NATIONAL MEETING
SUMMER / CHICAGO
2024 NAIC

Draft date: 8/7/24

2024 Summer National Meeting
Chicago, Illinois

CASUALTY ACTUARIAL AND STATISTICAL (C) TASK FORCE
Tuesday, August 13, 2024
2:00 – 3:30 p.m.
McCormick Place Convention Center—S105—Level 1

ROLL CALL

D.J. Bettencourt, Chair	New Hampshire	Anita G. Fox	Michigan
Chlora Lindley-Myers, Vice Chair	Missouri	Grace Arnold	Minnesota
Mark Fowler	Alabama	Eric Dunning	Nebraska
Lori K. Wing-Heier	Alaska	Justin Zimmerman	New Jersey
Barbara D. Richardson	Arizona	Alice T. Kane	New Mexico
Ricardo Lara	California	Judith L. French	Ohio
Andrew N. Mais	Connecticut	Glen Mulready	Oklahoma
Karima M. Woods	District of Columbia	Andrew R. Stolfi	Oregon
Michael Yaworsky	Florida	Michael Humphreys	Pennsylvania
Gordon I. Ito	Hawaii	Alexander S. Adams Vega	Puerto Rico
Amy L. Beard	Indiana	Michael Wise	South Carolina
Doug Ommen	Iowa	Cassie Brown	Texas
Vicki Schmidt	Kansas	Kevin Gaffney	Vermont
Timothy J. Temple	Louisiana	Mike Kreidler	Washington
Robert L. Carey	Maine	Allan L. McVey	West Virginia
Joy Y. Hatchette	Maryland		

NAIC Support Staff: Kris DeFrain/Roberto Perez

AGENDA

1. Consider Adoption of its July 9, June 17, May 7, March 20, and Spring National Meeting Minutes—*Christian Citarella (NH)* — Attachment One, Attachment Two, Attachment Three, Attachment Four, Attachment Five
2. Consider Adoption of its Working Group Reports
   A. Actuarial Opinion (C) Working Group—*Miriam Fisk (TX* — Attachment Six
   B. Statistical Data (C) Working Group—*Sandra Darby (ME)* — Attachment Seven

3. Hear a Presentation from the Casualty Actuarial Society (CAS) on Race and Insurance—*Ken Williams (CAS)* — Attachment Eight

4. Discuss the Private Flood Insurance Supplement
—*Michael McKenney (PA)*

5. Hear Liaison Reports—*Christian Citarella (NH)*

6. Consider Exposure of a Draft White Paper Appendix on Penalized Regression (Lasso and Ridge)—*Sam Kloese (NAIC)* — Attachment Nine

7. Hear Activity and Research Updates from Professional Actuarial Associations—*Christian Citarella (NH)* — Attachment Ten

8. Discuss Any Other Matters Brought Before the Task Force
—*Christian Citarella (NH)*
   A. Report on the Status of the Schedule P Instructions Proposal
   —*Kris DeFrain (NAIC)*

9. Adjournment—*Christian Citarella (NH)*

Draft: 7/31/2024

Casualty Actuarial and Statistical (C) Task Force
Virtual Meeting
July 9, 2024

The Casualty Actuarial and Statistical (C) Task Force met July 9, 2024. The following Task Force members participated: Chlora Lindley-Myers, Vice Chair, represented by Cindy Amann and Julie Lederer (MO); Lori K. Wing-Heier represented by Sian Ng-Ashcraft (AK); Mark Fowler represented by Charles Hale (AL); Barbara D. Richardson represented by Tom Zuppan (AZ); Ricardo Lara represented by Mitra Sanandajifar (CA); Andrew N. Mais represented by Wanchin Chou (CT); Karima M. Woods represented by David Christhilf (DC); Michael Yaworsky represented by Virginia Christy (FL); Doug Ommen represented by Travis Grassel (IA); Amy L. Beard represented by Patrick O'Connor and Larry Steinert (IN); Vicki Schmidt represented by Nicole Boyd (KS); Timothy J. Temple represented by Arthur Schwartz (LA); Robert L. Carey represented by Sandra Darby (ME); Anita G. Fox represented by Kevin Dyke (MI); Eric Dunning represented by Michael Muldoon (NE); Justin Zimmerman represented by Sam Sackey (NJ); Alice T. Kane represented by Christian Myers (NM); Judith L. French represented by Tom Botsko (OH); Glen Mulready represented by Andrew Schallhorn (OK); Michael Humphreys represented by Michael McKenney (PA); Cassie Brown represented by Miriam Fisk (TX); Kevin Gaffney and Rosemary Raszka (VT); Mike Kreidler represented by Eric Slavich (WA); and Allan L. McVey (WV).

1. Heard an Update on Rate Filing Issues

Lederer chaired the meeting and reported on two rate filing issues.

The first issue is that some companies are creating unknown or missing rates in inappropriate situations. For example, if the missing information is something that the company is responsible for collecting (e.g., limits, deductibles, mileage, vehicle model), many regulators find that the company is responsible for obtaining that information and recording it in the customer's file before rating.

The other issue is whether to apply a rating variable to a geographical area rather than at an individual insured level. For example, rating dog ownership in a geographical area is not considered appropriate. Dog ownership varies by individual household, and data should be collected to appropriately rate this at the household level. Just because data is available by geographical area does not mean that the insurance rating is appropriate.

2. Adopted the Report of the Actuarial Opinion (C) Working Group

Fisk said the Actuarial Opinion (C) Working Group met on June 25 in regulator-to-regulator session, pursuant to paragraph 3 of the NAIC Policy Statement on Open Meetings, for the annual discussion of observations resulting from regulators' review of the 2023 Statements of Actuarial Opinion. She said that in the coming weeks, the Working Group will begin discussing potential changes to the 2024 Regulatory Guidance document and the 2025 Statement of Actuarial Opinion instructions.

Fisk made a motion, seconded by Darby, to adopt the report of the Actuarial Opinion (C) Working Group. The motion passed unanimously.

3. Adopted the Report of the Statistical Data (C) Working Group

Darby said the Statistical Data (C) Working Group met May 30 to discuss proposed changes to the *Report on Profitability by Line by State* (Profitability Report) and the *Dwelling, Fire, Homeowners Owner-Occupied, and Homeowners Tenant and Condominium/Cooperative Unit Owner's Insurance Report* (Homeowners Report). She said no changes were adopted. The Working Group will meet again in late July.

Darby made a motion, seconded by Dyke, to adopt the report of the Statistical Data (C) Working Group. The motion passed unanimously.

4. Discussed the Academy's Cyber Risk Toolkit

Citarella, Lederer, Darby, Grassel, Schwartz, and Kris DeFrain (NAIC) began a presentation about the American Academy of Actuaries' (Academy's) Cyber Risk Toolkit at the Task Force's May 7 meeting. The presentation concluded during this meeting.

5. Heard a Report on the Cybersecurity (H) Working Group

Amann and Miguel Romero (NAIC) provided the Task Force with a presentation on activities of the Cybersecurity (H) Working Group (Attachment __).

6. Discussed Reporting for the Private Flood Insurance Supplement

McKenney said that there have been many reporting issues with the NAIC Private Flood Supplement to the Annual Statement. He said some insurers are interpreting "first dollar" to mean there cannot be any deductible, while others are interpreting it to mean "primary insurance regardless of the existence of any deductible." Other issues include writing flood insurance by endorsement combined with other perils and not completing the supplement, as well as not excluding flood on some residential insurance policies and not completing the supplement.

He said the entirety of the flood insurance coverage should be reported, and he did not think that was happening. He said he thought reporting should also be separated by owner-occupied private residential dwellings, renters, condo owners, mobile homeowners, and secondary/seasonal. Also, he recommended ensuring consistency with what is reported by alien insurers to the NAIC International Insurers Department (IID).

Schwartz, Chou, J'ne Byckovski (TX), and Jackie Horigan (MA) volunteered to join McKenney on a volunteer drafting group to write a proposal. The Task Force will consider whether to submit to the Blanks (E) Working Group a proposal to change the Private Flood Insurance Supplement. McKenney said other volunteers are welcome to participate.

Having no further business, the Casualty Actuarial and Statistical (C) Task Force adjourned.

SharePoint/NAIC Support Staff Hub/Member Meetings/C CMTE/2024_Spring/CASTF/070924 Minutes.docx

Draft: 7/9/24

Casualty Actuarial and Statistical (C) Task Force
E-Vote
June 17, 2024

The Casualty Actuarial and Statistical (C) Task Force conducted an e-vote that concluded June 17, 2024. The following Task Force members participated: D.J. Bettencourt, Chair, represented by Christian Citarella (NH); Chlora Lindley-Myers, Vice Chair, represented by Julie Lederer (MO); Lori K. Wing-Heier represented by Sian Ng-Ashcraft (AK); Barbara D. Richardson represented by Tom Zuppan (AZ); Andrew N. Mais represented by Wanchin Chou (CT); Karima M. Woods represented by David Christhilf (DC); Michael Yaworsky represented by Alexis Bakofsky (FL); Doug Ommen represented by Travis Grassel (IA); Amy L. Beard represented by Larry Steinert (IN); Vicki Schmidt represented by Nicole Boyd (KS); Robert L. Carey represented by Sandra Darby (ME); Grace Arnold represented by Phil Vigliaturo (MN); Justin Zimmerman represented by Sam Sackey (NJ); Judith L. French represented by Tom Botsko (OH); Glen Mulready represented by Andrew Schallhorn (OK); Andrew R. Stolfi represented by David Dahl (OR); Michael Humphreys represented by Michael McKenney (PA); Michael Wise represented by Will Davis (SC); Cassie Brown represented by Nicole Elliott (TX); Kevin Gaffney represented by Rosemary Raszka (VT); Mike Kreidler represented by Eric Slavich (WA); and Allan L. McVey (WV).

1. Adopted the *2022 Auto Insurance Database Average Premium Supplement*

The Task Force conducted an e-vote to consider adoption of the *2022 Auto Insurance Database Average Premium Supplement*. The motion passed unanimously.

Having no further business, the Casualty Actuarial and Statistical (C) Task Force adjourned.

SharePoint/NAIC Support Staff Hub/Member Meetings/C CMTE/2024_Summer/CASTF/061724 evote min.docx

Draft: 7/31/2024

Casualty Actuarial and Statistical (C) Task Force
Virtual Meeting
May 7, 2024

The Casualty Actuarial and Statistical (C) Task Force met May 7, 2024. The following Task Force members participated: D.J. Bettencourt, Chair, represented by Christian Citarella (NH); Chlora Lindley-Myers, Vice Chair, represented by Julie Lederer (MO); Lori K. Wing-Heier represented by Sian Ng-Ashcraft (AK); Mark Fowler represented by Charles Hale (AL); Ricardo Lara represented by Mitra Sanandajifar (CA); Andrew N. Mais represented by Susan Andrews and Qing He (CT); Michael Yaworsky represented by Peshala Disanayaka (FL); Gordon I. Ito represented by Randy Jacobson (HI); Doug Ommen represented by Travis Grassel (IA); Vicki Schmidt represented by Nicole Boyd (KS); Timothy J. Temple represented by Arthur Schwartz (LA); Kathleen A. Birrane represented by Walter Dabrowski and Bill Fawcett (MD); Timothy N. Schott represented by Sandra Darby (ME); Anita G. Fox represented by Kevin Dyke (MI); Grace Arnold represented by Phil Vigliaturo (MN); Eric Dunning represented by Michael Muldoon (NE); Justin Zimmerman represented by Sam Sackey (NJ); Alice T. Kane represented by Christian Myers (NM); Judith L. French represented by Tom Botsko (OH); Glen Mulready represented by Andrew Schallhorn (OK); Michael Humphreys represented by James DiSanto and Bojan Zorkic (PA); Michael Wise represented by Will Davis (SC); Cassie Brown represented by J'ne Byckovski and Miriam Fisk (TX); Kevin Gaffney represented by Zoie Y. Swaim (VT); and Mike Kreidler represented by Eric Slavich (WA).

1. Adopted the Report of the Actuarial Opinion (C) Working Group

Fisk said the Actuarial Opinion (C) Working Group plans to meet in June in regulator-to-regulator session, pursuant to paragraph #3: *Specific companies, entities or individuals, including, but not limited to, collaborative financial and market conduct examinations and analysis* of the NAIC Policy Statement on Open Meetings, to discuss observations and issues that have come up during the process of reviewing 2023 opinions. Next, the Working Group plans to discuss 2024 regulatory guidance and 2025 instructions.

Fisk made a motion, seconded by Dyke, to adopt the report of the Actuarial Opinion (C) Working Group. The motion passed unanimously.

2. Adopted the Report of the Statistical Data (C) Working Group

The Statistical Data (C) Working Group met April 25 in regulator-to-regulator session, pursuant to paragraph #3: *Specific companies, entities or individuals, including, but not limited to, collaborative financial and market conduct examinations and analysis* of the NAIC Policy Statement on Open Meetings, to discuss data for the *Auto Insurance Database Average Premium Supplement* (Auto Supplement) and the *Dwelling, Fire, Homeowners Owner-Occupied, and Homeowners Tenant and Condominium/Cooperative Unit Owner's Insurance Report* (Homeowners Report). The Working Group also discussed a work plan that includes building and updating Tableau dashboards with the statistical data collected for its statistical reports. NAIC staff will work on the initial buildout of the dashboards, and the Working Group will meet this summer to review and discuss the dashboards.

The Auto Supplement has been adopted by the Working Group and will be sent to the Task Force for review and adoption. The Working Group plans to meet in May in regulator-to-regulator session to continue reviewing the data for the Homeowners Report. The Working Group will also meet in open session to hear proposed updates to various statistical reports.

Darby made a motion, seconded by Botsko, to adopt the report of the Statistical Data (C) Working Group. The motion passed unanimously.

3. Heard a Presentation on Reserving Analytics for Regulators

Charlie Stone and Cat Drummond (InsurSight) presented on reserving analytics for regulators.

4. Discussed the Academy's Cyber Risk Toolkit

Citarella, Lederer, Darby, Grassel, Schwartz, and Kris DeFrain (NAIC) presented the American Academy of Actuaries' (Academy's) Cyber Risk Toolkit to the Task Force (Attachment __). The discussion of the toolkit was tabled.

Having no further business, the Casualty Actuarial and Statistical (C) Task Force adjourned.

SharePoint/NAIC Support Staff Hub/Member Meetings/C CMTE/2024_Summer/CASTF/050724 min.docx

Draft: 3/22/24

Casualty Actuarial and Statistical (C) Task Force
E-Vote
March 20, 2024

The Casualty Actuarial and Statistical (C) Task Force conducted an e-vote that concluded March 20, 2024. The following Task Force members participated: D.J. Bettencourt, Chair, represented by Christian Citarella (NH); Chlora Lindley-Myers, Vice Chair, represented by Julie Lederer (MO); Barbara D. Richardson represented by Tom Zuppan (AZ); Ricardo Lara represented by Mitra Sanandajifar (CA); Andrew N. Mais represented by Wanchin Chou (CT); Karima M. Woods represented by David Christhilf (DC); Doug Ommen represented by Travis Grassel (IA); Vicki Schmidt represented by Nicole Boyd (KS); Timothy J. Temple represented by Nichole Torblaa (LA); Timothy N. Schott represented by Sandra Darby (ME); Grace Arnold represented by Phil Vigliaturo (MN); Judith L. French represented by Tom Botsko (OH); Glen Mulready represented by Andrew Schallhorn (OK); Andrew R. Stolfi represented by TK Keen (OR); Michael Humphreys represented by Michael McKenney (PA); Michael Wise represented by Will Davis (SC); Cassie Brown represented by J'ne Byckovski (TX); Kevin Gaffney represented by Rosemary Raszka (VT); Mike Kreidler represented by Eric Slavich (WA); and Allan L. McVey and Juanita Wimmer (WV).

1. Adopted the Profitability and Competition Reports

The Task Force conducted an e-vote to consider adoption of the *Report on Profitability by Line by State* (Profitability Report). The motion passed with one abstention.

The Task Force conducted an e-vote to consider adoption of the *Competition Database Report* (Competition Report). The motion passed with two abstentions.

Having no further business, the Casualty Actuarial and Statistical (C) Task Force adjourned.

SharePoint/NAIC Support Staff Hub/Member Meetings/C CMTE/2024_Summer/CASTF/032024 Profit and Comp evote min.docx

Draft Pending Adoption

Draft: 3/20/24

Casualty Actuarial and Statistical (C) Task Force
Phoenix, Arizona
March 16, 2024

The Casualty Actuarial and Statistical (C) Task Force met in Phoenix, AZ, March 16, 2024. The following Task Force members participated: D.J. Bettencourt, Chair, represented by Christian Citarella (NH); Chlora Lindley-Myers, Vice Chair, represented by Cynthia Amann and Julie Lederer (MO); Lori K. Wing-Heier represented by Sian Ng-Ashcraft (AK); Barbara D. Richardson represented by Tom Zuppan (AZ); Ricardo Lara represented by Mitra Sanandajifar (CA); Andrew N. Mais represented by Wanchin Chou and Qing He (CT); Michael Yaworsky represented by Richie Frederick (FL); Doug Ommen represented by Travis Grassel (IA); Amy L. Beard represented by Patrick O'Connor (IN); Vicki Schmidt represented by Nicole Boyd (KS); Timothy J. Temple represented by Nichole Torblaa (LA); Kathleen A. Birrane represented by Bill Fawcett (MD); Robert L. Carey represented by Sandra Darby (ME); Anita G. Fox represented by Kevin Dyke (MI); Grace Arnold represented by Phil Vigliaturo (MN); Eric Dunning represented by Michael Muldoon (NE); Judith L. French represented by Tom Botsko (OH); Glen Mulready represented by Andrew Schallhorn (OK); Michael Humphreys represented by Michael McKenney (PA); Alexander S. Adams Vega represented by Glorimar Santiago (PR); Michael Wise represented by Will Davis (SC); Cassie Brown represented by J'ne Byckovski and Miriam Fisk (TX); Kevin Gaffney represented by Rosemary Raszka (VT); Mike Kreidler represented by Eric Slavich (WA); and Allan L. McVey represented by Ellen Potter (WV). Also participating was: Tomasz Serbinowski (UT).

