Punishment Past the Cell

An Analysis of Employment and Earnings of Ex-Offenders

David Stillerman

Economics Senior Integrative Exercise

February 28, 2014

Carleton College

Abstract:
This paper studies the labor market for ex-offenders and examines a potential source of statistical discrimination that stems from racial differences in rates of incarceration. Lacking perfect information about a candidate's criminal background, employers may instead rely upon visible characteristics, including race, to estimate the candidate's likelihood of having a criminal record. Theoretically, this would lead to employment and wage differentials among non-offenders of various racial groups. My empirical results are consistent with statistical discrimination theory, as differences in compensation and likelihood of employment are explained by race rather than the actual presence of an incarceration record.

I. Introduction

After reaching its peak of over 1.6 million in 2009, the United States prison population has been consistently shrinking (Carson and Golinelli, 2013).[1] As the number of incarcerated individuals declines, it will become increasingly important for policy-makers to assist ex-offenders in their transition back into the general population. Otherwise, it is likely that ex-inmates will revert to their old ways and end up back in jail. Government programs and non-profit organizations, such as the Office of Justice Programs and the Safer Foundation, seek to provide ex-offenders with the resources necessary to find employment or continue their education in the years following incarceration (Keilman, 2011). Additionally, groups have proposed "Ban the Box" legislation, which aims to prohibit employers from asking candidates about their criminal backgrounds (Harless, 2013). At the end of October 2013, the "Ban the Box" movement took a major step, as Target Corporation announced that it would remove all questions involving criminal history from its applications (Staples, 2013).

As the size of the prison population continues to fall, understanding the labor market for former inmates will become crucial. Because individuals of certain racial groups are overrepresented in the criminal justice system, it will be especially important to ensure that ex-offenders from all groups have an opportunity to reenter the labor force and receive compensation based upon their true productivity. My comps paper focuses on this particular aspect of the post-release labor market: the potential for racial discrimination. It analyzes the labor market implications of contact with the criminal justice system, incarceration specifically, in terms of both the post-release earnings and probability of employment of ex-offenders. It attempts to answer the following three questions: Does incarceration lead to a lower likelihood of

[1] As of 2012, the United States prisoner population was 1,571,013, which translates to 480 prisoners per 100,000 residents.

employment and, for those who are employed, lower earnings? Furthermore, does an individual's race alter the size of incarceration's impact on employment and earnings? Finally, do any residual effects of incarceration change the labor market outcomes of non-offending individuals from certain demographic groups?

Although not all employers have criminal history information about potential candidates, my data set contains incarceration status as a variable. This allows me to estimate relationships between incarceration and employment outcomes, but I must ground my findings in theory that considers the fact that not all employers have perfect information. In my analysis, I find that the presence of an incarceration record has an ambiguous relationship with the probability of employment of individuals from all racial groups in both 2006 and 2008. When considering hourly compensation, the data show a similar ambiguous result for white and Hispanic men. My findings do not suggest a statistically significant relationship between incarceration and earnings for individuals in these demographic groups. However, for black men, I find that those who have incarceration records have lower compensation levels. Supplemental analysis suggests that this discrepancy could exist because of differences in average lengths of incarceration. In other words, black males spend more time behind bars than white or Hispanic males, so they lose a greater amount of human capital. This leads to lower rates of hourly compensation for blacks. This explanation does not indicate preference-based discrimination, as individuals are paid their marginal product of labor.

Despite the absence of racial differences in employment outcomes of ex-offenders, there are clear employment and wage differentials between black and non-black individuals when considering the entire sample, not just those who have been incarcerated. My analysis suggests that one potential explanation for this trend is the use of race as a proxy for criminality in the absence of perfect information. This result implies that a difference in rates of incarceration among racial groups is another source of potential statistical discrimination. Economists have studied the general theory of statistical discrimination in multiple contexts. My paper focuses on one possible cause of this discriminatory behavior: the variance in rates of criminal behavior among individuals of different races.

The remainder of my paper is structured as follows: Section II summarizes relevant economic theory of the post-release labor market and previous empirical research on the topic, Section III presents my data set and model, Section IV describes my hypotheses, Section V discusses regression results and their implications, and Section VI describes potential limitations of my study and presents ideas for future research.

II. Relevant Theory and Empirical Work

Economists and other researchers have conducted a number of studies analyzing the impact of a criminal record on subsequent labor market outcomes, and their findings suggest a negative relationship between incarceration and employment results in the long run. Many of the studies summarizing the theoretical reasoning for the presence of crime, including Becker (1974) and Ehrlich (1996), as well as those inspecting the association between incarceration and employment, are written from an economic perspective. However, sociologists have conducted the majority of research involving discrimination in the post-release labor market. This leaves room for further examination by economists. My paper will expand on the ideas presented in these papers and ground them in economic reasoning, specifically the theory of statistical discrimination.

Schmitt and Warner (2011) and Raphael (2010) summarize the theoretical arguments for a negative relationship between incarceration and market outcomes such as employment and wages. When an individual is incarcerated, he or she must spend time away from the labor force. This leads to a loss of human capital. It can manifest itself in the deterioration of knowledge acquired via formal education, on-the-job skills, or intangibles such as organization and punctuality. The loss of these skills reduces the individual's marginal product of labor, which decreases a firm's willingness to pay for his or her services. This results in lower levels of employment and lower earnings for those who are employed.

Additionally, incarceration can sever an individual's social links. Without a strong network of connections, it can be difficult for an ex-inmate to navigate the job search process, which leads to a lower likelihood of finding employment. This theoretical outcome is supported by Calvó-Armengol and Zenou (2005). These researchers develop a stochastic model to describe employment outcomes in a network context. When an individual's current social network is smaller than a certain threshold level, Calvó-Armengol and Zenou determine that fewer connections lead to a lower probability of employment. Social links provide more avenues to learn about job opportunities. As long as an individual does not face congestion effects due to too many options, the size of the network is positively associated with likelihood of employment. Upon release from prison, ex-offenders likely have a small social network that could lead to legal employment. They may potentially have a large network of other ex-cons, but these connections would not lead to a higher likelihood of employment.

Finally, the presence of incarceration in an individual's history creates a stigma. Many employers are instinctively averse to hiring a candidate who has been incarcerated, as a criminal record indicates the potential for a laborer to be a liability for an employer. For instance, an individual who has served prison time for armed robbery could have a propensity for theft. This would be a trait to take into account when considering whether to hire someone for a job that involves handling money. Employers' hesitance to hire ex-offenders for such positions thus reduces the overall demand for them in the labor market.

Findings from empirical studies support the notion that an individual who has been incarcerated has a lower likelihood of employment and lower wages than an otherwise similar person with no criminal record. Freeman (1991) analyzes data from the National Longitudinal Survey of Youth (NLSY) of 1979 and finds that individuals who had been in jail or on probation in 1980 had a 19 percent lower probability of being employed in 1988 than those who had not been in contact with the criminal justice system.[2] Additionally, he indicates that spending time in jail reduced the number of weeks worked in a year by between 25 and 30 percent, while probation reduced it by between 10 and 15 percent. A part of this effect could be attributed to recidivism, or ex-inmates returning to prison following their initial release. Individuals who are in prison are technically not in the labor force. However, Freeman initially includes these cases in his analysis as working no weeks. This could lead to artificially low employment probabilities and weeks worked for ex-offenders. Freeman corrects for this by eliminating individuals in the data set who were incarcerated in 1988. This correction decreases the magnitude of incarceration's impact on employment, as Freeman concludes that individuals who had been in jail or on probation in 1980 had a 12 percent lower likelihood of being employed in 1988 than those with no criminal records. One final aspect of Freeman's analysis involves accounting for the endogeneity of incarceration. Unobserved factors that have an impact on the probability of incarceration are also likely correlated with the probability of employment. To account for this, Freeman employs an instrumental variables approach and still finds that incarceration has a

[2] Freeman controls for various demographic and lifestyle characteristics, including age, age-squared, race, marriage status, and substance use.

negative, statistically significant impact on likelihood of employment.[3]

In another analysis of the link between incarceration and employment outcomes, Raphael (2007) accounts for the endogeneity of a felony record in a different way than Freeman. He uses four separate specifications of a fixed-effects approach[4] and finds that people who have been to prison work between six and 14 fewer weeks per year than those who are otherwise similar but have never been incarcerated. He also observes a significant, negative effect of incarceration on wages, and the penalty ranges from 17 to 23 percent depending on the model specification. Raphael's findings are consistent with the research by Freeman and provide more support for the negative relationship between incarceration and both employment and wages.

While the empirical evidence discussed indicates that ex-offenders face a generally lower likelihood of employment and lower wages, Pettit and Lyons (2009) analyze whether or not these observed relationships change over time. They also examine how the associations vary with respect to the age at which one is admitted to prison. Using a fixed-effects approach[5], Pettit and Lyons find that incarceration has a negative relationship with wages for all ages of inmates, yet it has a positive effect on employment immediately following release for all but the inmates who were admitted at the youngest ages (20-24 years old). In other words, they determine that offenders are more likely to find work immediately following their release than before their initial contact with the criminal justice system. However, as time passes, the likelihood of

[3] Freeman examines longitudinal data and follows individuals before and after their period of incarceration. He addresses the endogeneity of 1980 incarceration and 1988 employment by using 1978 work experience as one instrument. He controls for 1979 work experience in the model, so including weeks worked in the previous year captures the unobserved qualities of an individual.

