

# Division of Insurance

## **Predictive Modeling – A Regulator’s Perspective: *Review of Property and Casualty Predictive Models in Nevada***

***Gennady Stolyarov II, FSA, ACAS, MAAA, CPCU, ARe, ARC, API, AIS, AIE, AIAF***

***Lead Actuary, Property and Casualty Insurance***

*Nevada Division of Insurance*

**Contact:** [gstolyarov@doi.nv.gov](mailto:gstolyarov@doi.nv.gov)

*Presentation for the  
Casualty Actuarial Society (CAS)  
Ratemaking, Product, and Modeling (RPM) Virtual Seminar  
July 29, 2020*



# Division of Insurance

## Commonly Used Abbreviations

- “DYA. YWBAQIYDN.\*” ~ [NVDOI](#)

\* Define your acronyms. You will be asked questions if you do not.

**NOTE:** All acronyms and abbreviations used in predictive models should be fully defined using complete English words. The use of undefined abbreviations or unexplained company-specific jargon will always subject a predictive model to an additional layer of detailed questioning and the corresponding elongation of the review timeframe. Comprehensively defining all shortened expressions is one of the easiest enhancements modelers can make to accelerate the review process.

CBIS = Credit-based insurance scoring

GLM = Generalized linear models / modeling

NRS = Nevada Revised Statutes

NCOIL = National Conference of Insurance Legislators

NVDOI = Nevada Division of Insurance

SERFF = System for Electronic Rate and Form Filing



# Division of Insurance

## *Types of Predictive Models We Review*

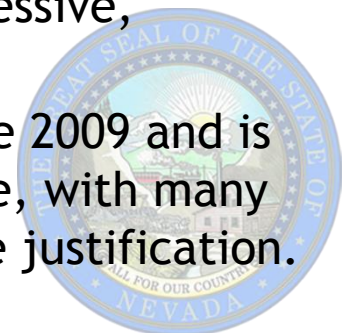
- **Credit-based insurance scoring (CBIS)** (majority of models)
- **Usage-based insurance (UBI)**
- **Vehicle history models** (only some characteristics allowed)
- **Location-based models** (use of geographical, demographic characteristics to develop proxies for territory)
- **Catastrophe and other peril-specific models** (earthquakes, wildfires, wind/hail)
- **“Price optimization” models:** Models which determine the extent to which a selected relativity moves toward an indicated relativity. These models may *only* consider characteristics related to the risk of insurance loss – ***not*** price elasticity of demand, tendency to complain, tendency to shop for insurance. (See [Bulletin 17-001](#).)
- **IMPORTANT:** We need to see the model and have it filed via SERFF in order to approve it.



# Division of Insurance

## Nevada's Regulatory Environment

- States vary in insurance laws, rate-regulatory regimes, and policy priorities. This presentation reflects Nevada's experience and environment.
- All rates and rating rules for personal lines of insurance must receive NVDOI approval prior to implementation.
- Predictive models must be filed with NVDOI by the individual insurers proposing to use them. Modelers may, at their discretion, require confidentiality for their models. However, confidentiality applies with respect to the general public, *not* with respect to regulators.
- **CBIS:** Greatest focus of model review to date. CBIS statutes are [NRS 686A.600-730](#), with the majority of the provisions in [NRS 686A.680](#).
- Statutes adopted in 2003, based on 2002 NCOIL Model Law.
- [NRS 686B.050-060](#): Standards for rates. Rates must not be excessive, inadequate, or unfairly discriminatory.
- Thorough review of all known Nevada CBIS models began in June 2009 and is ongoing. Over 45 models have been thoroughly reviewed to date, with many models receiving revisions to treatments which lacked adequate justification.



