APPENDIX B-RGLM INFORMATION ELEMENTS AND GUIDANCE FOR A REGULATOR TO MEET BEST ~~GLMS~~REGULARIZED GENERALIZED LINEAR MODELS)

This appendix identifies the information a state insurance regulator may need to review a ~~predictive~~ regularized general linear model used by an insurer to support a personal automobile or home insurance rating plan. Regularized Generalized Linear Models include lasso, derivative lasso, lasso credibility, ridge, elastic net, and accurate generalized linear models (AGLM). Other modeling approaches may fall within the category of regularized generalized linear models. The main distinguishing feature of regularized GLMs is that they have complexity penalty hyper parameter(s) (a.k.a. shrinkage factors) which put constraints on the model such that the coefficients are tempered from what they would be in a standard (unpenalized) Generalized Linear Model (GLM). Generally, if the complexity penalties in a regularized GLM are set to zero, the model indications will be identical to those achieved from a standard GLM. The list of information elements below 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 sufficient information that helps determine if the rating plan meets state-specific filing and legal requirements.

Documentation of the design and operational details of the model will help ensure the business continuity and transparency of the models used. Documentation should be sufficiently detailed and complete to enable a qualified third party to form a sound judgment on the suitability of the model for the intended purpose. The theory, assumptions, methodologies, software, and empirical bases should be explained, as well as the data used in developing and implementing the model. Relevant testing and ongoing performance testing 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

61aTom37 444.98 Tm0eg0. TET( )]TETQq0.00000912 0 612 792 reW*nBT/F2 7.176 Tf1 0 0 146l.

the information in Level 1, Level 2, and Level 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 by a particular state. It is typically requested only if the reviewer has serious concerns that the model may produce rates or rating factors that are excessive, inadequate, and/or unfairly discriminatory.

~~Lastly, although the best practices presented in this white paper will readily be transferrable to review of other predictive models, the information elements presented here might be useful only with deeper adaptations when starting to review different types of predictive~~ models. If the

Section	Information Element	Level of

Section	Information Element	Level of Importance to the Review	Comments

0 0 gq[(ca_ET_O_g_26.496 578.714599474 47.04 151.19 n_/Span2 7.176 Tf_1 0 0 1 220.973557.6247re_g_0 Glem)-_ET_O_g_210.22 578.714599474 47.04 151.19 n_/Span2 7.176 Tf_1 0

Section	Information Element	Level of Importance to the Review	Comments
3. Adjustments to Data			
A.3.a	Determine if premium, exposure, loss, or expense data were adjusted (e.g., on-leveled, developed, trended, adjusted for catastrophe experience, or capped). If so, how? Do the adjustments vary for different segments of the data? If so, identify the segments and how the data was adjusted.	2	

A.3.c	Ask for aggregated data (one dataset of pre-adjusted/scrubbed data and one dataset 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 are 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,

A.4.c	Identify material findings the company had during its 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

~~B.~~ BUILDING THE MODEL

~~C.~~B.

Section	Information Element	Level of Importance to the Reviews	Comments
1. High-Level Narrative for Building the Model			
B.1.a	Identify the type of model underlying the rate filing (e.g., ~~GLM, decision tree, Bayesian GLM, gradient boosting machine, neural network, etc.~~lasso regression, ridge regression, elastic net regression, etc). Understand the model's role in the rating system and provide thereasons why that type of model is an appropriate choice for		

Section

			Alternatively, the regulator could ask for a plot where the X-axis is the hyperparameter value and there are separate lines representing the coefficient value for each variable given the complexity hyperparameter.

Not Expanded by / Condensed by

~~For all variables (discrete or continuous), review the appropriate parameter values and relevant tests of~~
sim011 0 1Q58.296ţ4.52 7ion

B.4.b

~~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, p471 0.831 RG[ )ЈTETQq2

~~B.4.h~~	~~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 inthe 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.~~i~~d	Obtain a description how the model was tested for stability over time.	2	[redacted]

~~D.~~C. THE FILED RATING PLAN

Section	Information Element	Level of Importance to the Reviews	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 (i.e., 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 filed rating algorithm.	1	Models are

Section	Information Element	Level of Importance to the Review	Comments
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)		

Section	Information Element	Level of Importance to the Reviews	Comments
7. Consumer Impacts			
C.7.a	Obtain a listing of the top five rating variables that contribute the most to large swings in renewal premium, both as increases and decreases, as well as the top five rating variables with the largest spread of impact for both new and renewal business.	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 company performed sensitivity testing to identify significant changes in premium due to small or incremental change in a single risk characteristic. If such testing was		

Section	Information Element	Level of Importance to the Reviews	Comments
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 theinsured's risk profile based on "static" variables changes over ~~time~~time, but the rate charged, based on a new business insurance score or tier

Section	Information Element	Level of Importance to the s Review	Comments
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 anyadjustments 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.
9. Efficient and Effective Review of			