INSTITUTE FOR SOCIAL RESEARCH
UNIVERSITY OF MICHIGAN

CJARS Criminal Justice Administrative Records System

# Sources and Implications of Race and Ethnicity (Mis)measurement in the U.S. Criminal Justice System


Keith Finlay
U.S. Census Bureau

Elizabeth Luh
University of Michigan

Michael Mueller-Smith
University of Michigan & NBER


Summer 2023

*Any opinions and conclusions expressed herein are those of the authors and do not reflect the views of the U.S. Census Bureau. The U.S. Census Bureau reviewed this data product for unauthorized disclosure of confidential information and approved the disclosure avoidance practices applied to this release (Data Management System number: P-7500378, Disclosure Review Board (DRB) approval numbers: CBDRB-FY23-0178.).

# Race research in the justice system

CJARS Criminal Justice Administrative Records System

Significant concern among researchers, policy-makers, and the general public about racial disparities in the U.S. justice system

Underappreciated fact about criminal justice data:

- Race/ethnicity data often populated by law enforcement
- Used for serving arrests warrants, tracking population, and allocating institutionalized population
- Operational purposes take priority over downstream statistical use

# This project

CJARS Criminal Justice Administrative Records System

1. Document the degree of "mislabeling" in criminal justice records
   - Uniquely possible through individual-level linkage of CJARS data to Census Bureau household survey data
   - Demonstrate implications of mismeasurement for the accuracy of federal statistical series
2. Study what factors drive geographic and temporal variation in mislabeling rates
   - In progress; no preliminary results disclosed
3. Examine implications of mislabeling on individual outcomes in justice system
   - Unique and previously unexplored source of variation to study racial bias

# What is CJARS?

CJARS, founded in 2016, is a joint data infrastructure project between University of Michigan and the U.S. Census Bureau

Building a novel data platform to modernize research and statistical reporting on the U.S. criminal justice system, including:

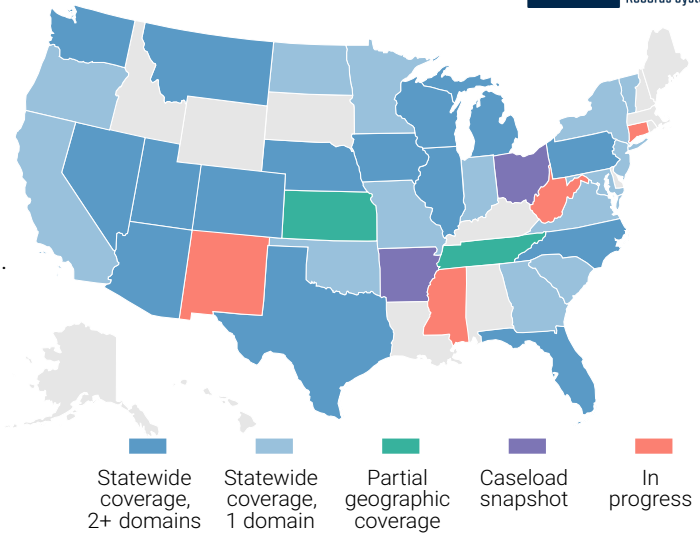
- Event-level criminal justice data with nationwide scope
- Tracking across key milestones in the justice system
- Capacity to link with individual-level survey and administrative data at the U.S. Census Bureau

Financial support from the NSF, Gates Foundation, Arnold Ventures, Robert Wood Johnson Foundation, Annie E Casey Foundation, the University of Michigan, and the Census Bureau

# CJARS data coverage

CJARS Criminal Justice Administrative Records System

- Data from ∼2,400 counties
- States with coverage of state court, state DOC, and/or state repository represent ∼84% of U.S. population
- >3b records
- >200m CJ events
- >41m unique individuals



# Accuracy of race & ethnicity data

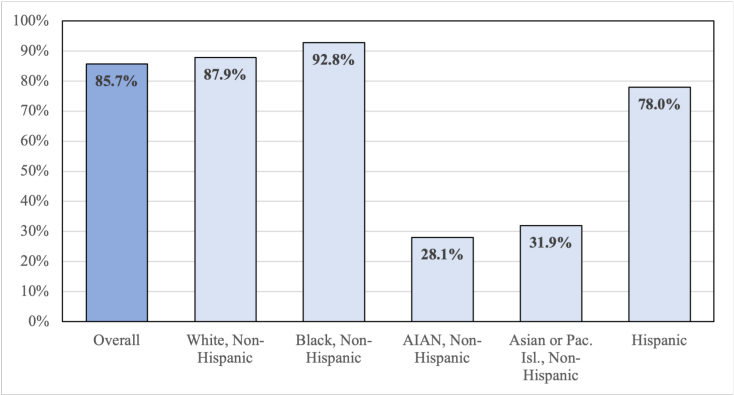



**Figure:** Share of justice-involved individuals correctly labelled by justice system, overall and by race/ethnicity using 2010 Decennial Census

