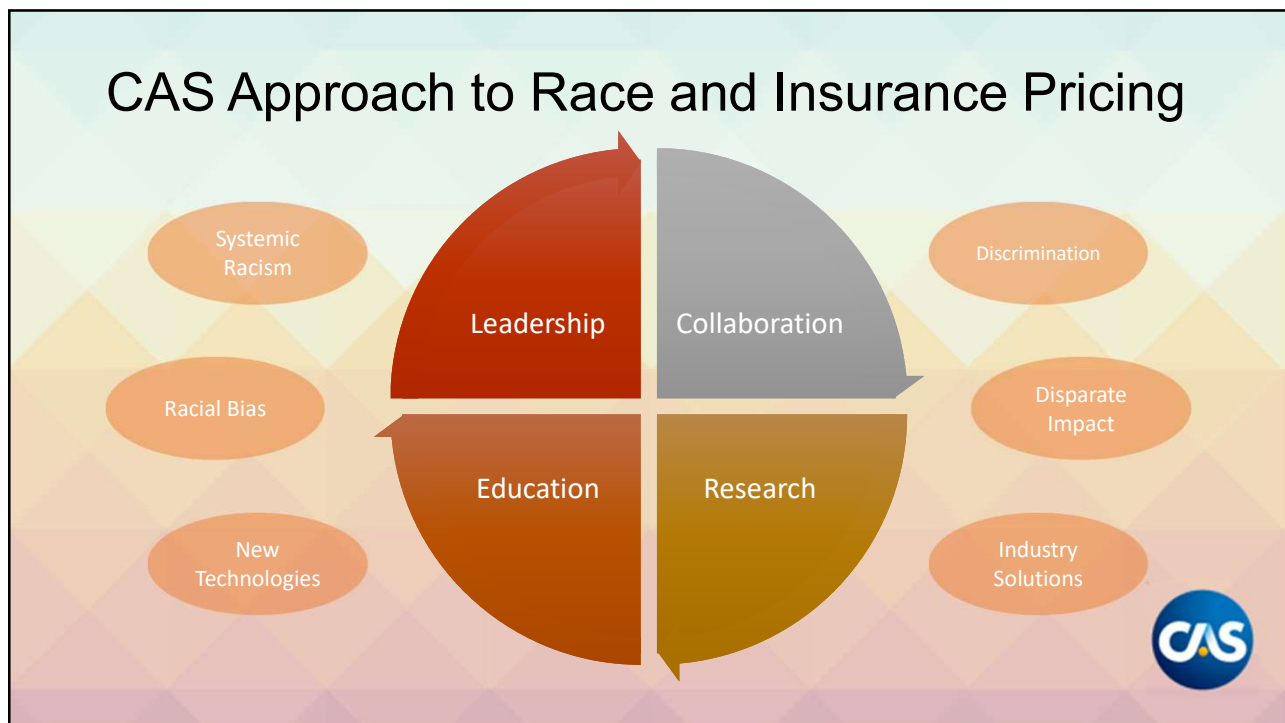




1



2

Race and Insurance Pricing Research

Today:

- Defining Discrimination in Insurance
- Methods for Quantifying Discriminatory Effects on Protected Classes in Insurance

Coming Soon:

- Approaches to Addressing Racial Bias in Financial Services
- Influences of Racial Bias on P&C Rating Factors



3

Defining Discrimination In Insurance

CAS Annual Meeting
November 8, 2021

Kudakwashe Chibanda, FCAS



4

Setting The Stage

01 Are You Sure You Know What **Protected Class** Is?

02 Revisiting **Unfair Discrimination**

03 The **Proxy Discrimination** Debate

04 What Is **Disparate Impact** Anyway?



5

Let's Start With A Question...

What makes a
society **fair**?


A. EQUALITY

B. EQUITY




6


Where Algorithms And Fairness Intersect




Deterministic
Supervised algorithms always produce the same result when given the same inputs



Auditable
Algorithms can be back-tested to determine whether they were right or wrong



Identity
Algorithms are like corporations...they have personas, but they are not people (and they make an easy target)



7

Protected Class & Unfair Discrimination

Protected Class

1886 A protected class is a group of people who share a **common characteristic**, for whom federal and state laws have created protections that prohibit against discrimination because of that trait.

