The Center for Economic Justice's Call to
Insurers and Insurance Regulators

To Address Societal Systemic Bias and Inherent Racism in Insurance
By Explicit Recognition of Disparate Impact as Unfair Discrimination in
Insurance

Submitted to the National Association of Insurance Commissioners'
Big Data Working Group
Artificial Intelligence Working Group
Market Regulation and Consumer Affairs Committee
Casualty Actuarial and Statistical Task Force
Accelerated Underwriting Working Group

June 18, 2020

Action, Not Just Words, Needed

The murder of George Floyd has led to widespread corporate recognition of and opposition to systemic bias and inherent racism in America. Corporate CEOs have spoken out, including major insurer CEOs.

> "In the coming days, I encourage each of us to step outside of our comfort zones, seek to understand, engage in productive conversations and hold ourselves accountable for being part of the solution. We must forever stamp out racism and discrimination." Those are the words of Kirt Walker, Chief Executive Officer of Nationwide.

> Floyd's death in Minneapolis is the latest example of "a broken society, fueled by a variety of factors but all connected by inherent bias and systemic racism. Society must take action on multiple levels and in new ways. It also requires people of privilege— white people—to stand up for and stand with our communities like we never have before," Those are the words of Jack Salzwedel, the CEO of American Family.

Perhaps this will be a turning point in insurer and regulatory practices, **but insurers have consistently opposed proposals to address systemic bias and inherent racism in insurance.** This opposition has come in two general themes – opposition to any responsibility by insurers or regulators to identify and minimize **disparate impact**[1] in insurance and opposition to any form of regulatory data collection to allow regulators and the public to assess market outcomes and thereby hold insurers accountable for their practices.

While insurers have been constant in opposing any responsibility to address systemic bias and inherent racism – in contrast to the recent public statements of insurer CEOs – most state insurance regulators believe they have the authority to stop proxy discrimination against protected classes. This belief, however, has never manifested itself, in regulatory standards, models laws or consistent approaches across states.

If insurers and insurance regulators truly want to address systemic bias and inherent racism in insurance, two long-overdue actions are needed.

1. Explicit recognition of disparate impact as unfair discrimination against protected classes in insurance coupled with responsibility for insurers and insurance regulators to identify such disparate impact and take steps to minimize this proxy discrimination within the overall regulatory framework of cost-based pricing.

2. Development of regulatory data collection and analysis infrastructure and capabilities for insurance regulators and the public to meaningfully monitor market outcomes for all consumers, to identify discriminatory outcomes and trace disparate impact to the causes.

---

[1] Disparate impact refers to practices that have the same effect as disparate treatment or intentional discrimination against protected classes. Protected classes reefer to those consumer characteristics which may not be the basis for discrimination and include, in most states, race, religion and national origin. Disparate impact is also known as disparate effect or proxy discrimination – discrimination against the protected class through a proxy for the protected class characteristic. Disparate impact as unfair discrimination has long been recognized under federal employment and housing laws. In 2015, the U.S. Supreme Court affirmed disparate impact as unfair discrimination under the Fair Housing Act which covers a variety of housing-related issues, including insurance, with Justice Kennedy writing, "Recognition of disparate-impact liability under the FHA plays an important role in uncovering discriminatory intent: it permits plaintiffs to counteract unconscious prejudices and disguised animus that escape easy classification as disparate treatment."

The mechanisms to accomplish these actions are straightforward.

1. Development of, and implementation by the states, through the National Association of Insurance Commissioners (NAIC)[2] of a model law addressing algorithmic bias including recognition of disparate impact as unfair discrimination against protected classes in insurance with guidance and safe harbors for insurers to identify and minimize disparate impact in marketing, pricing, claims settlement and anti-fraud efforts.

2. Development of, and implementation by the states, through the NAIC, of a market regulation data collection and analysis infrastructure to timely and meaningfully monitor consumer insurance outcomes – similar in scope and capability to what state insurance regulators and the NAIC currently have for monitoring the financial condition of insurers.

In the absence of the necessary actions by insurers and the states, Congress and federal agencies will eventually address these problems through civil rights legislation and enforcement.

In An Era of Big Data Analytics and Insurers' Rapidly Growing Use of Third-Party Data and Complex Algorithms, the Potential For Algorithmic Bias and Proxy Discrimination Has Grown Dramatically.

