



**The Challenges and Opportunities of Big Data:
Reforming State-Based Insurance Regulation in the 21st Century**

Auto Insurance Report Conference

Birny Birnbaum

Center for Economic Justice

April 24, 2017

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

Insurance Products Are Essential Financial Security Tools for Individual and Community Economic Development.

CEJ Works to Ensure Fair Access and Fair Prices for These Essential Products and Services, particularly for Low- and Moderate-Income Consumers.

Insurance is the Primary Institution to Promote Loss Prevention and Mitigation, Resiliency and Sustainability:

CEJ Works to Ensure Insurance Institutions Maximize Their Role in Efforts to Reduce Loss of Life and Property from Catastrophic Events and Promote Resiliency and Sustainability of Individuals, Businesses and Communities.

Big Data Defined

Insurers' use of Big Data has transformed the way they do marketing, pricing and claims settlement. Big Data means:

- 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.

The scoring models generated by data mining and predictive analytics are algorithms. Algorithms are lines of computer code that rapidly execute decisions based on rules set by programmers or, in the case of machine learning, generated from statistical correlations in massive datasets. With machine learning, the models change automatically.

What's So Big About Big Data?

1. There has been a revolution in insurance pricing, marketing and claims settlement resulting from insurers' use of Big Data -- massive databases of new insurance and non-insurance personal consumer information with associated data mining and predictive analytics and scoring.
2. Insurers' use of Big Data has huge potential to benefit consumers and insurers by transforming the insurer-consumer relationship and by discovering new insights into and creating new tools for loss mitigation.
3. Insurers' use of Big Data has huge implications for fairness and affordability of insurance and for regulators' ability to keep up with the changes and protect consumers from unfair practices

4. The current insurance regulatory framework in most states – particularly related to risk classifications and unfair discrimination – does not provide regulators with the tools to effectively respond to insurers’ use of Big Data. Big Data has massively increased the market power of insurers versus consumers and versus regulators.
5. Market forces alone – “free-market competition” – cannot and will not protect consumers from unfair insurer practices. So-called “innovation” without some consumer protection and public policy guardrails will lead to unfair outcomes.
6. Regulatory reform emphasizing regulatory collection, analysis and publication of consumer market outcomes – Regulatory Big Data – will yield more efficient and effective regulation for consumers, insurers and producers, will promote more competitive markets and will foster quicker adoption of innovative technologies that benefit consumers and fulfill public policy goals.

Personal Consumer Information in Big Data

- Telematics – Auto, Home, Wearable Devices
- Social Media
- Shopping Habits/Purchase History
- Hobbies and Interests
- Demographics/Household Data/Census Data
- Government Records/Property Records
- Web Tracking
- Vehicle Registration and Service Records
- Facial Analytics
- Mainstream Credit Files: Loans, Credit Cards
- Alternative Credit Data: Telecom, Utility, Rent Payment

Examples of Insurer Big Data Algorithms

Pricing/Underwriting:

- Price Optimization/Demand Models
- Customer Value Scores,
- Telematics,
- Credit Scores,
- Criminal History Scores,
- Vehicle Scores,
- FireLine Rating
- Accelerated Life Insurance Underwriting

Claims:

- Fraud Scores,
- Severity Scores
- Telematics

A 20th Century Regulatory Framework for 21st Century Challenges

Oversight of Inputs Based on Proposition That Preventing Bad Inputs Will Prevent Bad Consumer Outcomes

Regulatory Oversight of

- Insurer Data – Licensing/Oversight of Advisory Organizations and Statistical Agents; Promulgation of Statistical Plans
- Cost-Based Pricing – Rates Not Excessive, Not Inadequate, Not Unfairly Discriminatory; Oversight of Rate Manuals/Rating Factors; Prohibition on Explicit Use of Certain Characteristics of the Consumer.
- Unfair Discrimination – Treating Individuals of the Same Class and Same Hazard Differently.

Interlude

What do advisory organizations do and why are they licensed and regulated?

What are advisory loss costs and how are they created?

Why do states (and actuarial standards of practice) require insurance prices to be cost-based?

Why is Insurance Different from Other Consumer Products?

1. **The insurance is required** – by law and by lenders requiring protection of home or vehicle collateralizing the loan. Limits normal competition.
2. **Contract is a promise for future benefits** if an undesirable event occurs. If the product “fails” – the consumer learns the insurance policy won’t cover the loss – she is stuck and can’t purchase another policy that would protect her against a known loss. ***Consumers have little or no information about the insurers’ performance.*** Again, limits normal competition.
3. **Cost-based pricing is required and consumer challenges to prices are prohibited.** The requirement for cost-based pricing is to protect insurer financial condition and prevent intentional or unintentional unfair discrimination.

