Regulatory Oversight of Insurers' Use of Big Data

Birny Birnbaum
Center for Economic Justice

NAIC Consumer Liaison Committee
March 2015

The Center for Economic Justice

CEJ is a non-profit consumer advocacy organization dedicated to representing the interests of low-income and minority consumers as a class on economic justice issues. Most of our work is before administrative agencies on insurance, financial services and utility issues.

On the Web: www.cej-online.org

Why CEJ Works on Insurance Issues

Essential Financial Security Tool for Individual and Community Economic Development: CEJ Works to Ensure Access and Fair Prices for These Essential Products and Services, particularly for Low- and Moderate-Income Consumers.

Primary Institution to Promote Loss Prevention and Mitigation: CEJ Works to Ensure Insurance Institutions Maximize Their Role in Efforts to Reduce Loss of Life and Property from Catastrophic Events.

Outline of Presentation

1. Big Data Defined
2. Insurer Big Data Application: Lexis Nexis Claims Tools
3. Public Policy and Insurer Goals of Risk Classification
4. Regulatory Framework For Risk Classification
5. History of Insurer Use of Big Data for Risk Classification
6. Insurer Big Data Application: Price Optimization
7. Insurer Rationale for PO: "Not Risk Classification, But Management Judgment"
8. PO and "Demand Models" Are Prohibited Risk Classes
9. Existing Risk Class Regulatory Framework Out of Date
10. Regulatory Big Data to Monitor Market Outcomes

Big Data Defined

- Massive databases of information about (millions) of individual consumers

- Associated data mining and predictive analytics applied to those data

- Scoring models produced from these analytics.

Insurance Big Data Example: LexisNexis Claims Tools

More Data Earlier: The Value of Incorporating Data and Analytics for Claims Handling at http://www.lexisnexis.com/risk/insights/value-incorporating-data-analytics-claims-handling.aspx

For third-party bodily injury settlements, the study found that more data earlier resulted in:

- 15–25 percent lower severity payments*
- 25–49 percent lower attorney involvement
- 5–15 percent shorter cycle times

Similar results were obtained for third-party property damage claims:

- 10–15 percent lower severity payments
- 8–15 percent shorter cycle times

LexisNexis Claims Tools

LexisNexis (LN) seeks to provide a Single Point of Entry for delivering all of information directly back into a carrier's system whether from a marketing standpoint, underwriting process or especially the claims part.

LN has over 10,000 data sources that feed into its infrastructure each month and has contributed information from the industry.

"Claims Data Fill" – deliver data and analytics directly into claims system in the claims process regarding parties, vehicles and carrier information. Used to verify information provided to insurers and provide indicators beyond the data to identify whether a social security number is an indicator of fraud or whether an address provided is a good address.

LexisNexis Claims Tools

Has an analytic component at first notice of loss and throughout the claim, constantly monitoring the claim looking for fraudulent activities. Real time data verification and enhancement with fraud scoring and attributes

Example, insured was rear-ended, all I got was license plate:

Claims Data Fill takes that license plate, reach out to DMV to get vehicle registration to get VIN number, we have policy database and get the carrier and policy information, take the registered owner, go out to public records, pull back their address, date of birth, telephone number, social security, wrap that into a package and put it back into our system, 88% of the time done in less than 5 seconds.

LexisNexis Claims Tools

Take minimum information provided at first notice of loss, provide a fraud score at the initial notice of loss. Daily monitoring of claim every time new information comes in, able to run various scores: fraud scores, severity score

New contributory claims database, much deeper than prior claims databases – this is claims file submitted as new information added – allows us to track vehicles across carriers, medical providers across carriers – sharing of information much deeper than has been done before. Text mining, watch list mixed with LexisNexis data.

Take-Away: Many databases and scoring models with little or no transparency to consumers and regulators and outside the scope of consumer protection laws like the FCRA.