Race

Religion

Origin

Sex

Family


Age

Disability

Genetics

Veterans


Unfair Discrimination



*rates must not be excessive, inadequate, or unfairly **discriminatory**¹*

- Discrimination ~ Differentiation
- No protected class mention
- Most states define protected class as part of unfair discrimination, but not all!

1. Race was prohibited for the purposes of accepting a risk



8

Proxy Discrimination – The Issues

In general, it is intuitive to think of proxy discrimination as the use of characteristics that stand in for other variables (i.e. proxies) for the purposes of prejudicing a certain group



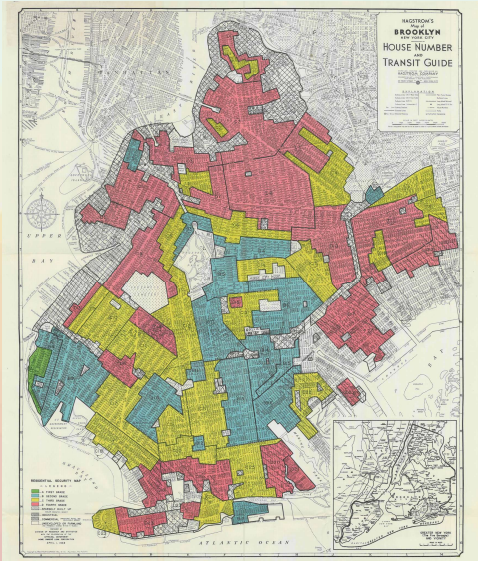
9

Proxy Discrimination

	FTC	NAIC	NCOIL	CEJ	APCIA
Definition	Whether an included variable acts in whole or in part as a statistical proxy for excluded variables such as race, ethnicity and income	Principles on AI: "AI actors should...avoid proxy discrimination against protected classes. AI systems should... avoid harmful or unintended consequences "	Proxy Discrimination means the intentional substitution of a neutral factor for a factor based on color, creed...for the purpose of discriminating against a consumer	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 based on prohibited class membership	"Proxy theory" was adopted by the courts as an element of disparate treatment discrimination to recognize a policy should not be allowed to use a technically neutral classification as a proxy to evade Title VII's prohibition against intentional discrimination
Similar Terms	Omitted Variable Bias		Defines proxy discrimination as a type of unfair discrimination	Disproportionate outcome	Disparate treatment
Intent Required?	Unknown	No	Yes	No	Yes
Notes / Issues	<ol style="list-style-type: none"> Are credit scores proxies for race? What happens when you control for race? 	<ol style="list-style-type: none"> Correlation vs. causation 	<ol style="list-style-type: none"> How do you identify intent? 	What is significant correlation?	Does proxy discrimination already have a legal definition?

10

Proxy Discrimination – An Example



What Is Redlining?

Classification of neighborhoods by desirability that was used by banks and insurers to determine eligibility for mortgage loans

How Was It Created?

The HomeOwners Loan Corporation (HOLC) categorized neighborhoods based on:

- Property Specific Characteristics
- Location Characteristics
- Borrower Characteristics

Boundaries were shown as Green, Blue, Yellow and Red

Why Was It Proxy Discrimination?

Race was not directly used, but it was clearly the target:

“If a neighborhood is to retain stability, it is **necessary that properties shall continue to be occupied by the same social and racial classes**. A change in social or racial occupancy generally contributes to instability and a decline in values



