



Comments of the Center for Economic Justice

To the NCOIL Special Committee on Race in Insurance Underwriting

Regarding Proxy Discrimination against Protected Class

March 5, 2021

The Center for Economic Justice (CEJ) appreciates NCOIL's efforts to examine the impact of systemic racism on insurer practices and insurance consumers.

However, the proposed revisions to the NCOIL P/C Insurance Modernization Act reflect a profound misunderstanding of how systemic racism affects insurance. By defining proxy discrimination only as intentional use of a proxy characteristic for a protected class characteristic, the revisions, if adopted, would memorialize insurer practices that discriminate indirectly on the basis of race, discourage insurers from examining the racial impact of their practices and would restrict current regulatory efforts to address such unfair discrimination.

Further, by the industry's own admission, the proposed revision do not enhance current regulatory authority. Rather, as industry is surely aware, the proposed revisions would hand cuff regulatory efforts to address proxy discrimination. The proposed revisions reflect such a misunderstanding of proxy discrimination and disparate impact that NCOIL will be inviting federal civil rights legislation aimed at insurance if the proposed revisions are adopted.

The proposed revisions reflect a profound misunderstanding of how systemic racism affects insurers' practices.

The proposed definition, with key parts highlighted, is

For purposes of this Act, as well as for the purpose of any regulatory material adopted by this State, or incorporated by reference into the laws or regulations of this State, or regulatory guidance documents used by any official in or of this State, "Proxy Discrimination" means the **intentional** substitution of a neutral factor for a factor based on race, color, creed, national origin, or sexual orientation **for the purpose of discriminating against a consumer to prevent that consumer from obtaining insurance or obtaining a preferred or more advantageous rate due to that consumer's race, color, creed, national origin, or sexual orientation**.

This definition misunderstands how systemic racism impacts society in general and insurance operations, specifically. Systemic racism involves institutional and personal biases that reflect and perpetuate historic discrimination. As Justice Kennedy wrote in the *Inclusive Communities* decision upholding disparate impact as unfair discrimination under the Fair Housing Act (which covers home insurance):

Recognition of disparate-impact claims is also consistent with the central purpose of the FHA, which, like Title VII and the ADEA, was enacted to eradicate discriminatory practices within a sector of the Nation's economy.

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.

There are two types of disproportionate outcomes based on race that can be found in insurance. The first is disproportionate outcomes tied to historic discrimination and embedded in insurance outcomes. One example is the disproportionate presence of certain health problems or shorter life expectancies in communities of color. Because the impacts of historic discrimination are embedded in insurance outcomes, it is a policy decision – and not a technical issue – to protect the groups experiencing the disproportionately poor insurance outcomes. As a policy issue, policymakers have addressed these outcomes by prohibiting discrimination on the basis of race in underwriting and pricing. So, regardless of actuarial indications, insurers cannot discriminate on the basis of race in health and life insurance.

The second type of disproportionate impact is tied to the use of proxies for the prohibited characteristic and not to the outcomes. We've previously described one such situation when insurers were using age and value of the home as underwriting factors for home insurance – with the result that communities of color were systematically denied home insurance because these communities were characterized by older, lower-value homes – a results directly tied to historic discrimination in housing and lending. When challenged, the insurers discovered that the factors they were using – age and value of the home – were more correlated with race than with insurance outcomes. As a result of the disparate impact challenge, the insurers moved to more relevant risk factors – such as the condition of the home and its systems – with the result that insurance became more available in communities of color as the insurers reduced the correlation between their risk classifications and race.

This second type of disparate impact involves intentional or unintentional unnecessary discrimination on the basis of race. It is unnecessary because the facially-neutral factor that is purportedly associated with the insurance outcome is, in whole or in part, a proxy for the protected class characteristic and predictive of that protective class characteristic. Stated differently, the facially-neutral factor has a spurious correlation to the insurance outcome and is really correlated to the protected class characteristic. Attached is a presentation I will be giving at the Casualty Actuarial Society's Ratemaking and Product Management Seminar that provides a more technical explanation of these concepts.

