



**Reimagining Market Analysis and Market Regulation
Through Enhanced Data Collection and Analytics**

**Presentation to NAIC Market Regulation and
Consumer Affairs (D) Committee**

July 27, 2021

Birny Birnbaum
Center for Economic Justice

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

About Birny Birnbaum

Birny Birnbaum is the Director of the Center for Economic Justice, a non-profit organization whose mission is to advocate on behalf of low-income consumers on issues of availability, affordability, accessibility of basic goods and services, such as utilities, credit and insurance.

Birny, an economist and former insurance regulator, has worked on racial justice issues for 30 years. He performed the first insurance redlining studies in Texas in 1991 and since then has conducted numerous studies and analyses of racial bias in insurance for consumer and public organizations. He has served for many years as a designated Consumer Representative at the National Association of Insurance Commissioners and is a member of the U.S. Department of Treasury's Federal Advisory Committee on Insurance, where he co-chairs the subcommittee on insurance availability. Birny is also a member of the U.S. Federal Reserve Board's Insurance Policy Advisory Committee.

Birny served as Associate Commissioner for Policy and Research and the Chief Economist at the Texas Department of Insurance. At the Department, Birny developed and implemented a robust data collection program for market monitoring and surveillance.

Birny was educated at Bowdoin College and the Massachusetts Institute of Technology. He holds Master's Degrees from MIT in Management and in Urban Planning with concentrations in finance and applied economics. He holds the AMCM certification.

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.

What is Transaction Data? What is Summary Data?

Transaction Data – separate data records for each transaction.

In insurance, transaction data means separate records for each quote/sales transaction and separate records for each claims transaction.

Insurers maintain sales and claims transaction databases to:

- Track interactions with the consumer; and
- Employ their analytics on transaction data sets.

With transaction data, new data elements can be added to the data record without disturbing the existing data elements.

With transaction data, the insurer can associate great detail – e.g. many characteristics and outcomes – to individual consumer records.

Summary Data – experience grouped into pre-defined categories, e.g. MCAS

Benefits of Transaction Data

More Effective for Market Analysis and Market Regulation

- More granular analysis / Discover, not confirm / Multi-variate
- More timely analysis – MCAS data are stale, compare to mortgage, workers comp¹
- Ability to employ predictive analytics / AI / Machine Learning
- Ability to identify proxy discrimination and disparate impact
- More consistent – compare STLDI and Travel MCAS Blanks
- More efficient and effective – compare STLDI transaction data collection with final summary data blank.

¹ See “Minnesota’s workers’ compensation response to COVID-19” at https://www.dli.mn.gov/sites/default/files/pdf/MN_work_comp_response_to_COVID-19.pdf and California Workers Compensation Institute Interactive App at <https://www.cwci.org/CV19claims.html>

Benefits of Transaction Data

More Efficient for Insurers and Regulators

Less Costly for Insurers

- Simpler Data Reports / Consistent with Other Insurer Uses
- Less Costly to Revise, Add Data Elements
- Dramatically Reduce Special Data Calls
- Dramatically Reduce Regulatory Inquiries

Less Costly for Regulators

- More Reliable Data / Less Time on Data Reporting Issues
- More Refined Market Analysis / Better Targeting of Regulatory Resources
- Fewer Special Data Calls
- Utilize Existing Statistical Agent Framework Instead of New Data Infrastructure

MCAS: 20th Century Technology Applied to 21st Century Industry Practices

Industry Data Sources and Uses circa 2004 versus Today

2004:

- Third-party data – predominantly consumer credit information
- Growing adoption of early predictive models using multiple regression of historical data
- Growing use of generalized linear models.

Today:

- Third-Party Data – dozens of non-insurance sources;
- Enhanced, real-time consumer insurance data;
- Real-time data access and decision-making;
- Micro-segmentation for marketing, pricing, claims settlement, anti-fraud
- Advanced data analytics -- data mining, generalized additive models, neural networks, machine learning, AI

MCAS: 20th Century Technology Applied to 21st Century Industry Practices

Market Regulator Data Sources and Uses circa 2004 versus Today

2004:

- Complaints
- Enforcement Actions
- News Reports
- Lawsuits
- MCAS for personal auto, homeowners, life and annuity

Today:

- Complaints
- Enforcement Actions
- News Reports
- Lawsuits
- MCAS for some additional lines

Contrast the absence of change in market regulation data collection with the significant increases in financial regulation data collection since 2004.

Transaction Data Collection on Consumer Market Outcomes is the Single Most Important Action to Fulfill the Promise of Market Analysis

Around 2000, the NAIC embarked on a “reinvention” of market regulation that promised to move away from the auditing model of market regulation to a more data-driven, analytic approach – market analysis. That promise has not been fulfilled. MCAS has become just another piece of the auditing approach.