11

Disparate Impact

Disparate impact is a legal term that has a very specific definition

1. Will the practice cause a discriminatory effect on a protected class?

Yes

2. Is there a necessary relationship to a legitimate interest?

No Disparate Impact

Yes

3. Alternate, less discriminatory practice?

No Disparate Impact

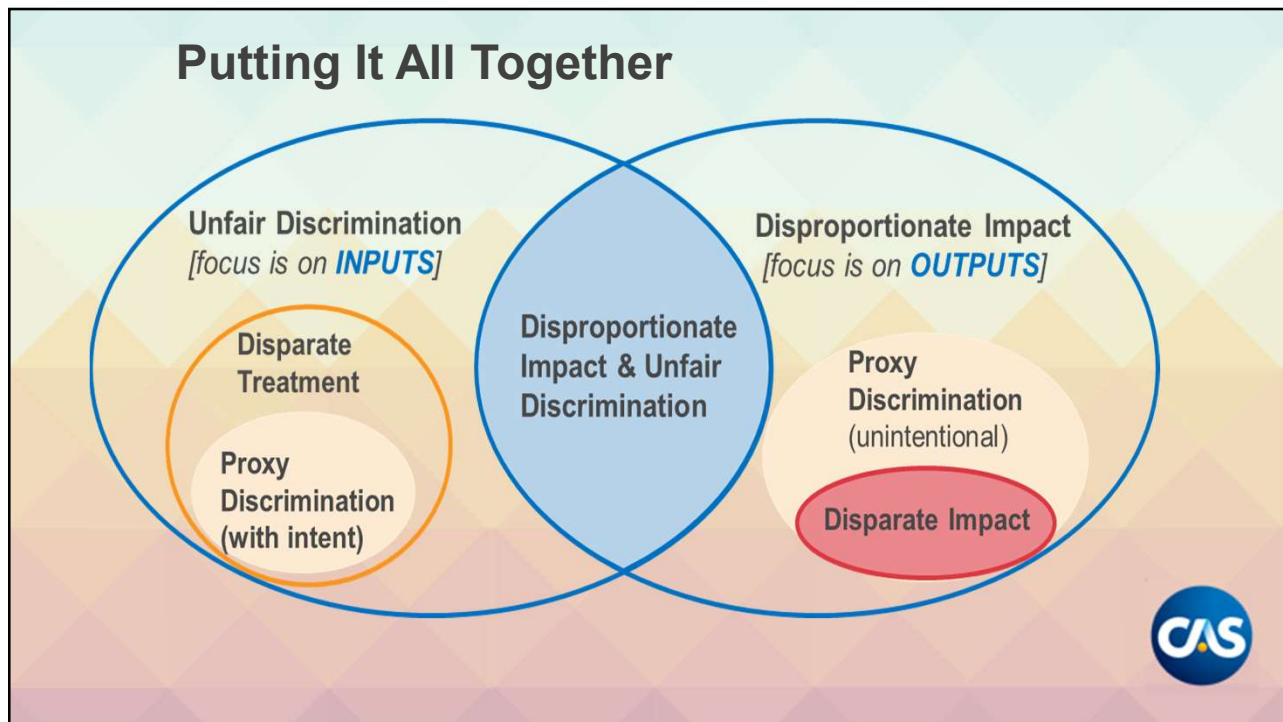
Yes

Disparate Impact Exists

No Disparate Impact



12



13

Let's Ask Again

What makes actuarial rating **fair**?

A. EQUALITY
B. EQUITY

CAS

14



15

PINNACLE
ACTUARIAL RESOURCES

Methods for Quantifying Discriminatory Effects on Protected Classes in Insurance

CAS Annual Meeting

November 8, 2021

Roosevelt C. Mosley, Jr., FCAS, MAAA, CSPA
Principal & Consulting Actuary

16

Methods for Quantifying Discriminatory Effects

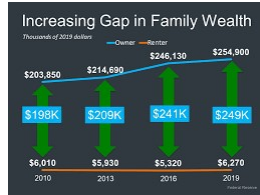
- Background
- Accusations of Bias in Insurance
- What is Unfairly Discriminatory?
- Approaches for Measurement and Mitigation

