Draft: 9/7/16

Big Data (D) Working Group
San Diego, California
August 26, 2016

The Big Data (D) Working Group of the Market Regulation and Consumer Affairs (D) Committee met in San Diego, CA, Aug. 26, 2016. The following Working Group members participated: Laura N. Cali, Chair (OR); Raymond G. Farmer, Co-Vice Chair (SC); Mike Kreidler, Co-Vice Chair (WA); Lori K. Wing-Heier (AK); Joel Laucher (CA); George Bradmer (CT); Sharon Shipp (DC); Franklin T. Pyle (DE); Susanne Murphy (FL); Anne Melissa Dowling (IL); Rhonda Fossitt (MI); Pam Gergen and Phillip Vigliaturo (MN); Angela Nelson (MO); Michael Kakuk (MT); Alex Feldvebel (NH); Richard J. Badolato (NJ); Michelle Brugh Rafeld (OH); Brian Gabbert (OK); Teresa D. Miller and Michael McKenney (PA); Elizabeth Kelleher Dwyer (RI); Leah Gillum (TX); and Kaj Samsom (VT). Also participating was: Bruce R. Ramge (NE).

1.   Adopted its May 19 and May 9 Minutes

Ms. Nelson made a motion, seconded by Director Farmer, to adopt the Working Group's May 19 (Attachment Four-A) and May 9 minutes (Attachment Four-B). The motion passed unanimously.

2.   Heard a Presentation from the CEJ on the Potential for Discrimination/Disparate Impact of Big Data and Predictive Analysis on Low-Income and Minority Consumers

Birny Birnbaum (Center for Economic Justice—CEJ) said 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, Mr. Birnbaum said the models change automatically.

Referencing a paper "Big Data's Disparate Impact" by Solon Barocas and Andrew D. Selbst, Mr. Birnbaum said traditional forms of data analysis return records or summary statistics in response to a specific query. Data mining attempts to locate statistical relationships in a dataset. Mr. Birnbaum said data mining automates the process of discovering useful patterns and revealing regularities upon which subsequent decision-making can rely. The accumulated set of discovered relationships is commonly called a "model," and these models can be employed to automate the process of classifying entities or activities of interest and predicting future outcomes. Mr. Birnbaum said by exposing so-called "machine learning" algorithms to examples of previously identified instances of fraud, the algorithm "learns" which related attributes or activities can serve as potential proxies for those qualities or outcomes of interest.

Mr. Birnbaum provided examples of insurer algorithms used for pricing, fraud and claim severity scores. He said personal information in big data includes social media, shopping habits, hobbies and web tracking. He said regulatory oversight is needed because correlation is not causation, and there might be spurious correlations. Mr. Birnbaum said a spurious correlation is a statistically valid association between variables that is not causally related. He said data mining and big data models in insurance are premised on correlation and not causation. Mr. Birnbaum said predictive models may reflect historic discrimination because of biased data, biased assumptions or faulty model specifications.

Mr. Birnbaum then provided an example of a LexisNexis claim tool that obtains public record information based upon a license plate number and provides a fraud score at the initial notice of loss. He said the foundational information may be biased for a variety of reasons, including historical scrutiny of claims submitted by minorities or biases of claims settlement personnel. Mr. Birnbaum said that historical biases in the data will be learned by the algorithm and that past discrimination will lead to future discrimination. He said insurers' use of big data has huge implications for fairness and affordability of insurance and for regulators' ability to keep up with the changes to protect consumers from unfair practices. Mr. Birnbaum said the use of big data has increased the market power of insurers versus consumers and regulators and that market forces alone will not protect consumers.

Commissioner Cali asked Mr. Birnbaum how insurance regulators can assess whether insurers are using prohibited data without also collecting this data. Mr. Birnbaum said it is not necessary to ask insurers to collect data on race and income, and he suggested census track data could be used as proxy for race and income.