1. Adopted its Feb. 14, 2024; Feb. 13, 2024; and 2023 Fall National Meeting Minutes

Citarella said the Task Force met Feb. 13, 2024, and conducted an e-vote that ended Feb. 14. During its Feb. 13 meeting, the Task Force took the following action: adopted a motion to allow the NAIC rate model review team to modify its workload process from a strict first-come, first-served process to one that also considers special circumstances (e.g., 30-day deadlines, a state has requested few reviews, a state has an emergency request). The Feb. 14 e-vote included adoption of the *2020/2021 Auto Insurance Database Report* (Auto Report).

The Task Force also met March 12, 2024; Feb. 20, 2024; Jan. 16, 2024; and Dec. 16, 2023, in regulator-to-regulator session, pursuant to paragraph 3 (specific companies, entities, or individuals) of the NAIC Policy Statement on Open Meetings, to discuss rate filing issues. The Task Force held its Predictive Analytics Book Club meetings Feb. 27 and Jan. 30. On Feb. 27, Dorothy Andrews (NAIC) presented "Where Does Bias Hide?" on behalf of the American Academy of Actuaries' (Academy's) Data Science and Analytics Committee. On Jan. 30, Matt Moore (Highway Loss Data Institute—HLDI) presented "ADAS, Marijuana, Teens, and Theft."

Botsko made a motion, seconded by Chou, to adopt the Task Force's Feb. 14, 2024 (Attachment One); Feb. 13, 2024 (Attachment Two); and Dec. 1, 2023, (*see NAIC Proceedings – Fall 2023, Casualty Actuarial and Statistical (C) Task Force*) minutes. The motion passed unanimously.

2. Adopted the Report of the Actuarial Opinion (C) Working Group

Fisk said the Actuarial Opinion (C) Working Group conducted an e-vote that ended Jan. 5 to adopt its proposed changes to the 2024 Statement of Actuarial Opinion (SAO) instructions for property and title. The proposal was exposed by the Blanks (E) Working Group for a public comment period ending April 23. The Actuarial Opinion (C) Working Group plans to meet in late spring or early summer to begin discussing potential changes to the 2025 SAO instructions and the 2024 regulatory guidance.

Fisk made a motion, seconded by Darby, to adopt the report of the Actuarial Opinion (C) Working Group, including its Jan. 5 (Attachment Three) minutes. The motion passed unanimously.

3. Adopted the Report of the Statistical Data (C) Working Group

Darby said the Statistical Data (C) Working Group met Dec. 11, 2023, to discuss changes to the *Report on Profitability by Line by State* (Profitability Report) and the *Competition Database Report* (Competition Report). These discussions focused on formatting already-adopted changes and updating language to match those formatting updates. During the Dec. 11 meeting, the Working Group received updates on the *2021 Homeowners Report* (Home Report) and the 2020/2021 Auto Report. Both reports have since been published.

Due to the changes adopted for the Profitability and Competition Reports, the reports were delayed but are now with the Task Force for consideration of adoption. Voting on the adoption of these reports ends March 20.

Currently, NAIC staff are checking data received for the 2022 Home Report. Data for the *2022 Auto Database Average Premium Supplement* has been checked and will be sent to the Working Group for review this month. The Working Group plans to meet in April to discuss the *2022 Auto Database Average Premium Supplement* and the 2022 homeowners' data and outline a work plan for the remainder of the year.

Darby made a motion, seconded by Grassel, to adopt the report of the Statistical Data (C) Working Group, including its Dec. 11, 2023, minutes (Attachment Four). The motion passed unanimously.

4. Discussed Schedule P

Kris DeFrain (NAIC) said the Task Force's proposal to require all data triangles in Schedule P to include 10 years of data has been adopted by the Blanks (E) Working Group for 2024 implementation. As a result, the financial groups will need to define short-tailed versus long-tailed lines of business rather than rely on Schedule P. Previously, the short-term lines were those in Schedule P with only two years of data, and "long-term" lines were those with 10 years of data.

DeFrain said the Blanks (E) Working Group received a comment letter when the proposal was exposed for comment. The comments were out of the scope of that proposal but contained suggestions to improve other Schedule P instructions. The Task Force agreed to draft a proposal to improve Schedule P instructions.

5. Adopted Comments on the Exposed ASOP No. 12, *Risk Classification*

Lederer said representatives from California, Connecticut, Louisiana, Missouri, New Hampshire, Oregon, and Pennsylvania discussed the exposed Actuarial Standard of Practice (ASOP) No. 12, *Risk Classification*, Feb. 22 and discussed draft comments March 5. She said the comments reflect the consensus reached by the representatives on areas of regulatory importance. Lederer said any individual can submit additional comments to the Actuarial Standards Board (ASB) if desired.

Lederer said the comments fall into two main categories: 1) a request for clarification; and 2) a recommendation to replace material removed from the current version of ASOP No. 12.

Chou made a motion, seconded by Davis, to adopt the comments (Attachment Five) and send them to the ASB. The motion passed unanimously.

6. Discussed the NAIC Rate Model Review Team's GLM Information Needs

Sam Kloese (NAIC) said the list of generalized linear modeling (GLM) information needs was first presented at the 2023 Fall National Meeting and discussed at the Task Force's Feb. 13 meeting. He said the plan is to use this list to help expedite NAIC rate model reviews. Kloese said the idea is that state insurance regulators would make sure companies have provided the standard list of items before sending the model review request to the team. This process could potentially remove one round of objections, which would help finish the model review sooner.

Citarella suggested gathering a small group of Task Force members to discuss the list with NAIC staff. The Task Force discussed the creation of a procedures handbook that could include this list of information items. DeFrain said the team is trying to find ways to improve processes and efficiency to get the queue shorter than it currently is (three or more months). She said at some point, maybe after the third objection, the state insurance regulator could arrange a call with the NAIC and the company.

Citarella added that he has found the Shared Model Database, where all NAIC reports are stored, to be quite valuable. He said he relies on it because it is unusual for New Hampshire to be the first state to review the model. Citarella said he can find the models in the database easily and finds the reviews useful for his work.

Dyke asked if states using the NAIC services would be required to change their checklist. DeFrain said it is not required, but it is encouraged. She said she encourages the list even for states that do not use the NAIC service. Muldoon said Nebraska updated its checklist by adding some items to this list. He said it would allow Nebraska to send a rate model to the team for review and make the NAIC's review similar to its own state review.

Serbinowski asked whether adding a list of modifications state insurance regulators required to the database would be possible. DeFrain said it would require regulatory action because the NAIC does not always know what modifications are required.

7. Heard Activity and Research Reports from Professional Actuarial Associations

The Academy, Actuarial Board for Counseling and Discipline (ABCD), ASB, Casualty Actuarial Society (CAS), and Society of Actuaries (SOA) provided reports on current activities and research.

8. Discussed Other Matters

Botsko said there are multiple ad hoc groups working on capital adequacy issues. He said there are four issues he wants to share: 1) there is the question of how to adjust risk-based capital (RBC) and/or financial examinations for companies that have risky geographic concentration; 2) there is a discussion about how investments should be incorporated or potentially incorporated into RBC, including questions about whether they should be separated and whether there should be an additional or separate charge for them; 3) the preamble to RBC was edited and a new section was added to clarify some things about how RBC is confidential; 4) based on the preamble and confidentiality, there is a question about whether some reporting should be removed from the financial statement. Botsko said there are RBC numbers in the five-year history that some wish to remove. He said some believe there is no problem because the data has been there for 30 years, and others say that information is too often misused.

Amann said the Cybersecurity (H) Working Group is close to finalizing the Cyber Event Reporting Plan (CERP). She said there have been recent cyber issues in the industry, and the Working Group will discuss whether its work plan needs to be adjusted. Chou said the Academy's Cyber Risk Committee is presenting, and he recommends participating. He said the Working Group now has charges on both cybersecurity and cyber coverage.

The Catastrophe Insurance (C) Working Group is moving forward with its charge on the risk mitigation plans. Amann said the Catastrophe Primer (formerly known as the Catastrophe Modeling Handbook) is being updated. Chou said catastrophe risk still includes climate change, and those interested should join the Catastrophe Risk (E) Subgroup to review some of the catastrophe models from vendors.

Having no further business, the Casualty Actuarial and Statistical (C) Task Force adjourned.

SharePoint/NAIC Support Staff Hub/C Cte/2024 Spring/Minutes CASTF 03162024

Draft: 8/8/24

Actuarial Opinion (C) Working Group
Virtual Meeting
July 23, 2024 / August 6, 2024

The Actuarial Opinion (C) Working Group of the Casualty Actuarial and Statistical (C) Task Force met July 23, 2024, and Aug. 6, 2024. The following Working Group members participated: Miriam Fisk, Chair (TX); Julie Lederer, Vice Chair (MO); David Christhilf (DC); Chantel Long (IL); Sandra Darby (ME); Tom Botsko (OH); Andy Schallhorn (OK); and Kevin Clark and James DiSanto (PA).

1. Discussed the Regulatory Guidance

During its July 23 meeting, the Working Group discussed potential changes to regulatory guidance and continued the discussion Aug. 6. The Working Group discussed the following changes:

- Because of a change in 2024, qualification documentation is only required at the initial appointment of an appointed actuary and no longer required annually thereafter.
- A new Actuarial Standard of Practice (ASOP) No. 36 version was modified, re-titled, and adopted by the Actuarial Standards Board (ASB) in 2024. "Other" reserves are added to the title. The new title is ASOP No. 36, *Statements of Actuarial Opinion Regarding Property/Casualty Loss, Loss Adjustment Expense, or Other Reserves.* Changes to Regulatory Guidance include: 1) removing some disclosures that are no longer required by ASOP No. 36 to be included in a public actuarial opinion; and/or 2) requesting specific disclosures regardless of whether required by an ASOP(s) in the public actuarial opinion or confidential actuarial report.

2. Discussed Instructions for the Actuarial Opinion

Fisk suggested some changes to 2025 instructions for the Property/Casualty Statement of Actuarial Opinion related to the Society of Actuaries modifying its educational program effective in 2025. Until state insurance regulators can review the educational program to determine whether it will meet regulatory requirements for Appointed Actuaries, the instructions for 2025 will need to mention a review will take place. The Working Group will also consider mentioning this forthcoming change in the 2024 regulatory guidance.

Discussion on the regulatory guidance and opinion instructions will continue in late August.

Having no further business, the Actuarial Opinion (C) Working Group adjourned.

SharePoint/NAIC Support Staff Hub/Member Meetings/C CMTE/2024_Summer/CASTF/AOWG/AOWG 072324 080624 min.docx

Draft: 8/7/24

Statistical Data (C) Working Group
Virtual Meeting
July 29, 2024

The Statistical Data (C) Working Group of the Casualty Actuarial and Statistical (C) Task Force met July 29, 2024. The following Working Group members participated: Sandra Darby, Chair (ME); Qing He, Vice Chair, and George Bradner (CT); David Christhilf (DC); Arthur Schwartz and Tom Travis (LA); Cynthia Amann (MO); Christian Citarella (NH); Alexander Vajda (NY); Tom Botsko (OH); David Dahl and Ying Liu (OR); and Nicole Elliot (TX).

1. Adopted its May 30 Minutes

The Working Group met May 30 to discuss proposed changes to the *Dwelling, Fire, Homeowners Owner-Occupied, and Homeowners Tenant and Condominium/Cooperative Unit Owner's Insurance Report* (Homeowners Report) and the *Report on Profitability by Line by State* (Profitability Report).

Qing He made a motion, seconded by Botsko, to adopt the Working Group's May 30 minutes (Attachment). The motion passed unanimously.

2. Discussed Proposed Changes to the Homeowners Report

Darby said the Working Group will continue the discussion from the last open meeting on Schwartz's proposed changes to the Homeowners Report.

Schwartz said the format of the Homeowners Report has not changed since 2003. He said his proposals would be a change to both the format and the data elements collected. He would like to see a table that shows the average premium for each state, along with the median insured value to account for the difference in home values in each state. This table would also include a column showing the premium per thousand dollars of median insured value. He said the report should also include a table that shows the ranking of each state for average premium and the ratio of premium to median insured value.

Birny Birnbaum (Center for Economic Justice—CEJ) said median insured value is a good metric to include, and this proposed table should also include median premium. He said using a ratio of premium to median insured value is not an accurate measure of cost of insurance, and he recommends against using this measure or using it to rank states.

Bradner said state insurance regulators want to see data on premium and insured value in different ranges within their state. He said regulators take exception to media reporting average premium values and comparing states without including caveats on the many factors that can impact pricing. Birnbaum said the report offers extensive explanations of these caveats. He said the best way to get the granular data that regulators are looking for is to update the NAIC Statistical Plan and mandate transaction-level data reporting. Bradner said if this report did include a ranking, the report should also describe metrics and methodology behind the ranking. Birnbaum said the NAIC has never endorsed ranking states. Brian Sullivan (Risk Information Inc.) said he publishes a table using HO-3 data from the Homeowners Report that includes home value distribution by state. Birnbaum said only using HO-3 data can be misleading because, for example, it would not include wind coverage in certain states.

Schwartz said the report currently does not capture data on manufactured and modular homes. He said these types of homes comprise a large percentage of homes in many states. He said the report should also look into

capturing data on additional dwelling units (ADUs). Birnbaum said he agrees the report should collect data on manufactured and modular homes. He said the report should also collect data on commercial coverages for homeowners associations and condominium associations. Libby Crews (NAIC) said data for manufactured homes written on HO-7 policies are not collected for this report. Bradner said not every state has HO-7 forms so the Working Group needs to investigate how data for manufactured homes is collected by the statistical agents and how that data can be reported for the Homeowners Report.

Schwartz said the report does not capture data on catastrophe losses, but that information is available in fast track reporting from statistical agents so it could be available for the Homeowners Report. He said capturing this data by state for the last 10 years would be important for all users of the report. He said he would also be interested in collecting data on losses by peril, separated by coverage type. Birnbaum said all of the proposals mentioned today do not distinguish between what a company is required to report versus what would be included in an NAIC statistical report. He said some data elements found in the proposals could be collected from the recent NAIC Property and Casualty Market Intelligence data call. He said a transaction-level data collection mechanism would allow regulators to answer questions about the market in real time.

Darby said the Working Group would investigate which data elements from the proposals for the Homeowners Report can be reported by statistical agents, and the Working Group would go over the findings during its next meeting.

3. Discussed Additional Proposed Changes to NAIC Statistical Reports

Sullivan said he would like to have further discussion on creating a profit margin metric in the Profitability Report, using return on net worth as a percentage of earned premium to net worth. Birnbaum asked how the profit measure Sullivan uses is different than the underwriting profit in the Profitability Report. Sullivan said the underwriting profit does not include investment gains on insurance transactions or tax on insurance transactions. Birnbaum said the profit on insurance transactions number published in the Profitability Report would be similar to the profit measure Sullivan is calculating.

Sullivan would like to see the Auto Insurance Database Report (Auto Report) include coverage limits. He said coverage limits are a useful proxy for understanding the relative cost of insurance.

Having no further business, the Statistical Data (C) Working Group adjourned.

SharePoint/NAIC Support Staff Hub/Member Meetings/C CMTE/2024_Summer/CASTF/SDWG/StatDataWGmin_0729

Draft: 7/24/24

Statistical Data (C) Working Group
Virtual Meeting
May 30, 2024

The Statistical Data (C) Working Group of the Casualty Actuarial and Statistical (C) Task Force met May 30, 2024. The following Working Group members participated: Sandra Darby, Chair (ME); Qing He, Vice Chair (CT); Charles Hale (AL); Arthur Schwartz and John Sobhanian (LA); Christian Citarella (NH); Tom Botsko (OH); David Dahl and Ying Liu (OR); and Nicole Elliot (TX).

1. Discussed Changes to the Profitability Report

Darby said the Working Group has been meeting over the last year to review proposals to update various statistical reports. She said it is important to continually review the usefulness of the reports for state insurance regulators and others using them. She said since the last meeting, Brian Sullivan (Risk Information Inc.) has reviewed the changes implemented in the *Report on Profitability by Line by State* (Profitability Report) and asked to share his thoughts.

Sullivan said this report is used to calculate and include an after-tax profit number, but that data element was removed from the report about 25 years ago. Sullivan said he has recreated this calculation on his own and uses this number to more accurately capture the high levels of reserves seen in the insurance industry. He said he would like to discuss whether the after-tax profit number or the profit on insurance transactions number in the Profitability Report would be more accurate. He said the return on net worth number can be misleading because some large companies carry a large amount of net worth, which can distort the comparison to other industries.

Sullivan said he would like to better understand the calculation and wide disparity across lines and states in earned premium to net worth.

Sullivan said he would like to see a 10-year weighted average of profit numbers included in the report. Aaron Brandenburg (NAIC) said the Profitability Report currently has a 10-year average for multiple data elements. Sullivan said he would like to see a weighted average because the premium can increase drastically in certain lines in a 10-year period. Darby asked NAIC staff to put together an example of what adding a weighted average column would look like in the report.

Sullivan said that many years ago, the Profitability Report had a section that explained and normalized Michigan personal auto insurance data. He said he would like to discuss including that in the report again. Brandenburg asked if this is still an issue with the reforms that happened a few years ago in Michigan. Sullivan said the Michigan Catastrophic Claims Association (MCCA) reports incurred losses but not premiums, affecting the incurred loss ratio.

Darby asked Sullivan to send more information regarding his calculations and suggestions to the Working Group so it can discuss them further on a future call.

2. Discussed Proposed Changes to the Homeowners Report

Schwartz said the *Dwelling, Fire, Homeowners Owner-Occupied, and Homeowners Tenant and Condominium/Cooperative Unit Owner's Insurance Report* (Homeowners Report) gets a tremendous amount of attention and media coverage. He said the report would benefit from having a table that shows the average

premium by state. He said this table should also show median home value because home values differ greatly between states. Darby asked where the median home value data would come from. Schwartz said he would have to check where he got the data but that each state had a range of home values, and he took the middle value in those ranges for each state. Darby said this report has not historically included a table with average premiums because it is hard to compare when the home values differ across states.

