My Paper on NLSY97 Data

Shawn Leavor

February 2022

1 Introduction

Taking data from the NLSY97 on Incarceration Status, I was able to find differences in incarceration rates based on race and gender. I analyzed...

We run the following regression:

$$y = \beta_1 + \beta_2 Hispanic + \beta_3 MixedRace(Non-Hispanic) + \beta_4 Non-Black/Non-Hispanic + \beta_5 Male + \varepsilon$$

2 Analysis

Wherein we do tables and graphs. To include the graph we made in ggplot, we create the `figure` environment. The 'H' option tells LaTeX to 'hold' the position of the figure instead of positioning it somewhere else. I use the `caption` command to add a caption—although I also put a title on the plot in ggplot so you would typically choose one or the other. I use the `label` command after the caption to add a label. Then in my paper I can use the `ref` command and LaTeX knows I am referring to Figure 1.

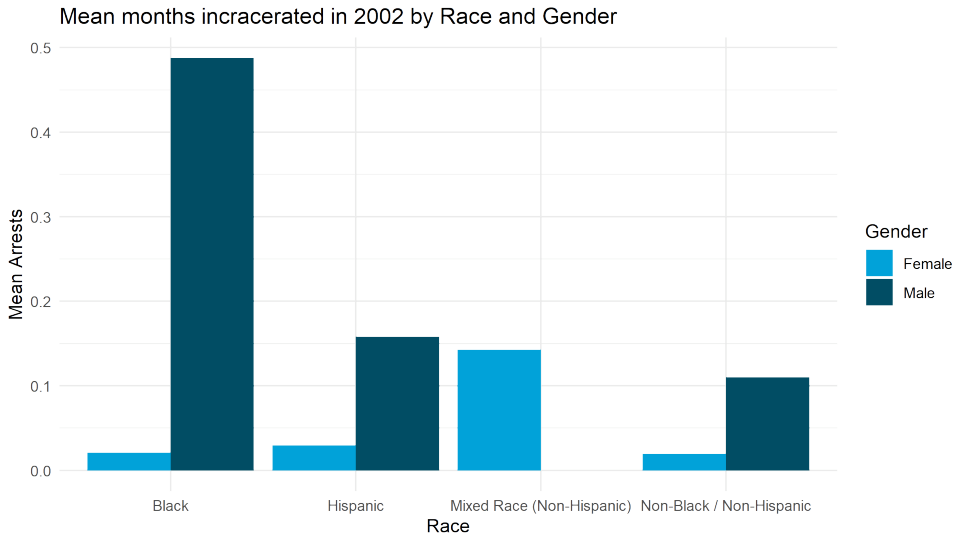

Figure 1: Mean Number of Months Incarcerated in 2002 by Race and Gender

The above graph shows the mean number of incracerations in 2002 based on race and gender. We see in every category except for Mixed Race (Non-Hispanic) that the mean number of incarceration months is higher for males than for females. We also see that this rate is much higher for black males than any other category. Interestingly, we see that male mixed race Non-Hispanic has an average incaceration months of 0.

Table 1: Mean months incracerated 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 1 is a numeric representation of Figure 1. This confirms that Black males spend the most average months incarcerated, while mixed race Non-Hispanic males spent no time incarcerated in 2002.

Table 2 is the result of a regression that shows the effects of race and gender based on the regression in Section 1. We find that the squared correlation is only 0.015 showing a poor fit of race and gender to the number of months incarcerated. However, we find that each of our p-values is significant, which tells us that race and gender have a large effect on the number of months incarcerated. On average, being male will increase the number of months incarcerated in 2002 by 0.194. We can also see that Black is the biggest factor in race, since it is the omitted variable and the other races have negative values. Our constant represents the average number of months incarcerated for a Black female.

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

	Dependent variable:
	Months incracerated 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