3.  Heard a Presentation from TransUnion on How Insurers Can Use Its Products to Better Rate and Underwrite Insurance Risks

Mr. Reynolds said TransUnion has 650 million consumer records and 450 million consumer files that are monitored daily. He said big data helps organizations make better decisions and provides the most complete and multidimensional information available about consumers. He said insurance companies are using more data to more accurately price and predict risk, increase objectivity in insurance underwriting, achieve increased scalability of rating systems, expand market availability, and reduce fraud. Insurance companies are using expanded credit data to help address gaps in traditional use of credit for underwriting. In addition, Mr. Reynolds said big data is being used as alternative data to supplemental credit data and improve pricing accuracy. Insurance companies use nontraditional/behavioral data sources, such as telematics, to improve risk evaluation and change insured behavior. Mr. Reynolds said insurance companies use data and analytics to prevent fraud prior to the issuance of a policy and during the claims process. Mr. Reynolds referenced a new vehicle history score, which assesses the risk of the vehicle independently of the driver.

Mr. Reynolds said policymakers should encourage the use of data from disparate sources since there is a net benefit to consumers. He said there should also be sufficient flexibility in the data reporting framework to allow for exceptions created by consumer needs to mitigate adverse impacts on consumers in special conditions. Mr. Reynolds said that service providers and insurers should also notify their customers before the inception of increased data use and that there are well-established laws that provide consumer rights covering access and dispute resolution.

Mr. Vigliaturo asked what protections are in place to protect insurers from using data that is prohibited by a state's Unfair Trade Practices Act. Mr Reynolds said TransUnion provides data to insurers but does not necessarily have knowledge of how insurers use certain data within their rating models. In response to a question from Director Dowling, Mr. Reynolds said that cybersecurity is a top priority for TransUnion and that TransUnion does not use any social media data because of its limited predictability. Ms. Nelson said she has concerns about the inability of both insurers and vendors to explain their rating models and the underlying data used in the models. Ms. Nelson said insurers often direct regulators to the third-party data vendors, and the vendors then direct regulators back to the insurer. In response to a question from Director Ramge, Mr. Reynolds said consumers can access data being used and that all data TransUnion supplies to insurers is subject to the Fair Credit Reporting Act.

Mr. McKenney said he appreciates TransUnion looking at 24-month history of credit data rather than a snapshot. Mr. McKenney asked about whether TransUnion is filing as an advisory organization. Mr. Reynolds said TransUnion is looking at creating a separate legal entity for filings with the states. The data filed with states would be available, but not all 650 million records of TransUnion would be accessible. Mr. McKenney said he is concerned with the rating of a vehicle because a person with lower economic means may purchase a damaged vehicle. Mr. Reynolds said TransUnion rates by make and model of vehicle, which increases risk segmentation.

Mr. Birnbaum said that insurance is based on risk pooling and that continued risk segmentation will not expand the market. Mr. Birnbaum said that there are also a very large number of data vendors and that insurance regulators will not be able to examine all of them. Peter Kochenburger (University of Connecticut) asked whether data unrelated to a claim is used in the claims settlement process. Mr. Reynolds said TransUnion looks for relationships between parties to eliminate fraud, such as a trying to identify whether two parties involved in an accident may have lived at the same residence in prior years.

4.  Discussed Potential Workstreams

Commissioner Cali suggested including the definition of "big data" in the future efforts of the Big Data (D) Working Group and moving toward a discussion of large data sets and sophisticated models rather than focusing on defining big data. Mr. McKenney suggested the Working Group should focus on the use of data to segment the market. Director Dowling said the Working Group should review the use of data for claim optimization. Mr. Samsom said insurance regulators have an obligation to know what data is being used, how it is being used and how its use affects consumers. Mr. Bradner said regulators may need to identify what data should not be used.

Commissioner Cali suggested three possible workstreams: 1) discussion of the regulatory framework focusing on how the regulatory framework might need to be adjusted to address how data is being used; 2) discussion of how regulators can know what data is being used in complex rating models and what expertise is needed to review these models; and 3) discussion of how regulators can better use available data and enhance their collaboration.

Mr. Laucher said the Big Data (D) Working Group has concerns about additional risk segmentation and the potential for large disparities between the lowest rated risk and the highest rated risk. Mr. Laucher said this may lead individuals to be

charged rates that are beyond their capacity to pay. Commissioner Cali said this may fall within the discussion about the regulatory framework.