17

Background

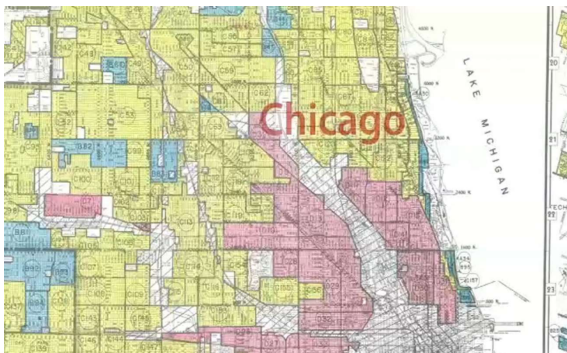
18

Importance of Insurance in Society



19

Redlining

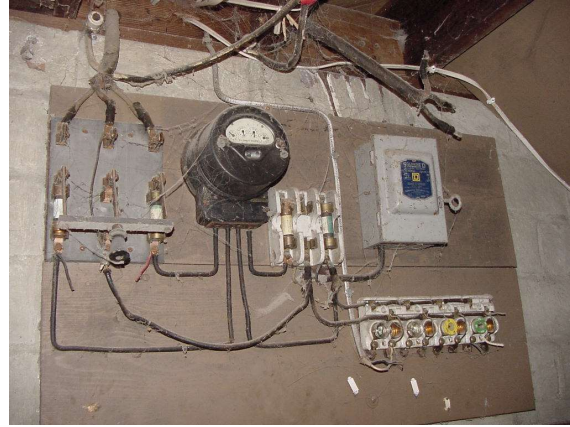


Without access to insurance and financial resources:

- Potential liability associated with operating a vehicle can lead to financial ruin
- Access to homeownership is significantly limited, thus limiting access to wealth
- Homes suffer from lack of investment in maintenance, upkeep

