Class (W205)	$200,000 / $600,000	$50,000	$200,000 / $600,000	$10,000	$2,500	$2,500

1 prior at-fault accident for 32-year-old female. Policy receives EFT discount and loyalty discount.

Primary impacts are the increases to the relativities for the age of insured, symbol for 2015 Mercedes-Benz C-Class, and increased-limit factors for Property Damage and Medical Payments coverages.

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

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

Attribute	**% Impact (Uncapped)**	**Dollar Impact (Uncapped)**
Insured Age (M/33)	3.15%	$80.48
Insured Age (F/32)	3.23%	$85.13
Vehicle Symbol (2015 Mercedes-Benz C-Class)	2.45%	$66.65
Increased-Limit Factor for PD	1.55%	$43.20
Increased-Limit Factor for MED	1.10%	$31.14
TOTAL	**12.00%**	**$306.60**

XV. APPENDIX D – INFORMATION NEEDED BY REGULATOR MAPPED INTO BEST PRACTICES

XVI. APPENDIX E – REFERENCES