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*

Presentation for the
Casualty Actuarial Society (CAS)
Ratemaking, Product, and Modeling (RPM) Seminar
March 21, 2018

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

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)
- **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.

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 40 models have been thoroughly reviewed to date, with many models receiving revisions to treatments which lacked adequate justification.

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 NVDOI, *but* we do have standards regarding information required as support for multivariate models (e.g., GLMs).
- NVDOI's univariate analysis tools: Excel, TI-83 Plus, Mind

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:** Provide the raw premiums, losses, and loss ratios that were used as inputs in the model. Specify the timeframe to which the data apply, the jurisdictional scope (state-specific, countrywide, etc.), and the books of business (private passenger automobile, home, etc., as well as specific companies).
- **Layer 2: Structure of Model:** Provide a thorough discussion of the underlying assumptions and modeling methodology and the reasons for the approaches selected. Include all mathematical formulas used.
- **Layer 3: Model Outputs:** These are typically indicated relativities, which should be compared/contrasted with the selected treatments.

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.**

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.

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 this timeframe, ***especially in newly developed models***, would raise serious concerns about obsolescence.
- **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 21st century. An immense growth (35%) in the Nevada population since 2000 was accompanied by a major decrease in catastrophe losses over the same timeframe.

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.)

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. All of the above are prohibited forms of **redlining**.
  - **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.

(LIST CONTINUES ON THE NEXT SLIDE.)

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
- 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

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