Predictive Modeling: Rules of Thumb for Communicators

Predictive Modeling Seminar
Insurance Marketing Communications Association
Chicago, IL
September 18, 2007

Download at
http://www.iii.org/media/presentations/predictivemodeling/

Robert P. Hartwig, Ph.D., CPCU, President
Insurance Information Institute ♦ 110 William Street ♦ New York, NY 10038
Tel: (212) 346-5520 ♦ Fax: (212) 732-1916 ♦ bobh@iii.org ♦ www.iii.org
PREDICTIVE MODELING: The Basics
Predictive Modeling: Communications Challenges

- Predictive Modeling Can Be Complex
  - Actuaries/Economists use a variety of statistical techniques
  - Understanding how they work requires formal statistical training
  - Underwriters apply them, usually as part of an already sophisticated and automated underwriting process
- Use of Some Predictive Factors/Models May Not be Intuitive
- Usage Often Not Explained or Even Revealed to Communicators
- Benefits Not Well Articulated to Communicators or Customers
- Failure to Recognize & Enlist Agents as Communicators
- Communications Obstacles in the Regulatory Context
  - Regulators may have difficulty understanding
  - Tendency is to react negatively
  - May seize on issue for political gain
- Models Maximize for Statistical Accuracy
  - Some May Feel Models Are Too Impersonal
  - Invasion of Privacy Concerns?
Predictive Modeling: What is It?

- **What is Predictive Modeling?**
  - While people (even within the insurance industry) tend to view it as new, it is in fact quite old—as old as insurance itself.
  - **DEFINITION:** Predictive modeling is a process used to create a statistical model of future behavior. In insurance, predictive models are primarily concerned with forecasting probabilities, trends and relativities.*
  - A predictive model is made up of a number of *predictors*, variable factors that are likely to influence future behavior or results.
  - In auto insurance, for example, a customer's gender, driving experience, type of vehicle, driving record, miles driven, etc., help predict the likelihood and cost of future claims. To create a predictive model, data is collected for the relevant predictors, a statistical model is formulated, predictions are made and the model is validated (or revised) as additional data becomes available. The models may employ a simple or extremely complex and employ a wide variety of statistical techniques.

- **Use of Some Predictive Factors/Models May Not be Intuitive**

*Adapted and modified by the Insurance Information Institute from [www.searchdatamanagement.com](http://www.searchdatamanagement.com) accessed Sept. 16, 2007.*
Predictive Modeling: Why Do We Hear So Much About it Today?

- Insurers rewrote their entire auto and homeowners book of business beginning in the later 1990s/early 2000s in response to huge losses in both of these key lines (which together account for nearly 50% of industry premiums)
- This re-underwriting process was effectively a re-evaluation of risk presented by each policyholder and the adequacy of the premium paid by the policyholder to transfer that risk.
- In most cases the premium was inadequate and premiums rose
- Re-underwriting process included the use of sophisticated new models designed to better match price with risk
- By definition, these models included more and better rating factors as well as new statistical methodologies for gauging interactions between these factors.
- Policyholders and regulators incorrectly associated new factors in the models as being solely responsible for the increase
- Credit-based “Insurance Scores” are the best known example
Private Passenger Auto (PPA) Combined Ratio

PPA is the profit juggernaut of the p/c insurance industry today.

Auto insurers have shown significant improvement in PPA after re-underwriting entire book of business in early 2000s.

Average Combined Ratio for 1993 to 2005: 101.0

Sources: A.M. Best; III
Predictive Modeling: Why Now?

- Predictive modeling is not new—big issue in most industries.
- Some form of it has been around since the earliest days of insurance—used in personal and commercial lines.
- In recent years the cost of data storage and acquisition have declined as has the cost of computing power.
- More data is available to insurers today at lower cost.
- Powerful computers make analysis (mining) of this data easier, faster, and more fruitful.
- Public and regulators have pushed for more individualized rates (and less reliance on factors like territory).
- Insurers responded by accelerating trend toward individual risk rating ➔ smaller pools of increasingly homogeneous individuals.
- Consequently, rating systems becoming fairer & more accurate.
- Implies that subsidies are being removed from system.
- Recipients of subsidies don’t like their removal nor do regulators who view insurance as an extension of the social welfare system.
Insurance Scores: The Perfect Example of a Communications Breakdown

- Insurers began to implement use of credit-based insurance score in the early/mid-1990s, but not on a large scale until late 1990s very early 2000s.
- Insurers had found that scores were among the most accurate of all rating factors for predicting future loss.
- Roll-out and use of credit was not communicated to most key personnel who come in contact with customers, regulators or media.
- Why credit works was not intuitive for most people (e.g., what does credit information have to do with my driving ability?)
- Agents dislike having to explain why premiums rose due to credit factors.
- Special cases warranted special treatment abounded: No credit, life-changing events, identity theft.
- Consumer protections formalized only later (e.g., NCOIL).
- Race issue became (and remains) big (but is red herring).
PREDICTIVE MODELING:
JUST PART OF THE RATEMAKING & UNDERWRITING PROCESS
Predictive Data Can Be Historical, Class or Individual Specific