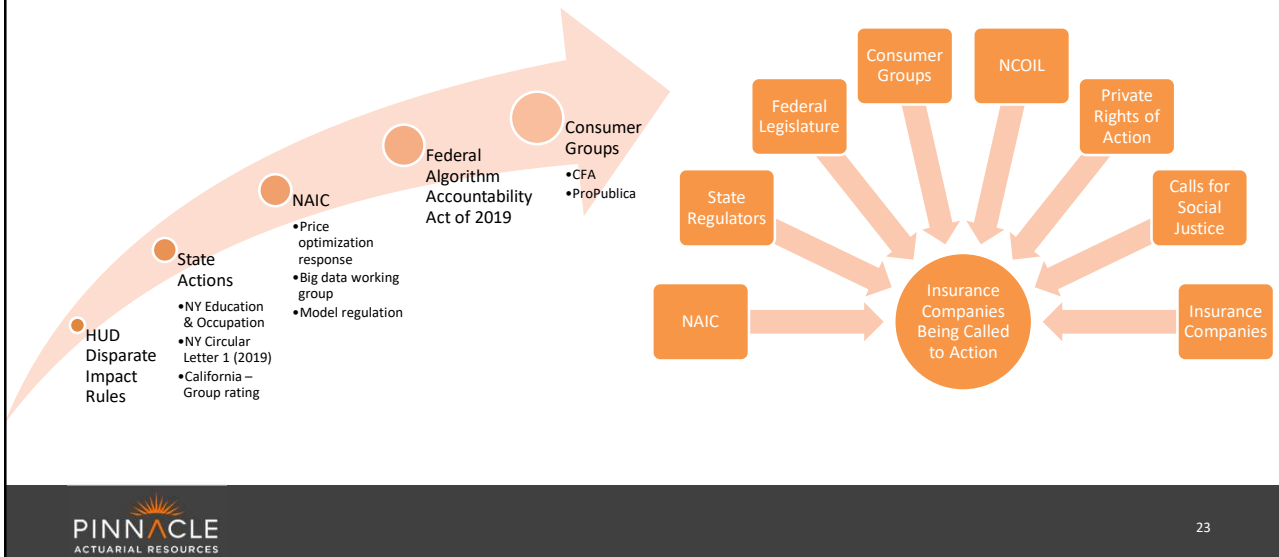
20

Examples of Persisting Impacts



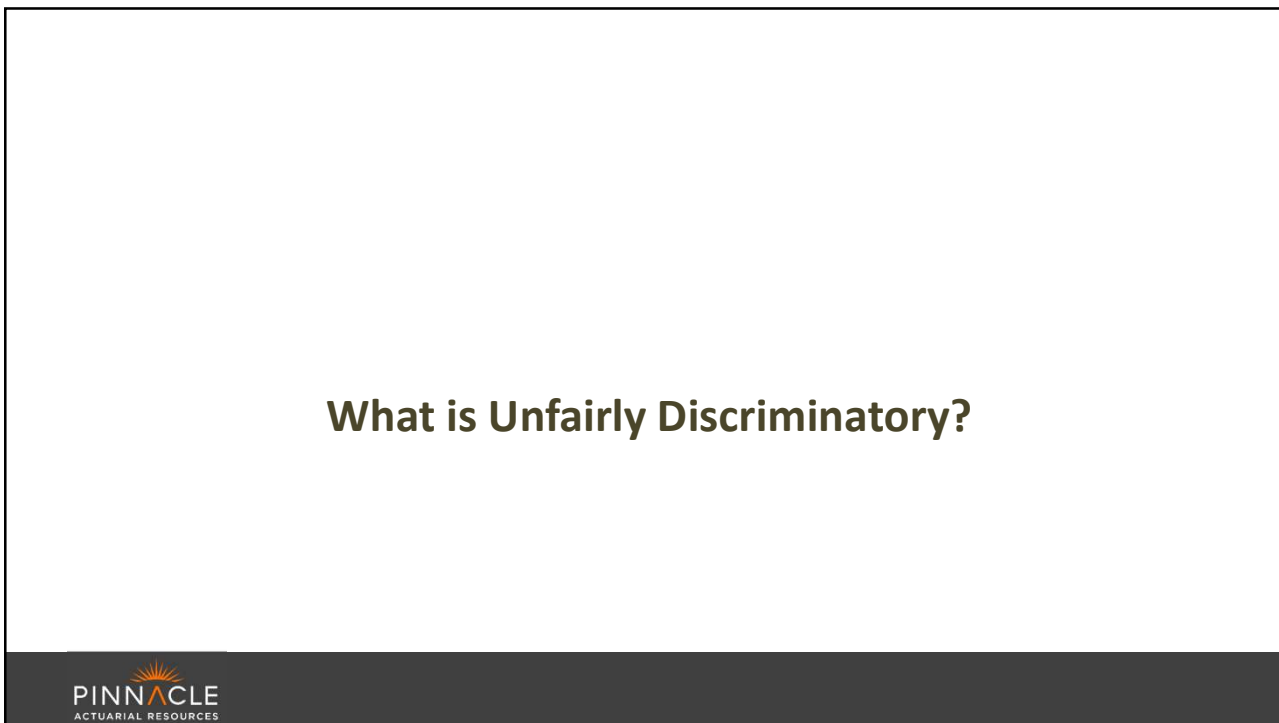
Accusations of Bias in Insurance

Rising Tide...And Mounting Pressure



23

What is Unfairly Discriminatory?



24

Applicable Ratemaking Guidance

- **State Rating Laws** – rates are to be not inadequate, not excessive, and not **unfairly discriminatory**
- **Statement of Principles Regarding Property and Casualty Insurance Ratemaking** – A rate is reasonable and not excessive, inadequate, or unfairly discriminatory if it is an **actuarially sound estimate of the expected value of all future costs associated with an individual risk transfer.**
- **Actuarial Standard of Practice No. 53** – “Estimating Future Costs for Prospective Property/Casualty Risk Transfer and Risk Retention”
- **Actuarial Standard of Practice No. 12** – “Risk Classification”
 - Rates within a risk classification system would be considered equitable if differences in rates reflect material differences in expected cost for risk characteristics. In the context of rates, the word fair is often used in place of the word equitable. (3.2.1)
 - While the actuary should select risk characteristics that are related to expected outcomes, it is not necessary for the actuary to establish a cause and effect relationship between the risk characteristic and expected outcome in order to use a specific risk characteristic (3.2.2)

25

Recent Guidance for Regulatory Review of Predictive Models

NAIC 2019 White Paper



Magnitude of premium disruption to individual policyholders and how the insurer will explain the disruption upon inquiry

Input variables should have a demonstrable relationship to expected losses or expense

Individual predictions from the predictive model and associated relativities are not unfairly discriminatory

26

Recent Industry Guidance for Actuaries

ASOP 56

3.1.3 Using the Model—When using the model, the actuary should make reasonable efforts to confirm that the model structure, data, assumptions, governance and controls, and model testing and output validation are consistent with the intended purpose.

27

Key Issue

Ultimately, the question of discriminatory effects (or unfair discrimination) on protected classes comes down to, at least in part, whether individual factors or combinations of factors derive their predictive power in full or in part from their correlation with a prohibited characteristic. If so, then it must also be determined whether this results in disproportionately higher or lower rates for certain groups within that protected class.