# Division of Insurance

## Recent Development: 1000-Page Rate-Filing Limitation

- Effective March 1, 2020 - Limitation of complex rate filings to 1000 company-submitted pages.
- See guidance at [http://doi.nv.gov/Insurers/Property\\_and\\_Casualty/Filing\\_Information/Filing\\_Information/](http://doi.nv.gov/Insurers/Property_and_Casualty/Filing_Information/Filing_Information/)
- For Rate/Rule filings, **split any large filing by subject matter**. Include **each predictive model in its own filing**.
- **Submit PDF tables as Excel documents**. Exhibits submitted as both Excel spreadsheets and PDF documents will not count twice towards the 1000-page limit, and submitting in Excel format will actually reduce the number of pages as one tab of the spreadsheet will count as one page, regardless of length.
- **Never include “placeholder” variables** for which there is no differentiation in rating treatment.
- Simplify the presentation of rate-related changes by providing **generally applicable formulas** and **verbally comprehensible decision rules** for assigning relativities to certain categories of insureds. **Graphical representations of relativity curves** can also help facilitate understanding.



# Division of Insurance

## Example: What to Avoid

This table is still manageable, but has many redundant elements. It would not scale well if there were many more categories of insureds receiving similar treatments.

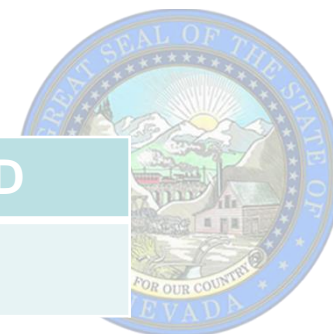
Insured Type	A	B	C	D
Relativity for Coverage X	1.00	1.00	1.00	1.35
Relativity for Coverage Y	1.00	1.00	1.00	1.35
Relativity for Coverage Z	1.00	1.00	1.00	1.35
Relativity for Coverage Q	1.12	1.14	1.16	1.16

## A Better Presentation

Make it as simple as possible but no simpler.

Insured Type	A, B, or C	D
Relativity for Coverages X, Y, and Z	1.00	1.35

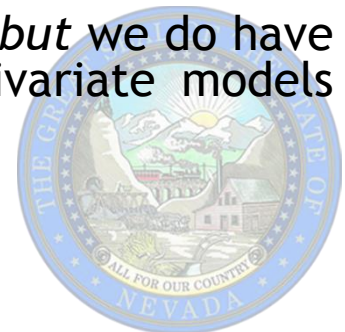
Insured Type	A	B	C or D
Relativity for Coverage Q	1.12	1.14	1.16



# Division of Insurance

## General Review Approaches

- Stand-alone in-depth analysis of predictive models (Tiering, company placement, relative weight compared to other variables outside the model are considered separately in individual insurer filings.)
- **We look at the details:** Overall correlation of the model as a whole with “deciles” or “vintiles” of the population is not sufficient.
- Consideration of the individual effects of each variable (comprehending the “language” in which the models are written).
- **Our aspiration is to “read” the model as we would read a book or essay, and achieve similar depth of insight.**
- Qualitative questioning regarding the rationales behind specific variables
- We prefer supporting **loss, premium, and loss-ratio** data by variable.
- Analysis of whether the variable treatments are supported by *loss-ratio* data (univariate statistical correlation)
- No direct multivariate analysis methods are available to NVDI, *but* we do have standards regarding information required as support for multivariate models (e.g., GLMs).
- NVDI’s univariate analysis tools: Excel, TI-83 Plus, Mind



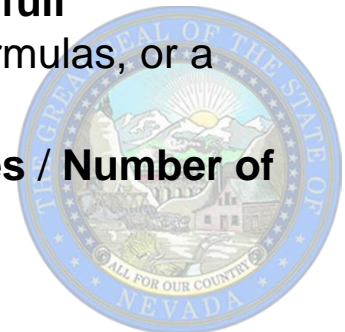
# Division of Insurance

## Multivariate Models: Required Support

Because we cannot replicate a specific multivariate modeling process (such as a generalized linear model) directly, we require three layers of support.