[4] The first specification includes fixed effects that adjust for changes in year-to-year characteristics that affect weeks worked and wages, the second adds individual fixed effects, and the third restricts the data set to contain only males who have complete survey information in each sample year. The fourth model builds upon the first three but includes only "at-risk" individuals, who are defined as those who eventually serve time in prison by the end of the survey in 1996.

[5] The approach by Pettit and Lyons employs individual fixed effects to account for "time invariant observed and unobserved characteristics that vary across individuals."

employment regresses to its pre-incarceration value and potentially falls below it. They provide two main explanations for this finding. First, upon release from prison, ex-offenders can participate in federal- and state-funded programs that provide employment resources, and the programs offer incentives to search for employment. Additionally, because individuals must endure a suboptimal, regimented lifestyle in prison, incarceration acts as a deterrent for future crime. Ex-offenders likely become more motivated to find a job following release because of a desire to avoid recidivism. This leads to an inflated probability of finding a job in the short-term. As time passes, these incentives disappear because the ex-offenders are no longer involved in post-release programs and memories of the conditions of incarceration are likely not as fresh in their minds. The loss of these two incentives causes the likelihood of employment to decline.

With respect to wages, Pettit and Lyons find a negative relationship with incarceration for all ages of ex-offenders. They control for work experience prior to conviction as well as conditions of confinement and conclude that the lower wages must be attributed to the stigma associated with spending time in prison. Previous work experience and confinement characteristics provide a proxy for the extent of human capital deterioration during incarceration. Therefore, a portion of the association between a criminal record and employment outcomes is left unexplained by differences in human capital levels before and after imprisonment. While Pettit and Lyons observe that the relationship between incarceration and employment and earnings is not uniformly negative, the findings are consistent over the long run with the theoretical framework set forth earlier in my paper.

Kling (2006) accounts for a potential confounding factor in his model of the post-release labor market. He examines whether the length of incarceration affects the relationship between incarceration and employment outcomes. This is relevant for two main reasons. First, longer stays in prison could lead to greater deterioration of human capital. Additionally, sentence length could indicate the severity of the crime committed, with longer sentences corresponding to more serious crimes. Both of these considerations suggest that length of incarceration has a negative association with wages and likelihood of employment. Empirically, Kling's results indicate that the length of an individual's stay in prison has no relationship with employment and earnings in the medium-term, which is seven to nine years after the individual was initially incarcerated. However, from one to two years after release, he finds that those who served longer prison sentences have higher levels of employment and higher earnings. These results are not intuitive with respect to the human capital aspect of the theoretical model of the post-release labor market. One potential explanation for the findings involves the post-release employment programs. If individuals receive resources and guidance to find employment, they may put forth more effort in the job search process. It could happen to be the case that individuals in the sample who spend more time in prison are the ones who are receiving these resources. The incentive effects provided by these programs could offset the loss of human capital, leading to the seemingly contradictory results.

Expanding upon the basic literature relating incarceration and employment outcomes, other researchers have examined the potential for discrimination in the post-release labor market. Using an audit study methodology, Pager (2003) explores the racial differences in the level of employer aversion to candidates with criminal backgrounds. She examines callback rates for pairs of candidates with the same qualifications and finds that the penalty for having a criminal record is 40 percent higher for blacks than for whites. Four male candidates applied for entry-level jobs in Milwaukee, two black and two white. From each of the pairs, one of the candidates presented himself as an ex-offender on a given application. Pager analyzes the proportion of applications from each demographic group that received further contact from the employer that could potentially lead to an offer of employment. Though she relies upon a small sample of only 350 observations, she finds that a criminal record reduces the likelihood of a callback by 50 percent for whites and by 64 percent for blacks. From this result, she concludes that employers prefer white candidates to black candidates and that employers penalize black candidates more for a criminal record than equally qualified white candidates.

Western (2006) expands upon the topics presented by Pager but uses longitudinal data instead of the audit study methodology. Western employs a fixed-effects approach to account for the endogeneity of incarceration, a popular technique in much of the literature.[6] He examines only data for men who are involved in crime at some point in the survey. This method is intended to control for the unobservable characteristics of the individuals, as those who are eventually involved in crime likely share certain traits that are difficult or impossible to quantify. Controlling in this way likely reduces the bias caused by endogeneity when estimating the relationship between incarceration and employment outcomes. With this data set, Western compares employment outcomes before and after an individual's prison term. He finds that incarceration reduces hourly wages by about 15 percent overall. The highest reduction occurs for Hispanics (24.7 percent), while the lowest occurs for blacks (12.4 percent). In further analysis, Western determines that incarceration reduces weeks worked by 15 percent for black and Hispanic men and 9.7 percent for white men.

In their papers, Western and Pager approach the subject of discrimination from the perspective of sociologists. The discussion involves animus-based discrimination, which is based upon the idea that employers inherently prefer candidates of a certain demographic group.

---

[6] Western's approach involves accounting for individual fixed effects, which are unobserved individual characteristics that are static over time. Additionally, he controls for age, education, industry of employment, and other exogenous characteristics.

This leaves room to incorporate economic analysis and interpretation, and I plan to address the topic through the theory of statistical discrimination. Phelps (1972) contributed the first model of statistical discrimination, which is based upon the concept of imperfect information in the labor market. Employers do not necessarily receive a precise representation of a candidate's characteristics and quality of work. They instead use their previous experiences with people of similar demographic groups as another indicator of performance.

Aigner and Cain (1977) refine the Phelps model, focusing on indicators of productivity and potential search costs associated with checking the validity of these indicators. The main premise involves a test score being a predictor of the quality of a candidate. The accuracy of the test score as an indicator of quality differs by race. Specifically, Aigner and Cain examine the case in which the test score is not as accurate a predictor of quality for blacks as it is for whites. In other words, the variance of quality about a given test score is higher for blacks than it is for whites. Therefore, employers incur greater search and screening costs to ensure that a black candidate is of high quality than they would if they were screening a white candidate. Because of this, when considering a black and a white candidate with equal productivity, employers offer lower wages to the black candidate.[7]

The statistical discrimination model is relevant for the examination of racial discrepancies in post-incarceration employment and income. Building upon the framework set forth by Aigner and Cain, consider the following situation from the perspective of an employer.

---

[7] One consideration for this model involves equal employment laws. If a court rules that an employer's test has adverse impact on individuals from a certain demographic group, then the test cannot be used. I do not believe this will affect the traditional statistical discrimination theory. Though some tests may be banned, employers use other signals (i.e. educational characteristics) to estimate the quality of a candidate. There is still variability in the predictive capabilities of these signals, which leaves room for employers to statistically discriminate.

Let E[δ] be the expected quality of work of a job candidate, Z be the individual's test score, and I be an indicator of whether an individual has been incarcerated. Consider the following model:

$$\mathrm{E}[\delta] = f(Z, I)$$

$$P(Employment) = g(\mathrm{E}[\delta], \mathrm{Var}[\delta \mid \mathrm{Z}])$$

Let the expected quality of work be positively related to the test score and negatively related to incarceration status. In this case, the test score can be an abstract concept, encompassing an individual's interview performance, previous experience, or performance on any employment tests administered by the firm. Let an individual's probability of employment have a positive relationship with expected quality and a negative relationship the variance of quality given a certain test score. For simplicity, I will examine the argument in relation to probability of employment. The same logic applies to earnings.

First, consider the case in which employers have full information about candidates' criminal histories. Assume equal average test scores across races. In this case, the probability of employment will only differ based upon incarceration status. Individuals who have been incarcerated will have a lower expected quality, so they will have a lower probability of employment. In terms of the model,

$$\mathrm{E}[\delta \mid \mathrm{Incarcerated}] < \mathrm{E}[\delta \mid \mathrm{Never\ Incarcerated}]$$

$$P(Employment \mid Incarcerated) < P(Employment \mid Never\ Incarcerated)$$

Now, consider the case in which employers have a test score but no information about criminal history. This presents the opportunity for statistical discrimination. The model must change to account for the imperfect information:

$$\mathrm{E}[\delta] = f(Z, E[I])$$

$$P(Employment) = g(\mathrm{E}[\delta], \mathrm{Var}[\delta \mid \mathrm{Z}])$$

In this new situation, the expected quality of work is again positively related to the test score but negatively related to the expected incarceration status. The probability of employment has a positive relationship with expected quality and a negative relationship with the variance of quality around a given test score.

Incarceration rates differ by race, and this affects two variables in the model. Blacks are incarcerated at a higher rate than whites so, in the absence of perfect information, employers will expect individuals from these demographic groups to have incarceration records. This leads to a lower expected quality for black individuals when compared to whites. It follows that the probability of employment will differ in the same way as expected quality. However, this does not necessarily indicate discrimination because individuals are hired at rates in accordance with their expected quality.