The Current Regulatory Framework is Stressed in an Era of Big Data

Old, Old School Big Data and The Current Regulatory Framework:

- Oversight of Statistical Plans and Data Collection
- Licensing and Oversight of Advisory Organization Providing Pricing Assistance to Insurers
- Filings and Statistical Data Contain and Reference Everything Insurers Use for Pricing
- Complete Transparency to Regulators

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, Limited Transparency to Regulators, No Transparency to Consumers

The Regulatory Framework Breaks Down in an Era of Big Data

- Insurers now using data not subject to regulatory oversight or the consumer protections of the FCRA. Regulators have no ability to ensure the accuracy or completeness of these new data sets.
- Concept of unfair discrimination – consumers of similar class and hazard treated differently – becomes meaningless when insurers submit rating plans with millions of rate classes.
- New risk classifications can be proxies for protected classes, but with no recognition of disparate impact, risk classifications that have the effect of discriminating against protected classes are permitted. Big Data amplifies this problem.
- Ratemaking is effectively deregulated. Regulators must rely on insurer representations, not independent analysis.

Big Data Algorithms Can Reflect and Perpetuate Historical Inequities

Barocas and Selbst: *Big Data's Disparate Impact*

Advocates of algorithmic techniques like data mining argue that they eliminate human biases from the decision-making process. But an algorithm is only as good as the data it works with. Data mining can inherit the prejudices of prior decision-makers or reflect the widespread biases that persist in society at large. Often, the “patterns” it discovers are simply preexisting societal patterns of inequality and exclusion. Unthinking reliance on data mining can deny members of vulnerable groups full participation in society.

Example: Pricing Model

TransUnion Criminal History Score

“TransUnion recently evaluated the predictive power of court record violation data (including criminal and traffic violations)

“While a court record violation is created during the initial citation, the state MVR is updated later and may be delayed depending on a consumer’s response to the citation. For example, if someone pleads guilty and pays a ticket immediately, the state MVR will be updated in approximately two months. If the ticket is disputed in court, in contrast, the state MVR may not be updated for 6–19 months or longer.

“Also, as court records are created when the initial citation is issued, they provide insight into violations beyond those that ultimately end up on the MVR—such as violation dismissals, violation downgrades, and pre-adjudicated or open tickets.”

Insurer Use of Big Data Scoring Models Lack Fundamental Consumer Protections

- Accuracy and Completeness of Data
- Oversight of Data Bases
- Disclosures to Consumer About Data Used, How Used and Privacy Protections
- Consumer Ability to Challenge False Information
- Regulators' Knowledge Of and Capability to Provide meaningful Oversight
- Prevent discrimination Against Low-Income and Minority Consumers and other protected classes
- Asymmetric Use of Data
- Greater Cybersecurity Danger for Consumers and Insurers

21st Century Insurance Regulation

1. Identify What Insurers are Doing
2. Monitor Market Outcomes
3. Develop a 21st Century Approach to Oversight of Risk Classifications, including Oversight of Modelers Acting as Advisory Organizations

Identify What Insurers Are Doing

To a great extent, regulators – and, of course, consumers and policy makers – do not know what types of information insurers are using and what they are using the information for and how they are using it.

A logical first step is to develop a template for states to use, with assistance from the NAIC for collection of requested information, to request from insurers the sources and uses of data for various insurance functions. For each source of data, the insurer would provide a name/description of the data, the source of the data and the use or uses of the data -- pricing (including underwriting), marketing, claims settlement, antifraud and other.

This periodic survey will provide regulators with the basic overview of what types of information are being used by insurers and what the information is being used for. This information is essential for regulators to respond to policy makers and to foster public discussion over potentially controversial types of data.

Monitor Market Outcomes

The old regulatory model of monitoring everything that goes into insurer marketing, pricing and claims settlement practices and models is not feasible in the era of Big Data. Big Data has massively increased the market power of insurers versus regulators and versus consumers

Regulatory Big Data is needed. The data needed for a robust market analysis – one that includes the ability to monitor the affordability and accessibility of insurance in underserved areas as well the ability to perform enhanced market analysis to focus regulatory resources on problem companies and problem markets – is transaction data on premium quotes, policies issued and claims. Insurance regulators lag other financial regulators in the ability to monitor market outcomes and the collection of granular data for market monitoring.

