

Studying Discrimination: Fundamental Challenges and Recent Progress

Kerwin Kofi Charles and Jonathan Guryan

Annu. Rev. Econ. 2011.3:479-511

James J. Heckman



Econ 350, Winter 2021

1. INTRODUCTION

2. ASSESSING HOW MUCH RACIAL DISCRIMINATION EXISTS IN THE LABOR MARKET

2.1. Basic Setup

Empirical work seeking to measure the extent of labor market racial discrimination usually starts with a simple statistical model in which an individual i with observable traits X_{it} and unobservable characteristics ε_{it} receives wages (or some other labor market outcome) y at time t given by

$$y_{it} = \alpha X_{it} + \delta B_i + \varepsilon_{it}. \quad (1)$$

Here the binary variable B_i indicates the person's race and equals 1 if he is black. The parameter δ answers the ceteris paribus question of how, given the statistical model, an individual's wages would be different if he were black instead of white but otherwise remained exactly the same.

a regression performed on Equation 1 for a sample of blacks and whites will produce a biased estimate of d unless race is independent of the unobservables e , conditional on the observables X , or

$$\text{Cov}(B_i, \varepsilon_{it} | X_{it}) = 0. \quad (2)$$

The implication of the requirement of the race variable's conditional independence is that, in general, estimates of discrimination based on comparisons using observational data will be biased unless all determinants of the outcome that remain unobserved after conditioning on observable controls are on average equal for blacks and whites, or

2.2. The Taxonomical Challenge

Table 1 Responses to multiple- and single-race questions in the Current Population Survey, 2002 and 2003

White	98.14	6.66	28.77	11.20
Black	0.48	91.36	1.83	0.84
American Indian/Aleut/Eskimo	0.23	0.27	54.55	0.20
Asian/Pacific Islander	0.28	0.35	1.42	75.22
Hawaiian/Pacific Islander	0.05	0.08	0.25	5.94
White-Black	0.09	0.35	0.00	0.05
White–American Indian/Aleut/Eskimo	0.58	0.01	13.01	0.08
White-Asian	0.09	0.02	0.00	1.57
White-Hawaiian	0.03	0.00	0.00	1.12
Black–American Indian/Aleut/Eskimo	0.01	0.59	0.08	0.00
Black-Asian	0.00	0.05	0.00	0.10
Black-Hawaiian	0.00	0.00	0.00	0.03
American Indian/Aleut/Eskimo-Asian	0.00	0.01	0.00	0.00

Table 1 Responses to multiple- and single-race questions in the Current Population Survey, 2002 and 2003

	Race report in 2002			
Asian-Hawaiian	0.00	0.00	0.00	1.80
White-Black–American Indian/Aleut/ Eskimo	0.01	0.23	0.00	0.05
White–American Indian/Aleut/ Eskimo–Asian	0.00	0.00	0.00	0.05
White-Asian–Hawaiian Pacific Islander	0.00	0.00	0.00	0.03
White-Black–American Indian/Aleut/ Eskimo–Asian	0.00	0.00	0.00	0.00
Two or three races	0.01	0.00	0.08	1.60
Four or five races	0.00	0.01	0.00	0.13
Same as in 2002	98.14	91.36	54.55	75.22
Different from 2002	1.86	8.64	45.45	24.78

Table 2 Relationship between indication of single race in 2003 and individual characteristics

	(1)	(2)	(3)	(4)
	Blacks		Whites	
Dependent variable	Just black in 2003		Just white in 2003	
Log(wage) ₀₂	-0.008	-0.021	0.004	0.005
	(0.009)	(0.011)	(0.001)	(0.001)
Log(wage) ₀₃ - Log(wage) ₀₂		-0.023		0.002
		(0.010)		(0.001)
Education/10	0.040	0.056	0.012	0.011
	(0.022)	(0.023)	(0.003)	(0.003)
Pot. experience/100	0.077	0.103	-0.004	-0.008
	(0.134)	(0.134)	(0.019)	(0.019)
Pot. experience ² /100	0.002	0.002	0.000	0.000
	(0.003)	(0.003)	(0.000)	(0.000)
R ²	0.005	0.006	0.002	0.002
N	4,777	4,777	48,279	48,279