As articulated by Aigner and Cain, discrimination can arise due to a discrepancy in the variance of quality about a given test score. The differential in incarceration rates leads to this discrepancy in variance. Specifically, $\mathrm{Var}[\delta \mid Z, \text{black}] > \mathrm{Var}[\delta \mid Z, \text{white}]$ because it is more likely that a black candidate will have a criminal record than a white candidate. When an employer can only consider test scores and not criminal history, individuals who have been incarcerated appear to be of higher quality than they actually are. For this reason, more black individuals have test scores that do not directly correspond to quality of work. It follows that the test score is a worse predictor of quality for blacks than it is for whites. In turn, employers face higher screening costs for black candidates. For this reason, discrimination is present. If a black and a white candidate have equivalent expected quality, the black individual will face a lower likelihood of employment.

Holzer et al. (2006) explore the connection between incarceration and statistical discrimination in more detail. They examine the impact of racial differences in incarceration rates among the entire population on the labor market prospects of those who have not been in contact with the criminal justice system. They base their model on the theory of statistical discrimination and analyze whether employer access to criminal history information affects the employment rate of non-incarcerated individuals from certain demographic groups. They find that employers who conduct criminal background checks are generally more likely to hire black candidates. They take this as evidence of statistical discrimination. When employers do not conduct such checks, they instead take demographic factors into account when evaluating the expected productivity of a candidate. If an employer extrapolates the average incarceration rate of a demographic group to an individual, providing the employer with true criminal background information would increase the likelihood of hiring a non-offending individual from that group. I will use some of the theoretical aspects of the model of statistical discrimination presented by Holzer et al. to support my analysis, but I will instead approach the topic using longitudinal data rather than a survey of employers.

III. Data and Methodology

To perform my study, I use data from the 1997 National Longitudinal Survey of Youth, which follows a cohort of young men and women (between ages 12 and 16 at the beginning of the survey) from 1997 until the present day. I include only the male respondents, as the sample size for incarcerated females is comparatively small and could skew results. I first isolate a single year, specifically 2006, in order to keep consistent the macroeconomic conditions that could affect the difference in employment rates between demographic groups. In particular, I examine a year prior to the onset of the most recent financial crisis because the data taken from the later survey rounds could be biased due to abnormal labor market conditions during the recession and recovery period. Fewer jobs were available in the years following the financial crisis and, for this reason, candidates likely underwent more intensive screening processes. In this case, I would expect that an incarceration record would have a greater impact on income and likelihood of employment than it would under typical labor market conditions. Following the initial analysis of 2006, I examine the data from 2008 to determine whether the financial crisis spurred structural changes to the post-release labor market. In both of those years, I exclude members of the survey who were enrolled in school, those who were incarcerated, those who were living in a foreign country, and those who were in the military, as these individuals technically were not in the United States civilian labor force.

III.i Regression Model

I model both employment rate and hourly compensation as a function of educational, demographic, geographic, job-specific, and criminal-history characteristics. The unobserved factors that have an impact on the probability of incarceration are also likely correlated with earnings and the probability of employment. To account for this endogeneity, I use the two-stage instrumental variable regression technique. Specifically, I run the following pairs of regressions separately for each year of data:

(i)

$$Logit(C_j) = \propto_0 + \propto_1 I_j + \propto_2 E_j + \propto_3 D_j + \propto_4 G_j$$

$$Logit(Employment) = \beta_0 + \beta_1 \widehat{Prob(C_j)} + \beta_2 \widehat{Prob(C_j)}\ x\ Race_j + \beta_3 E_j + \beta_4 D_j + \beta_5 G_j$$

(ii)

$$Logit(C_j) = \propto_0 + \propto_1 I_j + \propto_2 E_j + \propto_3 D_j + \propto_4 G_j + \propto_5 J_j$$

$$Log(Compensation) = \beta_0 + \beta_1 \widehat{Prob(C_j)} + \beta_2 \widehat{Prob(C_j)}\ x\ Race_j + \beta_3 E_j + \beta_4 D_j + \beta_5 G_j + \beta_6 J_j$$

Both regressions (i) and (ii) use the incarceration rate by race/ethnicity and region as an instrument (I_j) to account for the endogeneity of criminal justice-system contact (C_j), controlling for educational (E_j), demographic (D_j), and geographic (G_j) variables. The instrument should not be directly correlated with either compensation or likelihood of employment for the individual, but it has a clear relationship with the probability of incarceration. These characteristics make it a plausible choice as an instrument for the two-stage regressions.

As shown in the second part of regression (i), I estimate the odds of an individual earning income from a job in the previous year (which represents whether or not he was employed) as a function of the estimated likelihood of incarceration, an interaction between the estimated likelihood of incarceration and race, educational (E_j), demographic (D_j), and geographic (G_j) variables. In regression (ii), I estimate log(*Compensation*) where *Compensation* refers to the average hourly earnings an individual received from wages, salary, tips, and bonuses in the previous year. Because the distribution of *Compensation* is heavily right-skewed, I use a log-transformation. The variables involving criminal history, education, demographics, and geography remain the same as those in specification (i). However, I also include job-specific characteristics (J_j) in this equation to account for the wage differentials among various types of jobs.

One shortcoming of this initial model is the lack of a variable describing the amount of time an individual spends in prison. This is theoretically important because it addresses the human capital deterioration that an individual undergoes while incarcerated. Simply adding a variable indicating the amount of time an individual has spent behind bars is not possible. The length of incarceration is zero for all individuals who do not have a criminal record, and this creates perfect collinearity with the variable describing whether one has ever been incarcerated.

Instead, I create a second model that replaces the likelihood of incarceration variable with length of incarceration. I run the following two regressions for each year of data:

(iii)

$$Logit(Employment) = \beta_0 + \beta_1 L_j + \beta_2 L_j \, x \, Race_j + \beta_3 E_j + \beta_4 D_j + \beta_5 G_j$$

(iv)

$$Log(Compensation) = \beta_0 + \beta_1 L_j + \beta_2 L_j \, x \, Race_j + \beta_3 E_j + \beta_4 D_j + \beta_5 G_j + \beta_6 J_j$$

As shown above, I estimate hourly compensation and the likelihood of employment as a function of length of incarceration (L_j) and the same controls as in the original models. I do not employ the two-stage regression technique for this new model because the instrument is no longer valid. While the incarceration rate by race and region is related to the likelihood of incarceration, it does not necessarily have a relationship with the length of time that an individual spends behind bars. By only using single-stage regressions, I am ignoring potential endogeneity. Coefficient estimates may be biased, but they will still provide information about whether controlling for duration of incarceration vastly alters results.

III.ii Main Response and Independent Variables

The NLSY provides hourly compensation rates calculated from weekly data on earnings and hours worked. For each job, respondents report their base rate of weekly pay plus tips, bonuses, commissions, and overtime, as well as the total number of hours worked per week. From this, the NLSY calculates an hourly compensation value for each job. In my model, I create the compensation variable by weighting the hourly rate given in the NLSY by the total hours worked per week in each job. For example, if an individual works 15 hours per week at job A, which pays $12 per hour, and 25 hours per week at job B, which pays $10 per hour, then the hourly compensation is calculated as follows:

$$\left(12\ x\ \frac{15}{40}\right)+\left(10\ x\ \frac{25}{40}\right)=\$10.75\ per\ hour$$

The employment variable follows directly from compensation. Respondents who list hours worked, and therefore have positive earnings, are considered employed. Those with no hours worked, and consequently no compensation, are listed as unemployed. As mentioned previously, I correct for individuals being out of the labor force by removing those who are in prison, in the military, or enrolled in school in the given year. In 2006, my data yield an unemployment rate of 8.6 percent. The sample's 2008 unemployment rate is 11.1 percent, which supports a change in the overall labor market environment between 2006 and 2008.[8]

I create the three main independent variables (incarceration status, length of incarceration, and race) using information directly from the NLSY. Each month from 1992 until the present, respondents listed whether or not they were incarcerated. From this information, I create the variable that indicates whether an individual has spent any time behind bars as well as one that lists the total number of months he has been incarcerated. The proportion of the sample with incarceration records is 6.9 percent in 2006, while it increases to 9.0 percent by 2008. The race variable divides individuals into one of four categories: black, Hispanic, non-black/non-Hispanic, and mixed race (non-Hispanic). I consider only those in the first three groups as to remove any confusion associated with racial categorization. In both the 2006 and 2008 data sets, approximately 26 percent of respondents are black, 21 percent are Hispanic, and 52 percent are non-black/non-Hispanic.

Before controlling for confounding factors, it appears as though incarceration has a minimal impact on the probability of employment in many cases. In 2006, when considering

[8] Though these unemployment rates are higher than the national averages, the NLSY oversamples individuals from low-income backgrounds. Therefore, it makes sense that there would be a greater number of individuals without jobs.

each racial group separately, the employment rate of individuals with an incarceration record is no more than 2.7 percent lower than that of non-offenders. For black individuals, a criminal record actually increases the likelihood of employment from 86.3 percent to 87.7 percent in 2006. The 2008 data tell a different story. Black respondents with criminal justice-system contact face an 11.9 percent lower probability of employment than those with no criminal record. White and Hispanic ex-offenders also face a lower likelihood of employment than the non-offenders in the sample. Figure 1 displays employment rates broken down by race and incarceration status.