Sullivan said Risk Information Inc. publishes a similar table in its Property Insurance Report, using premiums and exposures from the $175,000 to $199,999 insured value range. He said the report clarifies that looking at the insured value changes the importance of the average. He said the report also includes a table with the distribution of home values by state. Darby said this kind of apples-to-apples comparison makes sense, but you cannot make the same comparison when the distribution is different by state. Sullivan said one solution would be to calculate a premium to medium household income ratio. Schwartz said he had considered including a column with a calculation of average premium divided by median home value in thousands, or cost per thousand dollars of insurance. Dahl said some states have standard fire policies, and others do not, which would increase the cost and add another comparison issue. Sullivan said the same problem exists in personal auto insurance. Dahl said the required limits in auto do not necessarily reflect what people are actually purchasing. Sullivan said he uses income data as a proxy because those making more money will likely buy nicer cars and carry higher liability insurance. Darby said the issue with using income is that you have to find a source for the data and then take an average for the state.

Schwartz said the Homeowners Report does not currently include data for manufactured or mobile homes, but it should be included due to the large number of these home types in the U.S.

Schwartz said he would like to see data on the effects of catastrophes. He said Fast Track data breaks down data by catastrophe losses, so the data is available from statistical agents. He said he would like the report to show the data, including and excluding catastrophes.

Schwartz said he would also like to discuss the idea of including the data by peril in the Homeowners Report.

Darby said the Working Group will continue this discussion during its next open meeting.

Having no further business, the Statistical Data (C) Working Group adjourned.

SharePoint/NAIC Support Staff Hub/Member Meetings/C CMTE/2024_Summer/CASTF/SDWG/StatDataWGmin_0530

Casualty Actuarial Society Race and Insurance Pricing 2024 Projects

Ken Williams
Staff Actuary – Chief of Advocacy
Casualty Actuarial Society

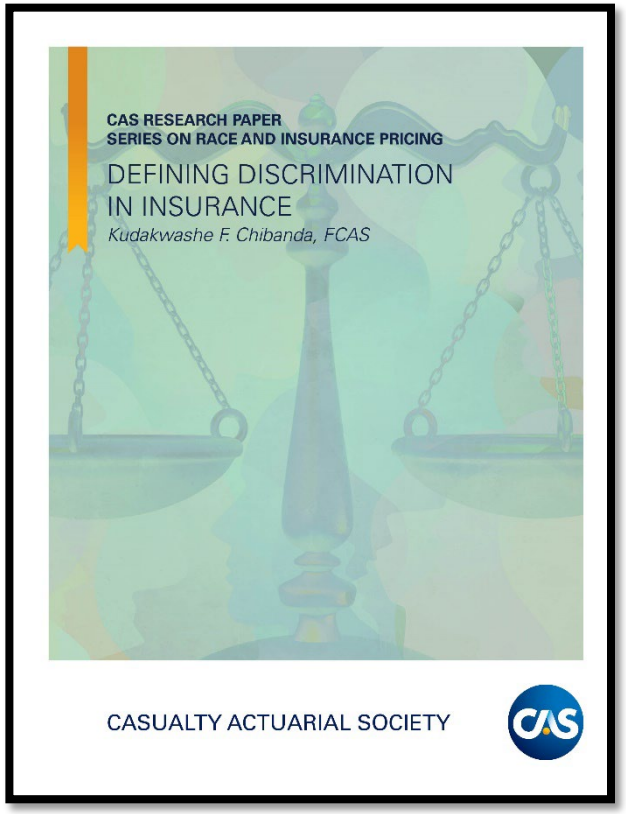
Introduction

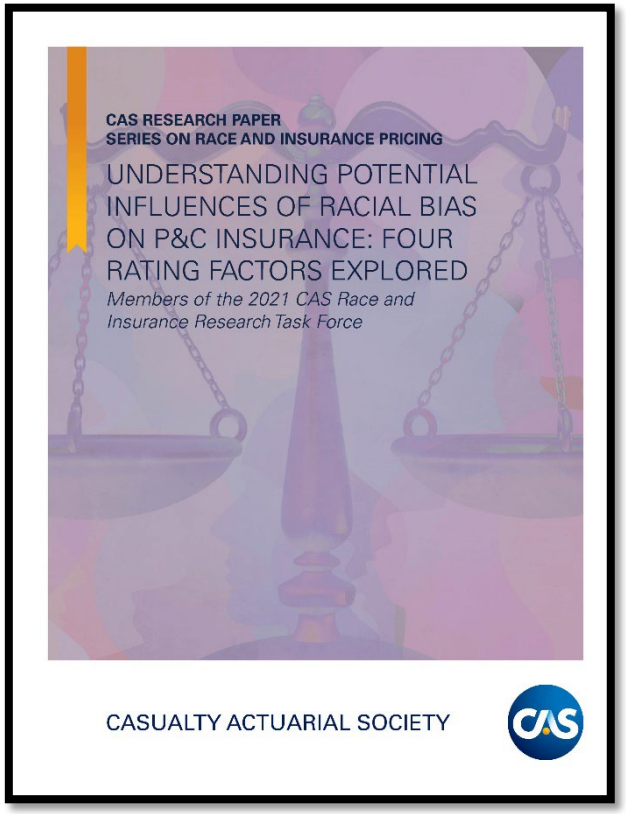
- Actuaries have a responsibility to examine the processes, systems and models they build to understand if the inputs and outcomes reflect fair and equitable practices.
- Casualty Actuarial Society (CAS) launched its Approach to Race and Insurance Pricing in February 2021, with activities in four key areas of focus:
  - **Collaboration** – to proactively engage and partner with regulators, insurers, actuarial organizations, consumer groups and other organizations addressing issues related to race and insurance and to ensure that diverse perspectives contribute to CAS-commissioned efforts.
  - **Research** – to develop methodologies that identify, measure, and address potential bias, to evaluate emerging technologies and prepare actuaries and insurers for potential regulatory actions, in alignment with the CAS Core Values of continual improvement and innovation.
  - **Basic and Continuing Education** – to provide members and candidates with a strong foundation in the historical issues of systemic bias and their potential impacts on insurance, covering concepts of disparate impact and discrimination, past and current research, and professionalism implications.
  - **Leadership and Influence** – to play a leading role in the discourse on potential racial bias in insurance pricing among our membership, the insurance industry and the public.

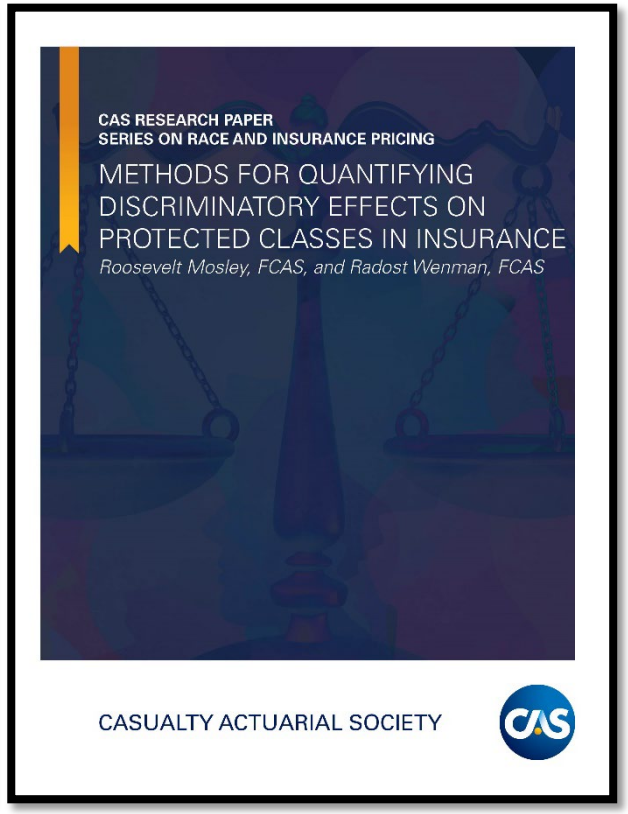
CAS Approach to Race and Insurance Pricing

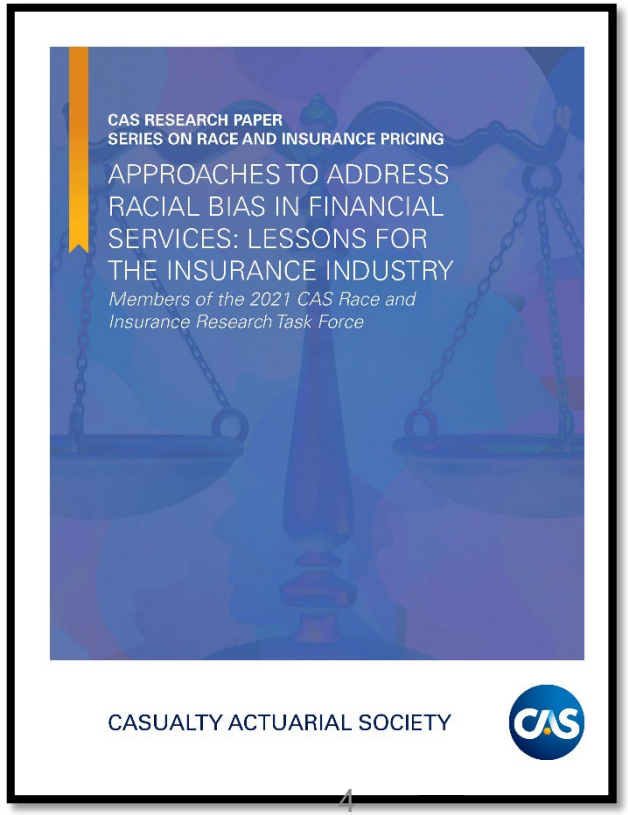
CAS Research Series on Race and Insurance Pricing – Phase I

https://www.casact.org/publications-research/research/research-paper-series-race-and-insurance-pricing









Why Actuaries Should Care

Regulators determined what is acceptable for rating

Actuaries determine rates based on regulatory guidance

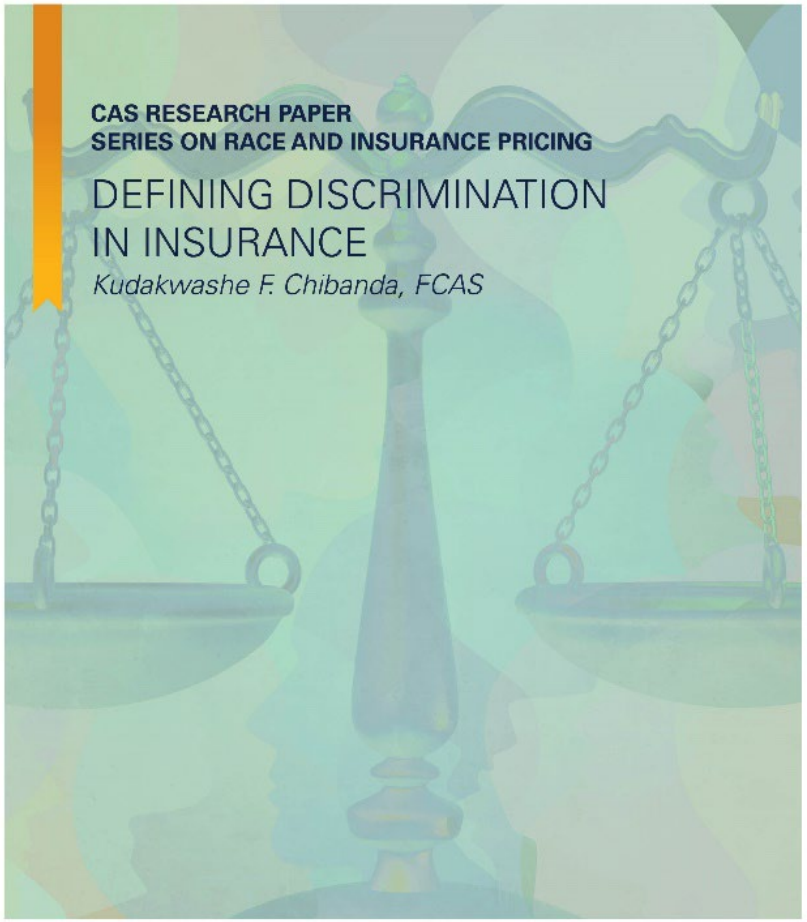
Confused consumers

Regulators contemplating new regs

Actuaries informing regulators on implications of insurance regulation

Better informed stakeholders (regs and consumers)

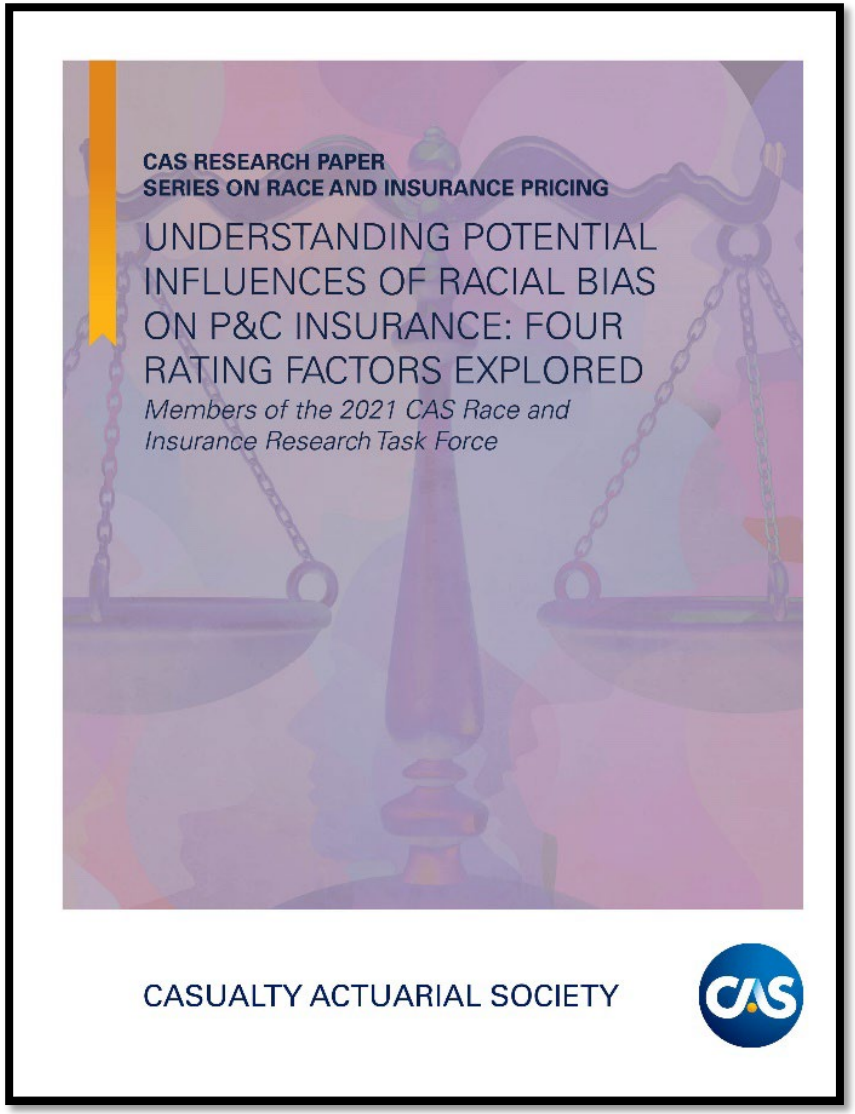
Defining Discrimination In Insurance



Setting The Stage

01	Are You Sure You Know What **Protected Class** Is?
02	Revisiting **Unfair Discrimination**
03	The **Proxy Discrimination** Debate
04	What Is **Disparate Impact** Anyway?

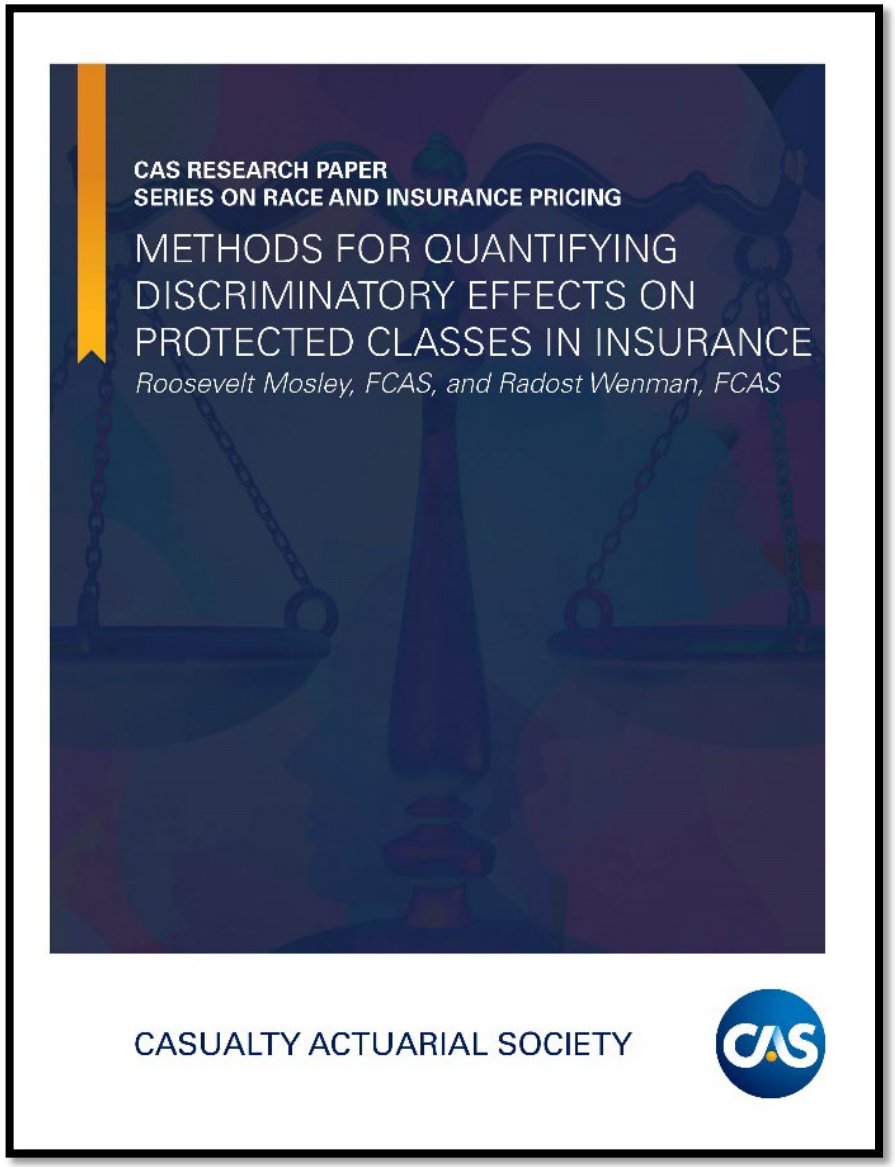
Understanding Potential Influences of Racial Bias on P&C Insurance: Four Rating Factors Explored