- **Layer 1: Raw Input Data:**

- The company should provide a **list of the input variables** that were used in the model.
- The company needs to specify the **timeframe** to which the data apply.
- The company needs to identify the **jurisdictional scope** of the data (state-specific, countrywide, etc.).
- The company needs to specify the **books of business included** in the model's input data (private passenger automobile, home, etc., as well as specific companies).
- If third-party data were used, the company should identify **who the third party is** and **how customers can obtain and correct any errors** that might be discovered in the data.
- If using the output of scoring algorithms, the company needs to provide **full documentation of any scoring algorithm used**, including sources, formulas, or a separate model used to generate the score.
- Underlying data for the model should include **Earned Premium / Losses / Number of Policies**.





# Division of Insurance

## Multivariate Models: Required Support

Because we cannot replicate a specific multivariate modeling process (such as a generalized linear model) directly, we require three layers of support.

- **Layer 2: Structure of Model:**

- The company should describe the type of model used - e.g., generalized linear model (GLM), decision tree, gradient-boosting machine (GBM), etc.
- The company should provide a thorough discussion of the **underlying assumptions** and **modeling methodology** and the **reasons for the approaches selected**.
- The company should include **all mathematical formulas** used as part of the model, including **formulas for any distribution function(s), link function(s), and numerical parameters** associated with any probability distribution.
- **NOTE:** If using a Tweedie distribution, the company **must** specify the numerical parameter  $\rho$  of the distribution.
- How did the company address the lack of credible data for any segments that had limited historical experience?



# Division of Insurance

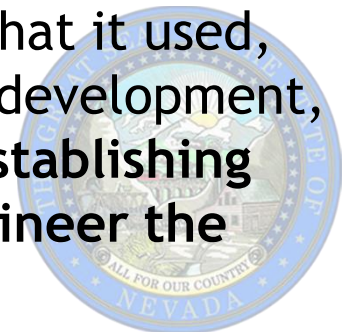
## Multivariate Models: Required Support

Because we cannot replicate a specific multivariate modeling process (such as a generalized linear model) directly, we require three layers of support.

- **Layer 3: Model Outputs:**

- These are typically **indicated relativities** and confidence intervals, which should be compared/contrasted with the **current** and **selected treatments** in a **side-by-side manner**.
- The company should provide an explanation for **filed rating values that deviate from the indicated values** and include **additional support for any values that are above indications and fall outside of any confidence intervals**. The additional support may be in the form of univariate analysis, loss-ratio analysis, or other similar detail that would support the deviation. Preferably, such support should be based on **Nevada-specific data**.

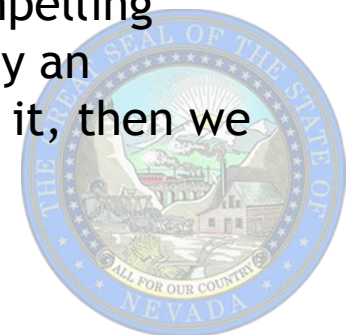
**Suggestion:** If an insurer enables us to see the same data that it used, analyzed via the same or similar tools as utilized in model development, this may greatly accelerate model review of approval by **establishing common ground and preventing the need to reverse-engineer the model justification**.



# Division of Insurance

## Logic and Common Sense: Going Beyond Correlation

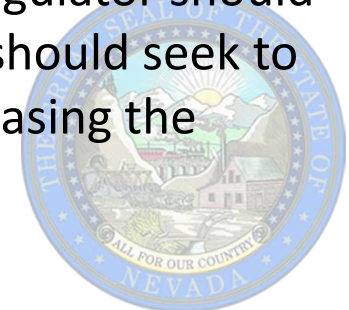
- Can the model variables be connected in any logical way with the underlying determinants of consumer risk, which they are supposed to measure/indicate (e.g., consumer financial responsibility or lack thereof in CBIS models)? If so, how? If not, why are they in the model?
- Do the variables reward or penalize financially reasonable, responsible behavior?
- Are there unintended consequences to the variables? (For instance, does an adverse treatment inadvertently encompass a highly favorable risk segment?) **Caution:** Are there unintended consequences to any contemplated changes to a treatment (e.g., massive premium disruption)?
- **We see the consumer side, too:** Could the NVDOI present a compelling justification for approving a particular treatment if challenged by an affected consumer or a legislator? If we cannot justify approving it, then we cannot approve it.