28

Approaches for Measuring and Mitigating Discriminatory Effects on Protected Classes

Model Fairness

The latest research in model fairness and model de-biasing is introducing an additional component to the **concept of model bias that transcends the purely statistical context**. The central theme in this additional dimension of bias detection and bias mitigation is attempting to provide practitioners of analytics with mechanisms and mathematical constructs to **minimize the social inequalities that their models may capture through data, and ensure that the model does not unfairly discriminate against certain protected classes**.

Table 1: Categories of Fairness Criteria

Independence	Separation	Sufficiency
$\hat{Y} \perp A$	$\hat{Y} \perp A Y$	$Y \perp A \hat{Y}$

A - protected attribute

Y - observed value of target variable

\hat{Y} - predicted value of target variable

Independence

- Requires that the predictions and the protected attribute be statistically independent
- Examples
 - Demographic Parity: requires that the model makes equal predictions for all levels of protected classes
 - $P(\hat{Y} = 1 | A = a) = P(\hat{Y} = 1 | A = b)$
 - Conditional Demographic Parity: requires that the model makes equal predictions for all levels of protected classes after controlling for permitted factors

31

Fairness through Unawareness

- Base case for machine learning
- Removes protected attributes from the data set
- Insufficient due to complex correlations among the variables

$$d(X = x, A = a) = d(X = x, A = a'), \quad \forall x \in X$$

32

Separation

- Separation is satisfied if the predictions and the protected attribute are statistically independent but conditional on the actual response
- Examples
 - Equal Opportunity: requires that the predicted outcomes are equal across the protected classes, but is conditional on the positive outcome being observed
 - $P(\hat{Y} = 1 | Y = 1 \ \& \ A = a) = P(\hat{Y} = 1 | Y = 1 \ \& \ A = b)$
 - Equalized Odds: requires that the protected classes have equal true positive rates and equal false positive rates
 - $P(\hat{Y} = 1 | Y = y \ \& \ A = a) = P(\hat{Y} = 1 | Y = y \ \& \ A = b), y \in \{0, 1\}$

33

Sufficiency

- Sufficiency is satisfied if the predictions and the protected attribute are statistically independent but conditional on the predicted values
- Examples
 - Calibration: requires that, conditional on the same predicted probability score p by the model, both the protected and unprotected classes have the same probability of actually belonging to the positive outcome
 - $P(Y = 1 | P = p \ \& \ A = a) = P(Y = 1 | P = p \ \& \ A = b), p \in [0, 1]$
 - Well-Calibration: adds an additional requirement that for a given predicted probability score p , the actually observed proportions should also equal p
 - $P(Y = 1 | P = p \ \& \ A = a) = P(Y = 1 | P = p \ \& \ A = b) = p, p \in [0, 1]$

34

Bias Mitigation – Machine Learning

Pre-Processing Bias in Data

- Attempt to mitigate bias in the training data

In-Processing Bias in Models

- Attempt to mitigate bias in the modeling phase

Post-Processing Bias in Predictions

- Attempt to mitigate bias in the modeling predictions

35

Bias Mitigation Techniques

Pre-process

- Reweighting
- Disparate Impact Remover
- Optimized Preprocessing
- Learning Fair Representations