There Has Been Transaction Data Reporting for Decades in Insurance and Other Financial Service Industries

Transaction data already reported in some insurance lines:

- Workers Comp – at least 1980s
- TX – since at least 1996 – monthly for homeowners, quarterly for personal auto and commercial lines
- ISO- and AAIS-reporting companies
- NAIC PBR for life insurance and, eventually, annuities and long-term care

Standard in other financial service regulators:

- CFPB and other members of FFIEC – HMDA
- OCC – Mortgage Metrics Report -- monthly
- SEC – Monitoring Securities Trading

Resources and Collectibles

1. CEJ July 21, 2021 Presentation to NAIC MIS R&D WG
2. CEJ 2011 Presentation at NAIC E-Reg Conference, “Data Mining and Predictive Analytics for Insurance Regulators: The Promise of Market Analysis”
3. CEJ 2008 Presentation to NAIC Consumer Liaison Committee, Centralized MCAS Data Collection
4. CEJ 2005 Presentation to NAIC E-Reg Conference, “Data Mining for Insurance Regulatory Market Analysis”
5. CEJ 2004 Proposal for Centralized Data Collection



**Opportunities for Artificial Intelligence Applications in
Insurance Regulatory Market Analysis**

**Presentation to NAIC Market Information Systems
Research & Development Working Group**

July 21, 2021

Birny Birnbaum
Center for Economic Justice

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

About Birny Birnbaum

Birny Birnbaum is the Director of the Center for Economic Justice, a non-profit organization whose mission is to advocate on behalf of low-income consumers on issues of availability, affordability, accessibility of basic goods and services, such as utilities, credit and insurance.

Birny, an economist and former insurance regulator, has worked on racial justice issues for 30 years. He performed the first insurance redlining studies in Texas in 1991 and since then has conducted numerous studies and analyses of racial bias in insurance for consumer and public organizations. He has served for many years as a designated Consumer Representative at the National Association of Insurance Commissioners and is a member of the U.S. Department of Treasury's Federal Advisory Committee on Insurance, where he co-chairs the subcommittee on insurance availability. Birny is also a member of the U.S. Federal Reserve Board's Insurance Policy Advisory Committee.

Birny served as Associate Commissioner for Policy and Research and the Chief Economist at the Texas Department of Insurance. At the Department, Birny developed and implemented a robust data collection program for market monitoring and surveillance.

Birny was educated at Bowdoin College and the Massachusetts Institute of Technology. He holds Master's Degrees from MIT in Management and in Urban Planning with concentrations in finance and applied economics. He holds the AMCM certification.

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.

Artificial Intelligence: Let's Get More Precise on Terms

AI vs. Data Mining vs. Machine Learning vs. Deep Learning

How Does AI work?

Simple – Automate human-designed rules – convert a manual premium calculation to an automated premium calculation – think of a spreadsheet.

More Complex – Implement human-designed rules to amounts of information difficult for humans to deal with – tools to identify quantitative outcomes – data mining and big data analytics.

Yet More Complex – Tools to make decisions based on desired outcomes or tools to identify qualitative outcomes – machine learning

Yet Even More Complex – Tools that teach themselves how to identify desired or qualitative outcomes – deep learning

Key distinction – moving from implementation of human-designed rules to algorithms that learn and change on their own.

Let's look at potential applications of AI to review of policy forms

Simple – tools to compare policy forms and show changes, tools to search for and identify key specific words.

More Complex – tools that apply human-designed tests to policy forms – Flesch tests / readability, identify specific wording / phrasing of concern.

Yet More Complex – tools that identify potentially misleading or deceptive terms and provisions.

Yet Even More Complex – tools that learn and change to better achieve regulatory goals.

Let's think about how these tools are possible and potential pitfalls

1. **Granular data and lots of it.** In the case of policy forms, you have thousands of individual words within a single policy form and thousands of policy forms from which to learn. The data you are analyzing – words, combinations or words and location of words – allows for granular analysis. You also have a rich database of accepted and rejected forms – you not only have the raw data, but you have outcome data.
2. **If the training data have errors or biases, the AI application will learn and perpetuate those errors and biases, unless steps are taken to identify and eliminate those biases.** Similarly, if the human-designed rules or defined qualitative and quantitative outcomes reflect human bias, the AI application will reflect and perpetuate those biases. While some bias may be offensive – intentional or unintentional racial bias – don't consider bias to be sinister – think about confirmation bias – where information and outcomes that support one's expectations are given great weight while information and outcomes that contradict one's expectations are given less weight by the analyst. One of the great benefits of data mining and AI is the ability to identify and mitigate biases in data and modelers.

3. Greater ability to analyze impacts on protected classes. AI has raised awareness of the potential for proxy discrimination and disparate impact in insurance. For those of you who attended the NAIC screening of *Coded Bias*, you learned how and why facial recognition technologies are biased against African Americans. You may have also read about Google’s attempt to use AI to develop an unbiased hiring algorithm – and the algorithm failed to hire any women because the training data reflected the historical bias in hiring white men. But, AI offers opportunities for improved identification of such biases.

But the other lesson is that you have to evaluate outcomes produced by AI to ensure the AI application is producing the intended results. Auditing an algorithm on the front-end is not sufficient – you must collect and analyze outcomes produced by the AI application.

With that lengthy introduction, let's look at opportunities for using AI in market analysis. The foundation of our inquiry are the following two points:

1. What outcomes are you trying to achieve?
2. What data are available / what data are needed?

AI for Market Analysis – What outcomes and what data?

Any AI application requires a clear articulation of the outcomes you are trying to achieve. Referring back to our policy form review example – the outcomes were identifying prohibited words, phrases and provisions and identifying misleading and deceptive provisions – the outcomes that would indicate a violation of statute or regulation.

The data available are not just thousands of policy forms but all the words in those policy forms as well as historical outcome data – accepted and rejected words and phrases. The historical data may go back many years, but recent data are also available.

AI for Market Analysis – What outcomes and what data?

When insurers utilize AI for, say, customer service, let's assume the intended outcomes are reduced insurer costs, higher return on investment, improved customer satisfaction.