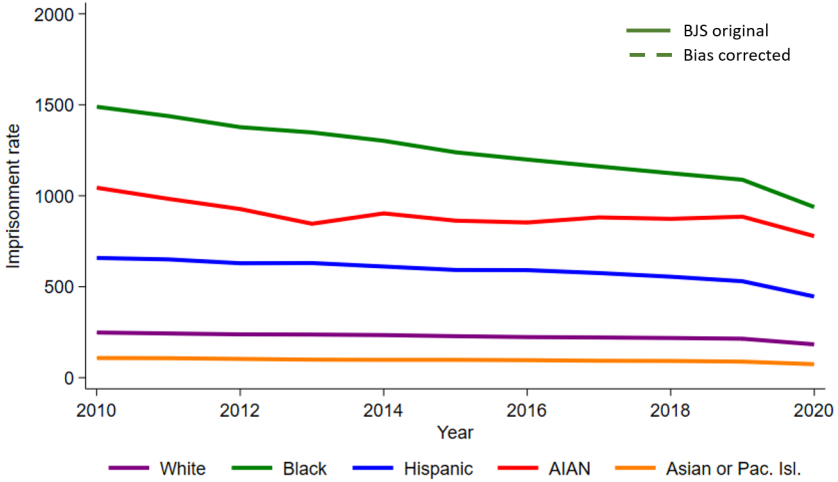
**Future work:**

- Other sources of race/ethnicity data
- Consistency in racial/ethnic data over time
- Changes in self-ID after justice contact

All results were approved for release by the U.S. Census Bureau, Data Management System number: P-7500378 and approval number CBDRB-FY23-0178.

# Correcting mismeasurement in BJS statistics

Imprisonment rate by demographic subgroup from the BJS report, *Prisoners in 2020*

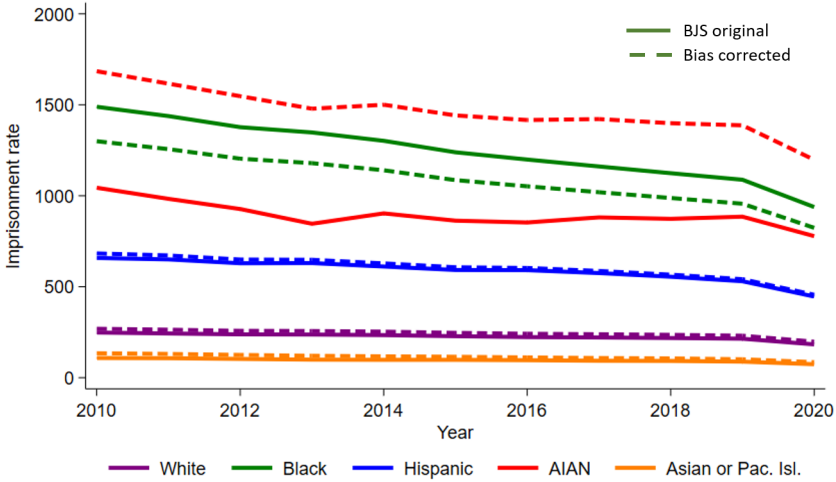



All results were approved for release by the U.S. Census Bureau, Data Management System number: P-7500378 and approval number CBDRB-FY23-0178.

# Correcting mismeasurement in BJS statistics

CJARS Criminal Justice Administrative Records System

Imprisonment rate by demographic subgroup from the BJS report, *Prisoners in 2020*




All results were approved for release by the U.S. Census Bureau, Data Management System number: P-7500378 and approval number CBDRB-FY23-0178.

