My Paper on NLSY97 Incarceration Data

David Scolari

Sping 2022

1 Analysis

To prepare this data, I used raw incarceration history data from the 2002 National Longitudinal Survey of Youths (NLSY). After aggregating the data to obtain total months of incarceration time for each race/gender group, I normalize by population to obtain per person incarceration days.

Table 1 shows per person incarceration days for each racial group in 2002. Mixed race is omitted from the analysis due to limited sample size. The table shows that for each racial group, males spend more days incarcerated per individual than females. However, the increase in incarceration time for males compared to females is several times larger for black males than other other race.

This fact is highlighted by Figure 1, which shows the male to female ratio of days incarcerated for the three racial groups of interest. We see again that black males spent nearly 30 days incarcerated for every 1 day a black female spent incarcerated.

Table 2 displays regression results that show how expected days in prison changes for different race and gender groups. The expectations obtained from the regression results are different from those displayed in Table 1 because the regression is not estimated with population normalized incarceration days. However, the signs do match the patterns shown in Table 1 that Black males spent the most time incarcerated in 2002.

2 Tables and Figures

Table 1: Per Person Incarceration Days

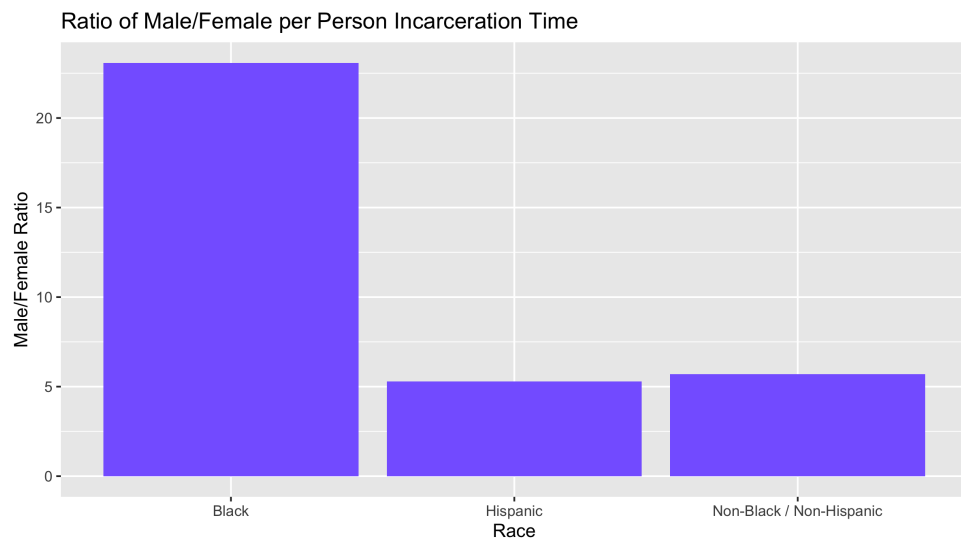| Race	Female	Male
Black	0.6426056	14.833333
Hispanic	0.9064570	4.804340
Non-Black / Non-Hispanic	0.5876265	3.344241

Figure 1: Ratio of Male/Female per Person Incarceration Time

Table 2: Regression Output. Omitted category is Black Females.

	Dependent variable:
	Per Person Incarceration Time in 2002 (days)
Hispanic	−4.823***
	(1.169)
Mixed Race (Non-Hispanic)	−5.300**
	(2.527)
Non-Black / Non-Hispanic	−5.737***
	(1.053)
Male	5.894***
	(0.671)
Constant	4.715***
	(0.777)
Observations	8,621
R^2	0.015
Adjusted R^2	0.014
Residual Std. Error	30.997 (df = 8616)
F Statistic	32.033*** (df = 4; 8616)
Note:	*p<0.1; **p<0.05; ***p<0.01