

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

**NAIC Big Data Working Group**

**Birny Birnbaum**  
Center for Economic Justice

August 6, 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](http://www.cej-online.org)

**Why CEJ Works on Insurance Issues**

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

CEJ works to ensure **fair access** and **fair treatment** for insurance consumers, 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 to **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, claims settlement and their approach to risk management. For purposes of my talk, 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. Coupled with the increased volume and granularity of data is the digital technology to generate, access, process, analyze and deploy big data algorithms in real time

## What's So Big About Big Data?

1. 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.
2. Insurers' use of Big Data has huge implications for fairness, access and affordability of insurance and for regulators' ability to keep up with the changes and protect consumers from unfair practices
3. The current insurance regulatory framework generally 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.
4. 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.

5. Regulators and policymakers must understand the economic and competitive implications of Big Data on insurance. Without public policy action, captive markets will no longer be limited to add-on products markets like credit-related insurance. Other insurance markets – whether personal or commercial lines – will become captive markets where control over access is with the data vendors and algorithms describing and scoring the individual consumer or business.
6. The insurance industry and insurance regulatory systems are at a crossroad. **One possible future is empowered consumers and businesses partnering with risk management and sustainability companies who also provide insurance.**

**Another choice is a small set of insurers, data brokers and consulting firms who control access to insurance through opaque algorithms.**

## 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

### Marketing/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

## Captive Insurance Markets: Who Are the Gatekeepers for Insurance Sales and Claims?

Today, there are a number of smaller insurance markets in which the consumer is captive to the intermediary:

- Consumer Credit Insurance / Payment Protection sold by Lenders
- Force-Placed Insurance, Private Mortgage Insurance “sold” by Lenders and Loan Servicers
- Travel Insurance Sold by Airlines, Travel Agents
- Rental Car Insurance Sold by Rental Car Companies

Reverse competition” means competition among insurers that regularly takes the form of insurers vying with each other for the favor of persons who control, or may control, the placement of the insurance with insurers.

## Captive Markets Spread

The Lender is the gatekeeper for a captive market – the lender as intermediary determines what products will be sold to which consumers and how much of the premium the lender will extract from the insurer as consideration for the lender’s market power to open the gate to its consumers for the insurer.

Over the past ten years, captive markets have spread to many other types of insurance product and markets, but **now, the gatekeepers are Big Data Algorithms.**

Without public policy action, captive markets will no longer be limited to add-on products markets like credit-related insurance. Other insurance markets – whether personal or commercial lines – will become captive markets where control over access is with the data vendors and algorithms describing and scoring the individual consumer or business.

### Big Data Algorithms as Insurance Market Gatekeepers

- Marketing: web searches and web advertising that pre-score and channel consumers to particular products, providers and price-levels.
- Pricing: pre-fill applications and pricing without the consumer providing information, pricing based not just on risk but on price optimization / consumer demand models, real-time competitive options and/or socio-economic characteristics
- Claims: automated, instant claim settlement proposals based on data generated by a vehicle, home telematics or wearable device and utilizing price optimization/consumer demand models to determine amount of claim settlement offer a particular consumer is likely to accept based on his or her personal data.
- Common characteristics – opaque algorithms, little or no disclosure or transparency to consumer, great potential to penalize most vulnerable consumers, limiting loss mitigation role of insurance

### Allstate CEO to Investment Analysts, May 2017<sup>1</sup>

The insurer’s “universal consumer view” keeps track of information on 125 million households, or 300 million-plus people, Wilson said.

“When you call now they’ll know you and know you in some ways that they will surprise you, and give them the ability to provide more value added, so we call it the trusted adviser initiative,” said Wilson.

### Allstate’s Data Analytics Subsidiary<sup>2</sup>

“Arity is a data company — an insight company, really — whether or not it’s data from fitness sensors or home sensors,” Hallgren says. “But everything out of the gate so far is focused on the connected car.” That’s because the company is benefiting from the wealth of data its parent company has gathered from its DriveWise programs and other telematics initiatives — 22 billion miles in total, according to Hallgren.

<sup>1</sup> “Allstate CEO: Agents Will Have Access to Data on 125 Million Households,” Best’s New Service, May 30, 2017

<sup>2</sup> “Allstate’s Arity Unit Navigates Rapidly Changing World of Data,” Digital Insurance, June 5, 2017

## How Insurance Is 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
4. **There is Profound Public Interest in Broad Coverage** – failure or inability of consumers and businesses to access insurance has implications not just for individual families and businesses, but for taxpayers, communities and the nation.

## A 20<sup>th</sup> Century Regulatory Framework for 21<sup>st</sup> 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.

## 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.

**A computer algorithm reflects historical biases of the data and the developers.**

## Large Insurer CEO to Investment Analysts in 2005

Tiered pricing helps us attract higher lifetime value customers who buy more products and stay with us for a longer period of time. That's Nirvana for an insurance company.

This year, we've expanded from 7 basic price levels to 384 potential price levels in our auto business.

Tiered pricing has several very good, very positive effects on our business. It enables us to attract really high quality customers to our book of business.

The key, of course, is if 23% or 20% of the American public shops, some will shop every six months in order to save a buck on a six-month auto policy. That's not exactly the kind of customer that we want.

### 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

## Competitive Markets?

### Do consumers have market power to discipline insurers?

- **Mandatory Purchase**
- **No Information on Actual Insurer Performance**
- **Lack of Transparency on Pricing / Use of Factors Consumers Don't Believe Should Be Used**
- **Arbitrary Pricing – Wide Variation in Impact of the Same Risk Characteristic Across Insurers Within a State and For the Same Insurer Across States**

## Insurance Regulation in an Era of Big Data

1. Articulate the Future of Risk Management, Sustainability, Resiliency and Insurance:

**Empowered consumers and businesses partnering with risk management and sustainability companies who also provide insurance.**

**Greater, not less, transparency in insurance pricing, sales and claims settlements.**

### 2. Innovation in Insurance Regulation – Enhanced Data Collection for Market Monitoring and for Development of New Tools for Consumers

- a. What data are insurers using for what purposes? Routine collection – and publication – by regulators of the types, sources and uses of data by insurers for marketing, sales, pricing, claims settlement and loss mitigation.
- b. What consumer outcomes are insurers producing? Routine collection and analysis by regulators of granular consumer insurance market outcomes, including transaction-detail data on quotes, sales and claim settlements.
- c. Public data to empower consumers. Routine publication of insurer-specific anonymized consumer market outcomes.

### 3. Innovation in Insurance Supervision – New Tools to Empower Consumers – A Future in Which Consumers Shop for Insurance Based Not Only on Price, But:

- a. What data about me are you collecting and how well are you protecting my personal information? Insurers' and producers' transparency about and use and protection of consumers' personal information;
- b. What is your actual history of treating consumers? Insurers' and intermediaries' performance based on actual market outcomes for consumers; and
- c. What types of tools and assistance do you offer to help me manage my risk and control my premium? Insurers' and intermediaries' tools and partnerships for loss mitigation, loss prevention and consumer empowerment for risk management to control premium costs

#### 4. Innovation in Insurance Supervision – Develop / Improve / Reinvalidate Capabilities for Economic Analysis of Markets, Competition and Anti-Trust.

Will future success in insurance market be determined by quality of products and services or by amount of consumer data insurer/intermediary controls?

Regulatory Intervention to align market forces with consumer interest, when needed.

#### 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.

## 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 21<sup>st</sup> century approach to oversight of risk classifications – more transparency and more accountability

## 21<sup>st</sup> 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.