8-26-2022

Sports participation and Labor Market Outcomes: A Correspondence Study

Laura Lorena Rodriguez Ortiz
Syracuse University, laural.rodriguezortiz@gmail.com

Abstract

Discrimination contributes to gaps in employment between Whites and Blacks and increases racial inequality. Young Blacks, ages 18 to 19, are twice more likely to be unemployed compared to Whites (Bureau of Labor Statistics, 2015). This dissertation is the first study, to my knowledge, to rely on an experimental design to examine whether participation in high school sports affects labor market discrimination. This dissertation uses a correspondence study in which 6,000 fictional resumes are sent to real job openings in Chicago, Dallas, and New York. Call-backs are recorded, and any differences in call-back rates are considered evidence of discrimination. All resume characteristics are randomly assigned, and race is signaled through names. Results suggest that participation in high school sports has a positive effect on call-back rates. The effect of sports participation is particularly beneficial to Black applicants, reducing some of the discrimination Black applicants experience in the labor market. This result is consistent with theories of statistical discrimination.

SPORTS PARTICIPATION AND LABOR MARKET OUTCOMES: A CORRESPONDENCE STUDY

by

Laura Rodriguez Ortiz

B.A., Pontificia Universidad Católica Madre Y Maestra, 2010
M.A., Syracuse University, 2013
M.P.A., Syracuse University, 2013

Dissertation
Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Public Administration.

Syracuse University
August 2022

Acknowledgement

None of this work would have been possible without the guidance and support of an incredible group of individuals. The unconditional support of my family. Mom, Laura Rodriguez Simons, Andrew Simons, and Bella Simons – thank you for your unconditional love and support throughout all the valleys and peaks that accompanied this process.

I cannot imagine finishing this dissertation without Emily Cardon. Thank you for listening to all the versions of this work and supporting me through all the breakdowns that got me here. To our boy, Ronan, who made this process sweet and exciting even in the most trying times.

Xiomara Alcántara, Pamela Hernandez, Sarah Sanchís and Yazmin Yeara, thank you for all your love, support and understanding throughout this process. Big chunks of this work were written in the comfort of your homes and fueled by your unwavering motivation – thank you!

To everyone who softened the burden of this process during my time at Syracuse: The Murchie clan (Jud, Biz, Bella & Vinny), Jeehee Han, Raghav Puri, Hannah Patnaik, Emily Gutierrez, Fabio Rueda, Boqian Jiang, Michelle Lofton and Yusun Kim.

Colleen Heflin, Sarah Hamersma and Shana Gadarian, thank you for your motivation, guidance, and feedback.

Thank you to all the wonderful staff at CPR!

I want to thank my committee members: Robert Bifulco, Kathy Michelmore, Johnny Yinger and Alfonsos Flores- Lagunes. Thank you for you all guidance, feedback, and patience throughout this process.

Last by not least, I want to thank my advisor, Leonard M. Lopoo. Thank you for being an instrumental piece in my academic journey since before starting the PhD program. Thank you for all your gracious guidance and support for all these years. This work would have not been possible nor completed, without your support and patience. Thank you, Len

Para Laura Herrera, quien, en su pecho, desde su mecedora y con sus historias de mujeres y jóvenes que rompían esquemas, me enseñó a soñar.

Te extraño, Maita.

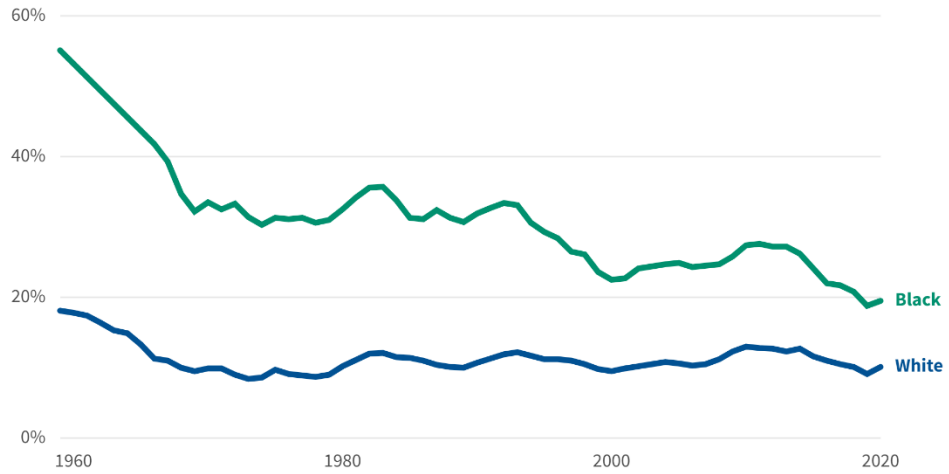
Table of Contents

Illustrative Content

Section 1. Background & Motivation

Race in America is a strong predictor of economic success, and Blacks underperform in all measures of economic success when compared to Whites, Council of Economic Advisers (1998). Understanding differences in economic outcomes between Blacks and Whites has been widely studied in academia. Important explanations, such as unequal opportunities between Blacks and Whites, have closed a large portion of the differences. However, even after controlling for such differences, racial gaps persist in the labor market. The presence of this unexplained gap is, to many researchers, evidence of discriminatory practices in the labor market that continue to perpetuate racial differences of economic success. Graph 1 shows the percent of people in poverty by race. The percent of Black population in poverty decreased substantially after the 1960s, however, a large racial gap remains today.