# Division of Insurance

## The Importance of a Rational Explanation

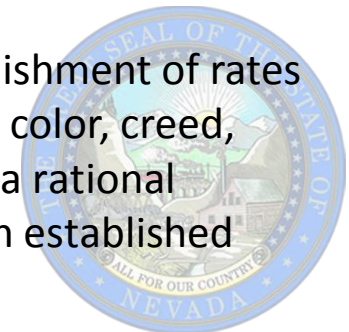
- The NAIC CASTF White Paper on Regulatory Review of Predictive Models *did not* originate the emphasis on rational explanation. **Regulators have been seeking rational explanations for newly introduced rating variables for decades. They will continue to seek rational explanations no matter what happens to the NAIC White Paper.**
- **Information Element 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.”
- **Comment in White Paper:** “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.”



# Division of Insurance

## “Rational Explanation” Defined

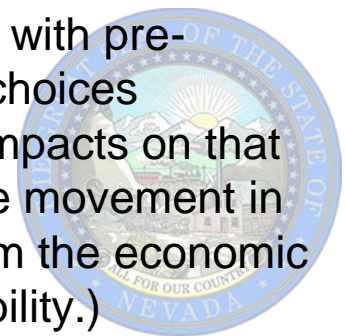
- **Rational Explanation** – A *“rational explanation”* refers to a plausible narrative connecting the variable and/or treatment in question with real-world circumstances or behaviors that contribute to the risk of insurance loss in a manner that is readily understandable to a consumer or other educated layperson. A *“rational explanation”* does not require strict proof of causality but should establish a sufficient degree of confidence that the variable and/or treatment selected are not obscure, irrelevant, or arbitrary.
- A *“rational explanation”* can assist the regulator in explaining an approved rating treatment if challenged by a consumer, legislator, or the media. Furthermore, a *“rational explanation”* can increase the regulator’s confidence that a statistical correlation identified by the insurer is not spurious, temporary, or limited to the specific data sets analyzed by the insurer.
  - NAIC CASTF White Paper on Regulatory Review of Predictive Models, Draft of June 12, 2020
- **In Nevada, we have requested rational explanations for various rating treatments for many decades. We will not approve a new rating variable without a rational understanding of why it works!**
- **NRS 686B.060(2):** “Risks may be classified in any reasonable way for the establishment of rates and minimum premiums, except that classifications may not be based on race, color, creed, national origin, sexual orientation or gender identity or expression.” If we lack a rational explanation, we cannot determine whether a classification treatment has been established based on reasonable considerations.



# Division of Insurance

## Common Issues: Inapplicable Data

- **Pre-economic-crisis data:** The 2007-2009 recession and consequent economic crisis constituted a paradigm shift in many areas of consumer life and financial behaviors. Nevada was especially affected. Use of any non-catastrophe data prior to or during this timeframe, *especially in newly developed models*, would raise serious concerns about obsolescence.
- **Note that even in the 2020 recession, the use of data from the prior recession and even before sometimes remains a problem!**
- **Failure to reflect COVID-19 impacts:** If the insurer solely relies on pre-COVID-19 data and fails to consider in some manner the shifts to consumer behaviors and constraints, as well as the changes to risk exposure resulting for massive disruptions to everyday life arising from COVID-19, then this would not be sufficient for demonstrating that the model would remain predictive during and after the pandemic.
- **Especially important:** Ensure that systemic adverse impacts from COVID-19 are not inadvertently reflected in variables for which indications were developed with pre-pandemic data. A variable that used to capture the impact of individual choices accurately when it was developed may no longer do so if the systemic impacts on that variable outweigh the individual-choice impacts. (**Example:** Any adverse movement in credit-based insurance scores during COVID-19 most likely resulted from the economic consequences of lockdowns and not from individual financial irresponsibility.)