	2006		2008	
	Never Incarcerated	Incarcerated	Never Incarcerated	Incarcerated
Non-Black/ Non-Hispanic	93.3%	90.6%	92.2%	88.4%
Black	86.3%	87.7%	82.4%	70.5%
Hispanic	93.5%	91.3%	91.2%	88.1%

Figure 1: Employment Rates by Race and Incarceration Status

When looking at compensation data, on first glance it appears as though incarceration has a negative relationship with hourly rate of pay. Additionally, this relationship seems to be stronger for individuals of certain races. The most extreme example is that of blacks in 2006. The mean wage for those who have never been incarcerated is over 8 dollars higher than that for ex-offenders. In the same year, whites who have been incarcerated receive a compensation penalty of less than a dollar. Figure 2 summarizes the initial data analysis of average hourly compensation broken down by race and incarceration status.

	2006		2008	
	Never Incarcerated	Incarcerated	Never Incarcerated	Incarcerated
Non-Black/Non-Hispanic	$19.80	$18.82	$23.30	$18.89
Black	$20.98	$12.94	$19.11	$18.16
Hispanic	$16.47	$13.65	$19.76	$15.96

Figure 2: Average Hourly Compensation Rates by Race and Incarceration Status

These initial values do not take confounding factors into account, and it is therefore difficult to draw any conclusions about the relationship among race, incarceration, and employment outcomes. For example, across races, respondents could have different average levels of education. Without controlling for this confounding variable, it could appear as though there is racial variance in employment rates when the discrepancies should actually be attributed to differences in education. Therefore, without adjusting for potential scenarios like this, it is not possible to draw inferences about the relationship between race and post-release employment outcomes. However, the figures provide a solid foundation for more comprehensive regression analysis.

III.iii Instrumental Variable

The incarceration rate by race and region instrument includes four regions (Northeast, North Central, South, and West) and separates each of the regional rates by race/ethnicity (black, Hispanic, non-black/non-Hispanic). I create this variable using data from both the Bureau of Justice Statistics' Prison and Jail Inmates at Midyear 2005 and the Census Bureau State Tables for 2006 and 2008. Prison and Jail Inmates at Midyear 2005 lists the number of inmates per

100,000 residents by state and race.[9] Though these statistics are not available for 2006 and 2008, the 2005 rates provide a reasonable estimate. The criminal justice system did not begin to decrease in size until 2009, so the general trend is likely consistent in the four years prior. The Census Bureau lists population estimates by state and race for each year, and I use the 2006 and 2008 versions to create the instrument.

Using these sources, I create a weighted average of incarceration rates by race and region. For each given race, I weight a state's incarceration rate using the proportion of individuals in the region who live in the state.[10] The incarceration rates vary greatly by race and region. In both 2006 and 2008, white individuals in the Northeast face the lowest incarceration rate of 225 per 100,000 residents. Blacks in the North Central region experience the highest rate. In 2006, 2,368 blacks were incarcerated per 100,000 residents. The weighted average is higher in 2008, with 2,375 per 100,000 black individuals incarcerated in the region.

III.iv Control Variables

I control for an individual's education by using his highest grade completed. This should theoretically have a positive relationship with both hourly compensation and probability of employment. A person with more education will have a higher level of human capital, which is positively correlated with the marginal product of labor. Individuals with high marginal products of labor should garner higher wages and likelihood of employment, as firms will have a greater willingness to pay for their labor.

[9] New Mexico and Wyoming are excluded from these data because they do not provide information about race.

[10] The Northeast region contains individuals from Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The North Central includes Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, Ohio, North Dakota, South Dakota, and Wisconsin. The South contains Alabama, Arkansas, Delaware, the District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Texas, Virginia, and West Virginia. The remaining states, Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, Oregon, Utah, and Washington, make up the West region.

The demographic variables are the highest grade completed for each of the respondent's parents, the relationship to the people with whom the individual was living at age 12, total number of biological children, age, and age-squared. The first two variables are meant to capture an individual's family background. Specifically, the relationship variable categorizes respondents in four ways: both biological parents, two parents with one biological, one biological parent, and other. Those who grow up in a two-parent household will likely have a stronger support system and network, which will lead to more job opportunities. For the same reason, parents' highest grade completed should have a positive association with wage and likelihood of employment. Including the individual's total number of children is meant to account for the outside incentives to find a job. If a person has a child to support, then it is more likely that he will work to find employment despite any potential search costs. The final demographic variable to consider is age. By itself, age typically has a positive association with employment and wages, while age-squared is expected to be negative. People begin by climbing up the employment ladder but eventually reach a peak. After this point, human capital begins to deteriorate, causing wages and likelihood of employment to fall (Becker, 1962). While these theoretical relationships hold over long periods of time, in my data, there is not a large amount of variation in age. In the 2006 set, the ages range from 22 to 26, while they span from 24 to 28 in 2008. Therefore, I do not necessarily expect to see the typical impact of age and age-squared. However, it is still important to control for this possibility.

With respect to the geographic characteristics, I control for region of residence, whether the respondent moved in the previous year, and whether the respondent lived in an urban or rural area. Region is comprised of the four categories mentioned previously: Northeast, North Central, South, and West. Employment rates and compensation could vary by region, so it is important to account for this in the model. Theoretically, earnings should be higher in areas with a high cost of living. For instance, it is more expensive to live in the Northeast than the South. Compensation patterns should reflect this with hourly rates being higher in the Northeast. The migration variable accounts for an individual's movement history. Specifically, it categorizes respondents based upon whether they have not moved in the last year, moved within the same state, moved between states, or moved from a foreign country. I hypothesize that an individual who has moved in the past year will have lower compensation and likelihood of employment. Unless the individual moved for a job (i.e. changing offices), he would face a transition period before finding employment and gaining recognition within the firm.

The urban/rural variable specifies whether an individual lived in a CBSA, which is equivalent to Metropolitan Statistical Area, as defined by the 2000 census. More jobs are available in urban areas, so I hypothesize that an individual living in a CBSA will be more likely to have a job than one living in a rural area. The data support this relationship. In Pennsylvania, the rural unemployment rate was over 1.5 percent higher than the urban rate for each year between 1976 and 2002 (Penn State University College of Agricultural Sciences, 2002). The cost of living is also higher in cities, which likely leads to wage differentials. Firms in urban areas must account for the cost of living, meaning wages will be higher for individuals in urban areas.

For the regression of hourly compensation, I also include job-specific characteristics to account for wage variability across occupations. In particular, I include a categorical variable that specifies the industry in which the respondent was employed based upon the first digit of the SIC code from the 2002 census. For respondents who listed multiple jobs, I use the industry of the one listed as their first occupation. I expect that individuals employed in manufacturing will, on average, earn less than those who work in white-collar occupations.

III.v Missing Variables and Data Set Consolidation

As is the case for many longitudinal studies, missing data plague the NLSY. The missing values take two main forms: item nonresponse and unit nonresponse. Item nonresponse occurs when an individual participates in a round of the survey but fails to answer one of the questions. Unit nonresponse refers to the situation in which an individual does not participate in one or more rounds of the survey (De Leeuw, 2001). Both types of missing responses influence the inference I can perform using the data. The linear model assumes that a sample is random and representative of the population. This assumption may not hold if observations are missing in a systematic fashion (De Leeuw, 2001). For example, if individuals from a certain race are more likely than others to be missing demographic information, removing them from the analysis could lead to biased results.

In my analysis, I correct for missing covariates in two ways. First, for a number of variables, I use the previous year's observation carried forward. It is likely that an individual's region, highest grade completed, number of children, and urban/rural status did not change over that year. Therefore, I estimate these missing values by using the recorded response from the previous year if it was listed. Despite this initial correction, a significant number of missing data points remain. A summary of missing variables in my data set following the observations carried forward can be found in Appendix A. For the remainder of the missing covariates, I use a multivariate imputation method using the 'MICE' package in R. This process employs Gibbs sampling to estimate each missing value using the other data as predictors. The package generates a probability distribution of possible values for each missing data point and assigns the value with the highest likelihood (Van Buuren and Groothuis-Oudshoorn, 2011). By relying upon the multivariate imputation method, I avoid the bias associated with simply removing observations.

While it is plausible to use imputation methods to address missing covariates, I do not use this same process for missing response variables. The 2006 data set is missing both employment and compensation data for 832 individuals and an additional 277 are missing only compensation rates. There are 783 respondents missing both in 2008, while only 167 are missing compensation alone. I choose not to impute values for the response variables. While it is statistically valid to do so, estimating values for the variable the model is built to predict is theoretically questionable. When attempting to isolate the relationship between the responses and given covariates, using imputation to guess values for earnings and likelihood of employment essentially assumes away what the model is meant to estimate. For this reason, I proceed using only the respondents with complete information about employment and earnings.

In addition to removing observations with missing values for employment outcomes, I also eliminate outliers in the compensation regression. A number of respondents list either extremely low or extremely high compensation rates. I remove all observations with compensation rates less than \$1.50 per hour and greater than \$1,000 per hour. I believe it is safe to assume that these values are miscoded. Even if they are the truly correct values, they heavily influence the regression results and could skew coefficient estimates away from the true population values. They also create problems with heteroscedasticity, which are completely cured after removing the outliers. A full summary of data set consolidation and outlier removal is provided in Appendix B.