It is this second type of disproportionate outcome on the basis of protected class membership that disparate impact analysis is intended to identify and minimize. It is this type of unnecessary racial discrimination that has been the concern of insurance regulators and that should be the focus of the NCOIL revisions. The proposed revisions to the NCOIL model would, however, define away this second type of unnecessary disproportionate outcome and memorialize some discriminatory practices by insurers.

The proposed revisions would undermine the effort to address systemic racism in two ways. First, by defining proxy discrimination only as intentional use of a proxy for a prohibited class characteristic, the revisions set an impossible bar. Even if an insurer were using a perfect proxy for race, under the proposed NCOIL revisions, the regulator could take no action without proving an intent to discriminate and substitute on the basis of race. Were this requirement in effect, state regulators could never have taken action in the past on race-based pricing in life insurance.

Second, the requirement to prove intent reflects a profound misunderstanding of how systemic racism is reflected in and perpetuated by unconscious institutional, cultural and personal biases. The purpose of a disparate impact standard and methodology to address unnecessary disproportionate outcomes on the basis of race and other protected classes is to reveal the unnecessary and unintentional bias through rigorous and objective analysis. The proposed revisions move in the opposite direction by focusing on intent instead of on outcomes.

To better reflect how systemic racism may impact insurance and to meaningfully address unnecessary racial discrimination caused by the use of proxies, any definition of proxy discrimination must not be limited to intentional actions and must cover intentional or unintentional actions that result in unnecessary disproportionate outcomes due to the use of the proxy factor. A definition of proxy discrimination that reflects this understanding might be:

Proxy discrimination is the use of a non-prohibited factor that, due in whole or in part to a significant correlation with a prohibited class characteristic, causes unnecessary, disproportionate outcomes on the basis of prohibited class membership.

Any effort to address systemic racism and define proxy discrimination should apply to all aspects of insurers' operations and not limited to pricing.

The proposed revisions are limited to proxy discrimination in pricing. Yet, systemic racism can lead to proxy discrimination in marketing, claims settlement, and anti-fraud efforts. It is not just underwriting and rating algorithms that risk the use of proxies for race. In fact, the potential for unnecessary proxy discrimination is as great or greater for marketing, claims settlement and antifraud as for underwriting and pricing.

Industry admits the proposed revisions add no new tools or resources for regulators.

During the December 9, 2020 meeting of the NCOIL Committee on Race, I asked Nat Shapo whether it was his position that if a regulator discovered an insurer using a perfect proxy for race, could the regulator take action to stop that discriminatory practice. Mr. Shapo offered the view that regulators already have authority to stop such a practice.

Given that view from industry, the proposed revisions not only fail to add any new tools or resources for regulators to address unfair and unnecessary racial discrimination, but actually restrict activities that insurance regulators have long engaged in to stop the use of blatant proxies. The proposed revisions would absolutely prohibit the regulator from taking action under the scenario posed because the regulator – despite being able to prove unnecessary disparate impact – could never prove intent.

In closing, we urge NCOIL to reject the proposed definition of proxy discrimination and related revisions to the NCOIL model act. We hope your intent is to address the impacts of systemic racism in insurance and, if that is the case, the proposed revisions accomplish just the opposite and would memorialize such unnecessary proxy discrimination.



Addressing Systemic Racism in Insurance

**Presentation to Casualty Actuarial Society
Ratemaking and Product Management Seminar**

March 16, 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 Information Does This Map of Omaha Nebraska Present?

- a. Concentration of Minority Population
- b. Home Insurance Rating Territories
- c. Rates of COVID Infections and Deaths
- d. Federal Home Loan Eligibility 1930's to 1960's

STATES, COUNTIES, CITY
AND
AUTOMOBILE MAPS
ALWAYS ON HAND

A horizontal scale bar labeled "SCALE IN MILES" below it. The bar has tick marks at 0, $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$, 1, and 2.