# Why is there mismeasurement?

CJARS Criminal Justice Administrative Records System

Plan to disaggregate mislabeling rates down to local jurisdiction (county/state) by year level

Study geographic/temporal differences to understand what drives variation in data quality

Hypothesized factors:

- Outdated/insufficient race and ethnicity codes in IT systems
- Demographic representation within law enforcement
- Economic and political incentives
- Local racial segregation
- Random chance

# Mislabeling and individual outcomes

CJARS Criminal Justice Administrative Records System

Mislabeled individuals might experience different outcomes in the justice system

Labels might:

- Prime judges, prosecutors, defense attorneys, police, etc.
- Resolve uncertainty in the face of racial/ethnic ambiguity
- Be only source of race/ethnic information when limited physical proximity occurs

Without taste-based or statistical discrimination, mislabels should have no impact on individual outcomes

# Empirical exercise

Preliminary findings focus on adjudication outcomes among Whites and Blacks:

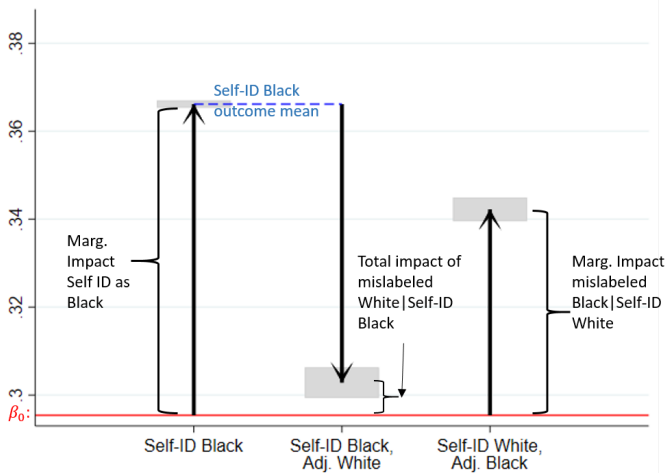
| | | Agency-recorded Race: $White$ | Agency-recorded Race: $Black$ |
|---|---|---|---|
| Self-identified Race | $White$ | Correct White | Mislabeled Black |
| | $Black$ | Mislabeled White | Correct Black |

With the following econometric specification:

$$Y_{i,j} = \beta_0 + \beta_1 I(Race_i^{Self-Report} = \text{Black}) + \beta_2 \underbrace{I(Race_i^{Self-Report} = \text{Black}) \times I(Race_i^{Agency} = \text{White})}_{\text{Mislabeled White}}$$
$$+ \beta_3 \underbrace{I(Race_i^{Self-Report} = \text{White}) \times I(Race_i^{Agency} = \text{Black})}_{\text{Mislabeled Black}} + X_i\delta + \epsilon_{i,j}$$

# Preliminary findings

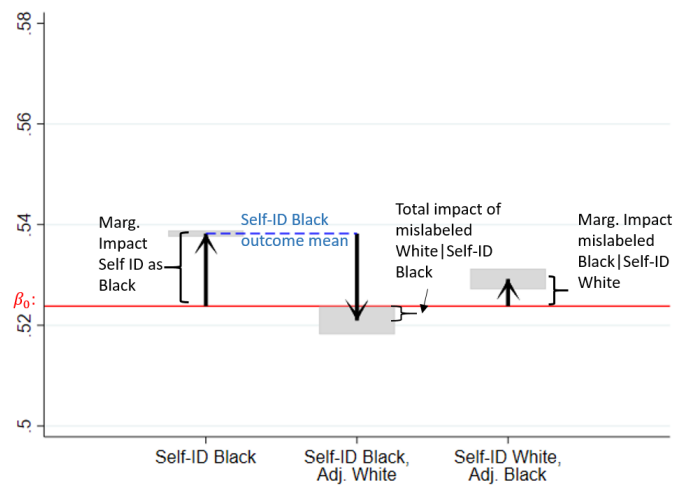
Impact of race/ethnicity discordance on likelihood of **felony-level charge**




All results were approved for release by the U.S. Census Bureau, Data Management System number: P-7500378 and approval number CBDRB-FY23-0178.

# Preliminary findings

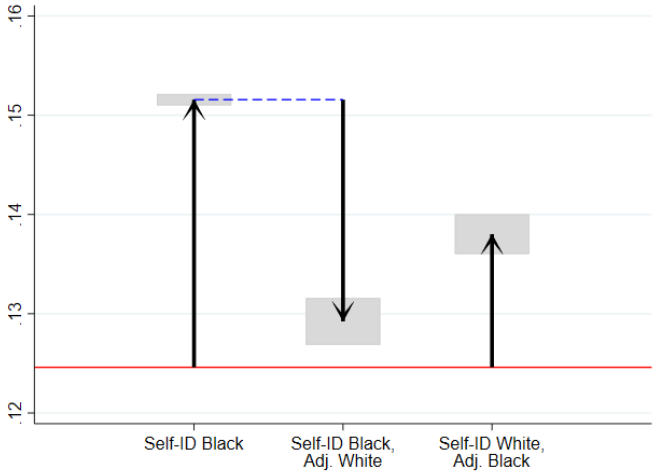
Impact of race/ethnicity discordance on likelihood of **conviction**




All results were approved for release by the U.S. Census Bureau, Data Management System number: P-7500378 and approval number CBDRB-FY23-0178.