The potential for big data, artificial intelligence, machine learning – implemented through rapid deployment of complex algorithms – has increased the potential for intentional or unintentional proxy discrimination through algorithmic bias. This potential is well recognized. Barocas and Selbst state the issue succinctly:[3]

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.

Most data sets of personal consumer information as well data sets of the built environment reflect historical discrimination against protected classes. For example, TransUnion has an insurance score used for pricing based on criminal violations filed with the courts – not just convictions, but all criminal filings regardless of the eventual outcome. TransUnion's marketing materials state:

---

[2] https://content.naic.org/index_about.htm
[3] http://www.californialawreview.org/wp-content/uploads/2016/06/2Barocas-Selbst.pdf

 "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."

It did not take the recent murders of Black Americans by police to recognize that this "criminal history score" will reflect historic discrimination in policing against Black Americans and perpetuate that discrimination in insurance. 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.

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.

One of the oft-cited benefits of big data analytics in insurance is greater personalization – the ability of insurers to develop products and pricing tailored to individual needs and characteristics. But the other side of personalization is exclusion. Insurers' use of algorithmic techniques called price optimization, claim optimization and customer lifetime value are examples of the flip side of big data personalization – differential treatment of groups of consumers that reflect and perpetuate inherent bias and systemic racism.

The TransUnion Criminal History Score is just one example – egregious and obvious – of algorithms that reflect and perpetuate historic discrimination against protected classes in insurance – algorithms that reinforce inherent bias and systemic discrimination. Others include:

- Employment categories and education levels for marketing, underwriting and pricing
- Price Optimization and Customer Lifetime Value Algorithms used for marketing, underwriting, pricing and claims settlement
- Facial analytics used in life insurance underwriting
- Household composition used for underwriting and pricing
- Credit scores for marketing, underwriting, pricing, claims settlement and anti-fraud efforts
- Fraud detection models based on biased learning data

Many of these practices have shown to discriminate unfairly against protected classes, generally, and Black Americans, specifically. A number of cities – as well as Google and IBM – have stopped using facial recognition technology because of the biases against Black Americans. After the New York Department of Financial Services developed a regulation permitting the use of employment and education characteristics in auto insurance pricing only if the insurer could demonstrate the practice did not unfairly discriminate against protected classes, insurers' use of the "risk" characteristics disappeared.

The Consumer Federation of America has produced a number of extraordinary studies of discriminatory market outcomes resulting from rating factors that reflect systemic racism.[4] Insurance industry trade associations have dismissed the CFA's discriminatory findings with the claim that insurers engage in cost-based, race-neutral practices – while refusing to both provide the data to back up these claims and refusing to recognize that systemic racism will show up as disparate impact.

If insurers and insurance regulators are serious about addressing inherent bias and systemic racism in insurance, then action is needed.   Fortunately, the insurance industry has the precise skill set needed to identify and minimize disparate impact and insurance regulators have the resources to develop the necessary guidance and infrastructure.

Disparate Impact Analysis is Straightforward and Particularly Suited to Insurance.

The mechanics of a disparate impact analysis in insurance are straightforward and use well-accepted statistical and actuarial methods. Any algorithm – whether for pricing, anti-fraud, claims settlement, lifetime customer value, price optimization or other – takes the basic form of an equation in which certain variables or factors – the explanatory factors – seek to explain or predict a particular outcome.

---

[4] https://consumerfed.org/cfa-studies-on-the-plight-of-low-and-moderate-income-good-drivers-in-affording-state-required-auto-insurance./

Consider the following general model.

$$b_0 + b_1X_1 + b_2X_2 + b_3X_3 + e = y$$

Say that $X_1, X_2 + X_3$ are explanatory variables used to predict y – the frequency of an auto claim, for example.

Let's assume that all three Xs are statistically significant predictors of the likelihood of a claim and the b values associate with each X are how much each X contributes to the explanation of claim.

b_0 is the "intercept" – a base amount and e is the error term – the portion of the explanation of the claim not provided by the independent variables.

When the algorithm or model is developed, the modeler will typically data mine some database of personal consumer information, built environment or natural environment for characteristics that are correlated with the desired outcome. These variables are combined into a model, but a variable that might be predictive on its own can lose its predictive capability when combined with other variables because the variables might be correlated with one another. In that event, the variable serving as the proxy for the other variable loses its individual explanatory power. In our example, above, if, say, $X_1 + X_2$ are highly correlated, when the two variables are used in the same algorithm, one of the variables will lose its predictive power.