IV. Hypotheses

My coefficients of interest will be those on the incarceration term, the race term, and the race/incarceration interaction term. I hypothesize that the first coefficient should be negative and significant, as ex-offenders should, on average, have lower rates of employment and lower incomes than those who are otherwise similar but have had no contact with the criminal justice system. The theory and previous empirical work suggest a negative relationship between contact with the criminal justice system and earnings. While the research by Pettit and Lyons finds a short-term post-release employment boost for most ex-offenders, this result is not visible for the youngest individuals (20-24 years old). The majority of my data set will consist of individuals in this age bracket, and therefore I expect to see results consistent with traditional theory.

I do not expect to see any racial differences in the impact of incarceration on post-release employment and weekly income. In other words, I hypothesize that the coefficient on the interaction term will be insignificant for two main reasons. First, assume that employers have full information about a candidate's criminal history. Thus, the incarceration variable itself should pick up the entire effect of an employer's response to an individual's criminal record. Second, assume that employers do not have perfect information about a candidate's incarceration history. Therefore, part of the relationship between a felony record and employment outcomes must be picked up elsewhere. Because average incarceration rates differ based upon race, one variable that could account for part of the incarceration effect is race. My data set likely contains some employers who do check criminal background information and others who do not. A survey conducted by the Society for Human Resource Management found that 14 percent of firms did not check the criminal history of potential employees (Society for Human Resource

Management, 2012). In this situation, I would not expect to see any additional response to appear on the interaction term that includes both an individual's race and incarceration record.

V. Results

As described in the methodology, I employ a two-stage logistic regression technique to test my hypotheses regarding likelihood of employment. After controlling for demographic, geographic, and educational characteristics, I examine whether contact with the criminal justice system, race, and an interaction of the two are related to an individual's likelihood of employment. Results involving the coefficients of interest for the 2006 and 2008 data sets are listed in Figure 3. Full results can be found in Appendix C.

	2006	**2008**
Ever Incarcerated	0.620	0.644
	(1.630)	(1.672)
Race - Black	-0.309	-0.544**
	(0.221)	(0.218)
Race - Hispanic	0.231	-0.151
	(0.285)	(0.271)
Black x Ever Incarcerated	-1.990	-1.427
	(1.574)	(1.381)
Hispanic x Ever Incarcerated	0.902	1.055
	(2.251)	(1.851)
Number of observations	3016	3152

Two-stage logit estimates
*p<0.10; **p<0.05; ***p<0.01

Figure 3: Two-Stage Regression Estimates for Likelihood of Employment

In the 2008 specification, the model returns a significant coefficient for only the *Race - Black* variable. The results imply that a black individual faces a lower likelihood of employment than an otherwise similar white or Hispanic individual. Specifically, after controlling for educational, demographic, and geographic factors, the odds of employment for black males are 42.0 percent than those for white or Hispanic males. Despite the initial racial differences in the likelihood of employment, the impact of a criminal record does not appear to differ by race. This is represented by the insignificant coefficients on the *Ever Incarcerated x Race* interaction terms. The analysis of the 2006 data set does not return any significant coefficients of interest, which could indicate a fundamental change in the labor market dynamics in the two years between the survey rounds.

The 2008 results support my original hypothesis and can be explained within the context of statistical discrimination. Although not all employers have criminal history information, my data set does. By quantifying the relationships among employment, race, and incarceration status, I can analyze the potential for statistical discrimination in the post-release labor market. Employers who lack perfect information about criminal history could instead rely upon observable characteristics, such as race, to provide information about an individual's likelihood of being an ex-offender. The incarceration rate for blacks is higher than that for whites and Hispanics. In the 2008 data set, 10.7 percent of black respondents have a criminal record, while the rates are 9.8 percent and 7.8 percent for Hispanics and whites, respectively. For this reason, in the absence of perfect information about criminal history, employers are more likely to believe that black individuals have incarceration records. In terms of my model:

$$E[I \mid \text{black}] > E[I \mid \text{white or Hispanic}]$$

Additionally, it is more likely that the test score will not be an accurate predictor of quality for blacks than it is for whites or Hispanics. It follows that:

$$\text{Var}[\delta \mid Z, \text{black}] > \text{Var}[\delta \mid Z, \text{white or Hispanic}]$$

As expected incarceration status and variance of quality about a given test score are both negatively related to the probability of employment, blacks face a lower likelihood of employment than whites or Hispanics.

Comparing the 2006 and 2008 results, a fundamental difference is visible with respect to the coefficient of the *Race-Black* variable. There is evidence of structural change in the post-release labor market in the two years that elapsed. One explanation for this difference falls within the statistical discrimination framework set forth earlier in the paper. It could be the case that an incarceration record does not matter to the employers in the 2006 data set because labor market conditions are such that screening is less important. More jobs are available and fewer individuals are competing for employment. Therefore, it is likely less crucial for employers to conduct intensive screening procedures. In terms of my model, it is now the case that:

$$P(Employment) = g(\mathrm{E}[\delta])$$

Because the unemployment rates is low and many jobs are available, employers only worry about a candidate's expected quality. Assuming quality to only differ by incarceration status, the racial differences in probability of employment will be captured by the *Ever Incarcerated* term. Though this term also lacks significance in my analysis, there are other potential explanations which I describe later in the paper.

The structural change between 2006 and 2008 creates more competition in the labor market. In 2008, the probability of employment can be defined in the original form:

$$P(Employment) = g(\mathrm{E}[\delta], \mathrm{Var}[\delta \mid \mathrm{Z}])$$

As in the typical statistical discrimination framework, differences in incarceration rates across races create gaps in employment outcomes.

Aside from the statistical discrimination argument, there is one other potential explanation for the fundamental difference between the 2006 and 2008 results. It is grounded in the idea of taste-based discrimination. As stated previously, screening is less important due to the widespread availability of jobs in 2006. In this case, even employers who have a preference for discrimination cannot do so because they need to fill the jobs. However, labor market conditions are more competitive in the 2008 data, so employers who have a taste for discrimination are able to fulfill this preference. They are not actively seeking employees, so they are able to discriminate based upon racial differences.

As mentioned earlier, another notable result of both models involving likelihood of employment is the insignificant coefficient of the *Ever Incarcerated* term. One explanation for this finding is employers' imperfect information. If a sufficient number of employers do not check criminal history, then the direct impact of an incarceration record will be ambiguous. Another potential explanation is the lack of information about post-release work programs. As articulated by Kling (2006), ex-offenders who participate in programs funded by the state or non-profit organizations can receive a short-term employment boost. The programs provide resources for ex-offenders and create incentives for them to find jobs. Without accounting for the importance of incentives that are provided only to those who have been incarcerated, it could appear as if a criminal record leads to a higher likelihood of employment. Though this situation is important to consider, not all ex-offenders receive post-release resources. Those who do not benefit from the programs' incentives face the lower probability of employment due to losses of human capital, decreased social network sizes, and the creation of a stigma. In my data sets, it is likely that some ex-offenders participate in the post-release programs and others do not. Therefore, two theoretically opposing forces are present: the positive impact of work programs and the negative impact of an incarceration record. Without controlling for program participation, it is impossible to isolate the true relationship between incarceration and employment. This is a potential explanation for the insignificant coefficient on the *Ever Incarcerated* term.

For the analysis of hourly compensation rates, I again employ an instrumental variable approach. As described in the methodology, the first stage is a logistic regression that estimates the probability of incarceration, while the second is a multiple least-squares model of hourly compensation. I control for the same covariates as in the model of employment rates, yet I also add job-specific characteristics to capture differences in rates of pay across industries and types of jobs. Results involving the coefficients of interest for the 2006 and 2008 specifications are listed in Figure 4. Again, full regression findings are listed in Appendix C.

	2006	**2008**
Ever Incarcerated	0.171	-0.293
	(0.286)	(0.315)
Race - Black	-0.055	-0.101**
	(0.039)	(0.042)
Race - Hispanic	-0.086**	0.030
	(0.040)	(0.043)
Black x Ever Incarcerated	-0.508	-0.572*
	(0.344)	(0.327)
Hispanic x Ever Incarcerated	0.242	-0.442
	(0.375)	(0.315)
Number of observations	2461	2625

Two-stage least-squares estimates
*p<0.10; **p<0.05; ***p<0.01

Figure 4: Two-Stage Regression Estimates for Hourly Compensation

In contrast to the examination of employment rates, the results of the hourly compensation model do not completely support the notion of statistical discrimination. In the 2008 specification, the *Race x Ever Incarcerated* interaction term is significant for blacks. The median compensation penalty for an incarceration record is 43.6 percent greater for black ex-offenders than it is for otherwise similar whites and Hispanics. There are a number of potential explanations for this result. First, it could indicate that there is another level of discrimination past that of simply the statistical sort. In this case, employers have a preference for candidates of a certain race, and the penalty associated with an incarceration record differs according to these preferences.

Another reasonable justification is the existence of racially-based differences in rates of human capital deterioration. The differences could be due to a number of factors. They could be related to differences in educational experiences across races before incarceration. Those who receive a higher quality education and connect more with the material likely retain knowledge more easily. If an individual learns solely through memorization of concepts with little connection to his everyday life, he will have a high level of human capital in the short-term. However, it will decline at a faster rate than that of others who connect more deeply with the material. Members of demographic groups with generally poorer quality of education could lose job-related skills at a faster rate than others, so incarceration would lead to a greater decrease in productivity.