Superintendent Dwyer said consumers need to know what data is being used and how it is being used so they know how they might change behavior for insurance purposes. Mr. Feldvebel said he is interested in discussing the disparity of knowledge between consumers and insurers and the need for greater market transparency. Mr. McKenney said it is difficult for regulators to review insurers' rating plans, which are not filed and change every of couple of years. Mr. McKenney said the organizations providing data to insurers meet the definition of advisory organizations and suggested regulators coordinate their efforts in reviewing the activities of advisory organizations.

Mr. Laucher said there is a need for more centralization of the review of rating models. As part of this review, Commissioner Kreidler said regulators need to know what data is used in the models and be able to explain this to consumers to ensure consumer confidence. Ms. Murphy said regulators said it might be appropriate to conduct a coordinated data call to obtain this information.

Commissioner Cali said the Big Data (D) Working Group will need to determine how to partner with other National Association of Insurance Commissioners (NAIC) task forces and working groups since these issues involve multiple lines of insurance. Commissioner Cali said she would like the Working Group to discuss regulator use of data and requested NAIC staff to develop draft charges for the Working Group to review. Commissioner Cali said the Working Group should discuss: 1) whether the current regulatory framework is sufficient; 2) how regulators review models; and 3) whether regulators have enough detail to communicate to the public on what data is being used by insurers and how it is being used.

Mr. Birnbaum said big data has an opportunity to benefit consumers and suggested the following charges: 1) develop a template for regulators to collect data sources and uses of data from insurers; 2) in coordination with the Casualty Actuarial and Statistical (C) Task Force and Market Analysis Procedures (D) Working Group, develop a proposal for NAIC resources to assist in the collection and review of data and develop a proposal for the collection of more granular data from insurers; 3) address insurers' responsibility and consumer protection regarding the use of big data; and 4) discuss ownership of consumer-generated data related to insurance.

Dave Snyder (Property Casualty Insurance Association—PCI) said the real world is different from what has been described today. Mr. Snyder said insurance markets are competitive and innovative with few complaints and small residual markets. Mr. Snyder said the PCI wants to make sure the efforts of the Big Data (D) Working Group are responsive to what is happening in the real world and wants to assist regulators with this effort. He said the Working Group should encourage innovation and competition, as well as protect the intellectual property of insurers. He said he does not believe a new legal framework is needed and that the current framework of prohibiting insurance rates from being excessive, inadequate or unfairly discriminatory has worked well for consumers and regulators.

Mark Smith (Insurance Services Office—ISO) said millennials are causing a shift in consumer culture, which desires mobile access to insurance products that are less expensive and innovative.

Commissioner Cali said the Big Data (D) Working Group would discuss TransUnion/CARFAX, which has previously been discussed by the Auto Insurance (C/D) Working Group. Commissioner Cali said the Auto Insurance (C/D) Working Group is concluding its work on this issue and is suggesting this issue falls within the scope of the Big Data (D) Working Group.

Having no further business, the Big Data (D) Working Group adjourned.

W:\National Meetings\2016\Summer\Cmte\D\BigData\BigDataWG Minutes 082616.docx

Draft: 10/05/16

**DRAFT CHARGES**
**RECOMMENDATION OF BIG DATA (D) WORKING GROUP**


A.  Consider and recommend improvements, as appropriate, to the regulatory framework regarding the use of consumer and non-insurance data. At a minimum, consider modifications to model laws/regulations, regulation of data vendors and brokers, regulatory reporting requirements, and consumer disclosure requirements.

B.  Propose a process for regulatory review of complex data models used for underwriting, rating, and claims. Create a means to share resources and coordinate review of these models through the NAIC.

C.  Establish data needs and required tools for regulators to appropriately monitor the marketplace and evaluate underwriting, rating, claims, and marketing practices. Include inventory of currently available data and tools, as well as additional data and tools needed, as appropriate. Propose a means to collect, house, and analyze needed data.