In-process

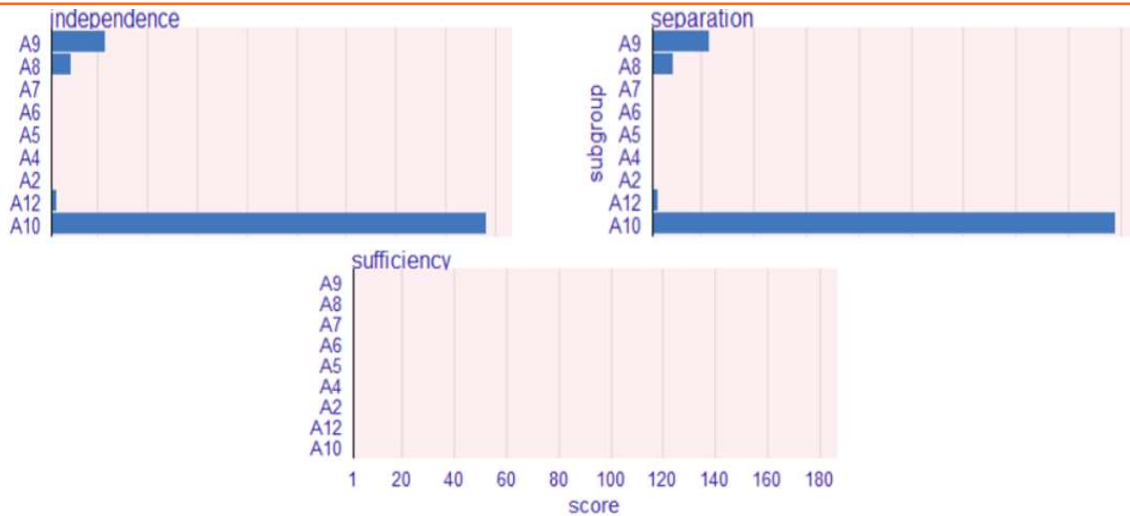
- Adversarial De-biasing
- Prejudice Remover
- Meta Fair Classifier

Post-process

- Reject Option Classification
- Equalized Odds
- Calibrated Equalized Odds

36

Example – Tracking Fairness Metrics (with Area in GLM)



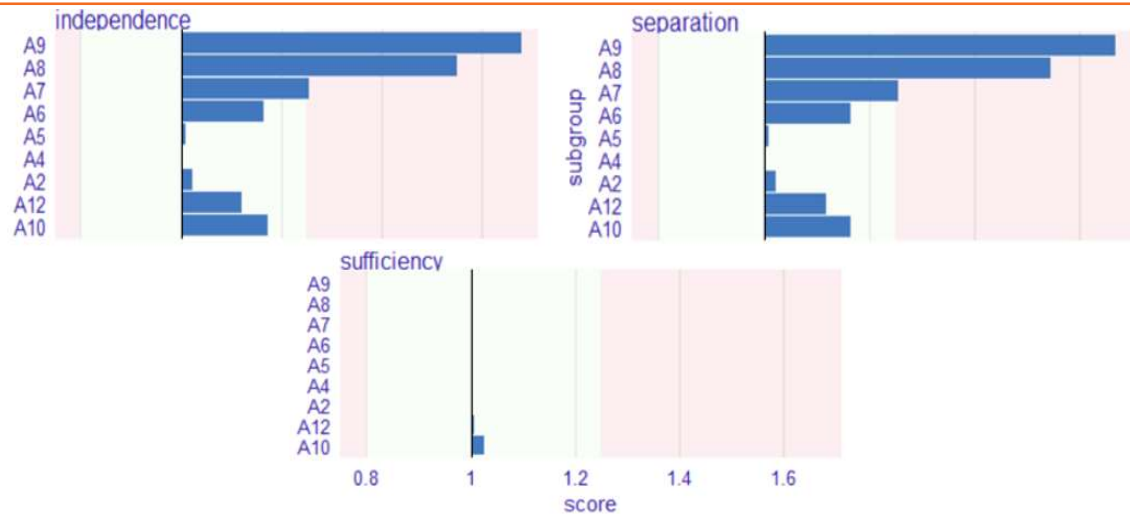
37

Example - Reweighting

	(1)	(2)	(3)	(4)	(5)	(6)
Area	with Area in GLM	with Area in GLM (wrt A3)	without Area in GLM	without Area in GLM (wrt A3)	with Area in Debiased GLM	with Area in Debiased GLM (wrt A3)
A2	0.062	1.000	0.064	0.955	0.070	1.014
A3	0.062	1.000	0.067	1.000	0.069	1.000
A4	0.067	1.081	0.067	1.000	0.068	0.986
A5	0.069	1.113	0.069	1.030	0.070	1.014
A6	0.076	1.226	0.074	1.104	0.073	1.058
A7	0.077	1.242	0.077	1.149	0.070	1.014
A8	0.092	1.484	0.080	1.194	0.072	1.043
A9	0.096	1.548	0.082	1.224	0.072	1.043
A10	0.125	2.016	0.074	1.104	0.067	0.971
A12	0.088	1.419	0.073	1.090	0.067	0.971