A final explanation for different rates of human capital deterioration is linked to potential differences in jail experience. The majority of the prison population is black. This could lead to a greater feeling of cohesion among black inmates. For example, while incarcerated, black inmates may spend more time with others and could face negative influences from prison gangs or other organizations. Individuals of other races may not face such strong pressure to become involved with these groups. Associating with such groups could cause an individual to lose a greater amount of human capital per unit of time spent behind bars. Differences in rates of human capital deterioration would not be indicative of any other discrimination but simply would be the result of a difference in the marginal products of labor.

The 2006 data tell a different story. The only significant coefficient is that on the *Race-Hispanic* term. It indicates that the median hourly compensation for Hispanic individuals is 8.2 percent lower than otherwise similar black and white individuals. This result does not align with the theory of statistical discrimination. The incarceration rate for blacks is higher than that for Hispanics, so individuals from both of these demographic groups should face a penalty when employers do not check criminal history records. It is unexpected to see Hispanics as the only group that faces lower levels of compensation, and this contradictory result could potentially be attributed to an omitted variable. I discuss this possibility in greater detail later in the paper.

Results for the second model accounting for length of incarceration differ slightly from the original findings. Looking at likelihood of employment, I find evidence of statistical discrimination in the specifications for both years. The only significant coefficient is that of the *Race-Black* term. Specifically, in 2006, the odds of employment are 37.2 percent lower for black individuals than it is for whites. In the 2008 data set, I find that this discrepancy increases to 50.4 percent. The lack of significance of the other terms provides evidence for the theory that, lacking perfect information, employers use visible characteristics to account for differences in incarceration rates. In this case, it appears that race is one of those characteristics. The visible change in the 2006 result from the original model could be due to the fact that I do not account for endogeneity in this new specification. Unobserved factors that increase the likelihood of incarceration likely decrease the probability of employment. Without accounting for this, there could be an unobserved characteristic that acts as a confounder and alters the significance of the *Race-Black* term. Results involving the coefficients of interest from the logit regressions for both 2006 and 2008 are listed in Figure 5. Full results can be found in Appendix C.

	2006	**2008**
Length of Incarceration	0.012	-0.014
	(0.032)	(0.014)
Race - Black	-0.465***	-0.702***
	(0.169)	(0.149)
Race - Hispanic	0.349	-0.028
	(0.229)	(0.194)
Black x Length of Incarceration	-0.027	-0.007
	(0.035)	(0.018)
Hispanic x Length of Incarceration	-0.045	-0.001
	(0.043)	(0.027)
Number of observations	3016	3152

Logit estimates
*p<0.10; **p<0.05; ***p<0.01

Figure 5: Incarceration Length Regression Estimates for Likelihood of Employment

Considering hourly compensation as the response variable, the results differ slightly from those of the original model. The new findings provide more information about potential explanations for what I discover in the original case. The coefficients for the *Race-Black* and *Race-Hispanic* terms are the only ones with any significance. From the 2006 data, I find that the median hourly compensation is 8.8 percent lower for blacks and 6.9 percent lower for Hispanics than it is for whites. The difference for blacks is greater when considering the 2008 data. In this case, black individuals face a median compensation value that is 13.3 percent lower than that for the other racial groups. Despite this initial discrepancy, for each extra month spent in prison, black individuals receive the same compensation penalty as white and Hispanic individuals. For this specification, estimates for the coefficients of interest and their standard errors are listed in Figure 6.

	2006	**2008**
Length of Incarceration	-0.006	-0.0005
	(0.006)	(0.004)
Race - Black	-0.092***	-0.143***
	(0.030)	(0.030)
Race - Hispanic	-0.071**	-0.002
	(0.033)	(0.033)
Black x Length of Incarceration	-0.002	0.005
	(0.007)	(0.005)
Hispanic x Length of Incarceration	-0.005	-0.006
	(0.011)	(0.006)
Number of observations	2461	2625

Ordinary least-squares estimates
*p<0.10; **p<0.05; ***p<0.01

Figure 6: Incarceration Length Regression Estimates for Hourly Compensation

These findings suggest that the rate at which an individual loses human capital does not differ by race. If this were the case, the marginal product of labor would decrease at a faster rate with each extra month spent in prison for individuals of certain races. Therefore, because hourly compensation is directly related to the marginal product of labor, compensation would decrease more rapidly for people of certain races. This would lead the *Race x Length of Incarceration* interaction term to be significant but, in my analysis, it is not.

Eliminating human capital deterioration leaves two potential explanations for the original finding from the 2008 data that an incarceration record decreases hourly compensation rates more for blacks than for whites or Hispanics. First, as described previously, employers could have a taste for discrimination. In this case, they prefer white and Hispanic candidates over blacks. Another potential explanation involves an omitted variable in the original model. Incarcerated black individuals spend a greater length of time in prison than white and Hispanic offenders. In 2008, considering only those who have spent any time behind bars, the average length of incarceration is 7.5 months for whites, 13.0 months for blacks, and 9.1 months for Hispanics. Without controlling for the extent of human capital loss, it could appear as though blacks are unfairly penalized by employers. In fact, this could be due to the greater losses of human capital due to longer lengths of incarceration. This final situation makes it appear as if the compensation penalty associated with an incarceration record differs by race even though it may not necessarily be the case.

Accounting for length of incarceration also provides more information about the unexpected finding from the 2006 data that Hispanics are the only racial group to face lower compensation levels. In the updated version, both the *Race-Black* and *Race-Hispanic* terms are significant. This fits into statistical discrimination theory. In the 2006 data set, 8.1 percent of black respondents and 7.3 percent of Hispanics have a criminal record, while the rate is 6.1 percent for whites. Lacking perfect information about criminal history, employers extrapolate the average incarceration rates for each group to individuals. Therefore, the expected quality of black and Hispanic workers is lower than that of whites. Also, firms will face search costs as the test score will be a less accurate indicator of quality for blacks and Hispanics than for whites.

Surprisingly, this is the only finding that suggests that Hispanics are at a disadvantage with respect to employment outcomes. There are a few potential reasons for this. First, there are more obvious differences in physical appearance between black and white individuals than between Hispanics and whites. If employers are discriminating based upon what they can see, then black individuals will bear a larger burden than Hispanics. Also, the mass incarceration of blacks has received a considerable amount of media attention. For this reason, employers may think of blacks as the only group with a high incarceration rate.

Considering all four models, the findings generally align with statistical discrimination theory. The majority of differences in employment outcomes appear to be captured by racial differences rather than incarceration status. Even in the 2006 likelihood of employment model, the lack of significance of the *Race-Black* term can be explained within the realm of statistical discrimination theory. Additionally, the extra penalty seen in the two-stage wage model can potentially be attributed to an omitted variable, namely length of incarceration. Taking my results as a whole, they indicate that statistical discrimination at least plays a role in the post-release labor market. It may not be the only type of discrimination occurring, but evidence suggests that it is partially responsible for gaps in employment outcomes between racial groups.

VI. Further Considerations and Conclusion

While this model provides an appropriate analysis of the post-release labor market, a few data issues persist throughout. The NLSY does not indicate the type of crime for each period of incarceration, which makes it difficult to control for severity of the offense. Ideally, I would include an indicator variable to control for whether or not the crime is of a violent nature. In general, I believe employers would be more averse to hiring individuals with a history of violent offenses, and this could lead to a stronger negative relationship between incarceration and employment outcomes. Also, as discussed previously, I do not have any information about participation in post-release employment programs. This could be a confounding factor, as the literature suggests that these programs increase the likelihood of employment in the short-term after release.

While my results fit within the context of statistical discrimination, my model is not structured in a way that allows me draw definite conclusions about the presence of statistical discrimination. Future researchers could address this topic more directly by conducting a study similar to that of Holzer, Raphael, and Stoll (2006). However, instead of only looking at the probability that a firm's most recent hire was black, they could examine more general hiring patterns. This would involve comparing probabilities of employment for non-offending individuals depending on whether the firm conducts criminal background checks. If statistical discrimination were present, the probability of employment for non-offending black males should be higher at firms that conduct background checks, and therefore have perfect information about criminal history. The findings of such a study could have policy implications. If it finds evidence of statistical discrimination, it indicates that policy-makers must weigh the costs and benefits of legislation, such as Ban the Box laws, that limit the information provided to employers. While these types of regulations improve the likelihood of employment of ex-offenders, there could also be externalities felt by non-offenders from demographic groups with higher average incarceration rates. These individuals would have lower compensation rates and likelihood of employment than they should given their expected productivity.

Despite their limitations, my results are useful in providing economic context in the examination of the post-release labor market. Sociological literature has generally found racial differences in the penalties associated with an incarceration record. My results suggest that discrimination can also manifest itself in the statistical form. In this case, individuals from all races face, on average, the same employment and compensation penalties for having a criminal record. Despite this, individuals from certain racial groups are still at a disadvantage with respect to employment outcomes. The penalty for different incarceration rates across races could potentially stretch beyond simply ex-offenders. Punishment does, in fact, span past the jail cell.

Acknowledgments

I first want to thank Jenny Bourne for her support throughout the entire comps process. She was helpful every step of the way, and I am grateful for having someone with whom I could discuss ideas from a critical perspective. I also want to thank Jeff Dsida for the time he took reading over my drafts and helping me work my paper into its final form.

