



**Insurance Regulation:  
The Challenge of Big Data in Insurance:**

**Casualty Actuarial Society  
Ratemaking & Product Management Seminar**

**Birny Birnbaum**  
Center for Economic Justice

March 20, 2018

## 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.

## **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.

## Reverse Competition and Captive Markets: Why a Problem?

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

## 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 and 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.**

## Current Regulatory Framework Challenged in 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

## **Current Regulatory Framework Challenged in 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.

## **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/Mobile Phone Tracking/Data Harvesting
- 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

## **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

## **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 by actuarial standards of practice and financial solvency.** 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.



# How to Keep Insurance Markets Competitive and Fair to Consumers and Improve Insurance Role for Economic Security, Loss Mitigation, Resiliency and Sustainability?

1. Articulate What the Future of Insurance Should Look Like.
2. Modernize Insurance Market Regulation.
  - a. Monitor Markets More Comprehensively and Efficiently
  - b. Develop Tools and Skills to Analyze Regulatory Big Data
  - c. Establish Consumer Disclosure, Access, Ownership and Protection Rules for Personal Consumer Information Used by Insurers
  - d. New Tools to Empower Consumers
  - e. Modernize Oversight of Risk Classification
    - i. Ethical Algorithms
    - ii. Emphasize Loss Mitigation
    - iii. Apply Disparate Impact Standard to Insurance
3. Assist, Not Criminalize, Low-Income Consumers to Obtain Essential Insurance.
4. Develop / Improve / Reinvigorate Capabilities for Economic Analysis of Markets, Competition and Anti-Trust.

## 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.

***“Before we choose our tools and techniques, we must first choose our dreams and our values, for some technologies serve them while others make them unobtainable.”*** Tom Bender

## 2. Modernize Insurance Market Regulation

### a. Monitor Market More Comprehensively and Efficiently

- i. 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.
- ii. 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.
- iii. Public data to empower consumers. Routine publication of insurer-specific anonymized consumer market outcomes.

## 2. Modernize Insurance Market Regulation

### b. Develop Skills and Tools to Analyze Regulatory Big Data

- i. NAIC resources to assist states with market outcome data collection, management and analysis comparable to NAIC tools for financial regulation and principles-based reserving.
- ii. Shift states' market regulation from primarily audit capability to primarily analytic capability by adding statisticians, economists, data scientists and big data modelers.

## 2. Modernize Insurance Market Regulation

### c. Consumer Disclosure, Access, Ownership and Protection Rules for Personal Consumer Information Used by Insurers

- i. Insurers' disclosures and consumer protections modeled after those in the Fair Credit Reporting Act – disclosure, consent, adverse action notice, access to data used, opportunity to correct erroneous data, life events exception
- ii. Ownership and consumer protections for consumer-generated data related to insurance – ownership by consumers and licensing to insurers of consumer-generated data, disclosure, affirmative opt-in, access, symmetrical use, transferability and standards for all industry databases, uses limited to agreed uses.

## 2. Modernize Insurance Market Regulation

### d. New Tools to Empower Consumers

- i. 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;
- ii. What is your actual history of treating consumers? Insurers' and intermediaries' performance based on actual market outcomes for consumers; and
- iii. 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

## 2. Modernize Insurance Market Regulation

### e. Modernize Oversight of Risk Classification

- i. Ethical Algorithms: Employ best practices to identify and eliminate disparate impact against protected classes. Commonly used by lenders and used by some insurance service organizations. Practices are consistent with cost-based pricing.
- ii. Emphasize Loss Mitigation: Deep commitment to cost-based pricing to ensure proper economic signals for cost of protection and loss mitigation investment. Emphasize risk classifications that empower consumers, prohibit use of socio-economic factors and credit scoring.
- iii. Apply Disparate Impact Standard to Insurance: If intentional discrimination against protected classes is prohibited, unintentional discrimination that has the same effect should be prohibited and minimized – see Ethical Algorithms.

## ***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.



New York Times, August 10, 2015: Algorithms and Bias: Q. and A. With Cynthia Dwork

Algorithms have become one of the most powerful arbiters in our lives. They make decisions about the news we read, the jobs we get, the people we meet, the schools we attend and the ads we see. Yet there is growing evidence that algorithms and other types of software can discriminate. The people who write them incorporate their biases, and algorithms often learn from human behavior, so they reflect the biases we hold.

Q: Some people have argued that algorithms eliminate discrimination because they make decisions based on data, free of human bias. Others say algorithms reflect and perpetuate human biases. What do you think?

A: Algorithms do not automatically eliminate bias. . . .Historical biases in the . . .data will be learned by the algorithm, and past discrimination will lead to future discrimination.

Fairness means that similar people are treated similarly. ***A true understanding of who should be considered similar for a particular classification task requires knowledge of sensitive attributes, and removing those attributes from consideration can introduce unfairness and harm utility.***

Q: Should computer science education include lessons on how to be aware of these issues and the various approaches to addressing them?

A: Absolutely! First, students should learn that design choices in algorithms embody value judgments and therefore bias the way systems operate. They should also learn that these things are subtle: For example, designing an algorithm for targeted advertising that is gender neutral is more complicated than simply ensuring that gender is ignored. They need to understand that classification rules obtained by machine learning are not immune from bias, especially when historical data incorporates bias.

Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*

America's poor and working-class people have long been subject to invasive surveillance, midnight raids, and punitive public policy that increase the stigma and hardship of poverty. During the nineteenth century, they were quarantined in county poorhouses. During the twentieth century, they were investigated by caseworkers, treated like criminals on trial. Today, we have forged what I call a digital poorhouse from databases, algorithms, and risk models. It promises to eclipse the reach and repercussions of everything that came before.

## Why Are Some Risk Classification Prohibited?

- Are Race, Religion and National Origin Prohibited Risk Classifications in Every State?
- If insurers can show a correlation between Race, Religion or National Origin and the costs of the transfer of risk, what is the basis for prohibiting these factors?
- What about prohibitions on the use of genetic test results, gender or consumer credit information – why does the Genetic Information Nondiscrimination Act of 2008 prohibit health insurers from denying coverage or different premiums based on genetic information? Why do some states ban the use of consumer credit information or gender as risk classifications?

## Why Are Some Risk Classification Prohibited?

Just over half of the states ban the use of race, religion and national origin in auto insurance risk classification. Just seven (7) states ban the use of race, religion and national origin for risk classification for auto, disability, life, health and property/casualty insurance.

Why are race, religion and national origin considered suspect classifications by the Supreme Court?

- (1) there is a history of discrimination against the group in question;
- (2) the characteristics that distinguish the group bear no relationship to the group members' ability to contribute to society;
- (3) the distinguishing characteristics are immutable; and
- (4) the subject class lacks political power.

Stated differently: utilizing race, religion and national origin as risk classification would reflect and perpetuate historical discrimination.

## **Ethical Algorithms: Minimizing Bias in Insurance Pricing and Claims Settlement Models**

### Common Industry Trade Argument:

Insurers don't consider race, religion or national origin, so there can be no unfair discrimination on the basis of these factors.

### Intentional Discrimination versus Disparate Impact:

If intentional discrimination against protected classes is prohibited, why would unintentional discrimination that has the same effect be permitted?

## **Example: Pricing Model**

### **TransUnion Criminal History Score**

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

“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.”

***What is the likelihood that TU Criminal History Scores have a disparate impact against African-Americans? Consider policing records in Ferguson, Missouri.***

## **US DOJ Investigation of the Ferguson Police Department**

Ferguson's approach to law enforcement both reflects and reinforces racial bias, including stereotyping. The harms of Ferguson's police and court practices are borne disproportionately by African Americans, and there is evidence that this is due in part to intentional discrimination on the basis of race.

Ferguson's law enforcement practices overwhelmingly impact African Americans. Data collected by the Ferguson Police Department from 2012 to 2014 shows that African Americans account for 85% of vehicle stops, 90% of citations, and 93% of arrests made by FPD officers, despite comprising only 67% of Ferguson's population.



## US DOJ Investigation of the Ferguson Police Department (2)

FPD appears to bring certain offenses almost exclusively against African Americans. For example, from 2011 to 2013, African Americans accounted for 95% of Manner of Walking in Roadway charges, and 94% of all Failure to Comply charges.

Our investigation indicates that this disproportionate burden on African Americans cannot be explained by any difference in the rate at which people of different races violate the law. Rather, our investigation has revealed that these disparities occur, at least in part, because of unlawful bias against and stereotypes about African Americans

## **Ethical Algorithms: Minimizing Disparate Impact in Insurance Models**

One Tool: Consider Prohibited Risk Classes in Model Development

***Step 1: Include race, religion and national origin – or proxies for these characteristics if actual individual characteristic unknown – as independent variables – control variables – in the model.***

By using the characteristics as independent variables in the development of the model, the remaining independent variables' contribution (to explaining the dependent variable) is shorn of that part of their contribution that is a function of correlation with the prohibited characteristics. For the independent variables other than race, religion and national origin, what remains is a more accurate picture of the remaining independent variables' contribution to the target outcome.

***Step 2: Omit race, religion and national origin when the model is deployed.***

## **Ethical Algorithms: Reasonable and Necessary for Insurance Pricing and Claims Settlement Models**

1. Minimizes Disparate Impact – Stop the Cycle of Perpetuating Historical Discrimination.
2. Promotes Availability and Affordability for Underserved Groups
3. Improves Cost-Based Insurance Pricing Models
4. Improve Price Signals to Insureds for Loss Mitigation Investments
5. Help Identify Biases in Data and Modelers / Improve Data Insights
6. Improve Consumer Confidence of Fair Treatment by Insurers

### **3. Assist, Not Criminalize, Low-Income Consumers' to Obtain Essential Insurance**

- Don't use insurance pricing to address affordability problems – no subsidies.
- Prohibit risk classifications that penalize consumers because of economic status.
- Put greater resources into assisting low-income consumers than to tracking, enforcement, penalizing, criminalizing and jailing consumers who cannot afford insurance.
- Create new product and pricing options to assist low-income consumers – low-cost auto product, pay-by-the-mile insurance.
- Federal, state and local government and insurer investments in resilient structures instead of subsidies to achieve affordable premiums.

#### **4. Develop / Improve / Reinvigorate 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 the insurer/intermediary/service organization controls?

The largest insurers – with the most data – have a profound competitive advantage over small- and medium-sized insurers because of far greater data assets.

Anti-competitive concerns raised by partnerships such as between largest insurers and automakers.

Regulatory Intervention to align market forces with consumer interest, when needed. Regulatory data and economic analysis skills need to meaningfully monitor structure and competitive nature of insurance markets. Fact-free sloganeering is not a substitute for substantive analysis.