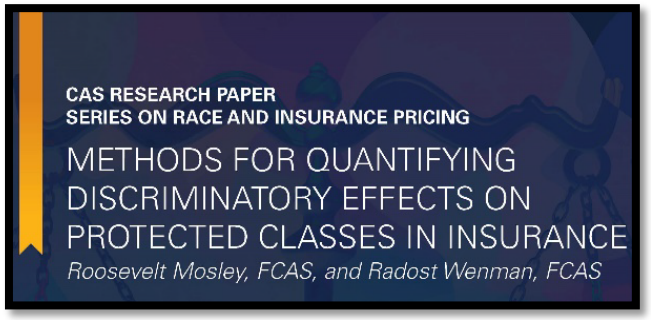


4 commonly used factors in personal lines

- Motor Vehicle Record (MVR)
- Credit Based Insurance Score (CBIS)
- Geographic location
- Homeownership

Methods for Quantifying Discriminatory Effects on Protected Classes in Insurance





Classification of Fairness Metrics

Equalized Odds

Calibration

Demographic Parity

Independence	Separation	Sufficiency
$\hat{Y} \perp A$	$\hat{Y} \perp A \mid Y$	$Y \perp A \mid \hat{Y}$

Predictive Equality

Statistical Parity

Group Fairness

A - protected attribute

Y - observed value of target variable

$\hat{Y}$ - predicted value of target variable

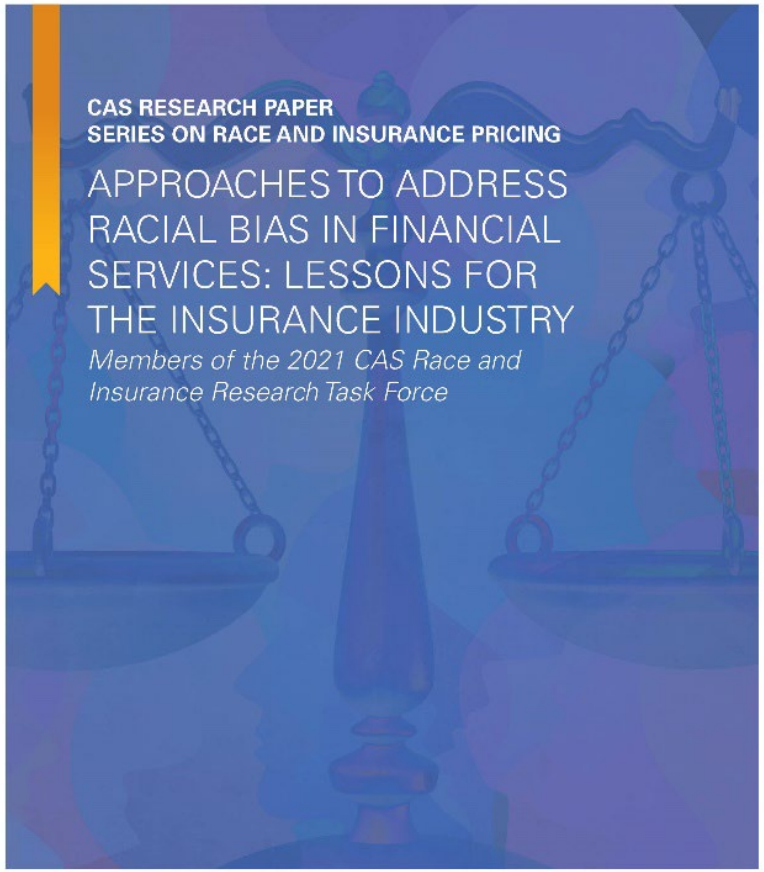
Predictive Parity

Balance Positive Class

Conditional Statistical Parity

Equal Opportunity

Approaches to Address Racial Bias in Financial Services: Lessons for the Insurance Industry

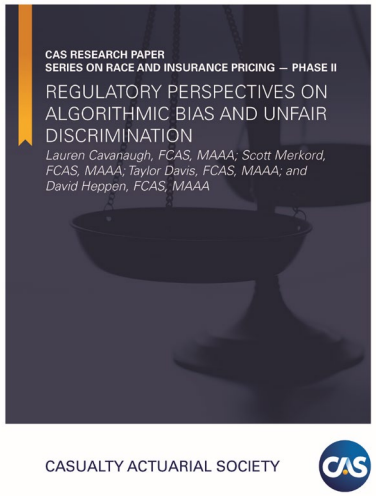


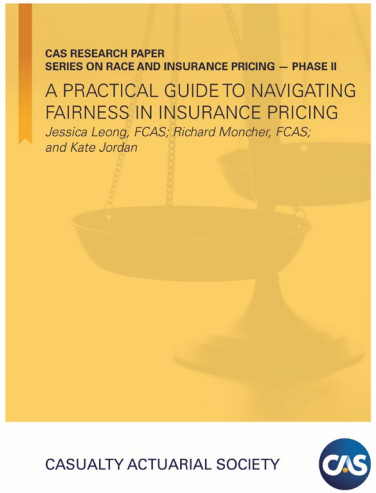
Financial Services Reviewed

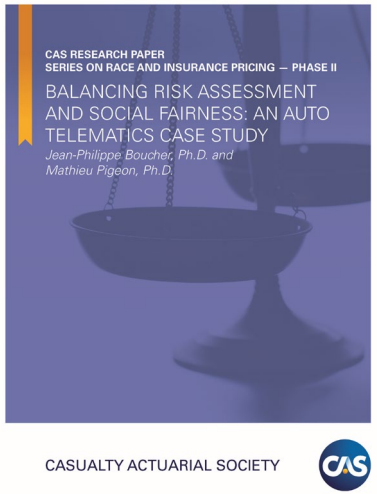
1. Mortgage Lending
2. Personal Lending
3. Commercial Lending
4. Credit Scoring

Phase 2 - Preparing for Tomorrow:

Regulatory Insights and Strategies for Mitigating Potential Bias in Insurance Pricing







Practical Applications of Bias Measurement and Mitigation Techniques

Potential Unintended Impacts of Bias Mitigation on Other Protected Classes

Comparison of Regulatory Frameworks for Non-Discriminatory AI Usage in Insurance

Comparison of Regulatory Frameworks for Non-Discriminatory AI Usage in Insurance

United States

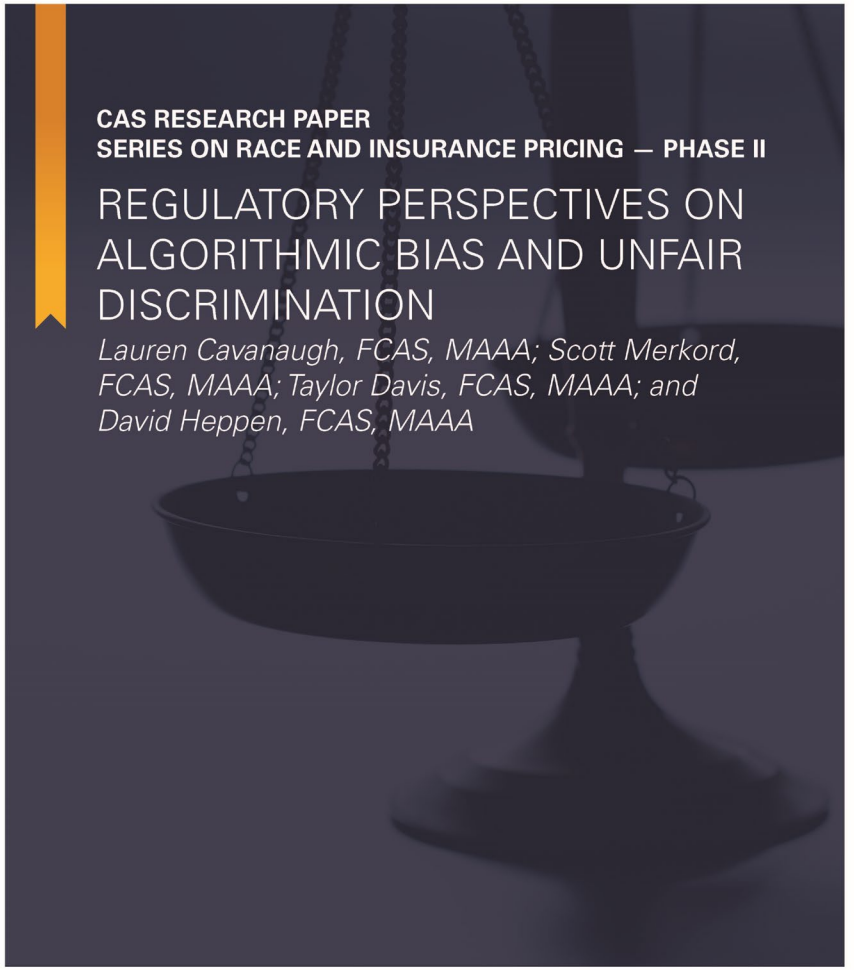
Canada

Europe

China

- Context: Driving Philosophy behind insurance and regulation
- Responsibility: Who regulates Insurance and AI?
- Action: Current Developments in AI and Bias in Insurance regulation

Regulatory Perspectives on Algorithmic Bias and Unfair Discrimination



CASUALTY ACTUARIAL SOCIETY

Regulatory Perspectives: Approach

Summary of Recent Regulatory Activity (U.S.)

Emphasis on collaborative efforts (e.g. NAIC)

Survey of U.S. State Insurance Departments

Considerations for Actuaries

Regulatory Perspectives: Plans & Perceptions

- Regulators were concerned about algorithmic bias but, few are engaged in activities to address algorithmic bias
- Most agree multiple bias testing methodologies should be used
- Mixed views on use of race/ethnicity for bias testing
  - AND most disagree with using race/ethnicity inference approaches

<u>Actuarial soundness does not satisfy discrimination concerns</u>

Regulatory Perspectives: Auto Rating Factors

More Concern

- Home ownership
- Occupation

Read the paper to see more!

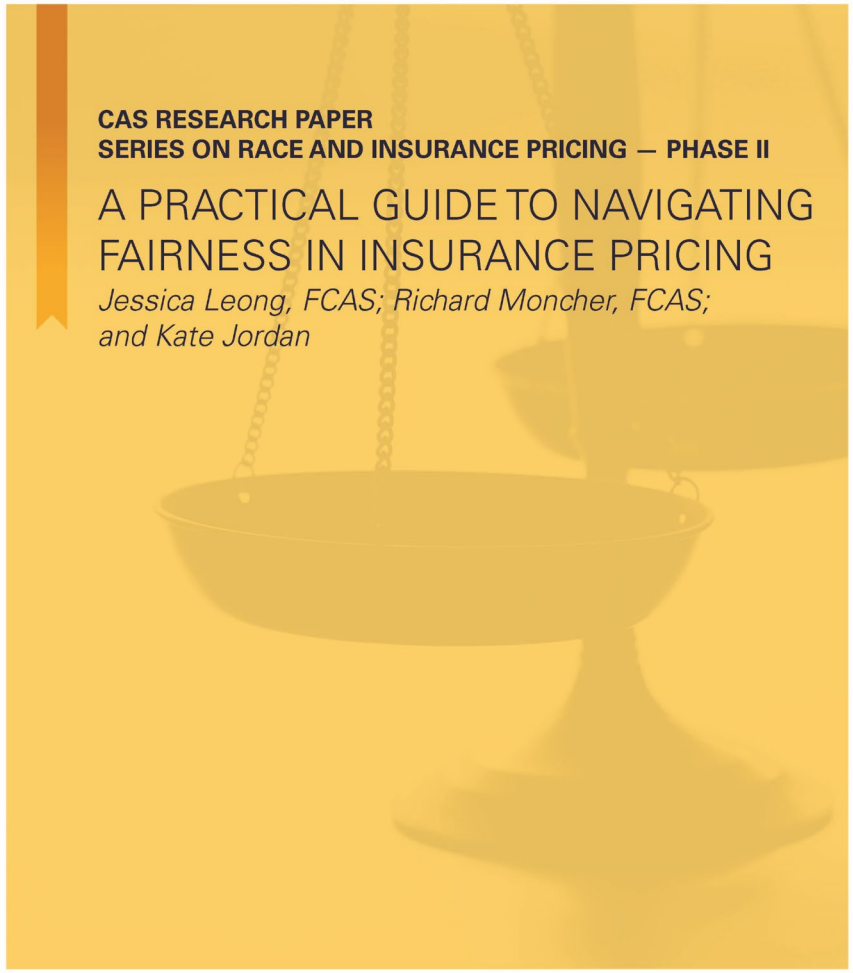
Less Concern

- Age
- Motor Vehicle Records

Read the paper to see more!

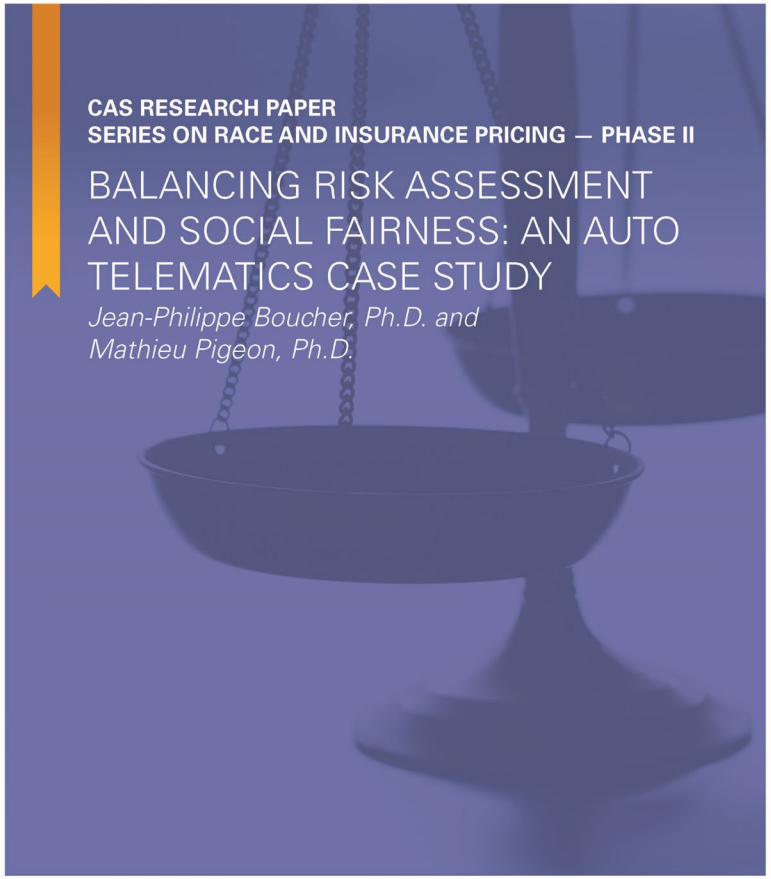
Mixed Perceptions: Geography, Read the paper to see more!

A Practical Guide to Navigating Fairness in Insurance Pricing



CASUALTY ACTUARIAL SOCIETY

Balancing Risk Assessment and Social Fairness: An Auto Telematics Case Study



Potential Unintended Impacts of Bias Mitigation on Other Protected Classes

Practical Applications of Bias Measurement and Mitigation Techniques

Q&A

Contact us: diversity@casact.org

www.casact.org/publications-research/research/research-paper-series-race-and-insurance-pricing

Casualty Actuarial Society
Race and Insurance Pricing 2024 Projects

8/13/2024

Introduction

Actuaries have a responsibility to examine the processes, systems and models they build to understand if the inputs and outcomes reflect fair and equitable practices. In February 2021, the Casualty Actuarial Society (CAS) launched its Approach to Race and Insurance Pricing, with activities in four key areas of focus:

- Collaboration — to proactively engage and partner with regulators, insurers, actuarial organizations, consumer groups and other organizations addressing issues related to race and insurance and to ensure that diverse perspectives contribute to CAS-commissioned efforts.
- Research — to develop methodologies that identify, measure, and address potential bias, to evaluate emerging technologies and prepare actuaries and insurers for potential regulatory actions, in alignment with the CAS Core Values of continual improvement and innovation.
- Basic and Continuing Education — to provide members and candidates with a strong foundation in the historical issues of systemic bias and their potential impacts on insurance, covering concepts of disparate impact and discrimination, past and current research, and professionalism implications.
- Leadership and Influence — to play a leading role in the discourse on potential racial bias in insurance pricing among our membership, the insurance industry and the public.

Phase 1: Four Introductory Papers

In 2022, during the first phase of this effort, the CAS published the first four papers in the CAS Research Paper Series on Race and Insurance Pricing. These reports were designed to guide the insurance industry toward proactive, quantitative solutions to identify, measure and address potential racial bias in insurance pricing. They are:

- Defining Discrimination in Insurance (and errata)
- Understanding Potential Influences of Racial Bias on P&C Insurance: Four Rating Factors Explored
- Methods for Quantifying Discriminatory Effects on Protected Classes in Insurance
- Approaches to Address Racial Bias in Financial Services: Lessons for the Insurance Industry

Phase 2: Six New Papers

As part of Phase 2 of the CAS Research Paper Series on Race and Insurance Pricing, the CAS is working on six new research projects. Papers are expected to be published in Summer / Fall 2024. An overview of each project with each title subject to change is below.

Comparison of Regulatory Frameworks for Non-Discriminatory AI Usage in Insurance
Produced in partnership with the Society of Actuaries

This brief will describe the legal and regulatory frameworks related to bias in predictive modeling and artificial intelligence across four jurisdictions — the United States, Canada, the European Union, and China. The report will illustrate the direction of laws and regulations internationally through description of existing regulations and will compare and contrast how each region has approached these issues.

Please reach out to Mallika Bender, mbender@casact.org, for more information on the Casualty Actuarial Society Research Paper Series on Race and Insurance Pricing.

Regulatory Perspectives on Algorithmic Bias and Unfair Discrimination
By Risk & Regulatory Consulting

Impact of recently proposed or enacted legislation/regulation related to algorithmic bias (including CO, DC, NAIC Model Bulletins, other proposed).