They might accomplish this with AI to reduce hold times on calls, reducing insurers' human interaction (by using a chat bot) or using AI to match the most appropriate customer service agent with a particular customer in real time.

The data available to the insurer are thousands of actual historical consumer interactions – which consumer with which service representative, specific issue of concern, time of day, hold time, length of call, success of initial menu selection, use of chat bot, success of chat bot, issue resolved or not, pushed to supervisor, customer satisfaction and more.

With these data, the insurer can use AI to develop tools to analyze the customer and issue in real time and route to the most suitable customer service representative and allow the tools to continue to learn and modify to achieve the intended outcomes.

Clearly-Identified Outcomes and Timely, Granular Data

Insurers' AI applications for marketing, pricing, claims settlement and anti-fraud rely on access to a variety of data sources accessible in real time to make decisions in real time based on specific customer or claim characteristics. Lots of data, accessed and analyzed in real time.

Let's turn to AI for Market Analysis – First, what are the outcomes we want to achieve with AI for market analysis?

- Speed up current market analysis processes?
- Eliminate false positives in current market analysis processes?

Or, re-imagine and re-engineer market analysis for more timely ways to:

- Identify individual insurer violations of laws or regulations?
- Identify individual insurer violations of laws and regulations that reflect a business practice?
- Identify industry trends that may violate laws or regulations or that may require new laws and regulations for consumer protection?
- Identify legal consumer outcomes that reflect consumer confusion or legal but unfair or deceptive practices?
- Identify legal or possibly illegal practices that unfairly disadvantage consumers?
- Identify anti-competitive practices, markets or market segments?

- Identify insurers with poor consumer outcomes? Industry segments with poor consumer outcomes?
- Identify producers with poor consumer outcomes?
- Identify opportunities / needs for regulatory engagement with consumers – improve complaint filing opportunities and process, improve consumer education and disclosures, improve DOI websites, target DOI media and social media activities?
- More timely identify problematic practices and poor consumer outcomes?
- What are some others – what do you want AI in market analysis to accomplish?

Data Needs for AI in Market Analysis

Current data available for market analysis:

- Complaints
- MCAS
- Exams and other Continuum Actions
- Financial statements
- Insurer SEC filings
- Insurer investor presentations
- Insurer form and rate filings
- Insurer press releases and news reports
- Lawsuits

Quite a bit of data and information, but two major problems for using these to develop AI for market analysis:

- Difficult to combine data for AI due to different time frames, different types / sources of data and lack of common analytic unit.
- No data for actual, granular consumer outcomes. MCAS woefully inadequate for market analysis, generally, and for implementing an AI tool for market analysis, specifically. Moreover, MCAS requires prior identification of potential problems which the summary data are intended to identify – so confirmation bias has an outsized impact on the selection of data elements and breakouts to be included.

AI for market analysis requires more timely and granular consumer outcome data – which means monthly or quarterly transaction data reporting. The single most important action necessary to develop AI for Market Analysis is to re-engineer current statistical agent reporting to a transaction reporting regime. Whatever market analysis outcomes for which you want assistance from an AI tool, you need the raw material of granular and timely data of consumer market outcomes – quote and sales transactions and claims transactions.

Let me offer some examples of the role of granular data in market analysis.

Mortgage forbearance as a result of the CARES Act – how much assistance was provided to whom and where?

Urban Institute – pdf

Black Knight -- pdf

Covid-related claims in Worker's Compensation – timely identification of impacts, ease of revision to data reporting.

CA -- pdf

MN -- pdf

WC stat plan revisions – existing data fields (“pandemic”) or add data field (“Covid-19”) as opposed to redesigning a summary reporting form to create new reporting categories and groupings.

Contrast WC with PPA – What type of market analysis was possible in April, May or June 2020 to

- Determine if insurers were complying with emergency orders?
- Determine if rates had become excessive or unfairly discriminatory?
- Determine if premium relief was sufficient to avoid excessive rates?
- Determine if premium relief was being provided fairly to consumers across insurers?

Unlike WC, PPA market analysts had no tools to perform timely market analysis because the available data didn't allow analysis of

- How did claim frequencies and claim severities change by week or month?
- How did new claims reported by week starting on March 1 compare with the same period in prior years?
- What was the volume of and who received premium payment grace periods?
- What was the volume of and who received premium relief and when was it delivered?
- What was timing on claim resolution, how did that compare to prior periods and how did the timing change over the course of 2020?

Resources:

FINRA Report on AI in the Securities Industry, June 2020:

<https://www.finra.org/sites/default/files/2020-06/ai-report-061020.pdf>

Speech by SEC's Scott Bauguess, June 2017, "The Role of Big Data, Machine Learning, and AI in Assessing Risks: a Regulatory Perspective:

<https://www.sec.gov/news/speech/bauguess-big-data-ai>

"SEC's advanced data analytics help detect even the smallest illicit market activity," June 30, 2017, Reuters: <https://www.reuters.com/article/bc-finreg-data-analytics/secs-advanced-data-analytics-helps-detect-even-the-smallest-illicit-market-activity-idUSKBN19L28C>

I'll conclude by re-stating the two major tasks necessary to introduce AI into insurance market analysis:

- Identify the outcomes you want the AI application to identify; and
- Get the raw material necessary to implement meaningful AI by developing more timely and granular insurer reporting of consumer market outcomes by re-engineering the current statistical agent data reporting to monthly or quarterly transaction data reporting.