-LEGEND-

RAILROADS
STEEL CAR LINES
SCHOOLS
BOULEVARDS - PARKS
CITY LIMITS

INDEX TO STREETS



NORTH FROM DODGE STREET

[illegible]

SOUTH OF DODGE STREET

| | |
|-----------------------------|---|
| 310--North and South. | 3380--Krug Ave., Bkno St. |
| 320--Dwight St. | 3477-- |
| 330--Furman St. | 3480--Herald St. |
| 430-- | 3500--E. 23d St., Carlton Lane |
| 500--Howard St. St. Mary's | 3521-- |
| 510-- | 3520--Lambert St., Grever St. |
| 520--Larkin St. (22 to 24). | 3530-- |
| 530-- | 3540--E. 23d St., Desmaris St. (17 to 18) |
| 540--Larkin St. (40 to 42). | 3550-- |
| 550-- | 3560--Van Camp Ave. (17-20). |
| 560-- | 3570--E. 23d St., Plunge St. |
| 570--Lancaster St. Stedley. | 3580--E. 23d St., Kussow St., Von |
| 580-- | 3590-- |
| 590-- | 3600-- |
| 600-- | 3610-- |
| 610-- | 3620-- |
| 620-- | 3630-- |
| 630-- | 3640-- |
| 640-- | 3650-- |
| 650-- | 3660-- |
| 660-- | 3670-- |
| 670-- | 3680-- |
| 680-- | 3690-- |
| 690-- | 3700-- |
| 700-- | 3710-- |
| 710-- | 3720-- |
| 720-- | 3730-- |
| 730-- | 3740-- |
| 740-- | 3750-- |
| 750-- | 3760-- |
| 760-- | 3770-- |
| 770-- | 3780-- |
| 780-- | 3790-- |
| 790-- | 3800-- |
| 800-- | 3810-- |
| 810-- | 3820-- |
| 820-- | 3830-- |
| 830-- | 3840-- |
| 840-- | 3850-- |
| 850-- | 3860-- |
| 860-- | 3870-- |
| 870-- | 3880-- |
| 880-- | 3890-- |
| 890-- | 3900-- |
| 900-- | 3910-- |
| 910-- | 3920-- |
| 920-- | 3930-- |
| 930-- | 3940-- |
| 940-- | 3950-- |
| 950-- | 3960-- |
| 960-- | 3970-- |
| 970-- | 3980-- |
| 980-- | 3990-- |
| 990-- | 4000-- |

LEGEND

-  - Best
- - Still Desirable
-  - Definitely Declining
- - Hazardous

What is Systemic Racism and Inherent Bias?

“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.

Why Do State and Federal Laws Prohibit Discrimination on the Basis of Race?

Justice Kennedy for the Majority in the U.S. Supreme Court's 2015 *Inclusive Communities* Opinion upholding disparate impact as unfair discrimination under the Fair Housing Act.

Recognition of disparate-impact claims is also consistent with the central purpose of the FHA, which, like Title VII and the ADEA, was enacted to eradicate discriminatory practices within a sector of the Nation's economy.

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.

Why Are Race and Other Protected Class Characteristics Carved Out of Fair Actuarial Discrimination?

The existence of historical, intentional discrimination based on these characteristics – discrimination that violates state and federal constitutions. But, also, the recognition that the historical discrimination has long-lasting effects that disadvantage those groups. Stated differently, you can't enslave a population for two hundred years and then expect the legacy of that enslavement will disappear overnight.

We continue to see those legacies of historical discrimination – systemic racism -- today both directly and indirectly in policing and criminal justice, housing, and the impacts of the Covid-19 pandemic.

How Can Systemic Racism Manifest Itself in Insurance – Whether for Marketing, Pricing or Claims Settlement?

1. Intentional Use of Race – Disparate Intent
2. Disproportionate Outcomes Tied to Historic Discrimination and Embedded in Insurance Outcomes
3. Disproportionate Outcomes Tied to Use of Proxies for Race, Not to Outcomes – Disparate Impact

Today's focus is on number 3 – practices that actuaries can identify and address.

Fair and Unfair Discrimination in Insurance

Provisions regarding unfair discrimination are generally found in two parts of insurance statutes: rating and unfair trade practices.

For life insurance, we look to the UTPAs and find two types of unfair discrimination:

- Actuarial – there must be an actuarial basis for distinction among groups of consumers; and
- Protected Classes – distinctions among groups defined by certain characteristics – race, religion, national origin – prohibited regardless of actuarial basis.

From the NAIC UTPA Section 4 G Unfair Discrimination

Actuarial Unfair Discrimination

(1) Making or permitting any unfair discrimination between individuals of the same class and equal expectation of life in the rates charged for any life insurance policy or annuity or in the dividends or other benefits payable thereon, or in any other of the terms and conditions of such policy.