Survey of insurance departments to determine what will likely be necessary for companies/actuaries in the future to demonstrate that insurers' use of models does not result in unfair discrimination. Issues surveyed include:

1. Enforcement of regulatory and legal requirements.
2. Use of protected class information and inference/imputation techniques.
3. Acceptable methodologies to identify and address bias.
4. Level of concern with specific private passenger automobile insurance rating variables.

The paper will also cover practical considerations and approaches for actuaries who must comply with regulatory and legislative requirements related to algorithmic bias, emphasizing applicable actuarial standards of practice when testing for algorithmic bias.

A Practical Guide to Navigating Fairness in Insurance Pricing
By Octagram Analytics

In the current insurance regulatory environment, a new practice is emerging around testing for bias and fairness. This paper first provides an overview of emerging insurance industry regulation at the U.S. State level, as well as Canada and the EU as well as relevant regulations in other industries such as housing, lending and hiring. It then provides a framework that insurers can use to minimize the chance of failing a test or address models that fail a test on unfair discrimination. This framework includes model governance, project planning, data preparation, model training, model evaluation and selection, model implementation, and monitoring.

Balancing Risk Assessment and Social Fairness: An Auto Telematics Case Study
By Prof. Jean-Philippe Boucher & Prof Mathieu Pigeon

After a brief description of telematics or usage-based insurance data and its adoption in the insurance market, this paper provides a case study illustrating the potential for telematics data to reduce insurer's reliance on protected information (such as gender and age) and sensitive information (such as marital status, geography, credit, which can be correlated with protected class status).

Utilizing a combination of real insurance data (which will not be shared) from Canadian insurer in Ontario and a synthetic dataset, this paper demonstrates the impact of utilizing telematics data in auto insurance pricing algorithms on model fairness with respect to protected classes. It also tests the impact of using traditional GLM modeling approaches versus more complex GBM modeling approaches in capturing the predictive power of the telematics information.

Potential Unintended Impacts of Bias Mitigation on Other Protected Class Dimensions
By Pinnacle Actuarial Resources

This report examines three types of regulatory actions aimed at improving fairness for one category of protected class and their respective potential impacts on other protected class categories.

1. Limiting the amount of differential for a rating variable
2. Prohibiting the use of a rating variable
3. Investigating model adjustment/reweighting/decorrelation approaches

These three regulatory scenarios will be applied to a simulated dataset with built-in bias scenarios and correlation between two different protected class dimensions, to illustrate the impacts of each regulatory approach.

Practical Application of Bias Measurement and Mitigation Techniques in Insurance Pricing
By the CAS Race and Insurance Task Force Subgroup

This paper builds on the Phase 1 paper titled Methods for Quantifying Discriminatory Effects on Protected Classes in Insurance. It begins with a brief discussion of various types of bias that may impact insurance data and processes, potentially contributing to unfair discrimination or disparate impact concerns. It then illustrates specific techniques that actuaries can apply to perform bias analyses on an insurance pricing model, including:

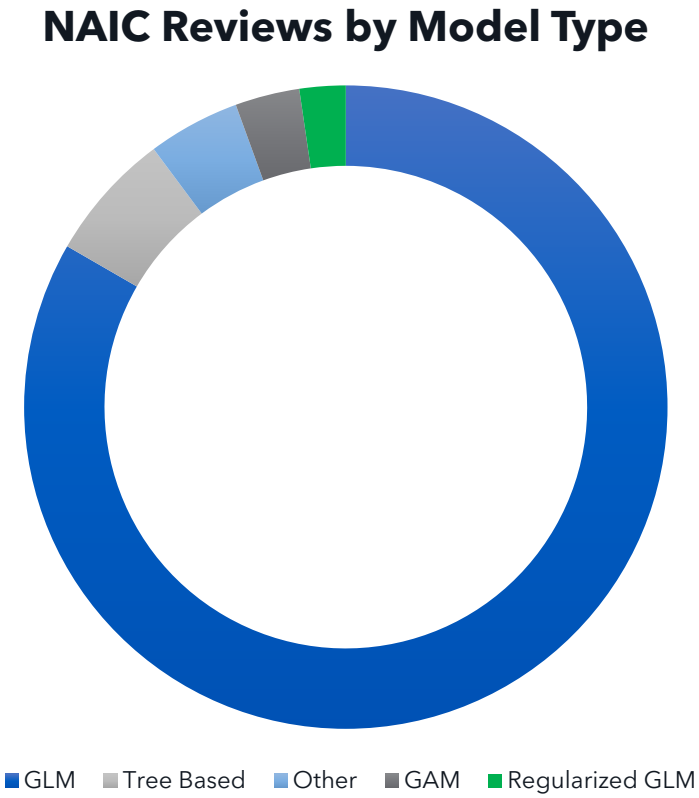
1. protected class imputation methods and alternative interpretations of method outcomes.
2. statistical model fairness tests with a comparison of their strengths and limitations
3. statistical mitigation methods and comparison of their strengths and limitations.
4. other considerations such as model purpose/use, underwriting/expense decisions and credibility.

Appendix B-RGLM Background Information

Sam Kloese, ACAS, MAAA, CSPA, CPCU, AIDA
Property/Casualty Rate Modeling Actuary
NAIC Research and Actuarial Services

CAS White Paper Appendices

- Regulatory Review of Predictive Models
  - (Adopted 9/15/2020)
  - Generalized Linear Models
- Tree-based Models Appendix
  - (Adopted Summer 2022)
  - Gradient Boosting Machines
  - Random Forest
- GAM White Paper Appendix
  - (Adopted Spring 2023)
  - Generalized Additive Models



Other includes neural net models on image data and minimum bias methods

Overview of Regularized GLM

- Definition
  - Extension of the traditional GLM that includes a penalty term
  - Adds a constraint to the model parameters, which help in controlling the model complexity
  - When the penalty is zero, the model returns similar indications to a GLM
- Benefits
  - Reduces risk of overfitting
  - Favors less complex model
  - Mitigates statistical bias in the model

Mastering regularization in machine learning - A 2023 guide (datasciencedojo.com)

Types of Regularized GLMs

- Lasso
  - Can penalize coefficients to zero
- Ridge
  - Can penalize coefficients to near zero
- Elastic Net
  - Combination of Lasso and Ridge
  - Can penalize coefficients to zero
- Derivative Lasso and Accurate GLM (AGLM)
  - Can penalize categorical variable coefficients to zero
  - Can group ordinal variable levels with adjacent levels
- Lasso Credibility
  - Can penalize toward a selected complement of credibility

Comparison to Other Types

	GLM	GAM	RGLM
P-values	✓	✓	✗
Smoothed Terms	✗	✓	✗
Complexity Parameters	✗	✓	✓

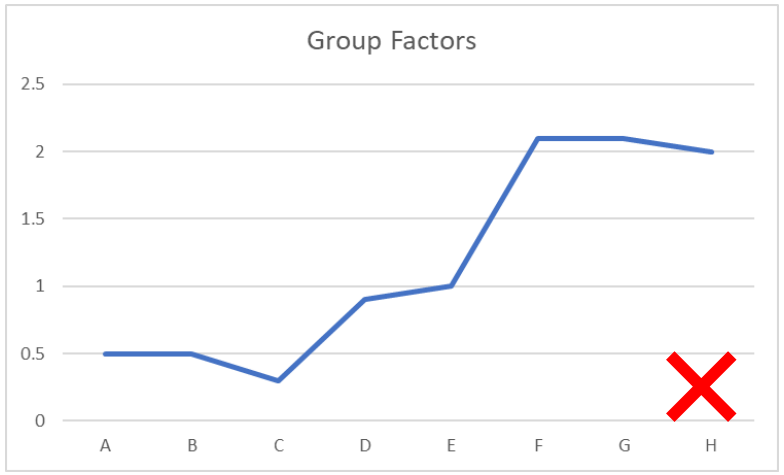
The p-values in a GAM are approximate for the smoothed terms

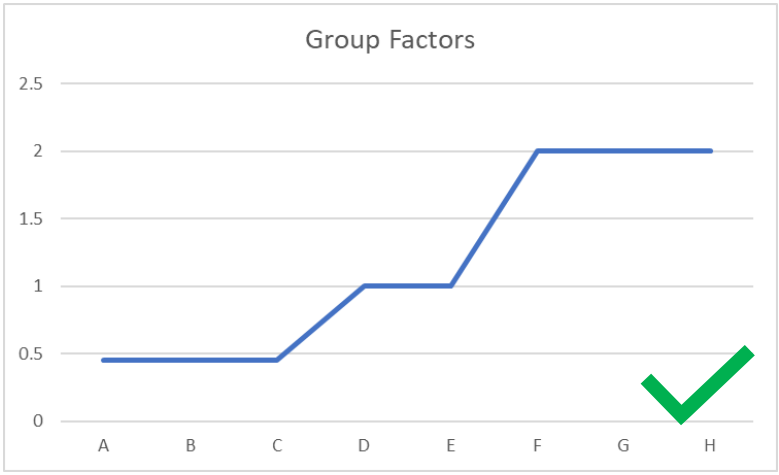
Notable Changes

- New Section B.1.b
  - Identify the credibility complement used
  - Discuss why it is reasonable
  - Applicable to Lasso Credibility Method
  - Ex: currently approved factors

Notable Changes

- B.2.g
  - Obtain value of model complexity hyperparameter
  - Obtain explanation on how it was chosen
    - Grid search to maximize Gini
    - Grid search to minimize Deviance
    - Assess variable plots to see if reversals are eliminated (applicable to derivative lasso)

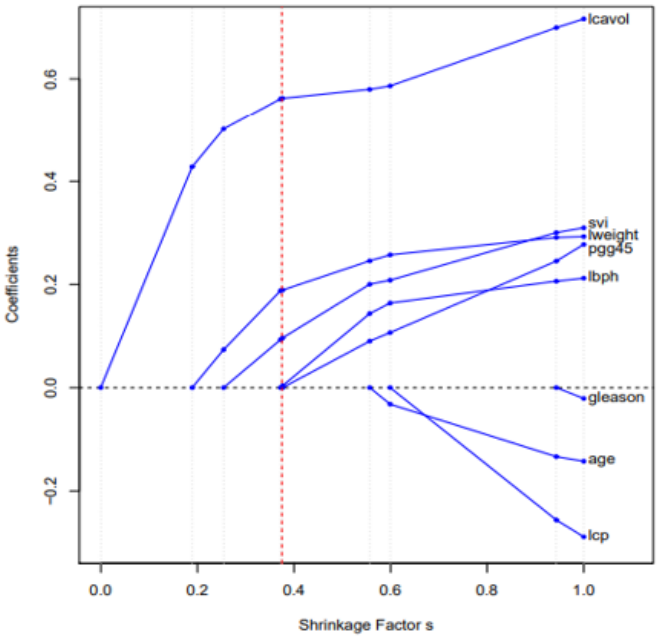




Notable Changes

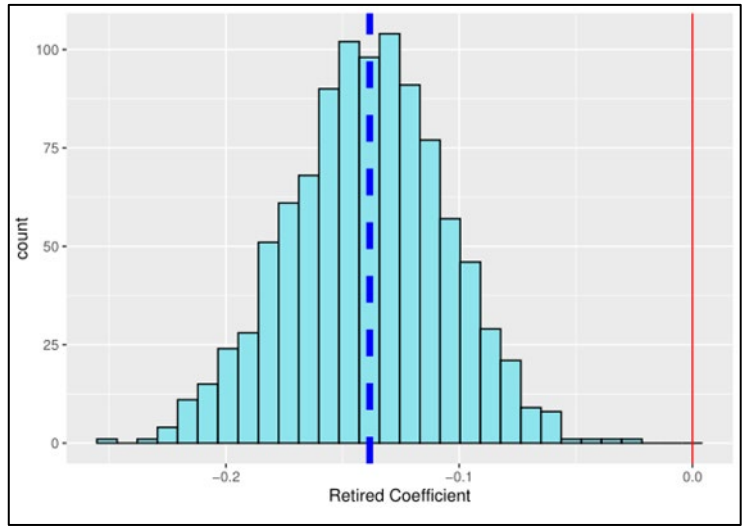
- B.2.h
  - *Level 4 Item*
  - Understand how output model would differ if other penalty parameter values are used.
  - Side-by-side sensitivity analysis
  - Plot showing coefficient value each variable given the complexity hyperparameter (aka shrinkage factor)

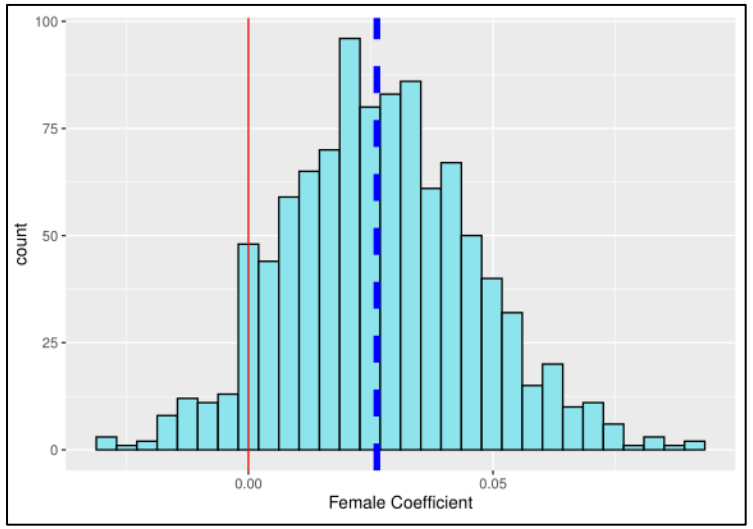
	Penalty Parameter		
	0.01	*0.02*	*0.03*
Group			
A	0.500	0.450	0.600
B	0.500	0.450	0.600
C	0.300	0.450	0.600
D	0.900	1.000	0.600
E	1.000	1.000	1.800
F	2.100	2.000	1.800
G	2.100	2.000	1.800
H	2.000	2.000	1.800



Notable Changes – P-Value Alternatives

- B.4.b
  - Bootstrapping: Do variations to the data result in radically different coefficients?





Notable Changes – P-Value Alternatives

- B.4.b
  - Bootstrapping: Do variations to the data result in radically different coefficients?
  - Cross Fold Validation: Are the coefficients consistent across folds?

		Elastic Net	Reference GLM	GLM p-value
Usage	All Trips	0.618	0.622	< 0.001
	Professional	0.239	0.239	0.002
	Retired	(0.138)	(0.142)	< 0.001
	Work Private	*Base*	*Base*	
Gender	Female	0.026	0.027	0.157
	Male	*Base*	*Base*	
Driver Age	16 - 20	0.161	0.170	0.398
	21 - 30	0.011	0.014	0.769
	31 - 40	*Base*	*Base*	
	41 - 50	(0.032)	(0.031)	0.320
	51 - 60	0.027	0.029	0.329
	61+	0.082	0.087	0.016

Notable Changes – P-Value Alternatives

- B.4.b
  - Bootstrapping: Do variations to the data result in radically different coefficients?
  - Cross Fold Validation: Are the coefficients consistent across folds?
  - GLM Reference Model: are the p-values from a similar GLM?

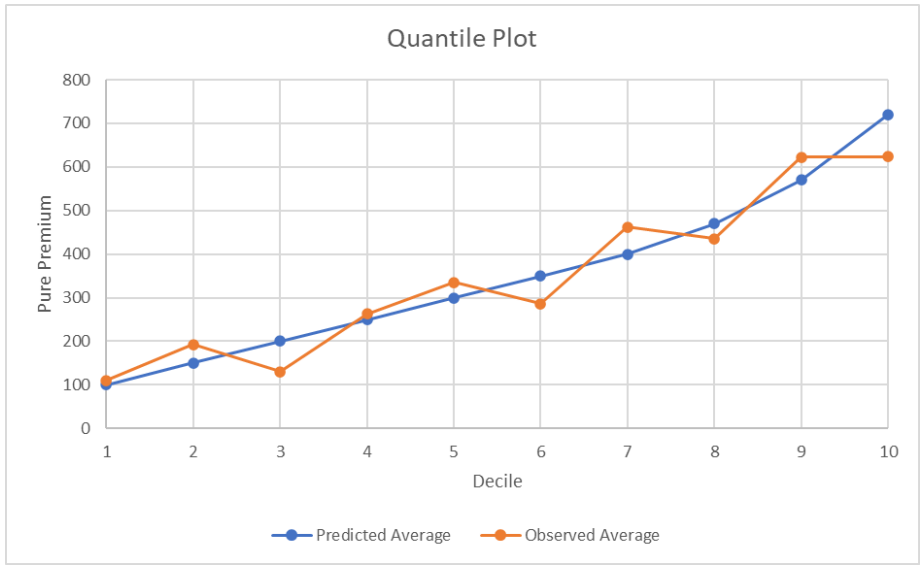
		1	2	3	4	5	Full Dateset
Usage	All Trips	0.530	0.518	0.560	0.755	0.690	0.618
	Professional	0.233	0.240	0.256	0.252	0.213	0.239
	Retired	-0.123	-0.134	-0.133	-0.152	-0.148	(0.138)
	Work Private			*Base*			
Gender	Female	0.039	(0.002)	0.034	0.020	0.041	0.026
	Male			*Base*			
Driver Age	16 - 20	0.132	0.124	0.185	0.080	0.271	0.161
	21 - 30	(0.037)	0.042	0.009	0.001	0.042	0.011
	31 - 40			Base			
	41 - 50	(0.039)	(0.043)	(0.042)	(0.022)	(0.013)	(0.032)
	51 - 60	0.028	0.009	0.005	0.049	0.043	0.027
	61+	0.067	0.063	0.063	0.104	0.112	0.082