NAIC E-Reg Conference 2011

Data Mining and Predictive Analytics for Insurance Regulators:
The Promise of Market Analysis

Birny Birnbaum
Center for Economic Justice

About Birny Birnbaum and the Center for Economic Justice

The Center for Economic Justice is a non-profit organization advocating on behalf of consumers for fair access to and fair treatment for basic services including insurance, credit and utility products.

Birny Birnbaum, CEJ's Executive Director, has spent the past 20 years working on insurance regulatory issues as a regulator, while Chief Economist and Associate Commissioner for Policy and Research at the Texas Department of Insurance, as a consulting economist to consumer and public agencies and as an advocate for consumers before administrative agencies and legislative bodies.

Among other responsibilities, Birny was responsible for the development of detailed statistical plans for market surveillance in Texas.

Data Mining, Predictive Analytics and Business Intelligence

Used by Insurers for

- Marketing and Sales
- Underwriting, Tier Placement, Rating
- Claims Settlement and Anti-Fraud Efforts

Example of Insurer Use of Data Mining and Predictive Analytics: Assurant Sales of Payment Protection

Sloan Management Review, Winter 2011

<http://sloanreview.mit.edu/the-magazine/2011-winter/52206/matchmaking-with-math-how-analytics-beats-intuition-to-win-customers/>

Payment Protection: Debt Cancellation Contract and/or Debt Suspension Agreements are Banking Product Version of Consumer Credit Insurance

Assurant Solutions manages product activation, claims, underwriting and customer retention for its bank customers.

Assurant Case Study (con't)

Problem: Increase retention rate for customers calling to cancel from industry standard of 16%

Solution: “Deep Analytics” – Brought in mathematicians (people who did not know anything about running call centers) to analyze mountains of data – details about every phone call received by the call center – every transaction recorded over four or five year period.

Result: Found that improving the customer operating experience – e.g., reduce call waiting time will improve customer satisfaction and improved satisfaction will result in higher retention rates – was wrong.

Assurant Case Study (con't)

Discovery: Some customer service reps much better dealing with certain types of customers. By mining mountains of individual transactions, learned to match specific in-calling customers with specific customer service reps – even if waiting time was longer.

Impact of Application of Predictive Analytics: Retention Rate doubled, but Saved Fee Rate Tripled as more high-value (higher fee) customers retained.

Assurant Case Study (con't)

Why Don't Other Organizations Do This?

1. We don't have the IT resources to do this now.
2. We already do skills-based routing

Assurant insight: Not skills-based routing, but evidence-based or success-based routing.

National Underwriter, June 7, 2010

“Top 10 Technologies Seen Most Likely to Have Impact on P&C Insurance”

Business Intelligence and Analytics:

Property and casualty insurers are increasing their focus in 2010 on improving how data is used and the value derived from past investments in data warehousing. Expanded use of data mining and analytics will assist p&c insurers in obtaining more granular insight into customers; greater operational performance; better claims performance; and reduced losses, leakage and underwriting risks.

The outcome will be improved risk management, compliance/reporting and customer intelligence that can be used for personalization and decisions regarding channels and products as well as provide insight into performance and process problems.

National Underwriter, June 7, 2010

“Top 10 Technologies Seen Most Likely to Have Impact on P&C Insurance”

Predictive Modeling Tools:

Traditional data analysis processes are typically historical (analysis and reporting of the past) or current day (analysis and reporting of current day position). Using predictive modeling, insurers can analyze data and create models that will enable them to predict future behaviors or outcomes.

Using predictive modeling will allow companies to look into the future and make predictions based on historical data analysis, pattern recognition and modeling. This insight will help companies avoid risks, understand future customer behavior so that they can intervene as needed, offer proactive customer service, project losses due to a catastrophe and improve underwriting profitability.

National Underwriter, June 7, 2010

“Top 10 Technologies Seen Most Likely to Have Impact on P&C Insurance”

Advanced Fraud Detection Solutions:

It is critical that insurers reduce losses and leakage to retain profitability. Better control of fraud is essential in accomplishing these goals.

Advanced tools analyze data (structured and unstructured) to identify fraudulent claims in real time at point of data entry. This will assist insurers in reducing losses that result in driving up operational costs, and may result in companies having to increase premiums based on these losses.

Insurers' Use of Predictive Modeling High and Still Growing

“Nearly 90 percent of U.S. insurance companies said the use of predictive modeling enhanced rate accuracy in 2010, according to the results of a survey conducted by Towers Watson.

Property Casualty 360.Com, February 1, 2011

Primer on Predictive Analytics:

“Predictive Analytics White Paper”

Charles Nyce

American Institute for CPCU/Insurance Institute of America

<http://www.aicpcu.org/doc/predictivemodelingwhitepaper.pdf>

Data Mining and Predictive Analytics for Market Analysis

The new paradigm for market regulation is to move away from comprehensive market conduct examinations as the fundamental – and largely only – tool for data collection and market regulation enforcement – to a system grounded in market analysis employing a variety of market regulation tools focused on particular issues and companies stemming guided by market analysis.

As market analysis is the foundation of more focused, efficient and effective market regulation, the availability of relevant data and information is the prerequisite for market analysis.

Data Mining and Predictive Analytics for Market Analysis (con't)

Summary vs. Transaction Data

What is Summary Data? Current MCAS

What is Transaction Data? Texas, ISO Statistical Plans

Why Can't Summary Data Be Used for Data Mining and Predictive Analytics?