38

Example – Tracking Fairness Metrics (without Area in GLM)



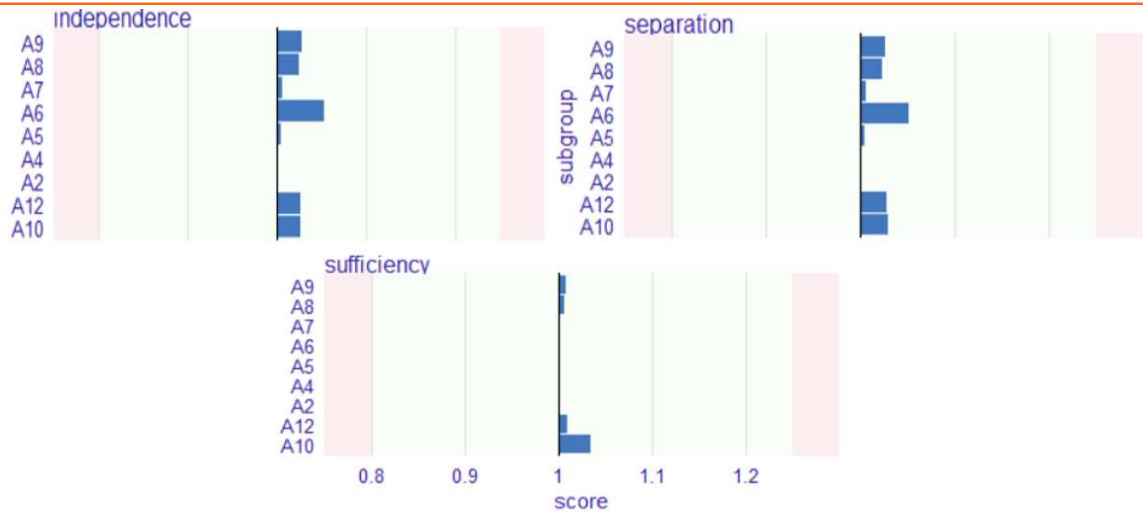
39

Example - Reweighting

	(1)	(2)	(3)	(4)	(5)	(6)
Area	with Area in GLM	with Area in GLM (wrt A3)	without Area in GLM	without Area in GLM (wrt A3)	with Area in Debiased GLM	with Area in Debiased GLM (wrt A3)
A2	0.062	1.000	0.064	0.955	0.070	1.014
A3	0.062	1.000	0.067	1.000	0.069	1.000
A4	0.067	1.081	0.067	1.000	0.068	0.986
A5	0.069	1.113	0.069	1.030	0.070	1.014
A6	0.076	1.226	0.074	1.104	0.073	1.058
A7	0.077	1.242	0.077	1.149	0.070	1.014
A8	0.092	1.484	0.080	1.194	0.072	1.043
A9	0.096	1.548	0.082	1.224	0.072	1.043
A10	0.125	2.016	0.074	1.104	0.067	0.971
A12	0.088	1.419	0.073	1.090	0.067	0.971

40

Example – Tracking Fairness Metrics (with Area in De-biased GLM)



41

Example - Reweighting

	(1)	(2)	(3)	(4)	(5)	(6)
	with Area in GLM	with Area in GLM (wrt A3)	without Area in GLM	without Area in GLM (wrt A3)	with Area in Debiased GLM	with Area in Debiased GLM (wrt A3)
Area						
A2	0.062	1.000	0.064	0.955	0.070	1.014
A3	0.062	1.000	0.067	1.000	0.069	1.000
A4	0.067	1.081	0.067	1.000	0.068	0.986
A5	0.069	1.113	0.069	1.030	0.070	1.014
A6	0.076	1.226	0.074	1.104	0.073	1.058
A7	0.077	1.242	0.077	1.149	0.070	1.014
A8	0.092	1.484	0.080	1.194	0.072	1.043
A9	0.096	1.548	0.082	1.224	0.072	1.043
A10	0.125	2.016	0.074	1.104	0.067	0.971
A12	0.088	1.419	0.073	1.090	0.067	0.971

42

Roosevelt C. Mosley, Jr., FCAS, MAAA, CSPA

309.807.2330

rmosley@pinnacleactuaries.com