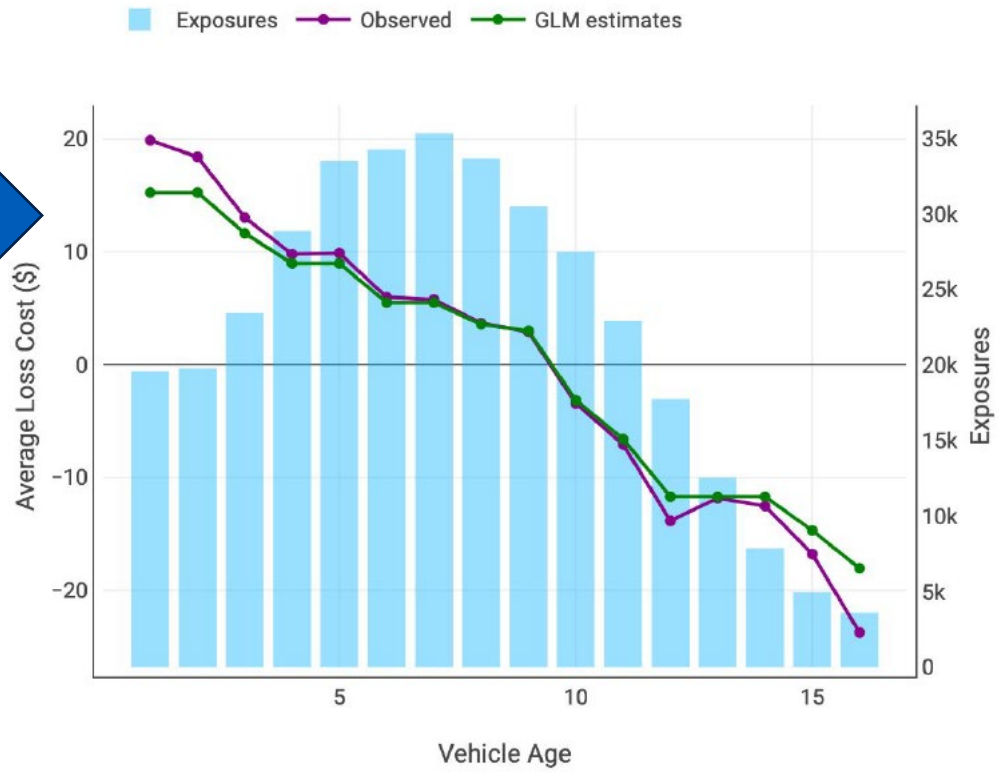
Notable Changes – P-Value Alternatives

- B.4.b
  - Bootstrapping: Do variations to the data result in radically different coefficients?
  - Cross Fold Validation: Are the coefficients consistent across folds?
  - GLM Reference Model: are the p-values from a similar GLM?
  - **The information element presents this as "either/or"**
  - **Certain p-value alternatives will be less applicable, depending on the scenario**

Notable Changes

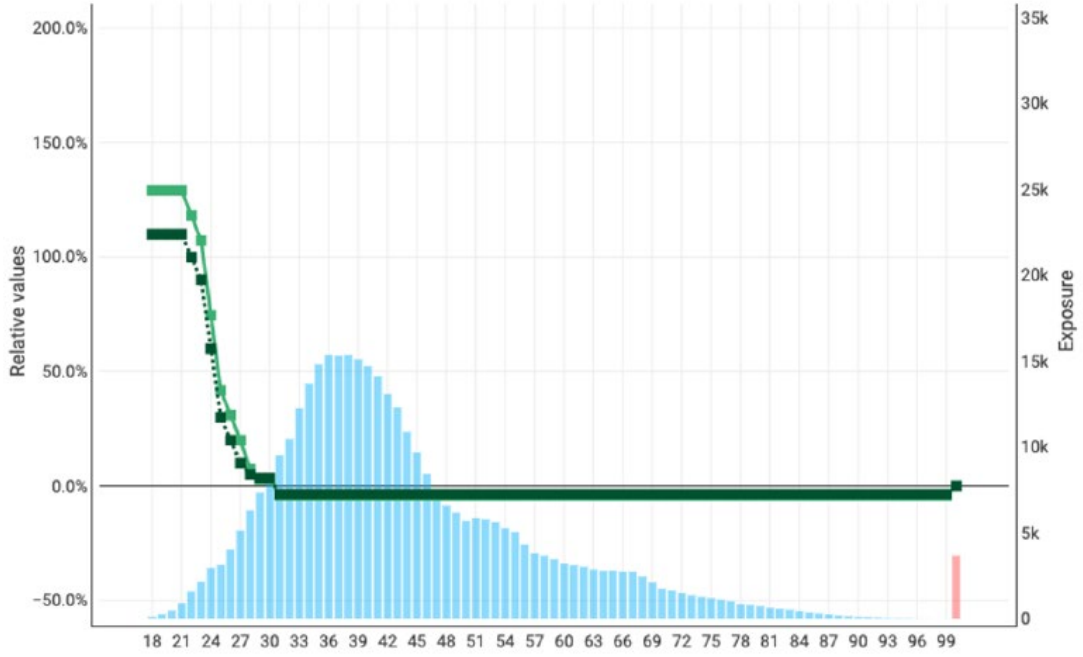
- B.4.c
  - Actual vs. Expected plots by variable
  - Quantile Lift Charts





Notable Changes

- B.5.e
  - Lasso Credibility includes a credibility complement
  - The model deviates from that starting point based on the data
  - It would be helpful to see a plot with BOTH the credibility complement factors and the new indicated factors



Past Book Club Presentations

- April 2024: Non-GLM Model Documentation
  - https://www.youtube.com/watch?v=SwB_E3wAB3Y
- October 2023: Derivative Lasso and Lasso Credibility
  - https://www.youtube.com/watch?v=QGa7Kf0BqRk
- August 2023: Regularized Regression with glmnet
  - https://www.youtube.com/watch?v=-H7bMFP4BSY
- October 2022: P-values and Alternatives
  - https://youtu.be/_V_z6f4L1qw

References

- Elastic Net, Lasso, Ridge
  - https://datasciencedojo.com/blog/regularization-in-machine-learning/
  - https://hastie.su.domains/ISLR2/ISLRv2_corrected_June_2023.pdf.download.html (Chapters 6.2 and 6.5)
- Derivative Lasso and Lasso Credibility
  - https://www.akur8.com/white-papers/derivative-lasso-credibility-based-signal-fitting-for-glms
- Accurate GLM (AGLM)
  - https://www.institutdesactuaires.com/global/gene/link.php?doc_id=16273

Questions

NATIONAL ASSOCIATION OF
INSURANCE COMMISSIONERS

APPENDIX B-RGLM – INFORMATION ELEMENTS AND GUIDANCE FOR A REGULATOR TO MEET BEST PRACTICES' OBJECTIVES (WHEN REVIEWING REGULARIZED GENERALIZED LINEAR MODELS)

This appendix identifies the information a state insurance regulator may need to review a 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 to be documented and shared with regulators in a timely and appropriate manner. Information technology (IT) controls should be in place, such as a record of versions, change control, and access to the model.[1]

Many information elements listed below are probably confidential, proprietary, or trade secret and should be treated as such, in accordance with state laws and/or regulations. Regulators should be aware of their state laws and/or regulations 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 compromise the model's protection.[2]

Although the list of information is 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 for a regulator (approximately 25%).

The "Level of Importance to the Regulator's Review" is a ranking of information a regulator may need to review which 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 Level 1 and Level 2. These data elements address even more detailed aspects of the model. This information does not necessarily need to be included with the initial submission, unless specifically requested by a particular state, as it is typically requested only if the reviewer has concerns that the model may not comply with state laws and/or regulations.

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 Level 1, Level 2, and Level 3. This most granular level of detail is addressing the basic building blocks of the model

[1] Bourdeau, M., 2016. "Model Risk Management: An Overview," The Modeling Platform, Issue 4, December. Accessed online at https://www.soa.org/globalassets/assets/library/newsletters/the-modeling-platform/2016/december/mp-2016-iss4.pdf.

[2] There are some models that are made public by the vendor and would not result in a hindrance of the model's protection.

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.

Appendix B-RGLM is focused on Regularized GLMs including lasso, derivative lasso, lasso credibility, ridge, elastic net, and accurate generalized linear models. This appendix should not be referenced in the review of other model types. This Appendix B-RGLM is intended to provide state guidance for the review of rate filings based on regularized GLMs.

A. 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 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).	1	Request details of 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 of 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 federal Fair Credit Reporting Act (FCRA). If the data is from an outside source, find out what steps were taken to verify the data was accurate, complete, and unbiased in terms of relevant and representative time frame, representative of potential exposures, and lacking in obvious correlation to protected classes. **Note**: 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. It is assumed that the data in the insurer's data banks is subject to routine internal company audits and reconciliation. "Aggregated data" is straight from the insurer's data banks without further modification (i.e., not scrubbed or transformed for the purposes of modeling). In other words, the data would not have been specifically modified for the purpose of model building. The company should provide some form of reasonability check that the data makes sense when checked against other audited sources.

Section	Information Element	Level of Importance to the Regulator' s Review	Comments
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	Many models are developed using a countrywide or a regional dataset. 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 explain why any states were excluded from the countrywide data. 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. However, 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. The company should provide a demonstration that the model fits well on the specific state or surrounding region.
2. Sub-Models			
A.2.a	Consider the relevance of (i.e., whether there is bias) of overlapping data or variables used in the model and sub-models.	1	Check if the same variables/datasets were used in the model, a sub-model, or as stand-alone rating characteristics. If so, verify the insurance company has processes and procedures in place to assess and address 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/accepted, that may reduce the extent of the sub-model's review. If approved, obtain the tracking number(s) (e.g., state, SERFF) and verify when and if it was the same model currently under review. Note: A previous approval does not necessarily confer a guarantee of ongoing approval, e.g., when statutes and/or 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 the prior decision needs to be revisited. In some circumstances, direct dialogue with the vendor could be quicker and more useful.

Section	Information Element	Level of Importance to the Regulator' s Review	Comments
A.2.c	Determine if the sub-model output was used as input to the Regularized GLM; obtain the vendor name, as well as the name and version of the sub-model.	1	To accelerate the review of the filing, it may be desirable to request (from the company), the name and contact information for a vendor representative. 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. Sub-model SMEs may need to be brought into the conversation with regulators (whether in-house or third-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	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. For example, it is important to know hurricane model settings for storm surge, demand surge, and long-term/short-term views.
A.2.e	Obtain an explanation of how catastrophe models areintegrated into the model to ensure no double-counting.	1	If a weather-based sub-model is input to the Regularized GLM under review, loss data used to develop the model should not include loss experience associated with theweather-based sub-model. Doing so could cause distortions in the modeled results by double-counting such losses when determining relativities or loss loadsin 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 or inclusion of freeze losses when using a winter storm model.
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. Depending on the result of item A.2.b, the importance of this item may be decreased.

Section	Information Element	Level of Importance to the Regulator' s 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	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 of 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 automobile comprehensive or home insurance. Premium should be brought to current rate level if the target variable is calculated with a premium metric, such as loss ratio. Premium can be brought to current rate level with the extension of exposures method or the parallelogram method. Note that the premium must be on-leveled at a granular variable level for each variable included in the new model if the parallelogram method is used. Statewide on-level factors by coverage are typically sufficient for statewide rate indication development but not sufficient for models that determine rates by variable level.
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	Pre-modeling binning may be unnecessary for ordinal variables in a lasso derivative or lasso credibility model, as the model will automatically set bins. Other regularized GLM approaches often group some variable levels with a base level during model fitting. However, if the insurer does bin variables or group levels before modeling, the reason should be understood.

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, null, or "not available" values in the data. For example, it would be helpful to the reviewer if the modeler were to provide a statement as to whether there is any systemic reason for missing data. If adjustments or recoding of values were made, they should be explained. It may also be useful to the regulator if the percentage of exposures and premium for missing information from the model data are 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 types of outliers and determine what adjustments were made to each type of 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 and/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, the company should provide support and a rational explanation for their use.

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 of how modeling analysis was adjusted and/or results were impacted.	1	"None" or "N/A" may be an appropriate response.

B. BUILDING THE MODEL

Section	Information Element	Level of Importance to the 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., lasso regression, ridge regression, elastic net regression, etc). Understandthe model's role in the rating system and provide thereasons why that type of model is an appropriate choice for that role. Understand why a Regularized GLM is preferable to a standard GLM for the specific modeling exercise.	1	It is important to understand if the model in question is a Regularized GLM and, therefore, these information elements 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 themodel (using the variables included in it) is appropriatefor the line of business. If by-peril or by-coverage modeling is used, the explanation should be by- peril/by-coverage. When a company is using a regularized GLM, it is helpful to understand why a penalized model is preferable to a standard GLM (without penalties for model complexity). **Note**: If the model is not a Regularized GLM, the information elements in this white paper may not apply in their entirety.
B.1.b	Identify the credibility complement used (if applicable). Lasso credibility is an example of a regularized generalized linear model which contains a credibility complement. Discuss why the selected complement is reasonable.	1	Many regularized generalized linear models are analogous in concept to a credibility weighted approach. Predictor variable values with low data volume will often result in coefficients that are closer to the credibility complement. For many regularized linear models, the implied credibility complement for each parameter is 0. However, in lasso credibility an alternate complement of credibility can be set. The alternate complement of credibility might be based on something like the currently approved rating factors. The regulator should determine if the complement of credibility is reasonable for use since it is not driven by the latest data.
B.1.c	Identify the software used for model development. Obtain the name of the software vendor/developer, software product, and a software version reference used in model development.	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 (e.g., a filing specialist) who can place the regulator in direct contact with the appropriate SME at the vendor. Open-source software/programs used in model development should be identified by name and version the same as if from a vendor.

Section	Information Element	Level of Importance to the Regulator' s Review	Comments
B.1.d	Obtain a description of how the available data was divided between model training, test, and/or validation datasets. The description should include an explanation why the selected approach was deemed most appropriate, 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 of why that came to occur. Obtain a discussion of whether the model was rebuiltusing all the data or if it was only based on the training data.	1	The reviewer should be aware that modelers may break their data into three or just two datasets. Although the term "training" is used with little ambiguity, "test" and "validation" are terms that are sometimes interchanged, or the word "validation" may not be used at all. It would be unexpected if validation and/or test data were used for any purpose other than validation and/or test, prior to the selection of the final model. However, according to the CAS monograph, "Generalized Linear Models for Insurance Rating": "Once a final model is chosen, … we would then go back and rebuild it using all of the data, so that the parameter estimates would be at their most credible." The reviewer should note whether a company employed cross-validation techniques instead of a training/test/validation dataset approach. If cross-validation techniques were used, the reviewer should request a description of how cross-validation was done and confirm that the final model was not built on any particular subset of the data, but rather the full dataset.
B.1.e	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.f	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.g	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.h	Obtain a description of the candidate variable selection process prior to the model building.	1	Candidate variables are the variables used as input to the modeling process. Certain variables may not end up used in the final model as some regularized GLM models (lasso, elastic net, etc.) will remove less significant variables. The narrative regarding the candidate 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 candidate variable selection. The modeler should comment on the use of automated feature selection algorithms to choose candidate predictor variables and explain how potential overfitting that can arise from these techniques was addressed.
B.1.i	In conjunction with variable selection, obtain a narrative on how the company determined the granularity of the rating variables during model development.	3	The narrative should include discussion of how credibility was considered in the process of determining the level of granularity of the variables selected.
B.1.j	Determine if model input data was segmented in anyway (e.g., 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.
2. Medium-Level Narrative for Building the Model			
B.2.a	At crucial points in model development, if selectionswere made among alternatives regarding model assumptions or techniques, obtain a narrative on the judgment used to make those selections.	3	
B.2.b	If post-model adjustments were made to the data andthe 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 of 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, including an explanation of the decision-making process to determine which interactions were included and which were not.	3	There should be a description of the 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. The number of interaction terms that could potentially be included in a model increases far more quickly than the number of "main effect" variables (i.e., the basic predictor variables that can be interacted together). Analyzing each possible interaction term individually can be unwieldy. It is typical for interaction terms to be excluded from the model by default, and only included where they can be shown to be particularly important. So, as a rule of thumb, the regulator's emphasis should be on understanding why the insurer included the

			interaction terms it did, rather than on why other candidate interactions were excluded. In some cases, however, it could be reasonable to inquire about why a particular interaction term was excluded from a model—for example, if that interaction term was ubiquitous in similar filings and was known to be highly predictive, or if the regulator had reason to believe that the interaction term would help differentiate dissimilar risks within an excessively heterogenous rating segment.
B.2.d	For the Regularized GLM, identify the link function used. Identifywhich distribution was used for the model (e.g., Poisson, Gaussian, log-normal, Tweedie). Obtain an explanation of why the link function and distribution were chosen. Certain distribution assumptions will involve numerical parameters, for example a Tweedie assumed distribution will have a p power value. Obtain the specific numerical parameters associated with the distribution. If changed from the default, obtain a discussion of applicable convergence criterion.	1	Solving the Regularized GLM is iterative and the modeler can check to see if fit is improving. At some point, convergence occurs; however, when it occurs can be subjective or based on threshold criteria. If the software's default convergence criteria were not relied upon, an explanation of any deviation should be provided. If the Regularized GLM did not reach convergence, an explanation should be provided.
B.2.e	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 toevaluate the predicted pure premium, relativity, or other value, for any real or hypothetical set of inputs.	2	
B.2.f	If there were data situations in which weights were used, obtain an explanation of how and why they were used.	3	Investigate whether identical records were combined to build the model.
B.2.g	Obtain the value of the applicable model complexity hyperparameter(s) and an explanation on how it was chosen.	2	Regularized GLMs have model complexity hyperparameters which can materially impact the final model parameters. The value of the model complexity hyperparameter determines whether the model is close to a standard GLM or is significantly tempered. For most regularized GLMs, tuning of the hyperparameter to maximize GINI on test data or minimize deviance on test data would be appropriate methods. For the derivative lasso method, it may be useful to review the plots of coefficients to determine if there is enough grouping of variable levels to remove reversals between adjacent variable levels.
B.2.h	Understand how the model would differ if different hyperparameter(s) were selected. Obtain a sensitivity analysis showing the coefficient output with higher and lower complexity hyperparameters or a plot showing coefficients by penalty value.	4	A regulator may decide they need more assurance that a reasonable value of complexity hyperparameter was selected. The regulator could ask for a sensitivity analysis showing how output model coefficients would differ if other hyperparameter values are used. 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.

3. Predictor Variables			
B.3.a	Obtain a complete data dictionary, including the names, data types, definitions, and uses of each predictor variable, offset variable, control variable, proxy variable, geographic variable, geodemographicvariable, and all other variables in the model used ontheir own or as an interaction with other variables (including sub-models and external models).	1	Data 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 its purpose 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 purpose of this requirement is to identify variables 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 a 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 greatest reduction in mean square error on test data, this would be a far, far weaker case forstatistical 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 Regularized 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 regulatory reviewer should understand what statistic was used to produce the matrix but should not prescribe the statistic.
B.3.d	Obtain a 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 may be 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 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 (usually linearly uncorrelated) transformed variables. In each instance, obtain a list of the pre- transformation and post-transformation variable names, as well as an explanation of how the results ofthe 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-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 multistate 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. **Note**: It may be useful to consider geographic stability measures for territories within the state.