W:\National Meetings\2016\Fall\Cmte\D\Big Data\Big Data Draft Charges 100516.docx

**Comments of the Center for Economic Justice to the**

**NAIC Big Data Working Group**

**August 19, 2016**

The Center for Economic Justice (CEJ) submits the following recommendations to the Big Data Working Group for the 2017 proposed charges called for in the Working Group's 2016 charge.

Regulatory Big Data and Regulatory Resources

1. Develop a template for states to use to request from insurers the sources and uses of data for various insurance functions by the 2017 Spring National Meeting. The template shall provide for reporting by the insurer of the name and description of each source of data, the source of the data and the use or uses of the data including pricing (including underwriting), marketing, claims settlement, antifraud and other. Develop a proposal for NAIC assistance to the states in the collection and compilation of the requested information reported by insurers.

2. In coordination with the Casualty Actuarial and Statistical Task Force and the Market Analysis Procedures Working Group, develop a proposal for an NAIC Resource to assist states in the collection and analysis of granular data for market analysis and in the analysis of complex pricing and claim evaluation/settlement models by the 2017 Summer National Meeting. The proposal shall be limited to the NAIC Resource providing data management and analysis to states and shall not provide any role for the NAIC in states' regulatory opinions.

3. In coordination with the Casualty Actuarial and Statistical Task Force and the Market Analysis Procedures Working Group, develop a proposal for reporting by insurers of granular data on market outcomes by the 2017 Fall National Meeting. The proposal shall be limited to personal lines property and automobile insurance and shall include proposed data elements and a proposed reporting mechanism.

<u>Insurer Responsibility and Consumer Protections for Insurers' Use of Big Data</u>

4. Develop a model law describing insurer responsibilities and consumer rights regarding insurers' use of big data by the 2018 Summer National Meeting: The model law shall include, at a minimum, requirements, for:
   a. insurers' submission of big data sources and uses to regulators,
   b. insurers' submission of big data algorithms and related insurer data to regulators,
   c. protection of personal consumer information and insurer trade secrets in the big data-related submissions to regulators;
   d. insurers' disclosure and consumer protections related to insurers' big data usage, modeled after the disclosure and consumer protection requirements for users of information provided by consumer reporting agencies pursuant to the Fair Credit Reporting Act including, but not limited to:
      i. disclosure by the insurer to the consumer of intended use of the data;
      ii. consent by the consumer to the insurer's use of the data
      iii. clear and specific disclosure by the insurer to the consumer of potential outcomes if the consumer declines permission to use the data;
      iv. adverse action notice by the insurer to the consumer if the use of the data results in an adverse action
      v. disclosure to the consumer of the actual consumer's data used for inspection by the consumer;
      vi. ability of the consumer to correct erroneous information;
      vii. reassessment of the data by the insurer if data are corrected; and
      viii. a provision for life events exceptions to use of the data if the life event may result in unfavorable data due to events outside of the consumer's control.

<u>Ownership and Consumer Protections for Consumer-Generated Data Related to Insurance</u>

5. Develop a model law regarding ownership and licensing of data generated by consumers, vehicles or properties and used by insurers, including, but not limited to, data generated by vehicle telematics, home or property telematics, mobile phones and wearable devices by the 2017 Fall National Meeting. The model law shall declare that such data is owned by the consumer and shall include provisions for licensing of such data by consumers to insurers in a manner that balances consumer protection and sovereignty with insurer and public needs for access to such data loss prevention and mitigation and innovation in insurance products and insurer-consumer interaction. Such licensing provisions shall include, but not be limited to:

a. consumer-tested disclosure to consumers of the sources and uses of the consumer-generated data;

b. affirmative agreements – opt-in – by the consumer for collection of the data and for each intended use of the data;

c. prompt access to the consumer-generated data by the consumer upon request with provision of the data by the insurer to the consumer in a format and medium which, among other things, allows consumers to take the data to another insurer;

d. prohibition against use of the consumer-generated data by the insurer for any purpose other than those agreed to by the consumer;

e. prohibition against sale of the consumer-generated data by the insurer to any third-party without affirmative consent by the consumer and compensation to the consumer by the insurer; and

f. standards for all-industry databases of specific types of consumer-generated data.