From a statistical and actuarial perspective, a disparate impact analysis does two things. First, it examines the amount of correlation between explanatory variables or factors and protected class characteristics to determine if any of the explanatory variables have significant correlation with, and thereby serve as proxies – in whole or in part – for protected class characteristics.

The second function of a disparate impact analysis is to remove the correlation between the explanatory variables and protected class characteristics with the result that the remaining explanatory power of the explanatory variables is the independent contribution – independent of correlation to protected class characteristics – of the explanatory variables relationship to the outcome.

Consider the following example. Suppose an explanatory factor was perfectly correlated with being a Black American. In statistical terms, this means a perfect or 100% correlation and the explanatory factor is a perfect proxy for being African-American. Assume that when used in an algorithm, this perfect proxy for being a Black American shows us as predictive of some outcome variable. Assume variable X_1 in our simple model above is the perfect proxy characteristic and further assume that the proxy variable shows a correlation to / is predictive of the outcome variable. Given our assumption that variable X_1 is a perfect proxy for being Black American, then the results of the model would be identical whether we used the proxy variable or used Black American explicitly. If the proxy variable is used, this would not be intentional discrimination – defined as explicit use of a protected class characteristic – even though it has

precisely the same effect. While most regulators believe they have the authority and obligation to stop the use of such proxies for protected class characteristics, the insurance industry view, as espoused by the American Property Casualty Insurance Association, is that even in this extreme case, there is no unfair discrimination against a protected class.

When the data are run through the model, variable X_1 shows some correlation to the outcome variable and is, therefore, "predictive." But, what it is really doing is simply standing in for being Black American and indirectly discriminating on the basis of race. This proxy factor is, in fact, simply reflecting and perpetuating discrimination against Black Americans.

One approach to disparate impact analysis – among many which generally try to remove the correlation between predictive variables and protected class characteristics – is to include a control explanatory variable for being Black American in the algorithm. Let's know add a new variable to algorithm – a specific variable for being Black American.

$$b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \mathbf{b_4R_1} + e = y$$

In statistical and actuarial terms, this is known as adding a control variable. The purpose of the control variable is to remove known correlations and biases in the other explanatory variables in order to better assess the independent and unique explanatory power of these other explanatory variables. For example, in personal auto pricing models, an insurer developing a national pricing model will utilize a control variable for State to remove the effects of / correlations with other explanatory variables of State-specific characteristics, such as different minimum liability limits, different tort and no-fault systems and different population distributions by age or other factors, among other things. In our example, our control variable $\mathbf{R_1}$ is being Black American.

Now, when the data are run through the model, explanatory variable X_1 – the perfect proxy for being Black American – shows no explanatory power and the control variable now shows the explanatory power that explanatory variable X_1 had in the original model. This is statistical evidence that explanatory variable X_1 was discriminating on the basis of race.

Let's consider two other examples – one in which there is a 50% correlation between variable X_1 and being Black American and a second in which there is a 0% correlation. In the 50% correlation, the variable X_1 may still show up as predictive of the outcome, but that predictive power will be different than from our first model without the control variable for being Black American. X_1's new contribution to explaining or predicting the outcome will now be its contribution independent of any correlation to being Black American. Consequently, disparate impact is recognized and minimized. Again, this is a common statistical and actuarial technique.[5]

---

[5] For example, the technique is explained is the chapter, ""Credit Scoring and the Fair Lending Issue of Disparate Impact, "in *Credit Scoring for Risk Managers*, Elizabeth May, editor, 2004.

In our third example there is 0% correlation between the variables X_1 and **R_1**. In this situation, the predictive power of X_1 remains the same as in the original model because there is no disparate impact.

As noted above, disparate impact analysis is particularly suited to insurance because the actuarial justification required for insurance risk classifications is a statistical test – is the characteristic correlated with risk of loss? The same statistical test can be used to evaluate and minimize disparate impact. Stated differently – if a particular correlation and statistical significance is used to justify, say, insurance credit scoring, those same standards of correlation and statistical significance are reasonable evidence of disparate impact and unfair discrimination on the basis of prohibited factors.

In addition, the ability of insurers to identify and minimize disparate impact can be easily built into the development of pricing, marketing or claim settlement models by including consideration of prohibited characteristics as control variables in the development of the model and then omitting these prohibited characteristics when the model is deployed. Again, this is one of many ways to remove the correlations between explanatory variables in algorithms and the protected class characteristics that result in reflection of and perpetuation of historic discrimination or disparate impact.