Public Policy Goals of Risk Classification

1. Protect Insurer Financial Condition by Minimizing Adverse Selection

2. Promote Loss Mitigation by Providing Incentives for Less Risky Behavior and Disincentives for More Risky Behavior

Foundation of Risk Classification is Cost-Based Pricing

Foundation of Statutory Standards for Rates – "Not Unfairly Discriminatory" – is Cost-Based Pricing

History of Insurer Big Data Use for Risk Classification

Old Old School Big Data: Advisory Organization Loss Costs. Oversight of Data, Advisory Organization, Analytic Techniques, Filings, Complete Transparency

Old School Big Data: Credit-Based Insurance Scores. Limited Consumer Protections for Completeness and Accuracy of Data via the FCRA, Limited Oversight of Modelers and Models, Limited Transparency

New School Big Data: Predictive Modeling of Any Database of Personal Consumer Information. No Consumer Protections for Completeness and Accuracy of Data, No Oversight of Modelers and Models, No Transparency to Consumers

Insurer Big Data Application: Price Optimization

Adjusting cost-based rate indications based on "demand models." Demand models are models of consumer price elasticity of demand and competitive alternatives. Price elasticity of demand is consumer willingness to pay in face of price change – how likely a consumer is to shop for new insurance in face of, say, 7% rate increase.

Insurer Justification for Price Optimization

1. Insurers have always deviated from indicated rates for a variety of competitive and business reasons, relying on management judgment for such deviations. PO is simply a more scientific, data-driven approach to employing such management judgment.

2. Rating factors are factors related to costs of transfer of risk – loss costs or expenses. Since PO is not related costs of transfer of risk, it is not a rating factor and, consequently, not subject to regulatory oversight.

3. There is a statistical confidence interval around the indicated rate and any selection based on management judgment within that confidence interval is actuarially sound.

Insurers' Historical Deviation from Indicated Rates

- Historical deviation from rates has typically been an insurer selecting a lower rate than the indicated rate.

- Regulators have not routinely approved insurer requests for, say, a 20% rate increase when the insurer's indication is for a 5% rate increase.

- Historical deviation from indicated rates has almost always been a lower selected than indicated rate and the lower selection has been across broad risk groups

Price Optimization is Risk Classification

Definition: A risk classification/rating factor is any characteristic of the consumer, vehicle or property utilized by the insurer to determine the premium charge.

Rating factors are risk classifications and, by statute, must be related to expected costs of the transfer of risk – expected losses or expenses to issue and administer the policy.

- PO is clearly a rating factor as it is based on individual consumer characteristics and is applied to individual consumers to determine the premium charge for that consumer. At once, it is now obvious that ***PO is an impermissible rating factor because it is not related to the cost of transfer of risk***.

"PO Not Applied to Individual Consumers, But to Risk Classes"

- Modeling of Rates and Ultra-Refined Risk Classification Has Created Tens of Millions of Rating Cells Within A State – Far More Rating Cells Than Policyholders

- Allstate Complementary Rating Group (CRG)includes factors based on birthdates – two consumers otherwise identical but born a day apart are treated differently. CRG factor based on rating territory, gender, years with prior carrier and birthdate.

Allstate CRG Rating Examples

- Drivers with same gender, rating territory and years with prior carrier:

Birthdate	Rate Relativity	Rate Impact
4/16/1943	0.9803	
4/16/1943	1.0510	+7.2%
12/7/1980	0.9374	
12/8/1980	1.0252	+9.4%
7/4/1983	1.1784	+13.2%
7/7/1983	1.0406	

Small Midwest Insurer Auto Filing: Big Data Run Amok

Rating Plan with Millions of rating cells for a book of business of 25,000 policyholders.

"Geo-Demographic Data" for Creating ZIP Code Factors initially based on "1,044 Raw Demographic and 600 industry NAICS variables." Factors include:

Medicare Payments	Artificial sweeteners
Housing 1 Unit Detached	Bathroom Linens
Weather Index	Blue Collar Profile
Quality of Life Index	Dating Services
Motor Vehicle Theft Index	Hospital Room and Services
Manufacturing Employment	Margarine
Alcoholic beverage at home	School Lunches

“Adjustments Are Within the Confidence Interval”

- A confidence interval is created around the output of a statistic or statistical model. The size and nature of the confidence interval is determined by inputs chosen by the modeler, including the type of probability distribution used and the size of the data set used (e.g., number of observations), among many other factors.