- **Historical Information**: Used to identify trends in data
  - Actuaries use a variety of statistical techniques; get base rate

- **Class Rating**
  - Data are adjusted for geographic, industry-specific factors or other factors statistically correlated with risk of future loss
  - E.g. Urban zip codes = greater accident frequency
  - E.g. Occupation in workers comp

- **Individual Risk Rating**
  - Policyholder-specific risk factors are taken into account
    - E.g., Model of car; wood frame vs. masonry home; office vs. construction worker
    - Credit profile
    - “Black box” data;
    - FUTURE: GPS Tracking (on voluntary basis)

- **Experience Rating**
  - Adjustments made to premium based on policyholder’s past claim filing activity
UNDERWRITING:

Key to Accurate Risk Assessments & Rates
What is Underwriting?

- Underwriting
  - Process by which insurer determines whether policy should be issued and on what terms
- Complex Process
  - Many market and individual factors considered
  - All relate to riskiness/likelihood of loss
- Insurers All Use Underwriting Guidelines
  - Helps keep insurers focused, disciplined, profitable, solvent
  - E.g., no writing risks within 5 miles of coast, no high-rise construction risks, no limits above $1 million, no sportscars
- Underwriting Tools
  - Objective is to improve accuracy of loss forecasts
  - Creates a more fair, equitable rating system for all
  - Premium is more closely associated with risk
RATING FACTORS

Helping to Match
Premium Charged to
Risk Assumed
Categories of Typical Auto Insurance Rating Factors/Criteria

- Vehicle Type Factors
- Use of Vehicle Factors
- Location (Territorial) Factors
- Driving History
- Prior Insurance
- Personal Factors
- Other
Typical Auto Insurance Rating Criteria

- Vehicle Type Factors
  - Number of vehicles to be insured on policy
  - Number of operators in household
  - Make, model & body style of each vehicle
  - Age of vehicle (model year)
  - Safety features (e.g., airbags, anti-lock brakes)
  - Anti-theft devices

- Use of Vehicle Factors
  - Distance driven annually
  - Commuting distance
  - Number of days per week used to commute
  - Who drives vehicle the most?
  - Years of driving experience (youthful operator?)
  - Use of vehicle for business purposes
Typical Auto Insurance Rating Criteria

- **Location (Territorial) Factors**
  - Location where vehicle is kept
  - Garage or street parking

- **Driving History**
  - Accidents
  - Moving violations
  - Convictions (e.g., DUIs)
  - Personal claims history

- **Prior Insurance Factors**
  - Currently insured?
  - Number of years with current insurer?
  - Current Bodily Injury limits
Typical Auto Insurance Rating Criteria

• **Driving History**
  - Accidents
  - Moving violations
  - Convictions (e.g., DUIs)
  - Personal claims history

• **Prior Insurance Factors**
  - Currently insured?
  - Number of years with current insurer?
  - Current Bodily Injury limits
Typical Auto Insurance Rating Criteria

- **Personal Factors**
  - Marital Status
  - Gender
  - Occupation
  - Education
  - Student?
  - Homeowner?

- **Other Factors**
  - Information from credit reports
  - Drivers education, defensive driving course taken
Examples of Relationships Between Underwriting Criteria & Losses
Example 1:

GENDER & AUTO INSURANCE
Sex of Drivers Involved in All Auto Crashes, 1994-2003

Males are involved in 50% more accidents on average

Fatality Rate by Sex of Drivers
Involved in Auto Crashes, 1994-2003

Males are involved in 61% more likely to be killed in an auto accident.

Example 2:

**DRIVER AGE**
Accidents by Age of Driver, 2003

Teens account for just 5% of drivers but 22% of accidents! But people 35-44 represent 21% of drivers but just 16% of accidents.

Teens are by far the most likely to be involved in accident than the elderly (but elderly more likely to die in crash).