References

Aigner, Dennis, J. and Glen G. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review* 30 (2): 175-187.

Becker, Gary S. 1962. "Investment in Human Capital: A Theoretical Analysis." *The Journal of Political Economy* 70 (5): 9-49.

Becker, Gary S. 1974. "Crime and Punishment: An Economic Approach." *Essays in the Economics of Crime and Punishment*: 1-54.

Carson, E. Ann and Daniela Golinelli. 2013. "Prisoners in 2012 – Advance Counts." *Bureau of Justice Statistics*. Available at http://www.bjs.gov/content/pub/pdf/p12ac.pdf.

Calvó-Armengol, Antoni and Yves Zenou. 2005. "Job Matching, Social Network, and Word-of-Mouth Communication." *Journal of Urban Economics* 57 (3): 500-522.

De Leeuw, Edith. 2001. "Reducing Missing Data in Surveys: An Overview of Methods." *Quality and Quantity* 35 (2): 147-160.

Ehrlich, Isaac. 1996. "Crime, Punishment, and the Market for Offenses." *Journal of Economic Perspectives* 10 (1): 43-67.

Freeman, Richard B. 1991. "Crime and the Employment of Disadvantaged Youths." *NBER Working Paper Series*.

Harless, William. "Ban the Box Laws Make Criminal Past Off-Limits." *Wall Street Journal*, August 3, 2013.

Holzer, Harry J., Steven Raphael, and Michael A. Stoll. 2006. "Perceived Criminality, Criminal Background Checks, and the Racial Hiring Practice of Employers." *Journal of Law and Economics* 49 (2): 451-480.

Keilman, John. "Foundation Puts Ex-Cons to Work." Chicago Tribune, October 16, 2011. Available at http://articles.chicagotribune.com/2011-10-16/news/ct-met-holiday-giving-safer-20111016_1_safer-foundation-area-nonprofits-drug-conviction.

Kling, Jeffrey R. 2006. "Incarceration Length, Employment, and Earnings." *American Economic Review* 96 (3): 863-876.

Pager, Devah. 2003. The Mark of a Criminal Record. *American Journal of Sociology* 108 (5): 937–975.

Penn State College of Agricultural Sciences. 2002. "Pennsylvania's Rural Economy: An Analysis of Recent Trends." Available at http://pubs.cas.psu.edu/FreePubs/pdfs/ua368.pdf.

Pettit, Becky and Christopher J. Lyons. 2009. “Incarceration and the Legitimate Labor Market: Examining Age-Graded Effects on Employment and Wages.” *Law & Society Review* 43 (4): 725-756.

Phelps, Edmund S. 1972. “The Statistical Theory of Racism and Sexism.” *American Economic Review* 62 (4): 659-661.

Raphael, Steven. 2007. “Early Incarceration Spells and the Transition to Adulthood” in *The Price of Independence: The Economics of Early Adulthood*. S. Danziger and C. Rouse, eds., New York: Russell Sage Foundation.

Raphael, Steven. 2010. “Incarceration and Prisoner Reentry in the United States.” *Institute for Research on Poverty Discussion Series*.

Staples, Brent. “Target Bans the Box.” *New York Times Opinion Pages*, October 29, 2013. Available at https://talkingnote.blogs.nytimes.com/2013/10/29/target-bans-the-box/.

Schmitt, John, and Kris Warner. 2011. “Ex-Offenders and the Labor Market.” *WorkingUSA: The Journal of Labor and Society* 14: 87-109.

Society for Human Resource Management. 2012. “SHRM Finds Fewer Employers Using Background Checks in Hiring.” Available at http://www.shrm.org/about/pressroom/PressReleases/Pages/BackgroundChecks.aspx.

Van Buuren, Stef and Karin Groothuis-Oudshoorn. 2011. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3), 1-67.

Western, Bruce. 2006. *Punishment and Inequality in America*. New York: Russell Sage Foundation.

Appendix A: Summary of Missing Variables

Variable - 2006	Num. Missing
Employment	0
Wage	535
Age	0
Race	0
Incarceration Status	0
Highest Grade	0
Father's Highest Grade	623
Mother's Highest Grade	218
Relationship at Age 12	356
Number of Children	2
Region	0
MSA	0
Industry of Job	258
Migration in Past Year	6
Total	1774

Variable - 2008	Num. Missing
Employment	0
Wage	513
Age	0
Race	0
Incarceration Status	0
Highest Grade	31
Father's Highest Grade	642
Mother's Highest Grade	243
Relationship at Age 12	379
Number of Children	0
Region	0
MSA	2
Industry of Job	345
Migration in Past Year	3
Total	2158

Appendix B: Summary of Data Set Consolidation

Total Observations - 2006	8984
(-) Female	*4385*
	4599
(-) Enrolled in School	*524*
	4075
(-) Out of U.S.	*20*
	4055
(-) Mixed Race	*34*
	4021
(-) Currently Incarcerated	*155*
	3866
(-) No Employment Record	*832*
	3034
(-) Military	*18*
Employment Data Set – 2006	3016
(-) Unemployed	*259*
(-) Missing Comp. Data	*277*
(-) Compensation Outliers	*19*
Compensation Data Set - 2006	2461

Total Observations - 2008	8984
(-) Female	*4385*
	4599
(-) Enrolled in School	*439*
	4160
(-) Out of U.S.	*21*
	4139
(-) Mixed Race	*39*
	4100
(-) Currently Incarcerated	*143*
	3957
(-) No Employment Record	*783*
	3174
(-) Military	*22*
Employment Data Set – 2008	3152
(-) Unemployed	*349*
(-) Missing Comp. Data	*167*
(-) Compensation Outliers	*11*
Compensation Data Set - 2008	2625

Appendix C: Full Regression Results

Two-Stage Regressions with Incarceration Status:

	Likelihood of Employment		Hourly Compensation	
	2006	2008	2006	2008
Ever Incarcerated	0.620	0.644	0.171	-0.293
	(1.630)	(1.672)	(0.286)	(0.315)
Race - Black	-0.309	-0.544**	-0.055	-0.101**
	(0.221)	(0.218)	(0.039)	(0.042)
Race - Hispanic	0.231	-0.151	-0.086**	0.030
	(0.285)	(0.271)	(0.040)	(0.043)
Black x Ever Incarcerated	-1.990	-1.427	-0.508	-0.572*
	(1.574)	(1.381)	(0.344)	(0.327)
Hispanic x Ever Incarcerated	0.902	1.055	0.242	-0.442
	(2.251)	(1.851)	(0.375)	(0.315)
Highest Grade	0.114**	0.161***	0.035***	0.036***
	(0.053)	(0.048)	(0.008)	(0.007)
Mother Highest Grade	0.048	0.049*	0.005	0.010*
	(0.032)	(0.028)	(0.005)	(0.005)
Father Highest Grade	-0.008	-0.053**	0.004	0.008*
	(0.029)	(0.025)	(0.005)	(0.005)
log(Total Children + 1)	0.00006	0.043	0.043	0.101***
	(0.154)	(0.130)	(0.026)	(0.025)
Age	2.009	-0.638	0.344	-0.097
	(2.035)	(1.881)	(0.321)	(0.340)
Age-squared	-0.041	0.013	-0.006	0.003
	(0.043)	(0.036)	(0.007)	(0.007)
Relationship - One Bio Parent	-0.434***	-0.315**	-0.063**	-0.023
	(0.166)	(0.151)	(0.025)	(0.027)
Relationship - Two Parents, One Bio	-0.370	-0.480*	0.007	0.005
	(0.341)	(0.281)	(0.056)	(0.057)
Relationship - Other	-0.441	-0.595**	-0.101	-0.008
	0.274	(0.238)	(0.048)	(0.051)
Region - Northeast	-0.395	-0.336	0.015	0.074**
	(0.247)	(0.213)	(0.037)	(0.036)
Region - South	-0.420**	-0.279	-0.046	-0.002
	(0.199)	(0.177)	(0.030)	(0.030)
Region - West	-0.379	-0.090	0.117***	0.127***
	(0.237)	(0.208)	(0.035)	(0.034)

Migration - To or from Foreign Country	-2.375***	-2.156***	-0.166	-0.055
	(0.565)	(0.641)	(0.208)	(0.280)
Migration - Move within State	0.568**	0.828***	0.145***	-0.139**
	(0.281)	(0.294)	(0.050)	(0.059)
Migration - No Move	0.628***	0.837***	0.103**	-0.072
	(0.234)	(0.212)	(0.043)	(0.050)
Not in CBSA	-0.240	0.140	-0.094*	-0.075
	(0.274)	(0.303)	(0.050)	(0.055)
Industry - 2	-	-	0.245***	0.171**
	-	-	(0.079)	(0.081)
Industy - 3	-	-	0.077	0.067
	-	-	(0.075)	(0.078)
Industry - 4	-	-	0.074	-0.012
	-	-	(0.068)	(0.066)
Industry - 5	-	-	-0.048	-0.067
	-	-	(0.073)	(0.073)
Industry - 6	-	-	0.264***	0.231***
	-	-	(0.068)	(0.067)
Industry - 7	-	-	0.198***	0.100*
	-	-	(0.062)	(0.061)
Industry - 8	-	-	-0.024	-0.084
	-	-	(0.065)	(0.063)
Industry - 9	-	-	0.124	0.139*
	-	-	(0.080)	(0.076)
Intercept	-23.780	7.835	-2.668	2.911
	(24.420)	(24.442)	(3.856)	(4.425)
Number of Observations	3016	3152	2461	2625