2.3. Limitations of Experimental Estimates

- Applications of this method relax the assumption in Equation 1 that the coefficient β on the control vector X is constant across races and assume instead that the labor market outcome received by a person of a given race β_i may be represented by a race-specific regression

$$y_{it}^B = \alpha_B X_{it}^B + \varepsilon_{it}^B, \quad B_i = \{0, 1\}, \quad (4)$$

$$E[\bar{y}_{it}^1 - \bar{y}_{it}^0] = \hat{\alpha}_1 \bar{X}_{it}^1 - \hat{\alpha}_0 \bar{X}_{it}^0 = \bar{X}_{it}^0 (\hat{\alpha}_1 - \hat{\alpha}_0) + \hat{\alpha}_1 (\bar{X}_{it}^1 - \bar{X}_{it}^0) \quad (5)$$

or

$$E[\bar{y}_{it}^1 - \bar{y}_{it}^0] = \hat{\alpha}_1 \bar{X}_{it}^1 - \hat{\alpha}_0 \bar{X}_{it}^0 = \bar{X}_{it}^1 (\hat{\alpha}_1 - \hat{\alpha}_0) + \hat{\alpha}_0 (\bar{X}_{it}^1 - \bar{X}_{it}^0), \quad (6)$$

where $\hat{\alpha}$ are the estimated regression coefficients.

*2.4. Limitations of Regression-Based
(or Selection on Observables)
Methods of Measuring Discrimination*

2.5. Is There Consensus About Existing Discrimination?

3. TESTING MODELS OF DISCRIMINATION

3.1. The Two Main Types of Discrimination Models

- Assuming blacks and whites to be perfect substitutes in production, an employer with prejudice $d_j \geq 0$ chooses black and white labor (L_b and L_a) to maximize

$$U_j = f(K, L_a + L_b) - w_a L_a - w_b L_b - d_j L_b, \quad (7)$$

where f is a constant returns to scale production function, and w_b and w_a are the wages for black and white labor, respectively.

3.2. Statistical Discrimination Models

- Firms know the distribution of true skill by race; in particular, suppose that for persons of race $x = \{a, b\}$ it is known that

$$q_x \sim \Omega\left(\bar{q}_x, \sigma_q^2\right), \quad (8)$$

where Ω is the cumulative distribution function of the normal distribution.

$$\tilde{q}_i = q_i + \eta_i \quad (9)$$

with the error of the signal, η , distributed

$$\eta_i \sim \Omega\left(0, \sigma_{\eta x}^2\right). \quad (10)$$

Firms pay wages equal to expected productivity, meaning that each person's wage y is given by

$$y_{ix} = E[q_{ix} | \tilde{q}_x, x] = \bar{q}_x \left(\frac{\sigma_\eta^2}{\sigma_q^2 + \sigma_\eta^2} \right) + \tilde{q}_x \left(\frac{\sigma_q^2}{\sigma_q^2 + \sigma_\eta^2} \right). \quad (11)$$