Protected Class Unfair Discrimination

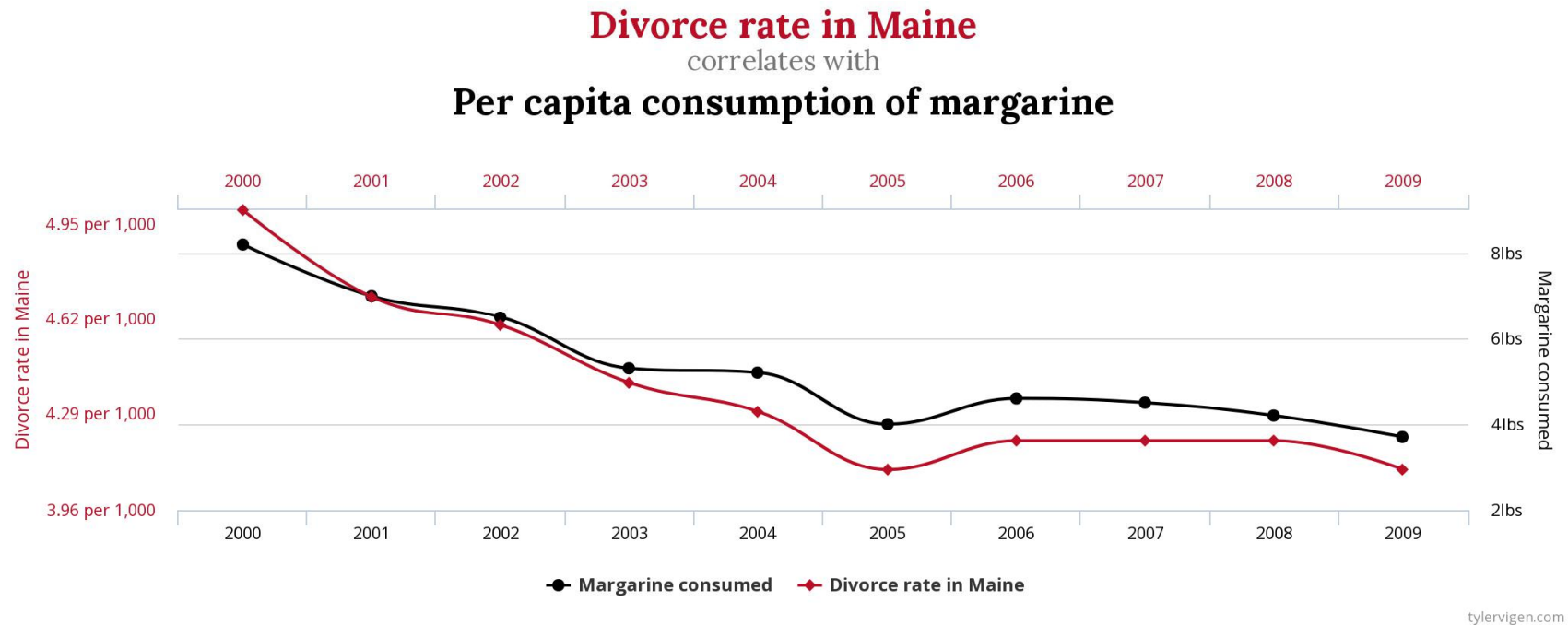
(6) Refusing to insure, refusing to continue to insure, or limiting the amount of coverage available to an individual because of the sex, marital status, race, religion or national origin of the individual;

Polling Question – Which best describes your experience?

- a. My company has a policy to examine the development and impact of our algorithms for racial bias
- b. My company has been receptive to suggestions to examine the development and impact of our algorithms for racial bias
- c. My company has not been receptive to suggestions to examine the development and impact of our algorithms for racial bias
- d. I would not suggest examining the development and impact of our algorithms for racial bias because of the reaction I think such a suggestion would create.
- e. My company doesn't use, collect or consider racial characteristics so there is no need to test our practices for racial impact.

Correlation is Not the Standard for Fair Actuarial Discrimination

Statutes and actuarial standards don't refer to correlation, but demand a more robust relationship. Why? Here's an example of an almost perfect correlation – over 99%.