*p<0.10; **p<0.05; ***p<0.01

One-Stage Regressions with Length of Incarceration:

	Likelihood of Employment		Hourly Compensation	
	2006	**2008**	**2006**	**2008**
Length of Incarceration	0.012	-0.014	-0.006	-0.0005
	(0.032)	(0.014)	(0.006)	(0.004)
Race - Black	-0.465***	-0.702***	-0.092***	-0.143***
	(0.169)	(0.149)	(0.030)	(0.030)
Race - Hispanic	0.349	-0.028	-0.071**	-0.002
	(0.229)	(0.194)	(0.033)	(0.033)
Black x Length of Incarceration	-0.027	-0.007	-0.002	0.005
	(0.035)	(0.018)	(0.007)	(0.005)
Hispanic x Length of Incarceration	-0.045	-0.001	-0.005	-0.006
	(0.043)	(0.027)	(0.011)	(0.006)
Highest Grade	0.109***	0.148***	0.031***	0.046***
	(0.035)	(0.027)	(0.006)	(0.005)
Mother Highest Grade	0.049	0.054**	0.004	0.007
	(0.032)	(0.028)	(0.005)	(0.005)
Father Highest Grade	-0.008	-0.056**	0.004	0.010**
	(0.029)	(0.025)	(0.005)	(0.005)
log(Total Children + 1)	-0.015	0.057	0.045*	0.081***
	(0.148)	(0.120)	(0.025)	(0.024)
Age	1.924	-0.858	0.302	-0.005
	(1.926)	(1.850)	(0.312)	(0.338)
Age-squared	-0.039	0.017	-0.005	0.0007
	(0.040)	(0.036)	(0.006)	(0.006)
Relationship - One Bio Parent	-0.414**	-0.273*	-0.058**	-0.036
	(0.164)	(0.144)	(0.025)	(0.025)
Relationship - Two Parents, One Bio	-0.360	-0.463*	0.014	-0.022
	(0.333)	(0.268)	(0.055)	(0.054)
Relationship - Other	-0.418	-0.555**	-0.090*	-0.037
	(0.270)	(0.222)	(0.047)	(0.048)
Region - Northeast	-0.390	-0.344*	0.010	0.084**
	(0.239)	(0.203)	(0.036)	(0.036)
Region - South	-0.416**	-0.293*	-0.048	0.010
	(0.197)	(0.169)	(0.030)	(0.029)
Region - West	-0.377	-0.095	0.113***	0.131***
	(0.236)	(0.207)	(0.034)	(0.034)
Migration - To or from Foreign Country	-2.447***	-2.180***	-0.186	-0.038
	(0.553)	(0.638)	(0.207)	(0.280)

Migration - Move within State	0.573**	0.884***	0.139***	-0.154***
	(0.279)	(0.272)	(0.050)	(0.058)
Migration - No Move	0.615***	0.852***	0.094**	-0.080
	(0.228)	(0.208)	(0.043)	(0.050)
Not in CBSA	-0.264	0.159	-0.092*	-0.097*
	(0.261)	(0.293)	(0.049)	(0.053)
Industry - 2	-	-	0.237***	0.221***
	-	-	(0.078)	(0.079)
Industsy - 3	-	-	0.068	0.123
	-	-	(0.073)	(0.075)
Industry - 4	-	-	0.064	0.022
	-	-	(0.066)	(0.065)
Industry - 5	-	-	-0.058	-0.018
	-	-	(0.073)	(0.071)
Industry - 6	-	-	0.256***	0.268***
	-	-	(0.067)	(0.066)
Industry - 7	-	-	0.193***	0.124**
	-	-	(0.062)	(0.061)
Industry - 8	-	-	-0.032	-0.073
	-	-	(0.065)	(0.063)
Industry - 9	-	-	0.115	0.179**
	-	-	(0.079)	(0.075)
Intercept	-22.721	10.718	-2.112	1.629
	(23.039)	(23.989)	(3.735)	(4.383)
Number of Observations	3016	3152	2461	2625

*p<0.10; **p<0.05; ***p<0.01

Appendix D: Data Appendix

Variable	Name in Code	Source	Description
Compensation	Wage06 Wage	NLSY	Weighted average of compensation rates based upon hours worked in each job
Employment	Employment06 Employment	NLSY	"Employed" if Compensation > 0 "Not Employed" otherwise
Incarceration	EverIncarcerated06 EverIncarcerated	NLSY	"Yes" if individual has been incarcerated "No" otherwise
Length of Incarceration	In.Total.Month06 In.Total.Month	NLSY	Number of months an individual has spent in prison
Race	Race06 Race	NLSY	"Black", "Hispanic", or "Non-Black/Non-Hispanic"
Incarceration Rate by Race/Region	IncRateRaceRegion06 IncRateRaceRegion	BJS Census	Weighted average of state incarceration rates by race
Highest Grade	HighestGradeNew06 HighestGradeNew	NLSY	Highest grade completed by the respondent
Mother Highest Grade	MotherHighestGrade06 MotherHighestGrade	NLSY	Highest grade completed by the respondent's mother
Father Highest Grade	FatherHighestGrade06 FatherHighestGrade	NLSY	Highest grade completed by the respondent's father
Relationship at Age 12	RelUpdate06 RelUpdate RelUpdateWage206 RelUpdateWage2	NLSY	"Both Biological Parents", "Two Parents, One Biological", "One Biological Parent", or "Other"
Biological Children	TotalChildrenNew06 TotalChildrenNew	NLSY	Total number of biological children (sum of in-household and non-resident)
Age	Age06 Age	NLSY	Calculated by taking the number of days between an individual's birth month and January 1 of either 2006 or 2008, depending on the data set, and dividing by 365.26.

Region	RegionNew06 RegionNew	NLSY	"Northeast", "North Central", "South", or "West"
Migration	Migration06 Migration	NLSY	In the past year, categorized as "Not Moved", "Moved Within State", "Moved Between States", or "Moved From a Foreign Country"
Urban/Rural	MSANew06 MSANew	NLSY	Categorizes whether respondent lived in "CBSA" or "Not in CBSA"
Industry	Industry06 Industry	NLSY	First digit of SIC code for job respondent listed first

Key:

1. If two names are listed as "Name in Code," the first is the code for the 2006 data set and the second is for 2008. If four names are listed, the first is the code for the 2006 employment data set, the second is for the 2008 employment data set, the third is for the 2006 wage analysis, and the fourth is for the 2008 wage analysis.

Final Data Set Names:
Employment06: Used for 2006 likelihood of employment analysis
Employment08: Used for 2008 likelihood of employment analysis
Compensation06: Used for 2006 hourly compensation analysis
Compensation08: Used for 2008 hourly compensation analysis

Data Sources:

1. Bureau of Labor Statistics, U.S. Department of Labor. National Longitudinal Survey of Youth 1997 cohort, 1997-2010 (rounds 1-14) [computer file]. Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH: 2012.

   Web Access: https://www.nlsinfo.org/investigator/pages/login.jsp

2. Bureau of Justice Statistics, U.S. Department of Justice. Prison and Jail Inmates at Midyear 2005. Produced by Paige M. Harrison and Allen J. Beck. 2006.

   Web Access: http://www.bjs.gov/content/pub/pdf/pjim05.pdf

3. United States Census Bureau, U.S. Department of Commerce. State Tables.

   Web Access:
   2006 – http://www.census.gov/popest/data/historical/2000s/vintage_2006/state.html
   2008 – http://www.census.gov/popest/data/historical/2000s/vintage_2008/state.html

Summary of Data Work:

I gathered the raw data for my 2006 and 2008 sets from responses to the 2007 and 2009 rounds of the NLSY survey, respectively. Individuals answer questions about the previous year, so the corresponding responses refer to 2006 and 2008 activities. To obtain my final data sets, I restructured the raw data using R. I first transformed all categorical variables from their numerical values using the NLSY codebook. One variable that I manually transformed was Race. Race-Black was coded as 1 in the original data, so I cycled through all data points to change this notation to "Black." I completed this process for the other categories of race, as well as a number of other variables, including Region, MSA, Relationship, Employment, Migration, and Incarceration status. Additionally, for all variables, I coded all missing values as such. The original data set lists them as -3 or -5, and I adjusted these entries to clarify.

After I completed this original data manipulation process, I updated missing values using the previous year's observation carried forward. A more explicit description of this process is provided in the paper. Next, I removed respondents who are female, enrolled in school, in the military, live outside of the U.S., are mixed race, or are incarcerated in the given year (2006 or 2008). Additionally, I removed individuals who do not list information about employment. For the covariates that were still missing, I employed multivariate imputation by chained equations using the 'MICE' package in R.

Once I accounted for the missing values, I added the Incarceration Rate by Race and Region variable. A more detailed derivation of this variable is described in the paper. This left me with the final data set for my analysis of employment rates. To obtain my data sets to analyze compensation trends, I removed outliers in the process described in the paper. My four final data sets are entitled "Employment06.csv", "Employment08.csv", "Compensation06.csv", and "Compensation08.csv."