My Paper on NLSY02 Data

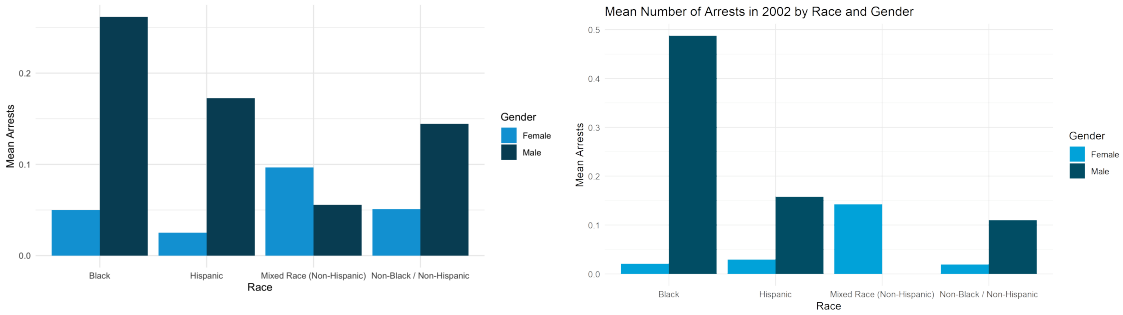
Seungwoon Shin

Sping 2022

1 Brief Analysis

The main objective of this paper is to compare the mean number of arrests between 1997 and 2002 by race and gender.

Figure 1: Mean Number of Arrests by Race and Gender



(a) 1997

(b) 2002

Table 1: Mean arrests in 1997 by Race and Gender

Gender	Black	Hispanic	Mixed Race Non Hispanic	Non Black Non Hispanic
Female	0.0500481	0.0251497	0.0967742	0.0510659
Male	0.2617230	0.1725441	0.0555556	0.1443401

Table 2: Mean arrests in 2002 by Race and Gender

Gender	Black	Hispanic	Mixed Race Non Hispanic	Non Black Non Hispanic
Female	0.0211268	0.0298013	0.1428571	0.0193192
Male	0.4876712	0.1579509	0.0000000	0.1099476

As we have seen Figure 1 and the Tables, the male is more likely to commit crimes than the female regardless of race and year except for mixed-race(non-Hispanic). The largest mean number of arrests is from Black male regardless of years. The mean number of arrests increased over time for Black males whereas it decreased for the rest of the other males. The mean number of arrests has increased for females regardless of the race except for Blacks.

2 Regression

Table 3: Regression Output. Omitted category is Black Females.

	Dependent variable:
	Arrests in 1997
Hispanic	-0.055^{***}
	(0.019)
Mixed Race (Non-Hispanic)	-0.084^{*}
	(0.043)
Non-Black / Non-Hispanic	-0.056^{***}
	(0.017)
Male	0.134^{***}
	(0.012)
Constant	0.087^{***}
	(0.013)
Observations	7,692
R^2	0.018
Adjusted R^2	0.017
Residual Std. Error	0.526 (df = 7687)
F Statistic	34.624^{***} (df = 4; 7687)
Note:	*p<0.1; **p<0.05; ***p<0.01

The sign of dependent variables indicates that Male with a positive coefficient is likely to commit crimes. Namely, it leads an increase in the mean number of arrests in 1997. In contrast, the rest induces a decrease in the mean number of arrests. The dependent variable 'Male' also partially implies the largest mean number of arrests from Blacks.

Table 4: Regression Output. Omitted category is Black Females.

	Dependent variable:
	Arrests in 2002
Hispanic	-0.159^{***}
	(0.038)
Mixed Race (Non-Hispanic)	-0.174^{**}
	(0.083)
Non-Black / Non-Hispanic	-0.189^{***}
	(0.035)
Male	0.194^{***}
	(0.022)
Constant	0.155^{***}
	(0.026)
Observations	8,621
R^2	0.015
Adjusted R^2	0.014
Residual Std. Error	1.019 (df = 8616)
F Statistic	32.033^{***} (df = 4; 8616)
Note:	*p<0.1; **p<0.05; ***p<0.01

Unlike previous Table 3, the absolute value of all dependent variable coefficients has increased. This indicates that some of the categories such as Black male remarkably increased in the mean number of arrests in 2002 while the other categories such as Mixed race (non-Hispanic) male decreased a lot.