3.3. Testing Between Two Types of Theories

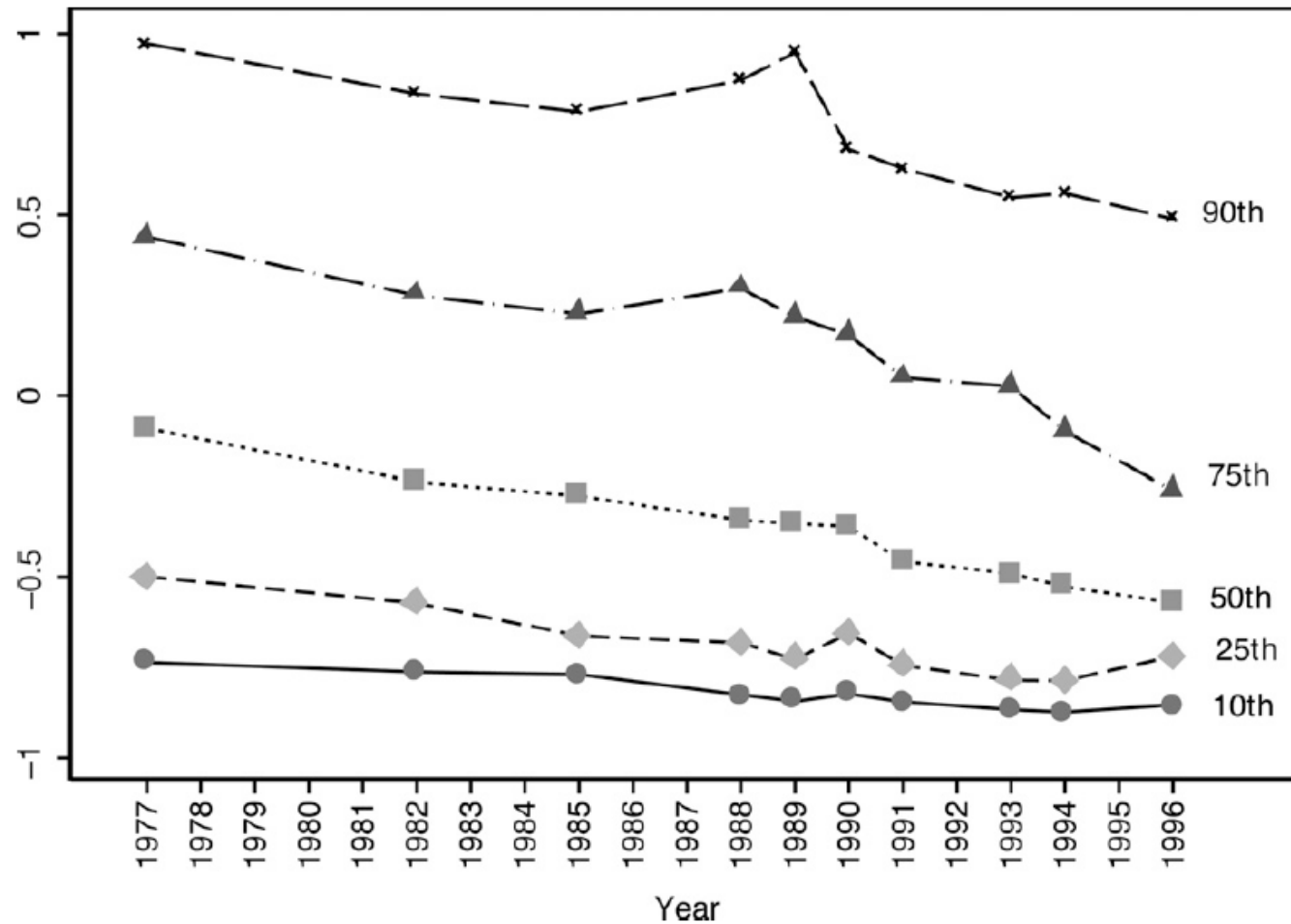


Figure 1

Trends in various percentiles of racial prejudice in the United States measured using data from the General Social Survey in certain years between 1977 and 1996. To calculate the 10th percentile of prejudice in the United States in a given year, the 10th percentile is first calculated for each state in that year. The weighted average is then taken for that year across all states, where the weight is the state population from the 1990 census. Each of the other measures (25th percentile, median, 75th percentile, and 90th percentile) is calculated analogously.