- Can only answer pre-determined questions
- Cannot discover new relationships – no multivariate analysis, no ability to analyze several characteristics simultaneously

Limitations of Summary Data: MCAS Examples

- Life: Replacements by Specific Product and Policy Size, Death Claims Denied by Specific Product and Policy Size
- Annuity: Replacements by Specific Product and Policy Size
- PPA/HO: Cancellations at 59 Days by Product Type and Neighborhood; Claims Closed Without Payment or Settled after 180 days by Type of Claim and Neighborhood
- All: No ability to model outcomes

Limitations of Summary Data: MCAS Examples

- Data Quality – Limited Ability for Auditing, Reasonability Testing
- Data Quality – Comparisons Across Reporting Companies Greatly Limited by Differing Data Definitions
- Evaluating Results – Giving Answers to the Test Ahead of Time

Capabilities of Transaction Data

- High Data Quality, Great Ability for Data Auditing
- Differing Company Data Definitions a Non-Event
- Analyze Multiple Characteristics Simultaneously
- Add External Data
- Answer Unlimited Questions / Questions Not Previously Anticipated
- Efficient for Changes

Barriers to Data Mining for Market Analysis

- Regulators – Time, Resources, Expertise

How do financial regulators analyze mountains of financial statement data?

- Industry – Contradictory Positions

Effective Market Analysis means reduced regulatory costs for the vast majority of insurers: Regulators employing focused tools on bad actors and reducing comprehensive exam – what insurers have asked for. Yet insurers fight the necessary expanded data collection necessary for market analysis.

Addressing Barriers to Data Mining for Market Analysis

- Ideal: Centralized Collection of Transaction Data with Associated Data Mining and Predictive Analytics Expertise
- Next Best: State-Specific Initiatives

Data Mining and Predictive Analytics for Market Analysis:

Regulatory Tools to:

- Enable Regulators to be Pro-Active
- Match Industry Practice
- Create More Effective and More Efficient Market Surveillance and Consumer Protection

For Further Information

Birny Birnbaum
Center for Economic Justice

www.cej-online.org

birny@flash.net

512 448 3096

Comments of the Center for Economic Justice

Proposal for Market Conduct Annual Statement Centralized Data Collection

Before the NAIC Consumer Liaison Committee

May 31, 2008

The Center for Economic Justice (CEJ) is one of many consumer organizations supporting the proposal from the NAIC Market Regulation (D) Committee for centralized collection of Market Conduct Annual Statement (MCAS) data.

In the strongest possible terms, CEJ urges the NAIC members to support the D Committee's fair and balanced recommendation for centralized data collection of MCAS and public availability of the currently-collected data elements. The NAIC's action on this proposal will tell insurance consumers if regulators are committed to modernizing insurance regulation in the public interest or if regulators are unable to take an action opposed by the industry you regulate.

Data Collection is Foundation for Market Analysis – The New Paradigm for Market Regulation

Insurance market regulation -- monitoring marketplace practices of insurers and ensuring fair and legal treatment of consumers in marketing, sales, underwriting, rating, and claims settlement -- must be dramatically improved. Insurance regulators have historically been among the last to recognize serious market problems -- life insurance replacements, churning, financed single premium credit insurance, and many other instances.

The modernization of insurance market regulation is based on vastly improved market analysis -- the ability of regulators and the public to monitor the marketplace performance of insurers. And the foundation of market analysis is data: market analysis assumes the availability of robust and comprehensive data.

While consumers want market regulation to become vastly more effective, we agree with insurers that market regulation must become more efficient and that the paradigm of market regulation must change from principally comprehensive market conduct exams to focused regulatory actions informed by market analysis and targeted at the worst problems and worst players. Improved effectiveness and efficiency depends on robust market analysis data collection.

The proposal before you is specific to the data elements currently being collected. This is a very modest number of data elements. At the point where the MCAS grows in scope to hundreds of data elements, we recommend a competitive bidding process of which the NAIC and industry statistical agents could compete for the data collection responsibilities. But, at this point in time and with this limited set of data elements, the proposal for an annual statement supplement is the reasonable and logical approach.

Centralized Data Collection of MCAS Has Been Under Discussion for Several Years

The proposal before the EX Committee has been under discussion for five years. It has been clear that centralized data collection is necessary for at least three things. One, improved data quality and analysis. With centralized data, there is much greater opportunity for data quality review and consistency across states. Two, centralized data is far more efficient for insurers in the same way that the interstate compact is more efficient for insurers' product filings. Three, centralized data is more efficient for regulators and allows states with limited resources to participate in the market conduct data collection -- and, subsequently, the market analysis -- program.

As evidence that centralized data collection has been under discussion for several years, we point to a proposal submitted by CEJ in December, 2004 for centralized MCAS data collection. Complaints by industry and some legislators that the NAIC has moved "hastily" on this project are simply inaccurate. There has been nothing unusual or speedy about the development and adoption of this proposal.

Industry argues that they support uniformity and consistent of regulation and support centralized data collection "in concept," but simply do not support this proposal. This claim rings hollow given that in the five years this issue has been before the NAIC, the industry has never provided any proposal for centralized data collection.

It is surely contradictory for industry to, on the one hand, tell regulators to modernize by becoming more uniform or risk a federal takeover of insurance regulation and on the other hand, tell the NAIC not to act on a proposal which accomplishes the very uniformity, consistency and efficiencies industry has called for.

Public Access to Data

CEJ supports the D Committee's recommendation that the currently-collected data elements be public information in the strongest terms. This is not to say that any or all additional data elements to be collected as part of the MCAS would be public. Rather, we are talking specifically about these data elements, which are clearly not trade secrets and whose public disclosure would provide useful information to consumers for evaluating insurers.