# Division of Insurance

## Common Issues: Inapplicable Data

- **Countrywide data or data solely applicable to other states:** The NVDOI does consider relevant countrywide data, but asks that Nevada-only data be presented as a basis for comparison wherever possible. However, due to Nevada's unique profile when it comes to major perils (no hurricane risk, negligible tornado risk, generally much lower other catastrophe losses than surrounding states), the NVDOI does **not** accept the use of countrywide, regional, or any other non-Nevada information with regard to **catastrophe losses or trends**.
  - **NOTE:** The argument that "Nevada data are not fully credible" does not justify non-reliance on Nevada data. Lack of full credibility may, however, justify some manner of credibility weighting of non-catastrophe data.
- **Catastrophe data:** The use of long historical timeframes for *Nevada-specific* catastrophe data is understandable. However, it is important to consider Nevada's changing risk profile during the 21<sup>st</sup> century. An immense growth (54%) in the Nevada population since 2000 was accompanied by a major decrease in catastrophe losses over the same timeframe.

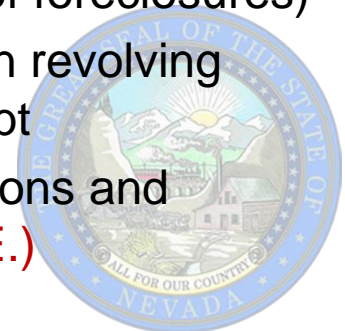


# Division of Insurance

## Common Issues: Unsupportable Variables

In the course of years of reviewing tens of major predictive models, the NVDOI has found the following variables to be lacking adequate support across the board. These variables generate outcomes which are adverse to responsible consumers, for whom the presence of such characteristics does not indicate increased insurance risk. These variables are considered unfairly discriminatory pursuant to NRS 686B.050 and, in the case of credit-related variables, are recognized to “lead to unfair or invidious discrimination” pursuant to NRS 686A.680(1):

- Any treatment whereby the absence of an automobile loan (e.g., the choice of a consumer to purchase a vehicle outright), a student loan, or other non-mortgage installment loan is treated more adversely than the presence of such a loan
- Any treatment whereby a “Missing” attribute is treated more adversely than the most adverse possible *known* attribute for a variable (e.g., treating the “Missing” category for foreclosures more adversely than the known presence of foreclosures)
- Any treatment whereby a \$0 outstanding credit balance on an open revolving account is treated more adversely than the presence of revolving debt
- Any treatment that rewards late payments / delinquencies / collections and penalizes their absence **(LIST CONTINUES ON THE NEXT SLIDE.)**