Source: National Safety Council; Insurance Information Institute
Example 3:

INSURANCE SCORING (CREDIT)
# Importance of Rating Factors by Coverage Type

<table>
<thead>
<tr>
<th>Coverage</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI Liability</td>
<td>Age/Gender</td>
<td>Ins. Score</td>
<td>Geography</td>
</tr>
<tr>
<td>PD Liability</td>
<td>Age/Gender</td>
<td>Ins. Score</td>
<td>Geography</td>
</tr>
<tr>
<td>PIP</td>
<td>Ins. Score</td>
<td>Geography</td>
<td>Yrs. Insured</td>
</tr>
<tr>
<td>Med Pay</td>
<td>Ins. Score</td>
<td>Limit</td>
<td>Age/Gender</td>
</tr>
<tr>
<td>Comprehensive</td>
<td>Model Year</td>
<td>Age/Gender</td>
<td>Ins. Score</td>
</tr>
<tr>
<td>Collision</td>
<td>Model Year</td>
<td>Age/Gender</td>
<td>Ins. Score</td>
</tr>
</tbody>
</table>

Texas Auto: Average Loss per Policy
(by Credit Score Decile, Total Market)

Interpretation:
Those with poorest credit scores generated incurred losses 65% higher than those with the best scores.

Source: University of Texas, Bureau of Business Research, March 2003.
Interpretation:

Those with poorest credit scores had loss experience 33% above average while those with the best scores had loss experience that was 19% below average.

Source: EPIC Actuaries, June 2003
Example: Credit Discount Can Save $100s per Year*

- Credit discount lowered annual premium by 14.7%
- Policyholder saved nearly $300
- Credit was single largest discount
- Opponents of credit will force people to pay more for coverage

Total Annual Savings from Discounts: $820

- Credit-Related Discount 36%
- Safety/Anti-Theft Discount 19%
- Good Driver Discount 24%
- Multipolicy Discount 21%

$296
$174
$196
$154

*Annualized savings based on semi-annual data from example

Source: Insurance Information Institute
Example 4:

WORKER AGE

(A Workers Comp Example)
THE AGEING WORKFORCE

Age Could be Used a Predictor of Occupational Injury and Loss, But it is Not
The median age of US workers as the Baby Boomer begin to retire is about 41 years. Immigration will hold this number down and may even lower the figure.

Fatal Work Injury Rates Climb Sharply With Age

Fatal Work Injuries per 100,000 Workers (2006)

Fatality rates for workers 65 and older are triple that of workers age 35-44. The workplace of the future will have to be completely redesigned to accommodate the surge in older workers.

Age is not used as a an underwriting factor in WC—should it be?

Example 5:

WORKER WEIGHT

(Another Example Relevant to Workers Comp that is Not Used)
THE OBESITY EPIDEMIC

Major Cost Driver that WC Has Yet to Address
The most obese workers file twice as many WC claims and 13 times more lost workdays than healthy weight workers.

Obesity is not a rating factor, but it is an identifiable cost factor.

Medical & Indemnity WC Claims Costs by BMI

Med claims costs are 6.8 times higher for the most obese workers and indemnity costs are 11 times higher.

Example 6:

TERRITORY
BI Liability costs in Baltimore are more than double (2.11 times) the state overall (i.e., 111% higher).

PD Liability costs in Baltimore are 47% higher than the state overall.

PIP costs in Baltimore are triple the state overall (200% higher).

*Baltimore Relativity to State Loss Cost, 2001-2003*

*ISO territories 33, 35, 36 and 39. Source: ISO.*
Baltimore Relativity to State Loss Cost, 1988

Costs in Baltimore were well above average back in 1988 too—still are today and will be in the future. This is permanent feature of most major urban auto insurance markets.

*ISO territories 33, 35, 36 and 39. Source: ISO.
Are There Limits to What Predictive Modeling Can or Should Do?

• Predictive Modeling Increases Accuracy, Equity in Rates
  ➢ Incumbent on insurers to use this information subject to limits imposed by policymakers

• Advances in Data Storage, Retrieval, Computing Will Advance the Frontier of Predictive Models

• Concern that Individual Risk Rating Will Replace Risk Pooling is Absurd
  ➢ No model will ever be 100% accurate
  ➢ Some degree of pooling will always exist

• Societal Boundaries Will Always Exist
  ➢ Predictive modeling will never be used to its full potential
  ➢ Privacy/”Big Brother” concerns
Predictive Modeling: 6 Rules of Thumb for Communicators

1. **EDUCATE:** Educate Yourself to Develop Understanding of How Products Work
   - Get to know actuaries and underwriters in your company

2. **PARTICIPATE:** Get Communications (not just Marketing) Involved at a Much Earlier Stage of Product Cycle

3. **ANTICIPATE:** Potential Communications Challenges *Before* Rollout

4. **IDENTIFY:** Subject Area Experts as Technical Resources

5. **DISSEMINATE:** Create Plan to Help Employees with Customer, Regulator & Media Contact Understand How Product Operates

6. **COORDINATE:** Ensure Marketing, Government Affairs, Customer Service, Agents all Operating from Same Playbook
Insurance Information Institute On-Line

WWW.III.ORG

If you would like a copy of this presentation, please give me your business card with e-mail address, or download at:

http://www.iii.org/media/presentations/predictivemodeling/