Industry has rolled out the same tired arguments against public disclosure of the current MCAS data used for every proposal for public information in recent years – medical malpractice data, climate change disclosures, geo-coded sales data. It is a knee-jerk reaction to any request for public disclosure of data and the arguments are always:

- "The data are trade secrets."
- "The data will be misunderstood and misinterpreted by the public."
- "The data will lead to massive litigation."

There is no evidence – absolutely none – in support of these claims. In fact, industry and some legislators have claimed the information is trade secrets without even knowing what data elements are proposed to be public information!

We ask regulators to look at the data elements, which are attached. Putting aside that much or all of this information is already available to the public in some form, what in these data elements qualifies as a trade secret? Any reasonable review of these data elements must conclude that the data are not trade secrets, that public disclosure will not harm consumers and that public disclosure will provide important information for consumers to better evaluate the market performance of insurers and contribute to a more competitive insurance market.

For example, in the EX Committee conference call, one industry party claimed that a competitor would look at the company's change in policies in force to determine whether the company was trying to expand its business in the state and that the competitor might respond to this information. Let us put aside that the competitor already has this information with the change in insurance premiums as reported on the state page of the Annual Statement. Rather, let consider whether a competitor would wait a year or more to respond to a competitive challenge – a competitive challenge that the company would know about in real time because their front-line agents would be reporting back whether consumers were coming to the agent or non-renewing in great numbers. The industry scenario of alleged harm is truly without merit.

Industry arguments about the alleged poor quality of the data or that such data would be misinterpreted or misused are without merit. Insurers can ensure accurate quality data with accurate reporting. And insurers should not be the ones who decide how data should be interpreted. The essence of open and accountable government is that government does not decide what is good for the public to know and what is not. Rather, public records law protects individual privacy and corporate trade secrets; it is not intended to censor what the public should know. Further, state regulators and NAIC staff have routinely sought comments from industry about improving data definitions to ensure better data quality. Centralized data collection will improved data quality! It is hypocritical for industry to argue that the data are poor quality and argue against an approach -- centralized collection -- which will improve data quality

Industry arguments that they were promised confidentiality are similarly without merit. We believe that regulators misused the market exam authority to collect and hold confidential MCAS data. Confidential treatment should be afforded to true trade secrets of insurers and market conduct exam authority should not be used to circumvent public records law.

We urge the regulators to focus on the data elements at question. The proposal is not for unlimited public release of all MCAS data to be collected at anytime in the future. The proposal is for this specific and limited set of data elements to be public information because the data elements are not trade secrets and do not warrant confidentiality.

Other Industry Complaints

The industry argument that the statutory annual statement is only for solvency regulation is incorrect. There are several exhibits of and supplements to the statutory annual statement which have no direct role in solvency regulation, including the state page, the insurance expense exhibit, the credit insurance experience exhibit and others. The annual statement is not only a data document; it is a framework for uniform and consistent data collection for a variety of regulatory purposes.

Industry has threatened to take legal action if the NAIC adopts this proposal. Let us suggest, that that the NAIC should welcome such a challenge – that the industry which calls for more uniformity and consistency of state insurance regulation is suing to stop state regulators from implementing exactly this type of action.

Industry has also incorrectly argued that public disclosure of MCAS will violate state laws concerning confidentiality of market conduct exam materials. This is incorrect and based on a faulty equation that market analysis is the same as a market conduct examination. MCAS is used for market analysis, which is a general review of the insurers and the market. There are many types of data used in market analysis, including consumer complaint data. By the industry logic, all data used for market analysis would be confidential – even complaint data which is currently public information. Simply stated, the industry arguments about conflicts with state laws are a red herring because market analysis is not a market conduct examination.

The industry complains that regulators made “commitments” to industry that MCAS data would be public. This argument fails because regulators do not have the right or the authority to promise confidentiality of information. Confidentiality of information is governed by public records law and the purpose of public records law is to make government accountable to the public and not to give government the right to decide what information is or is not good for the public to see. The MCAS was initially collected pursuant to market conduct examination authority as a convenience to regulators; it is inappropriate to use market conduct examination authority as a means to keep public information confidential.

Perhaps most troubling has been the industry argument that the MCAS has been a collaboration between industry and regulators and the current proposal departs from this collaborative approach. This sounds like the same type of threat industry used years ago to withhold funding from the NAIC unless state regulators eased up on pro-consumer initiatives.

The Bottom Line

The bottom line is that improved market regulation requires robust market analysis data collection. When the industry argues against centralized data collection and this proposal, industry reveals that it is not serious about improving the efficiency of market regulation because robust data collection is the prerequisite for focused market regulatory activities and lessened reliance on comprehensive market conduct exams.

If EX committee members take a step back, you will see that the current proposal is a modest one that helps move the new vision of market regulation along. Were the NAIC to reject this modest proposal, it will be difficult for consumer groups to defend state-based regulation as independent from the regulated insurance industry.

Data Mining for Insurance Regulatory Market Analysis

aka

Stump the Chump!



making progress . . . together

Birny Birnbaum
birny@flash.net
512 448 3096

What is Data Mining?

“A new discipline lying at the interface of statistics, data base technology, pattern recognition, and machine learning, and concerned with secondary analysis of large data bases in order to find previously unsuspected relationships, which are interest or value to their owners”

Hand, American Statistician, 1998

What is Data Mining?