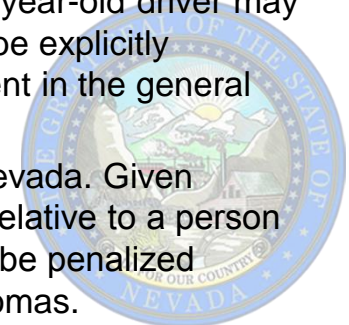




# Division of Insurance

## Common Issues: Unsupportable Variables (continued)

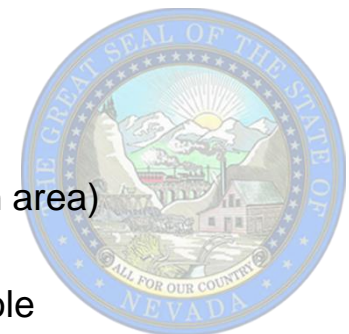
- Any treatment that penalizes a consumer for having paid off a loan (e.g., a mortgage or an automobile loan)
- Any worse-than-neutral treatment of credit “no hits” and “thin files” (a neutral treatment is the presumed baseline in NRS 686A.680(5)(b))
- Any treatment that adversely rates any policyholder in an area solely because of the prevalence of vacant housing units, a certain proportion of owner-occupied units, a certain income level in the area, a certain prevalent household composition in the area, certain prevalent education levels or occupational classifications in an area, or certain median / statistically prevalent ages of other residents in the area – irrespective of the risk characteristics of the individual policyholders in question. **This restriction is important for geodemographic models.**
  - **Example 1:** Age-based rating of *individuals* is allowed in Nevada. For instance, an 18-year-old driver may be surcharged relative to a 50-year-old driver. However, a 50-year-old driver may not be explicitly surcharged for sharing the road with a larger proportion of 18-year-olds than are present in the general population.
  - **Example 2:** Rating based on an *individual's* education or occupation is permitted in Nevada. Given adequate supporting data, a person with a bachelor's degree may receive a discount relative to a person with a high-school diploma only. However, a person with a bachelor's degree may not be penalized specifically for living in an area where most other residents only have high-school diplomas.



# Division of Insurance

## Common Issues: Unsupportable Variables (continued)

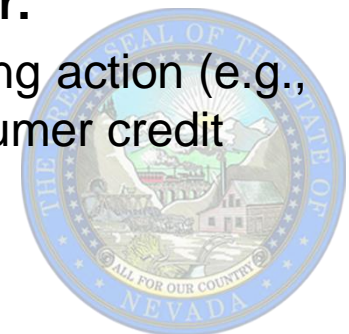
- A given vehicle's history of being stolen
- The mere fact that a vehicle was mentioned in a police accident report
- The behavior of a vehicle's prior owner
- The number of times a vehicle or home changed ownership
- The mere fact that a vehicle was present in an area that suffered a natural disaster
- Consumer's history of shopping for products that were not purchased on an installment basis
- Consumer's history of responses to advertising offered in print, in person, or online
- Criminal history of anyone other than the insured
- Price elasticity of demand, tendency to complain, tendency to shop for insurance  
(See [Bulletin 17-001](#).)
- Length of residency, except to offer discounts at new business
- Social-media usage habits (other than phone use while driving)
- Any demographic attributes of an area, including, without limitation:
  - Health status of residents (e.g., obesity rates for workers' compensation)
  - Residency characteristics (e.g., lengths of residency or typical occupancies in an area)
  - Prevalent occupations or educational levels, income, or wealth
  - Proportions of credit-based delinquencies and foreclosures for the area as a whole



# Division of Insurance

## Common Issues: Unsupportable Variables (continued)

- Statistics about an area that are irrelevant to insured behavior or geographic / climate features that affect risk of loss.
  - **Example 1:** Information about vegetation, slope, annual rainfall *may* be germane to wildfire risk for home insurance.
  - **Example 2:** However, concentrations of particulate matter – PM<sub>2.5</sub>, PM<sub>10</sub>, Ozone – are not germane to driving behavior or risk of loss for automobile insurance.
- **Cross-line pricing:** Use of automobile loss experience in home insurance and home loss experience in automobile insurance has been repeatedly shown by Nevada-specific data not to be predictive of risk in Nevada. There is also a concern that, if the same insurer provides both policies, then that insurer will double-charge the consumer for the impacts of the same incident or claim.
- Some data on accounts for telecommunications and utilities – e.g., number of inquiries for such accounts, number of such accounts, how recently those accounts were opened. **This information does not indicate risky behavior.**
- Treatments that result in premium increases or adverse underwriting action (e.g., declination, cancellation, nonrenewal) due to deterioration in consumer credit information arising out of the COVID-19 pandemic.