Recognition by regulators and insurers of disparate impact as unfair discrimination in insurance against protected classes and requirements to identify and

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

What NAIC Committees and Working Groups Should Be Doing

The NAIC has spread work streams related to Big Data Analytics over a number of groups. With the exception of the Artificial Intelligence Working Group, none of these groups' work efforts address systemic bias in insurance.

Artificial Intelligence Working Group

The NAIC Artificial Intelligence (AI) Working Group is developing insurance-specific principles for the governance and use of AI in insurance. While there are a number of consumer protection issues associated with insurers' use of AI (or Big Data Analytics, generally), such as protection of personal data and transparency and accountability to consumers and regulators, the most important consumer protection is establishing a responsibility for insurers and regulators to identify and minimize algorithmic bias and proxy discrimination. Recognition of disparate impact and responsibility of insurers and regulators to minimize such systemic bias must be a core AI insurance principle.

Big Data Working Group

The NAIC Big Data Working Group is examining big data analytics issues across a variety of insurance operations and lines of business. The two actions called for by CEJ regarding disparate impact and data collection should be at the core of all the working group's inquiries and activities. The Big Data Working Group should be developing the model law or revisions to existing model laws regarding explicit recognition of disparate impact, guidelines for identify and minimizing proxy discrimination and safe harbors for insurers.

Market Regulation and Consumer Affairs Committee

The NAIC Market Regulation and Consumer Affairs Committee is the parent committee for a number of working groups related to insurance market regulation, including data collection for market regulation, market surveillance, market conduct examinations and antifraud efforts. The Committee should be a contributor to the development of model laws regarding disparate impact, but must take the lead on market regulation data collection – both to identify the types of data and algorithms used by insurers and what these data are used for and to re-engineer market regulation data collection to match the granularity and frequency of financial regulation data collection.

Casualty Actuarial and Statistical Task Force

The NAIC Casualty Actuarial and Statistical Task Force deals generally with actuarial issues in property casualty lines of insurance. The Task Force is currently developing a white paper to provide best practices for regulatory review of complex pricing models used by insurers to justify rates. The current draft does not incorporate identification and minimization of systemic bias or disparate impact, but simply lists it as another consideration. Insurance rate standards include rates not being excessive, not being inadequate and not being unfairly discriminatory.

The use of complex predictive models for pricing by insurers is focused on risk segmentation and the development of risk classifications and rating factors. Traditional actuarial techniques – not complex predictive models – are generally used for overall rate level indications – the metric for assessing whether rates are excessive or inadequate. The overwhelming reason

for close scrutiny of complex predictive models by regulators is to assess whether the risk classifications are fair or unfairly discriminatory. It is an understatement to say that the current draft white paper has a massive whole because of the failure to address proxy discrimination and disparate impact. Guidance to insurance regulators for regulatory review of complex insurance predictive models should prioritize the identifications and minimization of systemic bias and disparate impact.

Accelerated Underwriting Working Group

The NAIC Accelerated Underwriting Working Group continues the NAIC's multi-year examination of life insurers' use of Big Data analytics and predictive models in place of traditional actuarial practices for underwriting and pricing life insurance. While the predictive models now used by life insurers have the same function as those used in auto, home and other property casualty lines of insurance – namely, using non-traditional data and an algorithm to predict claims (or other outcomes of value to the insurer). While there are requirements for property casualty insurers to file these predictive models for regulatory review for some purposes – justifying rates – and special laws and provisions governing property casualty insurers' use of consumer credit information, there are no similar regulatory requirements for life insurers. The time is long overdue for this working group to develop the model laws for regulatory guidance and consumer protections to ensure consumer protections in the face of life insurers' growing use of non-traditional, non-insurance data and complex algorithms. And the core of such models laws and regulatory guidance must be identification and minimization of disparate impact and systemic racism.

Conclusion

Recent events have highlighted a long-standing gaps in insurer and insurance regulatory practices – the failure to monitor consumer market outcomes for discriminatory impacts against protected classes and the failure to incorporate identification and minimization of proxy discrimination in insurers' development of predictive models for all aspects of their operations and regulators' review of these algorithms. The tools are available to address these problems – analysis of disparate impact and improved data collection. CEJ calls on insurers and regulators to match their statements of outrage over systemic racism with the actions needed to identify and minimize such unfair discrimination in insurance.