B.4.b	For all variables, review the appropriate parameter values and relevant demonstrations of stability. Relevant demonstrations of stability may be provided as either plots by variable of indicated factors which also show upper bound and lower bound values (95th percentile and 5th percentile) on bootstrapped datasets, coefficient ranges across dataset folds, or p-values from a comparable standard GLM.	1	Statistical confidence intervals and p-values are often not available for Regularized GLMs. However, there are other ways to demonstrate model stability. The model could be run 100+ times on bootstrapped datasets to determine the stability of model parameters. If the bootstrapped models produce a narrow range of coefficient values, this implies the model is stable. Extra scrutiny should apply if the range of coefficient values includes negative and positive values. If the bootstrapped models produce a wide range of coefficient values, this implies the model is less stable. The range could be represented visually for each predictor variable by showing a plot with predictor variable values on the X-axis, and three separate lines representing mean indicated factors, the 95th percentile factors, and the 5th percentile factors. If the model was built with k-fold cross validation, the range of coefficients could be reviewed in a similar fashion. Narrower ranges represent a more stable model. The results may be less meaningful if more than 20 folds were used, since each model run would be based on significantly similar datasets. Variable stability can also be approximated by looking at the p-values from a comparable standard GLM which contains the same predictor variables as the Regularized GLM in question.

B.4.c	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	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 a Regularized GLM. The regulator should not assume to know what the company did and ask, "How?" Instead, the regulator should ask what the company did and be prepared to ask follow-up questions. For a Regularized GLM, such evidence may be available using observed vs. predicted average plots by variable and overall model lift charts. The regulator should ask for the company to provide exhibits or plots that show how the fitted average makes sense when compared to the observed average for variables of interest. Regulators would ideally review this comparison for every variable, but time constraints may limit the focus to just variables of interest. Variables of interest should include variables with high potential impacts on consumers (steep discounts or surcharges), variables without an intuitive relationship to loss, or variables that may be proxies for a protected class attribute. Lift charts such as quantile plots demonstrate the overall model fit. The risks in the modeling data are bucketed into quantiles with equal volume representing different levels of predicted risk. Quantile plots graph the predicted averages versus the observed averages by quantile. The quantile plots should have at least 10 quantiles to demonstrate predictive accuracy across different risk levels.

B.4.d	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. Obsolescenceis 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 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.e	Obtain a narrative on how potential concerns with overfitting were addressed.	2	
B.4.f	Obtain support demonstrating that the overall Regularized GLM assumptions are appropriate.	3	A 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 Regularized 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.g	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 of 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	The regulator 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 Lorenz curves or Gini coefficients were compared and obtain a narrative on the conclusion drawn fromthis comparison.	3	This information element requests a comparison of the Lorenz curve and Gini coefficient from the prior model to the Gini coefficient of proposed model. It is expected that thereshould be improvement in the Gini coefficient.A higher Gini coefficient indicates greater differentiation produced by the model and how well themodel fits that data. This is relevant when one model is being updated or replaced. The regulator should expect to see improvement in the new class plan's predictive ability. One example of a comparison might be sufficient. **Note**: This comparison is not applicable to initial model introduction. Reviewer can look toCAS monograph, "Generalized Linear Models for Insurance Rating."
B.5.c	Determine if double-lift charts were analyzed and obtain a narrative on the conclusion drawn from this analysis.	3	One example of a comparison might be sufficient. **Note**: "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 of 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	It is useful to differentiate between old and new variables, so the regulator can prioritize more time on variables not yet reviewed.
B.5.e	If using a credibility complement, obtain variable plots which visualize the credibility complement and the model indicated as separate lines. Lasso credibility is an example of a regularized generalized linear model which contains a credibility complement.	2	It is useful to see the coefficients as originally specified in the credibility complement, and how the model indicates these initially set coefficients should change based on the modeling data.
6. Modeler Software			
B.6.a	Request access to SMEs (e.g., modelers) who led the project, compiled the data, and/or built the model.	4	The filing should contain a contact that can put the regulator in touch with appropriate SMEs and key contributors to the model development to discuss the model.

C. THE FILED RATING PLAN

Section	Information Element	Level of Importance to the 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 (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 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. The regulator should 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 inputto 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.

Section	Information Element	Level of Importance to the Regulator' s 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) for the type of insurance productbeing 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 rational relationship to cost, and model results should be consistent with the expected direction of the relationship. **Note**: 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 relativities filed in the rating plan.	1	The documentation should include explanations for the necessity of any such adjustments and each significantdifference 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. **Note**: 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 regarding how each characteristic/variable was tempered or adjusted to account for possible overlap or redundancy in what the characteristic/variable measures.	2	Modeling loss ratios 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 inthe 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	The regulator should determine at what level of granularity credibility is applied. If modeling was by-coverage, by-form, or by-peril, the company should explain how these were handled when there was not enough credible data by coverage, form, or peril to model.

Section	Information Element	Level of Importance to the Regulator' s Review	Comments
C.4.b	If the rating plan is less granular than the model, obtain an explanation of why.	2	This is applicable if the company 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 of why.	2	A more granular rating plan may imply that the company had to extrapolate certain rating treatments, especially at the tails of a distribution of attributes, in a manner not specified by the model indications. It may be necessary to extrapolate due to data availability or other considerations.
5. Definitions of Rating Variables			
C.5.a	Obtain a narrative regarding 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 of 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 of whether it is raw or adjusted and, if the latter, obtain a detailed explanation for the adjustments.		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-level data should be considered when state indications differ from modeled results based on a broader dataset. However, the relevance of the broader dataset to the risks being priced should also be considered. Borderline reversals are not of as much concern. If multivariate indications perform well against the state-level data, this should suffice. However, credibility considerations need to be taken into account as state-level segmentation comparisons may not have enough credibility.

Section	Information Element	Level of Importance to the Regulator' s Review	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 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 renewal business 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 toindividual 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 D 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 D for an example of an exposure distribution.

Section	Information Element	Level of Importance to the Regulator' s Review	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, 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 chargeda 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. The 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**: This information may be proprietary. For the rating plan, the rate order of calculation rule may be sufficient. However, it may not be feasible for a regulator to get all the input data necessary to reproduce a model's output. Credit and telematics models are examples of model types where model output would be readily available, but the input data would not be readily available to the regulator.
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 toinquire as to how consumers can verify their data and correct errors.	1	If the data is from a third-party source, the company should provide information on the source. Depending on the nature of the data, it may need to be documented with an overview of who owns it. The topic of consumer verification may also need to be addressed, including how consumers can verify their data and correct errors.

Section	Information Element	Level of Importance to the Regulator' 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 Rate Filing			
C.9.a	Establish procedures to efficiently review rate filings and models contained therein.	1	"Speed to market" is an important competitive concept for insurers. Although the regulator needs to understand the rate filing before accepting the rate filing, the regulator should not request information that does not increase his/her understanding of the rate filing. The regulator should review the state's rate filing review process and procedures to ensure that they are fair and efficient.
C.9.b	Be knowledgeable of state laws and regulationsin order to determine if the proposed rating plan (and models) are compliant with state laws and/or regulations.	1	This is a primary duty of state insurance regulators. The regulator should be knowledgeable of state laws and regulations and apply them to a rate filing fairly and efficiently. The regulator should pay special attention to prohibitions of unfair discrimination.
C.9.c	Be knowledgeable of state laws and regulations in order to determine if any information contained in the rate filing (and models) should be treated as confidential.	1	The regulator should be knowledgeable of state laws and regulations regarding confidentiality of rate filing information and apply them to a rate filing fairly and efficiently. Confidentiality of proprietary information is key to innovation and competitive markets.

APPENDIX B-RGLM – INFORMATION ELEMENTS AND GUIDANCE FOR A REGULATOR TO MEET BEST PRACTICES' OBJECTIVES (WHEN REVIEWING ~~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 to be documented and shared with regulators in a timely and appropriate manner. Information technology (IT) controls should be in place, such as a record of versions, change control, and access to the model.[1]

Many information elements listed below are probably confidential, proprietary, or trade secret and should be treated as such, in accordance with state laws and/or regulations. Regulators should be aware of their state laws and/or regulations 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 compromise the model's protection.[2]

Although the list of information is 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 for a regulator (approximately 25%).

The "Level of Importance to the Regulator's Review" is a ranking of information a regulator may need to review which 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 Level 1 and Level 2. These data elements address even more detailed aspects of the model. This information does not necessarily need to be included with the initial submission, unless specifically requested by a particular state, as it is typically requested only if the reviewer has concerns that the model may not comply with state laws and/or regulations.

Level 4 – This information is necessary to continue the review of a model where concerns have been raised and not resolved based on

[1] Bourdeau, M., 2016. "Model Risk Management: An Overview," The Modeling Platform, Issue 4, December. Accessed online at https://www.soa.org/globalassets/assets/library/newsletters/the-modeling-platform/2016/december/mp-2016-iss4.pdf.

[2] There are some models that are made public by the vendor and would not result in a hindrance of the model's protection.

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 model is not a GLM, some listed items might not apply; e.g., not all predictive models generate p-values or F tests. Depending on the model type, other considerations might be important but are not listed here. When information elements presented in this appendix are applied to lines of business other than personal automobile and home insurance or other type of models, unique considerations may arise. In particular, data volume and credibility may be lower for other lines of business. Regulators should be aware of the context in which a predictive model is deployed, the uses to which the model is proposed to be put, and the potential consequences the model may have on the insurer, its customers, and its competitors. This white paper does not delve into these possible considerations, but regulators should be prepared to address them as they arise.~~

Appendix B-RGLM is focused on Regularized GLMs including lasso, derivative lasso, lasso credibility, ridge, elastic net, and accurate generalized linear models. This appendix should not be referenced in the review of other model types. This Appendix B-RGLM is intended to provide state guidance for the review of rate filings based on regularized GLMs.

A. 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 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).	1	Request details of 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 of 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 federal Fair Credit Reporting Act (FCRA). If the data is from an outside source, find out what steps were taken to verify the data was accurate, complete, and unbiased in terms of relevant and representative time frame, representative of potential exposures, and lacking in obvious correlation to protected classes. **Note**: 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. It is assumed that the data in the insurer's data banks is subject to routine internal company audits and reconciliation. "Aggregated data" is straight from the insurer's data banks without further modification (i.e., not scrubbed or transformed for the purposes of modeling). In other words, the data would not have been specifically modified for the purpose of model building. The company should provide some form of reasonability check that the data makes sense when checked against other audited sources.

Section	Information Element	Level of Importance to the Regulator' s Review	Comments
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	Many models are developed using a countrywide or a regional dataset. 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 explain why any states were excluded from the countrywide data. 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. However, 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. The company should provide a demonstration that the model fits well on the specific state or surrounding region.
2. Sub-Models			
A.2.a	Consider the relevance of (i.e., whether there is bias) of overlapping data or variables used in the model and sub-models.	1	Check if the same variables/datasets were used in the model, a sub-model, or as stand-alone rating characteristics. If so, verify the insurance company has processes and procedures in place to assess and address 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/accepted, that may reduce the extent of the sub-model's review. If approved, obtain the tracking number(s) (e.g., state, SERFF) and verify when and if it was the same model currently under review. Note: A previous approval does not necessarily confer a guarantee of ongoing ~~approval;~~approval, e.g., when statutes and/or 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 the prior decision needs to be revisited. In some circumstances, direct dialogue with the vendor could be quicker and more useful.

Section	Information Element	Level of Importance to the Regulator' s Review	Comments
A.2.c	Determine if the sub-model output was used as input to the ~~GLM~~Regularized GLM; obtain the vendor name, as well as the name and version of the sub-model.	1	To accelerate the review of the filing, it may be desirable to request (from the company), the name and contact information for a vendor representative. 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. Sub-model SMEs may need to be brought into the conversation with regulators (whether in-house or third-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	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. For example, it is important to know hurricane model settings for storm surge, demand surge, and long-term/short-term views.
A.2.e	Obtain an explanation of how catastrophe models areintegrated into the model to ensure no double-counting.	1	If a weather-based sub-model is input to the ~~GLM~~ Regularized GLM under review, loss data used to develop the model should not include loss experience associated with theweather-based sub-model. Doing so could cause distortions in the modeled results by double-counting such losses when determining relativities or loss loadsin 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 or inclusion of freeze losses when using a winter storm model.
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. Depending on the result of item A.2.b, the importance of this item may be decreased.

Section	Information Element	Level of Importance to the Regulator' s 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	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 of 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 automobile comprehensive or home insurance. Premium should be brought to current rate level if the target variable is calculated with a premium metric, such as loss ratio. Premium can be brought to current rate level with the extension of exposures method or the parallelogram method. Note that the premium must be on-leveled at a granular variable level for each variable included in the new model if the parallelogram method is used. Statewide on-level factors by coverage are typically sufficient for statewide rate indication development but not sufficient for models that determine rates by variable level.
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	Pre-modeling binning may be unnecessary for ordinal variables in a lasso derivative or lasso credibility model, as the model will automatically set bins. Other regularized GLM approaches often group some variable levels with a base level during model fitting. However, if the insurer does bin variables or group levels before modeling, the reason should be understood.

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, null, or "not available" values in the data. For example, it would be helpful to the reviewer if the modeler were to provide a statement as to whether there is any systemic reason for missing data. If adjustments or recoding of values were made, they should be explained. It may also be useful to the regulator if the percentage of exposures and premium for missing information from the model data are 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 types of outliers and determine what adjustments were made to each type of 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 and/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, the company should provide support and a rational explanation for their use.

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 of how modeling analysis was adjusted and/or results were impacted.	1	"None" or "N/A" may be an appropriate response.

~~B.~~ BUILDING THE MODEL

~~C.~~B.

Section	Information Element	Level of Importance to the 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., ~~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 that role. Understand why a Regularized GLM is preferable to a standard GLM for the specific modeling exercise.	1	It is important to understand if the model in question is a ~~GLM~~ Regularized GLM and, therefore, these information elements 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 themodel (using the variables included in it) is appropriatefor the line of business. If by-peril or by-coverage modeling is used, the explanation should be by- peril/by-coverage. When a company is using a regularized GLM, it is helpful to understand why a penalized model is preferable to a standard GLM (without penalties for model complexity). **Note**: If the model is not a ~~GLM~~Regularized GLM, the information elements in this white paper may not apply in their entirety.
B.1.b	Identify the credibility complement used (if applicable). Lasso credibility is an example of a regularized generalized linear model which contains a credibility complement. Discuss why the selected complement is reasonable.	1	Many regularized generalized linear models are analogous in concept to a credibility weighted approach. Predictor variable values with low data volume will often result in coefficients that are closer to the credibility complement. For many regularized linear models, the implied credibility complement for each parameter is 0. However, in lasso credibility an alternate complement of credibility can be set. The alternate complement of credibility might be based on something like the currently approved rating factors. The regulator should determine if the complement of credibility is reasonable for use since it is not driven by the latest data.
B.1.c~~b~~	Identify the software used for model development. Obtain the name of the software vendor/developer, software product, and a software version reference used in model development.	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 (e.g., a filing specialist) who can place the regulator in direct contact with the appropriate SME at the vendor. Open-source software/programs used in model development should be identified by name and version the same as if from a vendor.

Formatted Table

Section	Information Element	Level of Importance to the Regulator' s Review	Comments
B.1.~~c~~d	Obtain a description of how the available data was divided between model training, test, and/or validation datasets. The description should include an explanation why the selected approach was deemed most appropriate, 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 of why that came to occur. Obtain a discussion of whether the model was rebuiltusing all the data or if it was only based on the training data.	1	The reviewer should be aware that modelers may break their data into three or just two datasets. Although the term "training" is used with little ambiguity, "test" and "validation" are terms that are sometimes interchanged, or the word "validation" may not be used at all. It would be unexpected if validation and/or test data were used for any purpose other than validation and/or test, prior to the selection of the final model. However, according to the CAS monograph, "Generalized Linear Models for Insurance Rating": "Once a final model is chosen, … we would then go back and rebuild it using all of the data, so that the parameter estimates would be at their most credible." The reviewer should note whether a company employed cross-validation techniques instead of a training/test/validation dataset approach. If cross-validation techniques were used, the reviewer should request a description of how cross-validation was done and confirm that the final model was not built on any particular subset of the data, but rather the full dataset.
B.1.e~~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.ef	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~~g	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~~h	Obtain a description of the candidate variable selection process prior to the model building.	1	Candidate variables are the variables used as input to the modeling process. Certain variables may not end up used in the final model as some regularized GLM models (lasso, elastic net, etc.) will remove less significant variables. The narrative regarding the candidate 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 candidate variable selection. The modeler should comment on the use of automated feature selection algorithms to choose candidate predictor variables and explain how potential overfitting that can arise from these techniques was addressed.
B.1.i~~h~~	In conjunction with variable selection, obtain a narrative on how the company determined the granularity of the rating variables during model development.	3	The narrative should include discussion of how credibility was considered in the process of determining the level of granularity of the variables selected.
B.1.i~~j~~	Determine if model input data was segmented in anyway (e.g., 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 that 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 selectionswere made among alternatives regarding model assumptions or techniques, obtain a narrative on the judgment used to make those selections.	3	
B.2.b	If post-model adjustments were made to the data andthe 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 of 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, including an explanation of the decision-making process to determine which interactions were included and which were not.	3	There should be a description of the 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. The number of interaction terms that could potentially be included in a model increases far more quickly than the number of "main effect" variables (i.e., the basic predictor variables that can be interacted together). Analyzing each possible interaction term individually

			can be unwieldy. It is typical for interaction terms to be excluded from the model by default, and only included where they can be shown to be particularly important. So, as a rule of thumb, the regulator's emphasis should be on understanding why the insurer included the interaction terms it did, rather than on why other candidate interactions were excluded. In some cases, however, it could be reasonable to inquire about why a particular interaction term was excluded from a model—for example, if that interaction term was ubiquitous in similar filings and was known to be highly predictive, or if the regulator had reason to believe that the interaction term would help differentiate dissimilar risks within an excessively heterogenous rating segment.
B.2.d	For the ~~GLM~~Regularized GLM, identify the link function used. Identify which distribution was used for the model (e.g., Poisson, Gaussian, log-normal, Tweedie). Obtain an explanation of why the link function and distribution were chosen. Certain distribution assumptions will involve numerical parameters, for example a Tweedie assumed distribution will have a p power value. Obtain the ~~formulas for the distribution and link functions, including~~ specific numerical parameters ~~of~~ associated with the distribution. If changed from the default, obtain a discussion of applicable convergence criterion.	1	Solving the ~~GLM~~ Regularized GLM is iterative and the modeler can check to see if fit is improving. At some point, convergence occurs; however, when it occurs can be subjective or based on threshold criteria. If the software's default convergence criteria were not relied upon, an explanation of any deviation should be provided. If the Regularized GLM did not reach convergence, an explanation should be provided.
B.2.e	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 toevaluate the predicted pure premium, relativity, or other value, for any real or hypothetical set of inputs.	2	
B.2.f	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.
B.2.g	Obtain the value of the applicable model complexity hyperparameter(s) and an explanation on how it was chosen.	2	Regularized GLMs have model complexity hyperparameters which can materially impact the final model parameters. The value of the model complexity hyperparameter determines whether the model is close to a standard GLM or is significantly tempered. For most regularized GLMs, tuning of the hyperparameter to maximize GINI on test data or minimize deviance on test data would be appropriate methods. For the derivative lasso method, it may be useful to review the plots of coefficients to determine if there is enough grouping of variable levels to remove reversals between adjacent variable levels.
B.2.h	Understand how the model would differ if different hyperparameter(s) were selected. Obtain a sensitivity analysis showing the coefficient output with higher and lower complexity hyperparameters or a plot showing coefficients by penalty value.	4	A regulator may decide they need more assurance that a reasonable value of complexity hyperparameter was selected. The regulator could ask for a sensitivity analysis showing how output model coefficients would differ if other hyperparameter values are used.