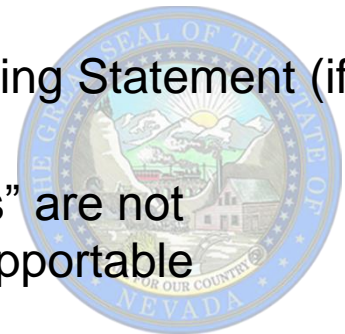


# Division of Insurance

## **Common Pitfalls:** *Rate Filings Involving Predictive Models*

**Each of the following pitfalls may delay approval of a filing by several months!**

- Failure to identify the existence of or any changes proposed to a predictive model. If we notice this, or see a cursory, single-sentence reference without further support for the model or change in question, the entire scope of filing review will expand significantly.
- Failure to include a copy of predictive-model algorithm (may be filed as confidential in SERFF).
- Refusal to provide any model or algorithm to the Division. (If we cannot review the model document, then the model may not be used in Nevada.)
- Objecting to the request for a rational explanation of a proposed rating treatment.
- Reference to older filings which contain obsolete predictive models (especially models subsequently revised at the Division's request), or which contain no model but references to still-older filings.
- Failure to provide the required informational items in the Credit Scoring Statement (if the company is using credit-based insurance scoring).
- Failure to provide variable-specific support. Overall model "lift charts" are not sufficient. An overall-predictive model may still contain several unsupportable categories and treatments.



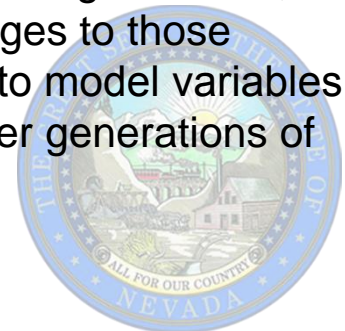
# Division of Insurance

## Common Pitfalls:

### *Rate Filings Involving Predictive Models*

**Each of the following pitfalls may delay approval of a filing by several months!**

- For predictive models involving multivariate techniques, failure to provide any of the three layers of support: **Layer 1: Raw Input Data**, **Layer 2: Structure of Model**, **Layer 3: Model Outputs**.
- Any worse-than-neutral treatment of credit no-hits and thin files (a neutral treatment is the presumed baseline in NRS 686A.680(5)(b); we will **not** approve more adverse treatments).
- Use of non-Nevada catastrophe data (for catastrophe models).
- Use of obsolete (especially pre-Great Recession) data.
- Sole reliance on countrywide experience when Nevada experience is qualitatively different.
- Introducing mid-filing major changes to predictive models and rating variables, except at the NVDOL's request.
- Implementing a transition from a prior generation of a predictive model to a new generation, where large premium impacts would occur to existing business without changes to those insureds' risk profiles or behavior, and without apparent significant changes to model variables. (We have commonly observed this issue in insurer attempts to move to newer generations of CBIS models. It is best not to re-score existing insureds in such situations.)



# Division of Insurance

## Communications with Modelers

- **We are extremely open** to communicating with modeling entities prior to any model development or submission, including meeting in person or via a teleconference, in order to convey expectations and/or give feedback as to how a particular treatment would be reviewed and what revisions (if any) and support would likely be requested. **Please contact us** if you have any questions whose resolution could accelerate a future model-review process.
- We are also very open to dialogue via various channels during the course of a formal model review. In addition to the objection-and-response mechanism in SERFF, we can receive supplementary documentation (e.g., Excel-based model score calculators or detailed spreadsheets of supporting data) via e-mail. We are happy to schedule conference calls if an issue could be more effectively discussed by telephone.
- **Key takeaway:** Open communication, thorough support, and avoidance of common pitfalls → More rapid model review and increased likelihood of approval



# Division of Insurance

## Closing Thoughts

- Offer the regulator the same courtesy of comprehensible narrative explanation that you would offer your company CEO. A CEO who trusts you will let your project proceed; a CEO who is confused by your presentation or concerned about adverse consequences likely will not. The regulator will have a similar degree of influence over whether your project moves forward and how.
- The P&C insurance industry response to COVID-19 has illustrated that insurers *are* capable of concise, straightforward, substantive filings that can be approved quickly. Now apply that skill set to predictive-model design and filings!