The process of analyzing data to identify patterns or relationships.
(Iomega)

Searching large volumes of data looking for patterns that accurately predict behavior in customers and prospects.
(Adobe)

What is Data Mining?

A class of database applications that look for hidden patterns in a group of data. For example, data mining software can help retail companies find customers with common interests. The term is commonly misused to describe software that presents data in new ways. True data mining software doesn't just change the presentation, but actually discovers previously unknown relationships among the data. (Webpoedia.com)

What is Data Mining?

Data mining is derogatory. It means sorting through a huge volume of data, extracting decision rules that seem to favor one team over another, but without regard to whether or not there is any cause-and-effect relationship. Data mining is the sports-betting equivalent of sitting a huge number of monkeys down at keyboards, and then reporting on the monkeys who happened to type actual words. (Basketballfreesports.com)

Insurers' Data Mining

- Underwriting and Rating
 - Credit Scoring
 - Risk Segmentation
- Marketing
- Claims

Importance of Market Conduct Data Collection

- MC Annual Statement is a Landmark Project in Insurance Regulation
- Fantastic Work of Ohio and other Pilot State Departments

Insurance Regulatory Market Analysis Today

- Complaints
- Summary Financial Data
- Market Conduct Annual Statement
- Regulatory Databases

Financial Analysis vs. Market Analysis

Data Available:

Financial Analysis: Annual
Statement, Thousands of
Data Elements

Market Analysis: 23 Data
Elements and Complaints

Insurance Regulatory Market Analysis

Summary Data

vs.

Transaction Data

MC Annual Statement

Claim Activity

- Type of Claim Count Indicator – Occurrence or Claimant
- Claims open beginning of period
- Claims opened during period
- Claims closed during period with payment
- Claims closed during period without payment

MC Annual Statement

Time from Claim Date to Final Payment

- Median days to final payment
- Number of claims settled 0 to 30 days
- Number of claims settled 31 to 60 days
- Number of claims settled 61 to 90 days
- Number of claims settled 91 to 180 days
- Number of claims settled 181 to 365 days
- Number of claims settled beyond 365 days

MC Annual Statement

Claim Date of Accident to Date of Report

- Median days to date of report

Claims in Suit

- Suits open at beginning of period
- Suits closed during the period
- Suits open at close of period

MC Annual Statement

Underwriting

- Number of dwellings that have policies in force at the end of the period
- Number of policies in force at the end of the period
- Number of new business policies written during the period
- Dollar amount of direct premium written during the period
- Number of non-renewals during the period
- Number of cancellations that occur 60 days or more after effective date, excluding those for either non-pay or at the insured's request.
- Number of cancellations that occur in the first 59 days after effective date, excluding those for either non-pay or at insured's request.

Limitations of Current MC Annual Statement

Related to Summary Data

- Questions That Can Be Answered
- Data Quality Assurance
- Changing Data Elements Over Time
- Interaction Among Data Elements
- Univariate Analyses Only
- Ability to Add External, Non-Insurance Data
- Burden on Companies to Collect Data Not Already Collected

Limitations of Current MC Annual Statement

Related to Limited Scope of Data

- Analysis of Data by Groups of Consumers, e.g. sub-state territories
- Limited Use Over Time

Transaction Data Reporting

- **Plan Code**
- **NAIC Company Number**
- **MGA Indicator Accounting**
- **Date**
- **Record Type Code**
- **Transaction Identifier Code**
- **Type of Business Code**
- **Coverage Code**
- **Sub-Coverage Code**
- **Annual Statement Line of Business**
- **Territory Code**
- **Classification Code**
- **Five-Digit ZIP**

Transaction Data Reporting

- **Plus Four ZIP**
- **Certified Risk**
- **Prior Compliance with Texas Financial Responsibility Laws**
- **Principal Operator**
- **Date of Birth of Rated Driver**
- **Gender of the Rated Driver**
- **Marital Status of the Rated Driver**
- **Vehicle Use**
- **Miles to Work or School**
- **Estimated Annual Mileage**
- **Date First Insured with Company**
- **Driving Record Surcharges**
- **Number of Major Convictions**
- **Number of Minor Convictions**

Transaction Data Reporting

- **At-Fault Accidents**
- **Driver Training Discount**
- **Defensive Driver Discount**
- **Student Away at School Discount**
- **Anti-Lock Brake Discount**
- **Renewal Discount**
- **Two or More Autos Credit**
- **Anti-Theft Discount**
- **Passive Restraint Discount**
- **Drug/Alcohol Education Discount**
- **Good Student Discount**
- **Policy and Membership Fees**

Transaction Data Reporting

- **Model Year**
- **Symbol Group**
- **Vehicle Identification Number**
- **Policy Identifier**
- **Policy Limit Per Claimant**
- **Policy Limit Per Occurrence**
- **Deductible Amount**

Premium Specific Data Elements

- **Transaction Effective Date**
- **Transaction Expiration Date**
- **Direct Written Premium**

Transaction Data Reporting

Claim Specific Data Elements

- **Occurrence Date**
- **Catastrophe Code**
- **Type of Loss**
- **Claim Count**
- **Loss Amount**
- **Occurrence Identifier**
- **Claimant Identifier**
- **Accident State**

Transaction Data: ISO, Texas Stat Plans

- Answer Unlimited Number of Questions
- Answer Questions Not Previously Anticipated
- Multivariate Analysis
- Add External Data