Why isn't a simple correlation relied upon or sufficient?

Because a predictive characteristic (or variable) may not be correlated in whole or in part to the outcome, but may also be correlated to other predictive variables.

Consider the difference between an outcome – say, mortality – and one predictive variable versus an outcome and multiple predictive variables.

Age to Mortality, Gender to Mortality, Tobacco Use to Mortality

Each of these represents a one-to-one – or univariate – relationship. But each predictive variable may be replicating part of another variable because of correlation between the predictive variables. Tobacco Use may be correlated with age or gender.

Eliminating Correlation among Predictive Variables: Multi-variate Analysis

The issue of correlation among predictive variables has become more important in life insurance as insurers have started to use new data and predictive variables.

Over the last 30 years, insurers and actuaries have developed new techniques to address the problems with univariate analysis.

Insurers use a variety of techniques to eliminate correlations among predictive variables in order to isolate each individual predictive variable's unique contribution to explaining the outcome.

How Does Multi-Variate Analysis Work?

Here's a simple illustration of a multivariate model. Let's create a simple model to predict the likelihood of an auto claim:

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

$X_1, X_2 + X_3$ are the predictive variables trying to predict y .

Say that $X_1, X_2 + X_3$ are age, gender and credit score and we are trying to predict y – mortality (or predicting the decision produced by traditional underwriting).

Let's assume that all three X s are statistically significant predictors of the likelihood of a claim and the b values are how much each X contributes to the explanation of claim. The b values can be tested for statistical significance – how reliable are these estimates of the contribution of each X ?

By analyzing these predictive variable simultaneously, the model removes the correlation among the predictive variables.

Use of Control Variables in Multivariate Insurance Models

Suppose an insurer want to control for certain factors that might distort the analysis? For example, an insurer developing a national pricing model would might want to control for different state effects like different age distributions, different occupation mixes, different frequencies of accidental accidents or differences in jurisprudence. An insurer would add one or more control variables.

$$b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \mathbf{b_4C_1} + e = y$$

C_1 is a control variable – let's say for State. By including State as a control variable, the correlation of the X s to State is statistically removed and the new b values are now the contribution of the X s, independent of their correlation to State, to explaining the likelihood of a claim. When the insurer deploys the model, it still only uses the X variables, but now with more accurate b values.

Disparate Impact as Both a Standard and a Methodology

Let's go back to multi-variate model, but now use Race as a control variable:

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

R_1 is a control variable – by including race in the model development, the correlation of the X s to race is statistically removed and the new b values are now the contribution of the X s, independent of their correlation to race, to explaining the likelihood of a claim

What if X_1 is a perfect proxy for Race?

Then once we add the control variable for Race, X_1 no longer has any predictive value because all it was doing was predicting race, not the outcome y .

What if X_1 is both predictive of mortality and correlated to Race? Then, the model still shows X_1 's (now different) predictive value, but shorn of its correlation to Race, leaving the unique contribution of X_1 to explaining mortality.

Disparate Impact Analysis Improves Cost-Based Pricing

There is a long history and many approaches to identifying and minimizing disparate impact in employment, credit and insurance. But, the general principle is to identify and remove the correlations between the protected class characteristic and the predictive variables.

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

What if X_1 , X_2 and X_3 are not perfect proxies for Race, but still have high correlation? Then, the disparate impact analysis – and our simple model – removes that correlation and the remaining values for b_1 , b_2 and b_3 are the unique contributions of each predictive variable to explaining the outcome. The result is more – not less – accurate cost-based or risk-based analysis.

Why is it Reasonable and Necessary to Recognize Disparate Impact as Unfair Discrimination in Insurance?

1. It makes no sense to permit insurers to do indirectly what they are prohibited from doing directly. If we don't want insurers to discriminate on the basis of race, why would we ignore practices that have the same effect?
2. It improves risk-based and cost-based practices.
3. In an era of Big Data, systemic racism means that there are no “facially-neutral” factors. From Barocas and Selbst:

*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.***

Consider Criminal History Scores

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

“Also, as court records are created when the initial citation is issued, they provide insight into violations beyond those that ultimately end up on the MVR—such as violation dismissals, violation downgrades, and pre-adjudicated or open tickets.”