Regulatory Big Data for Monitoring Consumer Market Outcomes

The regulatory framework of monitoring inputs in hopes of ensuring good consumer outcomes is no longer feasible in the era of Big Data.

Insurers' use of Big Data requires an approach that focuses on collecting and analyzing granular data on consumer market outcomes.

Regardless of what an insurer tells a regulator in filings, regulatory Big Data allows a regulator to monitor what is actually happening to consumers – both in terms of insurers' pricing and claim settlement intent and broader public policy issues.

- Do consumer market outcomes reflect the rate filing representations?
- Are consumers with similar claims receiving similar claim settlements regardless of income or race?
- What types of consumers in what locations are facing severe affordability problems and why?
- How can insurers, government and consumers partner for greater loss mitigation and resiliency?

More Granular Data Reporting by Insurers and Analysis by Regulators Produces Big Improvements in Efficiency and Effectiveness of Market Analysis.

More granular data reporting allows more refined market analysis.

More refined market analysis means fewer contacts with insurers to explain non-problems.

More granular data reporting means a huge reduction in special data calls because regulators already have the data in almost all cases.

More granular data reporting means more focused regulatory investigations and inquiries.

More granular data reporting means identifying insurers with good consumer market outcomes and leaving them alone.

More granular data reporting means regulatory involvement in company management policies and procedures if there is a problem, not as a routine practice.

This is Not State-of-the-Art Technology –

It is Technology from the 1990's

State-of-the-Art Technology would be Data Visualizations and Predictive Analytics of Market Outcome Data by Regulators.

There Are Opportunities to Build On Existing Data Reporting

Examples include:

- Transaction reporting for Principles-Based Reserving – eventually for life insurance, annuities and long-term care.
- Transaction reporting already in place for workers' compensation and significant portions of the industry in personal lines property and casualty insurance.

Benefits of Regulatory Big Data

- Calibrate inputs/processes to outcomes
- Assist financial regulation
- More efficient/effective market regulation
- Information to evaluate insurer and producer practices
- Information to evaluate existing and proposed public policies
- Information to meaningfully evaluate affordability and availability and competition issues
- **Information to empower consumers and promote competitive markets**

On the last point, consider how little information consumers have regarding the market performance of insurers and how publication of such information – frequency of claim denials, frequency of attorney involvement, average time to settle different types of claims and more – would empower consumers in the same way that information on car safety empowers consumers and motivates automakers.

Modernize Regulatory Oversight of Risk Classifications

Risk classification represents insurers' – and society's – decisions about how to group consumers for the purpose of assigning premium. Risk classification determines the affordability and availability of essential financial security tools – insurance – for consumers. In the vast majority of states, the only justification needed for a risk classification is a correlation. But, in an era of Big Data and the modeling of rates and risk classifications, such a test is arbitrary and opaque. We see this in the huge differences in rate impact of the same risk classification across insurers and even for an insurer across states. There is a need for a 21st century approach to oversight of risk classifications – more transparency and more accountability

21st Century Regulation of Insurance Risk Classification

Require More Than A Correlation:

- Risk-Based Pricing / Avoid Adverse Selection
- Promote Loss Mitigation / Degree of Consumer Control
- Transparency to Consumers/ Accuracy & Completeness of Data
- Honor the Risk-Spreading Purpose and Function of Insurance
- Avoid/Minimize Disparate Impact

Risk Classifications – whether called underwriting guidelines, tier placement rule or rating factors – should all be approved prior to use with opportunity for stakeholder input.

Regulatory Oversight of New Data – update statutes to treat organizations providing advisory organization services as advisory organizations. The key activity is collective pricing, not insurer ownership of the advisory organization.

Insurers Will Be Graded – and Consumers Will Base Purchases – on Treatment and Protection of Their Personal Information

- Disclosure – What data are you collecting and what specifically are you using it for?
- Consent – Are you obtaining my informed permission to access and collect this personal information?
- Access – Can I see the data you have gathered about me to see if it is correct and accurate?
- Error Correction – Do I have the ability to correct errors in the data and have my rate calculated or claim settled with correct data?
- Ownership – Do you agree that I own my data and can limit its uses and distribution?
- Portability – Can I take my personal data to another insurer?
- Symmetric Use – Are the data available to both you and me when I apply for coverage or file a claim?