Insurance Regulatory Market Analysis

- **Effective: Massive Increase in Regulatory Ability to Analyze Insurance Market and Insurer Market Performance**

Market Analysis Capabilities with Transaction Data

- Premium Calculation Audits
- Claim Settlement Patterns
- Impact of Rating and Underwriting Factors on Prohibited Classes

Insurance Regulatory Data Mining with Transaction Data

- Efficient for Insurers: One Data Report for Financial, Rate and Market regulation
- Efficient for Insurers: Eliminate Majority of Special Data Calls

Insurance Regulatory Data Mining with Transaction Data

- Efficient for Insurers and Regulators: Enable Regulators to Pinpoint Problems and Focus Attention on Problem Players

Insurance Regulatory Data Mining with Transaction Data

- Can Be Done Today
- Can be Done with Existing Data Collection Infrastructure

Insurance Regulatory Market Analysis

Data Mining Supplemented by:

- UW Guidelines
- Complaints
- Other External Data

Now, Stump The Chump!

Think of a Market Problem a Regulator Could Not Identify through the Transaction-Data-Mining-Supplemented-By-External-Data-Approach to Insurance Market Analysis?

Proposal for Centralized Data Collection and Baseline Analysis of Market Performance Data

Submitted by the Center for Economic Justice
December 1, 2004

Background

Over the past two years, several states have implemented a pilot program to collect data for market analysis. The Market Conduct Annual Statement pilot program has provided the participating states with very useful information to identify potential market problems and problem companies and to help focus further market conduct investigation and enforcement activities. The MC AS has proven the premise of market analysis – better market performance information and better analysis of that information by regulators can lead to more effective and more efficient market regulation.

The pilot program included a standard / uniform data request issued by the ten or so states participating in the pilot program. Each state was responsible for collecting and analyzing the data from insurers for their particular state, with a lot of assistance from the Ohio and Illinois departments. It became clear that more centralized data collection and baseline analysis would be more efficient for both insurers and regulators

The participating states also recognized that value of the MC AS would dramatically increase with universal participation by states. Again, a centralized data collection and baseline analysis program is necessary to involve all the states, since many states do not have the resources or other capability to issue a data call, collect the data and perform statistical analyses.

Basic Approach – Designation of Market Conduct Statistical Agents

Although centralized data collection and baseline analysis is the logical approach, concerns have been raised about data confidentiality and data sharing among regulators. To address these concerns, we suggest the following approach to implementing a centralized data collection program for market regulation data.

1. Develop a request for interest and qualifications (RFIQ) for organizations interested in designation as a statistical agent for market regulation data. This is the approach used to designate statistical agents in Texas for all property casualty lines of insurance in the mid 1990's and in Florida for workers' compensation statistical agents in the late 1990's. See attached RFIQ from Texas for an example.
2. Based upon the responses to the RFIQ and recommendations by a steering committee, discussed below, each state would designate a market regulation data statistical agent or agents and promulgate the market regulation statistical plan to be used by those designated agents.

3. The statistical agent or agents would collect the market regulation data, just as statistical agents collect statistical data today, and produce compilations and reports of the data, just as statistical agents do today. Each state would get a report from its statistical agent or agents and nationwide compilations would also be produced. These reports would be both compilations of the data submitted and the necessary baseline analyses. Thus, regulators would have be presented data and reports for immediate use in targeting further market conduct investigation and enforcement activities. In the event there are multiple MC statistical agents, a procedure would be in place for one of the MC statistical agents would be responsible for compiling the data from other MC statistical agents. Of course, one of the issues for the Steering Committee is whether one or multiple MC statistical agent.
4. By virtue of using the statistical agent designation approach, the issue of data confidentiality and data sharing goes away because those issues are already addressed in each state through existing law on statistical agents and data collection.

Possible Workplan and Timeline

The workplan for developing, issuing and evaluating responses to the RFIQ might go something like this:

1. Establish a Steering Committee that includes at least 9 state insurance regulators but allows as many states to participate as have interest and resources. Also include one representative each for personal lines insurers, commercial lines insurers, life insurers and health insurers and four consumer representatives. Representatives from insurer trade associations can fill the industry spots as long as the trade association is not also a statistical agent. The Steering Committee, with the assistance of NAIC staff would develop the RFIQ and supporting documents, including, for example, the MC statistical plan, required reports, and performance standards. Target Date: February 1, 2005
2. Develop and issue the RFIQ. The Steering Committee (SC), with the assistance of NAIC staff, would consult as necessary with existing statistical agents and other stakeholders to develop the RFIQ and supporting documents. The SC would identify key issues to be addressed and ensure that the RFIQ contains information requests to provide information necessary to address the key issues. For example, do any existing statistical agents already collect all the MC data desired by regulators in normal statistical reports made by insurers? Is it more efficient and cost effective for regulators and/or insurers to designate one MC statistical agent than multiple agents? Target Date: May 1, 2005
3. Solicit and answer questions about the RFIQ from interested organizations. Target Date for submission of questions and subsequent answers: June 1, 2005.

4. Deadline for Submissions. Target Date for Submissions: June 15, 2005.
5. Review of Submissions and Recommendations by Steering Committee for organization or organizations to be designated as MC statistical agents. Target Date: September 15, 2005.
6. Designation of MC statistical agents by individual states to ensure initial collection of MC data for experience beginning 1/1/06. Target Date: December 15, 2005. This gives individual states three months to act on the recommendations of the Steering Committee and have a MC statistical agent in place for 2006 data collection.