			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.

Formatted: Not Expanded by / Condensed by

3. Predictor Variables			
B.3.a	Obtain a complete data dictionary, including the names, data types, definitions, and uses of each predictor variable, offset variable, control variable, proxy variable, geographic variable, geodemographicvariable, and all other variables in the model used ontheir own or as an interaction with other variables (including sub-models and external models).	1	Data ~~T~~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 its purpose 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 purpose of this requirement is to identify variables 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 a 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~~greatest reduction in mean square error on test data, 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~~ Regularized GLMs accommodate collinearity, the correlation matrix provides more information about themagnitude of correlation between variables. The company should indicate what statistic was used (e.g., Pearson, Cramer's V). The regulatory reviewer should understand what statistic was used to produce the matrix but should not prescribe the statistic.
B.3.d~~e~~	Obtain a 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 may be 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 connection that variable has to increasing or decreasing the target variable.
B.3.e~~f~~	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 (usually linearly uncorrelated) transformed variables. In each instance, obtain a list of the pre- transformation and post-transformation variable names, as well as an explanation of how the results ofthe 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-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 multistate 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. **Note**: It may be useful to consider geographic stability measures for territories within the state.

B.4.b	~~For all variables (discrete or continuous), review the appropriate parameter values and relevant tests of significance, such as confidence intervals, chi-squaretests, p-values, or F tests. Determine if model development data, validation data, test data, or other data was used for these tests.~~ For all variables, review the appropriate parameter values and relevant demonstrations of stability. Relevant demonstrations of stability may be provided as either plots by variable of indicated factors which also show upper bound and lower bound values (95th percentile and 5th percentile) on bootstrapped datasets, coefficient ranges across dataset folds, or p-values from a comparable standard GLM.	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; 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; e.g., confidence intervals around each level of an AOI curve might be more than whatis needed.~~ Statistical confidence intervals and p-values are often not available for Regularized GLMs. However, there are other ways to demonstrate model stability. The model could be run 100+ times on bootstrapped datasets to determine the stability of model parameters. If the bootstrapped models produce a narrow range of coefficient values, this implies the model is stable. Extra scrutiny should apply if the range of coefficient values includes negative and positive values. If the bootstrapped models produce a wide range of coefficient values, this implies the model is less stable. The range could be represented visually for each predictor variable by showing a plot with predictor variable values on the X-axis, and three separate lines representing mean indicated factors, the 95th percentile factors, and the 5th percentile factors. If the model was built with k-fold cross validation, the range of coefficients could be reviewed in a similar fashion. Narrower ranges represent a more stable model. The results may be less meaningful if more than 20 folds were used, since each model run would be based on significantly similar datasets. Variable stability can also be approximated by looking at the p-values from a comparable standard GLM which contains the same predictor variables as the Regularized GLM in question.
~~B.4.e~~	~~Identify the threshold for statistical significance and explain why it was selected. Obtain a reasonable and appropriately supported explanation for keeping the variable for each discrete variable level where thep-values were not less than the chosen threshold.~~	~~1~~	~~The explanation should clearly identify the thresholds for statistical significance used by the modeler. 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.~~

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

			~~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.f~~	~~For overall discrete variables, review type 3 chi-square tests, p-values, F tests and any other relevant and material test. Determine if model development data, validation data, test data, or other data was usedfor 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; e.g., confidence intervals around each level of an AOI curve might be more than whatis needed.~~

Formatted: Right: 0.07"

B.4.~~gc~~	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	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 a Regularized GLM. The regulator should not assume to know what the company did and ask, "How?" Instead, the regulator should ask what the company did and be prepared to ask follow-up questions. For a ~~GLM~~Regularized GLM, such evidence may be available using ~~chi-square tests, p-values, F tests and/or other means.~~observed vs. predicted average plots by variable and overall model lift charts. The regulator should ask for the company to provide exhibits or plots that show how the fitted average makes sense when compared to the observed average for variables of interest. Regulators would ideally review this comparison for every variable, but time constraints may limit the focus to just variables of interest. Variables of interest should include variables with ~~h~~high potential impacts on consumers (steep discounts or surcharges), variables without an intuitive relationship to loss, or variables that may be proxies for a protected class attribute. Lift charts such as quantile plots demonstrate the overall model fit. The risks in the modeling data are bucketed into quantiles with equal volume representing different levels of predicted risk. Quantile plots graph the predicted averages versus the observed averages by quantile. The quantile plots should have at least 10 quantiles to demonstrate predictive accuracy across different risk levels. ~~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 a GLM.~~ ~~The regulator should not assume to know what the company did and ask, "How?" Instead, the regulator should ask what the company did and be prepared to ask follow-up questions.~~

~~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	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. Obsolescenceis 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 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.e~~j~~	Obtain a narrative on how potential concerns with overfitting were addressed.	2	
B.4.f~~l~~	Obtain support demonstrating that the overall ~~GLM~~ Regularized GLM assumptions are appropriate.	3	A 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~~ Regularized 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

Formatted: Right: 0.07"

			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.g~~n~~	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 of 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	The regulator 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 Lorenz ~~curvers~~curves or Gini coefficients were compared and obtain a narrative on the conclusion drawn fromthis comparison.	3	This information element requests a comparison of the Lorenz curve and Gini coefficient from the prior model to the Gini coefficient of proposed model. It is expected that thereshould be improvement in the Gini coefficient.A higher Gini coefficient indicates greater differentiation produced by the model and how well themodel fits that data. This is relevant when one model is being updated or replaced. The regulator should expect to see improvement in the new class plan's predictive ability. One example of a comparison might be sufficient. **Note**: This comparison is not applicable to initial model introduction. Reviewer can look toCAS monograph, "Generalized Linear Models for Insurance Rating."
B.5.c	Determine if double-lift charts were analyzed and obtain a narrative on the conclusion drawn from this analysis.	3	One example of a comparison might be sufficient. **Note**: "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 of 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	It is useful to differentiate between old and new variables, so the regulator can prioritize more time on variables not yet reviewed.
B.5.e	If using a credibility complement, obtain variable plots which visualize the credibility complement and the model indicated as separate lines. Lasso credibility is an example of a regularized generalized linear model which contains a credibility complement.	2	It is useful to see the coefficients as originally specified in the credibility complement, and how the model indicates these initially set coefficients should change based on the modeling data.
6. Modeler Software			
B.6.a	Request access to SMEs (e.g., modelers) who led the project, compiled the data, and/or built the model.	4	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.

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

Section	Information Element	Level of Importance to the 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 (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 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. The regulator should 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 inputto 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.

Section	Information Element	Level of Importance to the Regulator' s Review	Comments
2. Relevance of Variables and Relationship to Risk of Loss			
C.2.a	Obtain a narrative regarding how the character-istics/rating variables included in the filed rating plan relate to the risk of insurance loss(or expense) for the type of insurance productbeing 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 rational relationship to cost, and model results should be consistent with the expected direction of the relationship. **Note**: 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 relativities filed in the rating plan.	1	The documentation should include explanations for the necessity of any such adjustments and each significantdifference 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. **Note**: 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 regarding how each characteristic/variable was tempered or adjusted to account for possible overlap or redundancy in what the characteristic/variable measures.	2	Modeling loss ratios with these characteristics/ variables as control variables would account for possible overlap. The insurer should address this possibility or other ~~considerations;~~considerations, e.g., tier placement models often use risk characteristics/ variables that are also used elsewhere inthe 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	The regulator should determine at what level of granularity credibility is applied. If modeling was by-coverage, by-form, or by-peril, the company should explain how these were handled when there was not enough credible data by coverage, form, or peril to model.

Section	Information Element	Level of Importance to the Regulator' s Review	Comments
C.4.b	If the rating plan is less granular than the model, obtain an explanation of why.	2	This is applicable if the company 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 of why.	2	A more granular rating plan may imply that the company had to extrapolate certain rating treatments, especially at the tails of a distribution of attributes, in amanner not specified by the model indications. It may be necessary to extrapolate due to data availability or other considerations.
5. Definitions of Rating Variables			
C.5.a	Obtain a narrative regarding 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 of 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 of whether it is raw or adjusted and, if the latter, obtain a detailed explanation for the adjustments.		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 anattribute is 1.5 and the multivariate indicated relativity is 1.25, this is potentially a plausible application of themultivariate techniques. If, however, the univariateindicated relativity is 0.7 and the multivariate indicatedrelativity is 1.25, a regulator may question whether the attribute in question is negatively correlated with otherdeterminants of risk. Credibility of state-level data should be considered when state indications differ from modeled results based on a broader dataset. However, the relevance of the broader dataset to the risks being priced should also be considered. Borderline reversals are not of as much concern. If multivariate indications perform well against the state-level data, this should suffice. However, credibility considerations need to be taken into account as state-level segmentation comparisons may not have enough credibility.

Section	Information Element	Level of Importance to the Regulator' s Review	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 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 renewal business 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 toindividual 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 D 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 D for an example of an exposure distribution.

Section	Information Element	Level of Importance to the Regulator' s Review	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 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 chargeda 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. The 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**: This information may be proprietary. For the rating plan, the rate order of calculation rule may be sufficient. However, it may not be feasible for a regulator to get all the input data necessary to reproduce a model's output. Credit and telematics models are examples of model types where model output would be readily available, but the input data would not be readily available to the regulator.
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 toinquire as to how consumers can verify their data and correct errors.	1	If the data is from a third-party source, the company should provide information on the source. Depending on the nature of the data, it may need to be documented with an overview of who owns it. The topic of consumer verification may also need to be addressed, including how consumers can verify their data and correct errors.

Section	Information Element	Level of Importance to the Regulator' 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 Rate Filing			
C.9.a	Establish procedures to efficiently review rate filings and models contained therein.	1	"Speed to market" is an important competitive concept for insurers. Although the regulator needs to understand the rate filing before accepting the rate filing, the regulator should not request information that does not increase his/her understanding of the rate filing. The regulator should review the state's rate filing review process and procedures to ensure that they are fair and efficient.
C.9.b	Be knowledgeable of state laws and regulationsin order to determine if the proposed rating plan (and models) are compliant with state laws and/or regulations.	1	This is a primary duty of state insurance regulators. The regulator should be knowledgeable of state laws and regulations and apply them to a rate filing fairly and efficiently. The regulator should pay special attention to prohibitions of unfair discrimination.
C.9.c	Be knowledgeable of state laws and regulations in order to determine if any information contained in the rate filing (and models) should be treated as confidential.	1	The regulator should be knowledgeable of state laws and regulations regarding confidentiality of rate filing information and apply them to a rate filing fairly and efficiently. Confidentiality of proprietary information is key to innovation and competitive markets.

Casualty Practice Council Update

Casualty Actuarial and Statistical (C) Task Force Meeting

August 13, 2024

Rob Fischer
Policy Analyst, Casualty

American Academy of Actuaries

About the Academy

The American Academy of Actuaries is a 20,000-member professional association whose mission is to serve the public and the U.S. actuarial profession. For more than 50 years, the Academy has assisted public policymakers on all levels by providing leadership, objective expertise, and actuarial advice on risk and financial security issues.

The Academy also sets qualification, practice, and professionalism standards for actuaries in the United States.

For more information, please visit:
www.actuary.org

Recent Activity

Committee on Property and Liability Financial Reporting (COPLFR)
- Comment Letter to the NAIC Actuarial (C) Opinion Working Group

Property and Casualty Risk-Based Capital Committee
- An Introduction to P&C Risk-Based Capital Webinar

P/C Extreme Events and Property Lines Committee
- California Catastrophe Modeling Comment Letter

P/C Committee on Equity and Fairness
- Comment Letter on Washington, D.C. Market Conduct Examination Report on Unintentional Bias in Private Passenger Auto Insurance
- Insurance Fraud Issue Brief (Upcoming)

Looking Ahead

Academy

- Casualty Loss Reserve Seminar, joint sponsorship with CAS (Sep. 9-11 in San Francisco)
- Academy Annual Meeting (Oct . 15-16 in Washington, DC)

Cyber Risk

- Cyber Vendor Model Comparison Issue Brief (Q3)
- SEC Cybersecurity Disclosure Requirements (Q3)

Workers Compensation

- Diamonds in the Rough: A Discussion of Lesser-Known Workers' Compensation Resources for Actuaries (Aug. 20)
- Unusual Animals Issue Brief (Q4)

Looking Ahead

P/C Risk-Based Capital
- Diversification Across Lines of Business Report (Q3/Q4)

Equity and Fairness
- Marketing/Underwriting Issue Brief (Q4)

Extreme Events
- Compound Disasters Issue Brief (Q4)
- Flood Monograph Update (Q4)

COPLFR
- 2024 Seminar on Effective P/C Loss Reserve Opinions (Dec. 9-10 in New Orleans, LA)
- 2024 Statements of Actuarial Opinion On Property and Casualty Loss Reserves Practice Note (Dec. 2024)
- 2024 P/C Loss Reserve Law Manual (Dec. 2024)

Thank you

Questions?

For more information, contact:
Rob Fischer, casualty policy analyst
fischer@actuary.org

American Academy of Actuaries

Expertise. Insight. Solutions.®

Casualty Actuarial Society Research and Continuing Education Update

Race and Insurance Pricing Research

The Casualty Actuarial Society is excited to announce that its first four research papers, which are part of Phase II of its Race and Insurance Pricing research initiative, will be available later in August. They will be posted at: **casact.org/raceandinsuranceresearch**

Phase I of the research, released in 2022, built a foundation for exploring areas of unintended racial bias by the insurance industry. Phase II, which will include release of an additional three papers in the fall, investigates bias itself, including racial bias, and how it can be identified and addressed in data, models and machine learning/artificial intelligence. Besides looking into rating factors, papers also focus on current regulatory and legislative approaches to addressing potential bias in individual states and across the globe.

At the 59th annual Actuarial Research Conference, Research Manager Annmarie Geddes Baribeau and CAS Board Member Richard Moncher provided a sneak peek into the Race and Insurance Pricing papers. At a plenary discussion, Baribeau emphasized that the CAS provides free access to all research and publications. She also announced several CAS research initiatives including access to more than four million earned exposures for confidential research purposes only and the new Quick Start Research Grant Program to encourage readers with timely ideas to apply for funding and receive a response within six weeks. University of Wisconsin Professor Peng Shi presented on the value of researching and publishing with the CAS and Moncher also provided an overall CAS update during the banquet meeting.

Highlights of Recent CAS Research

- **Artificial Intelligence and Social Inflation**
    - The CAS just released E-Forum, which includes three papers on artificial intelligence and a novel approach to determining social inflation. The July/August issue of Actuarial Review also includes two articles concerning artificial intelligence as well.
- **Reserving and Technology Papers**
    - In the next few weeks, the CAS will be releasing at least seven papers on different ways technology can be applied to reserving. The papers cover the gamut from new software to Bayesian models.

CAS Events

Regulators Welcome!

The CAS is offering regulators a special incentive to attend its continuing education programs with reduced registration fees.

The CAS also welcomes session proposals from the regulatory community. The current open calls for presentations and due dates are as follows:

The CAS also welcomes session proposals from the regulatory community. The current open calls for presentations and due dates are as follows:

- Ratemaking, Product and Modeling Seminar .....September 3

Upcoming Professional Education Events

- **Large CAS Meetings/Seminars**
  - Casualty Loss Reserve Seminar (CLRS) — San Fransisco, Sept 9-11
  - 2024 Annual Meeting — Phoenix, November 3 – 6, 2024
  - Ratemaking, Product and Modeling Seminar — Orlando, March 9 - 12, 2025
  - 2025 Spring Meeting — Ontario, Canada, May 4 – 7, 2025
  - Reinsurance Seminar — Washington, DC, June 4 – 6, 2025
- **Other Offerings**
  - **2024 Crash Course on Vehicle Technology and Automation** October 7 – 8
- **Virtual Workshops and Seminars**
  - 2024 Virtual Predictive Analytics Bootcamp — September 16
  - CAS Virtual Workshop: Introduction to Python — September 24
  - 2024 CAS Virtual Underwriting/Pricing Seminar – December 11

Webinars (through 4Q 2024)

- Recent and Proposed Changes to Actuarial Standards — August 13
- CAS International Webinar: Risk Based Capital for General Insurance — August 21
- GBMs and How to Use Them in P&C Insurance — August 29
- Estimating Pure IBNR Counts Using Policy Level Information — September 19
- Unlock Your Potential: Elevate Soft Skills and Master Self-Leadership — September 26
- 2024 AI Fast Track — November 12
- Navigating Insurance Fairness in Rapidly Evolving Regulatory Environments — November 14