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

US DOJ Investigation of the Ferguson Police Department

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

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

US DOJ Investigation of the Ferguson Police Department (2)

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

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

Why is it Reasonable and Necessary to Require Insurers to Test for and Minimize Disparate Impact?

Insurer practices and algorithms do not necessarily use expected claims as the outcome variable. Sometimes the desired outcome is based on non-cost factors and these non-cost factors has disproportionate impact on communities of color.

In 2005, then CEO of Allstate, Ed Liddy told investment analysts about how credit scoring was helping Allstate avoid the wrong customers:¹

Tiered pricing helps us attract higher lifetime value customers who buy more products and stay with us for a longer period of time. That's Nirvana for an insurance company. That drives growth on both the top and bottom line.

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

¹ Transcript of Presentation to Edward M. Liddy, Chairman and CEO, The Allstate Corporation Twenty-First Annual Strategic Decisions Conference, Sanford C. Bernstein & Co., June 2, 2005.

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

The key, of course, is if 23% or 20% of the American public shops, some will shop every six months in order to save a buck on a six-month auto policy. ***That's not exactly the kind of customer that we want.*** So, the key is to use our drawing mechanisms and our tiered pricing to find out of that 20% or 23%, to find those that are unhappy with their current carrier, are likely to stay with us longer, likely to buy multiple products and that's where tiered pricing and a good advertising campaign comes in.

These statements were made in the Stone Age of Big Data – 2005. Since then, insurers' use of new, bigger and more granular personal consumer data has exploded.

Practices That Raise Concerns About Proxy Discrimination on the Basis of Race

Price Optimization and Consumer Lifetime Value Scores

By definition, these algorithms used by insurers utilize non-cost factors to differentiate among consumers and the factors and data reflect bias against communities of color.

Credit-Based Insurance Scores

The consumer credit information factors used in CBIS are highly correlated with race. The Missouri Department of Insurance found that the single best predictor of the average CBIS in a ZIP Code was minority population.

Criminal History Scores

Here, the problem is not just the legacy of historical discrimination, but ongoing discrimination in policing and criminal justice.

What are the Benefits and Costs of Requiring Insurers to Test For and Minimize Disparate Impact?

If racial and economic justice are a priority, if cost-based insurer practices are a priority, if closing the protection gap and making insurance more affordable and available in traditionally underserved communities is a priority, then the benefits of requiring insurers to test for and minimize disparate impact far, far outweigh the costs.

While there are examples of disparate impact claims brought against insurers under the federal Fair Housing Act that have resulted in improved risk-based pricing and improved insurance availability in communities of color – e.g., challenges against underwriting based on age and value of the home – industry has not been able to cite a single example of a successful disparate impact claim that has harmed risk-based pricing.

Why Do Efforts to Address Discrimination on the Basis of Race Require Explicit Consideration of Race?

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

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

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

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

Steve Bellovin, “Yes, ‘algorithms’ can be biased. Here’s why. A computer scientist weighs in on the downsides of AI.”²

This is what's important: machine-learning systems—“algorithms”—produce outputs that reflect the training data over time. If the inputs are biased (in the *mathematical* sense of the word), the outputs will be, too. Often, this will reflect what I will call “sociological biases” around things like race, gender, and class.

One thing is to exercise far more care in the selection of training data. Failure to do that was the likely root cause of Google Images labeling two African-Americans as gorillas. Sometimes, fixing the training data can help.

Of course, this assumes that developers are even aware of the bias problem. Thus, another thing to do is to test for biased outputs—and some sensitive areas, such as the criminal justice system, simply do not use these kinds of tools.

² <https://arstechnica.com/tech-policy/2019/01/yes-algorithms-can-be-biased-heres-why/>

There are several reasons to be wary of the "algorithmic" approach. One reason is that people put too much trust in computer output. Every beginning programmer is taught the acronym "GIGO:" garbage in, garbage out. To end users, though, it's often "garbage in, gospel out"—if the computer said it, it must be so. (This tendency is exacerbated by bad user interfaces that make overriding the computer's recommendation difficult or impossible.) We should thus demand less bias from computerized systems precisely to compensate for their perceived greater veracity.