# Preliminary findings

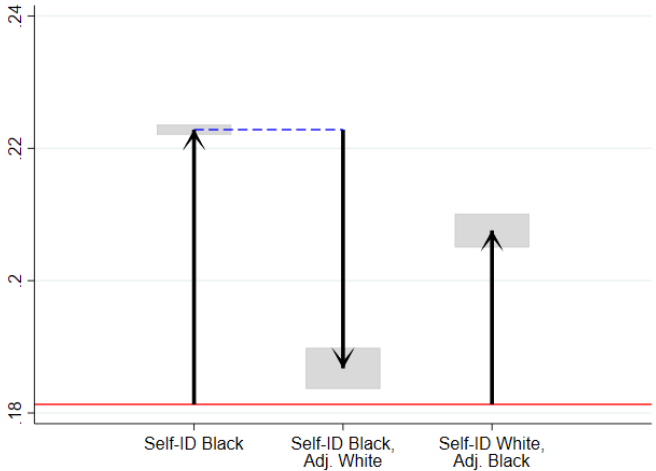
Impact of race/ethnicity discordance on likelihood of **violent charge**




All results were approved for release by the U.S. Census Bureau, Data Management System number: P-7500378 and approval number CBDRB-FY23-0178.

# Preliminary findings

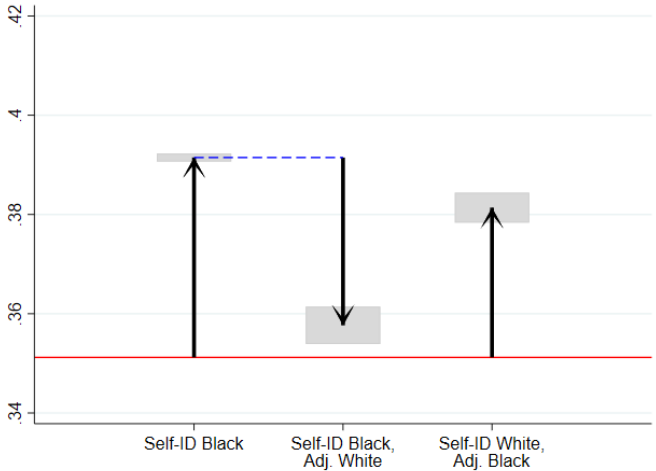
Impact of race/ethnicity discordance on likelihood of **drug charge**




All results were approved for release by the U.S. Census Bureau, Data Management System number: P-7500378 and approval number CBDRB-FY23-0178.

# Preliminary findings

Impact of race/ethnicity discordance on likelihood of **incarceration sentence**




All results were approved for release by the U.S. Census Bureau, Data Management System number: P-7500378 and approval number CBDRB-FY23-0178.

# Potential biases?

CJARS Criminal Justice Administrative Records System

1. PIK linkage errors:
   - "Mislabeling" is actually just mismatched records
   - Agency-recorded race is the better measure of true race
   - Evidence suggests this is not the case
     - Mislabeling uncorrelated with pre-existing IRS reported income levels

2. Law enforcement stereotyping based on physical appearance/socio-economic status:
   - Unclear if exercise captures racial discrimination or impact of socio-economic background on adjudication outcomes
     - Prediction of asymmetric effect of mislabeling by recorded race on pre-existing characteristics (criminal history, earnings, etc)
     - Exercise completed on IRE, but not yet disclosed

# Future Work

CJARS Criminal Justice Administrative Records System

1. Incorporate broader array of Census-based measures of race/ethnicity
   - 2000 Decennial, 2020 Decennial, SNAP/TANF/HUD/Medicaid admin data, Numident, Parental self-identification
2. Produce/disclose geographic/temporal variation in mislabeling rates and examine panel correlates
3. Extend individual impacts to post-adjudication outcomes: recidivism, employment, earnings
4. Extend individual impacts to other racial/ethnic categories, especially Hispanics and Native Americans

# Thanks!

CJARS Criminal Justice Administrative Records System

Questions, comments, and suggestions welcome.

- Mike Mueller-Smith, mgms@umich.edu, @Econ_Mike

To learn more about integrating CJARS into **your** research, check out our website:

- https://cjars.org