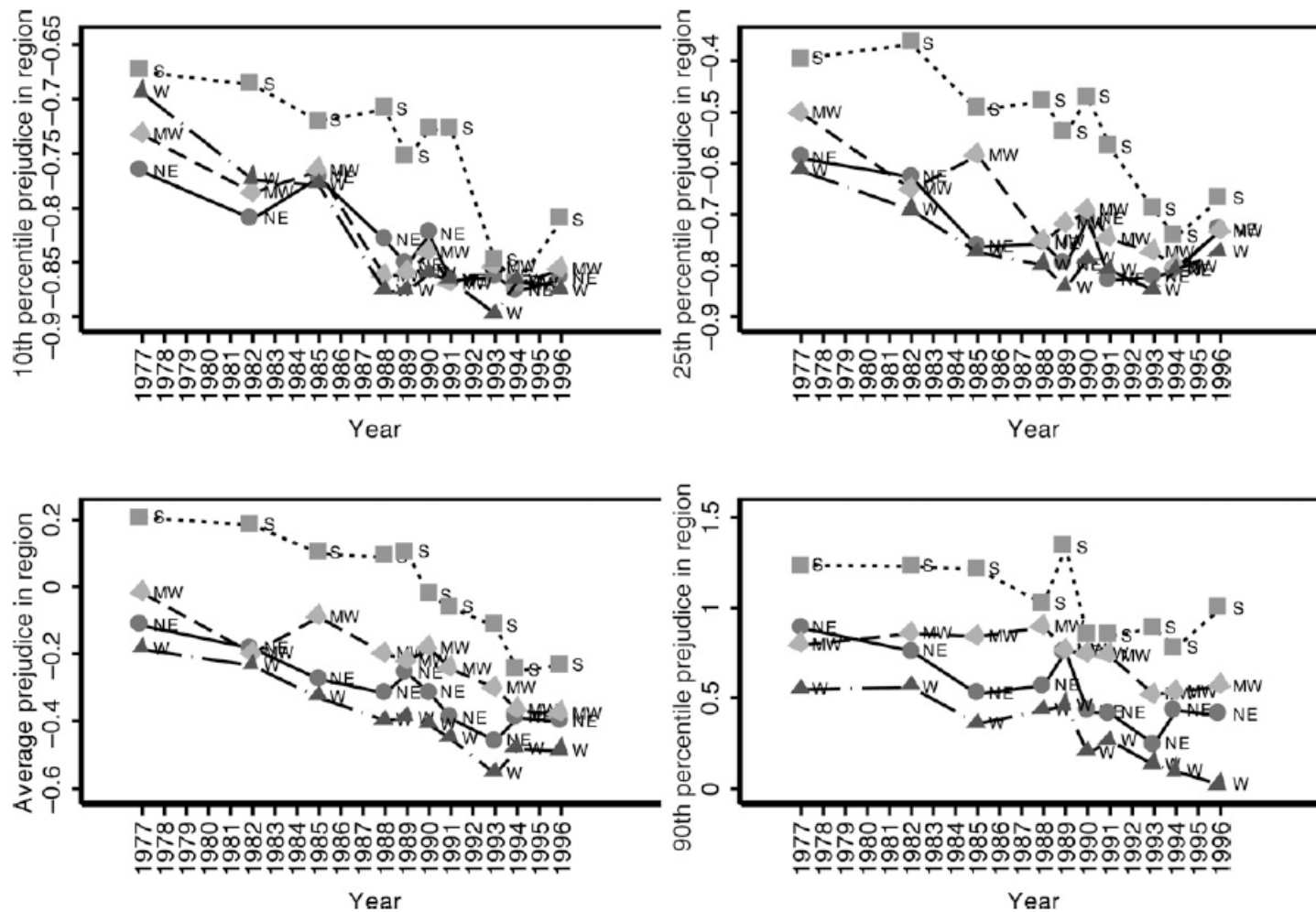


Figure 2

Trends in various measures of racial prejudice calculated from General Social Survey data for each of the four census regions in various years between 1977 and 1996. To calculate the 10th percentile of prejudice in a region in a given year, the 10th percentile is first calculated for each state in that year. The weighted average is then taken for that year across states within a census region, where the weight is the state population from the 1990 census. Each of the other measures (25th percentile, average, and 90th percentile) is calculated analogously. Abbreviations: MW, Midwest; NE, Northeast; S, South; W, West.

4. FINAL THOUGHTS