The second reason for caution is that computers are capable of doing things—even bad things—at scale. There is at least the perceived risk that, say, computerized facial recognition will be used for mass surveillance. Imagine the consequences if a biased but automated system differentially misidentified African-Americans as wanted criminals. Humans are biased, too, but they can't make nearly as many errors per second.

Our test, then, should be one called disparate impact. "Algorithmic" systems should be evaluated for bias, and their deployment should be guided appropriately. Furthermore, the more serious the consequences, the higher the standard should be before use.

“The Real Reason Tech Struggles with Algorithmic Bias”³

These are mistakes made while trying to do the right thing. But they demonstrate why tasking untrained engineers and data scientists with correcting bias is, at the broader level, naïve, and at a leadership level insincere.

No matter how trained or skilled you may be, it is 100 percent human to rely on cognitive bias to make decisions. Daniel Kahneman’s work challenging the assumptions of human rationality, among other theories of behavioral economics and heuristics, drives home the point that human beings cannot overcome all forms of bias. But slowing down and learning what those traps are—as well as how to recognize and challenge them—is critical. As humans continue to train models on everything from stopping hate speech online to labeling political advertising to more fair and equitable hiring and promotion practices, such work is crucial.

³ Yael Eisenstat at <https://www.wired.com/story/the-real-reason-tech-struggles-with-algorithmic-bias/>

Insurers Don't Collect Applicant's Race – How Can an Actuary Get Data on Race to Perform a Disparate Impact Analysis?

1. Assign a racial characteristic to an individual based on racial characteristic of a small geographic area – Census data at the census block level.
2. Utilize the Bayesian Improved Surname Geocoding Method, based on census geography and surname data.⁴
3. Reach out to data brokers and vendors for a new data service.

⁴ See Consumer Financial Protection Bureau, "Using publicly available information to proxy for unidentified race and ethnicity." <https://www.consumerfinance.gov/data-research/research-reports/using-publicly-available-information-to-proxy-for-unidentified-race-and-ethnicity/> and Yin Zhang, "Assessing Fair Lending risks Using Race/Ethnicity Proxies."

Ethical Algorithms -- Sources

Pauline T. Kim, “Auditing Algorithms for Discrimination”

<https://www.pennlawreview.com/online/166-U-Pa-L-Rev-Online-189.pdf>

Claire Whitaker, “Ethical Algorithms”

<https://www.kdnuggets.com/2019/03/designing-ethical-algorithms.html>

Erin Russel, “The Ethical Algorithm”

<https://www.cognitivetimes.com/2019/01/the-ethical-algorithm/>

Barocas and Selbst

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899

Kroll, et al, “Accountable Algorithms:

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765268

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

Selbst and Barocas, “The Intuitive Appeal of Explainable Machines

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3126971

Levy and Barocas, “Designing Against Discrimination in Online Markets

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3084502

New York Times, “Algorithms and Bias, Q and A with Cynthia Dwork,” 10 August 2015

<https://www.nytimes.com/2015/08/11/upshot/algorithms-and-bias-q-and-a-with-cynthia-dwork.html>

Martin, Kirsten E. M., What Is an Ethical Algorithm (And Who Is Responsible for It?) (October 21, 2017). Available at SSRN:

<https://ssrn.com/abstract=3056692> or <http://dx.doi.org/10.2139/ssrn.3056692>

Kirsten Martin, “Ethical Implications and Accountability of Algorithms”

<http://kirstenmartin.net/wp-content/uploads/2013/11/Martin-JBE-Ethics-and-Accountability-of-Algorithms.pdf>

Kirsten Martin, DATA AGGREGATORS, BIG DATA, & RESPONSIBILITY ONLINE

<http://kirstenmartin.net/wp-content/uploads/2013/11/AAPOR-Martin-Info-Value-Chain-v2.pdf>

AIandBigData: A blueprint for human rights, social and ethical impact assessment Alessandro Mantelero

<https://reader.elsevier.com/reader/sd/pii/S0267364918302012?token=3836947F0CAD3C145A1F273E3CBE6C38F67E777DD7E4D590548F481916130DAACA8D57BED4667BD1FE1F4D8FC80E7C56>