- Ratemaking has been transformed from actuarial analysis of historical experience into a modeling exercise. Modeling is highly subjective and the results of the underlying ratemaking model can be manipulated.

Ohio DOI Bulletin on Price Optimization

Price optimization, however, involves gathering and analyzing data related to numerous characteristics specific to a particular policyholder that are unrelated to risk of loss or expense. Though not an exhaustive list, the Department has been presented with factors such as: whether the policyholder has complained about his or her policy, the amount or percentage change of the policyholder's auto premium at renewal in prior years, and the amount or percentage change of the policyholder's homeowners premium at renewal in prior years. From this and similar data, insurers are able to determine the "price elasticity of demand," or how much of a premium increase a particular policyholder will tolerate before switching insurance carriers.

Ohio DOI Bulletin cont'd: Thus, price optimization techniques allow insurers to set premiums based on an analysis of individual policyholder behavior reflecting a willingness to pay higher premiums than others - a factor completely unrelated to risk of loss or expense.

The use of price optimization represents a departure from traditional cost-based rating and can result in two insureds with similar risk profiles being charged different premiums. Therefore, by its nature, price optimization involves "discriminat[ing] between individuals of the same class and of essentially the same hazard" based on factors which do not have a demonstrable "probable effect upon losses or expenses."

Consequently, the use of price optimization results in rates that are unfairly discriminatory . . ...

PO Undermines Public Policy Goals of Risk Classification

- Undermines Risk Classification as Tool to Assure Financial Condition of Insurer – Replaces traditional and proven actuarial analysis for rates with modeling of prices. Introduces modeling risk to financial condition of insurers.

- For example of modeling risk, AIG Financial Services risk modeling indicated a 98.15% probability that AIG would not lose money on credit default swaps.

- Undermines Loss Mitigation Role of Insurance by Making Pricing More Opaque to Consumers and Less Related to Activities a Consumer Can Take or Avoid to Impact Pricing.

PO, Other Insurer Big Data Models Lack Key Consumer Protections

- Accuracy and Completeness of Data
- Regulatory Oversight of Data Bases
- Disclosures to Consumer: Data Used and How Used
- Consumer Ability to Challenge False Information
- Discrimination Against Low-Income and Minority Consumers
- Exacerbate Availability and Affordability Issues
- Undermine Insurance Pricing Role in Loss Mitigation

Regulatory Oversight of Insurers' Use of Big Data: Existing Risk Class Regulation Doesn't Work

Existing risk class regulation based on old old school big data, where regulators have oversight of all factors going into pricing and the data underlying the risk class analysis of rating factors and relativities.

Today, regulators simply do not have the resources to monitor all the databases and scoring models used by insurers nor access to the data underlying these new models.

If it is unrealistic to expect regulators to provide the type of historical review of advisory loss costs to new pricing tools, what is the way forward?

Regulatory Oversight of Insurers' Use of Big Data:

The current approach of allowing insurers to use any factor they want unless specifically prohibited does not fit with current data availability and technology. Regulators and legislators need to consider an approach of pro-actively identifying permissible risk classifications based not only on actuarial considerations, but also public policy goals of loss mitigation and availability.

Regulatory Big Data for Monitoring Market Outcomes

If regulators' ability to monitor what goes into marketing, sales, pricing and claims practices is realistically limited, then monitoring market outcomes is essential:

- Who is offered what insurance products at what prices in what locations?

- How are different groups of consumers treated in claims settlement?

Regulatory Big Data as a tool and strategy to improve effectiveness, efficiency and uniformity of state-based insurance market regulation.

Regulatory Big Data Already Used/Planned By State Insurance and Other Financial Regulators:

- Home Mortgage Disclosure Act data on individual mortgage applications by state and federal banking regulators
- Statutory Annual Statement data on individual bonds and investments by insurance prudential regulators
- PBR Transaction data on life insurance, disability insurance, long-term care insurance and annuities by insurance regulators as part of principles-based reserving.
- FINRA Comprehensive Automated Risk Data System (CARDS) – data relating to securities and account transactions, holdings, account profile information (excluding personally-identifiable information and securities reference data.