

My Paper on NLSY97 Data

Taylor Neal

Spring 2022

1 Introduction

In this report, we will utilize publicly available NLS data (from the National Longitudinal Survey of Youth 1997) to analyze incarceration rates in the year 2002. This exercise will seek to describe patterns revealed when examining this data across both race and gender. Incarceration status for each individual in the survey is provided on a monthly basis. Our primary variable of interest will be mean number of months spent incarcerated in 2002.

2 Results and Discussion

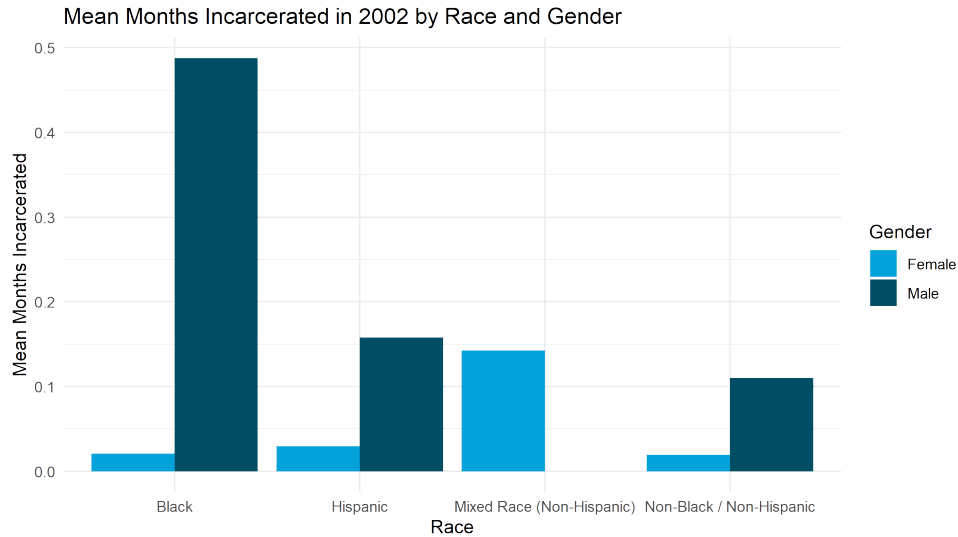


Figure 1: Mean Number of Months Incarcerated in 2002 by Race and Gender

The most striking aspect of the figure above is the relative height of the bar for black men. For individuals included in the survey, mean number of months incarcerated in 2002 was over twice as high for black men compared to the next highest race/gender (Hispanic men). The mean number of months black women spent incarcerated was dramatically lower than their male counterparts. This is generally true for the male/female comparisons in this data (with the exception of Mixed Race Non-Hispanic), but the distinction is greatest in the case of black women and men. An additional point to make is the Mixed Race (Non-Hispanic) bars for men and women both seem inconsistent when compared to the rest of the data. The female bar is significantly larger than the next largest bar for women. And the mean number of months men spent incarcerated shows as zero. This could indicate that mixed race men are incorrectly being coded as women in the data. Alternatively, this observed behavior might be due to under-representation of Mixed Race (Non-Hispanic) individuals in the survey making our observations prone to outliers. Note the table below displays numerical values for the figure on the prior page.

Table 1: Mean Months Incarcerated 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

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

	<i>Dependent variable:</i>
	Months Incarcerated 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)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

The table above displays results of an OLS regression with robust (to heteroskedasticity) standard errors. This simple model finds all statistically significant coefficients at the 0.01 level. This indicates that the differences we have observed between the mean months incarcerated for genders and races in this survey are highly statistically significant. Based on the positive coefficient for Male we find, as expected, that being male increases the expectation of mean months incarcerated. Additionally, being any race other than black decreases the expectation of mean months incarcerated. These regression result serve to increase our confidence in the significance of observed patterns plotted earlier in this section.