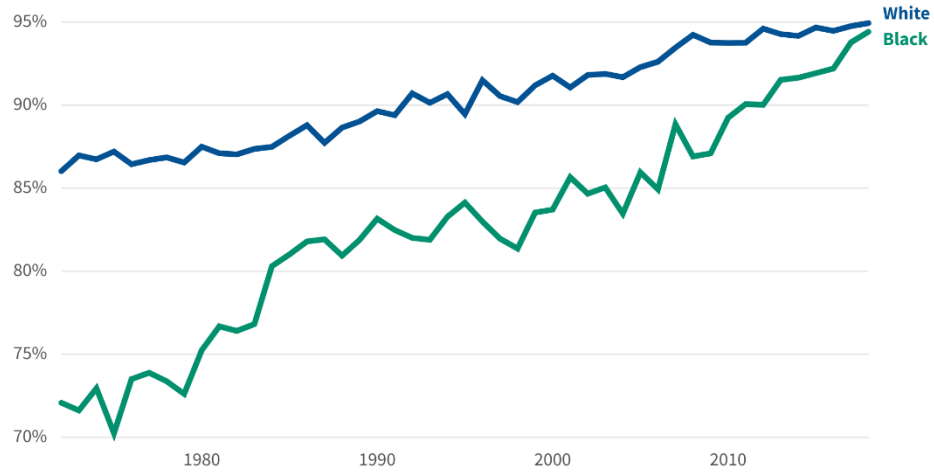
Graph 1. Percent of people in poverty by race



Source: Census Bureau (2021)

Starting with educational attainment, which is an important predictor of later life outcomes, studies have shown that between 1960s and 1970s some of the improvement in labor market outcomes for Blacks was due to improvements in educational attainment and school quality brought by the Civil Rights Act of 1964, Donohue and Heckman (1991). Graph 2 shows how the gap in educational attainment, at the high school level, has been consistently narrowing since the mid-1980s.

Graph 2. High School completion rates by race

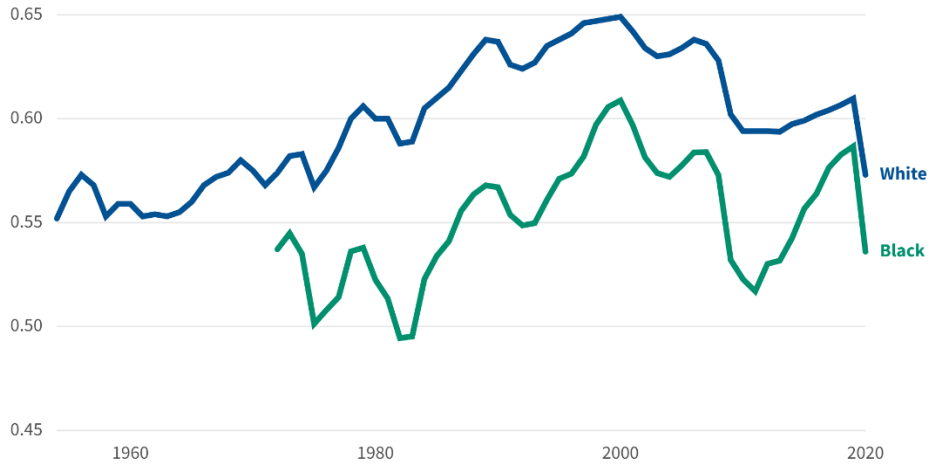


Source: National Center for Education Statistics (NCES, 2016)

However, important racial disparities persist in the quality of education, even after school finance reforms according to Card and Payne (2002). Highly segregated neighborhoods paired with school district financing system that rely heavily on local taxes has produced large inequalities in school funding across poorer and wealthier localities. Card and Payne (2002) argue this inequality in funding has been identified as a source of quality disparities, where richer districts have higher test scores compared to poorer districts. Even after school finance reforms aimed at equalizing school funding and quality, schools in the US remain highly segregated by racial lines (Orfield et

al., 2012). And predominantly minority high schools continue to have less access to resources and underperform on quality measures compared to predominantly White schools (NWLC and PRRAC, 2015).

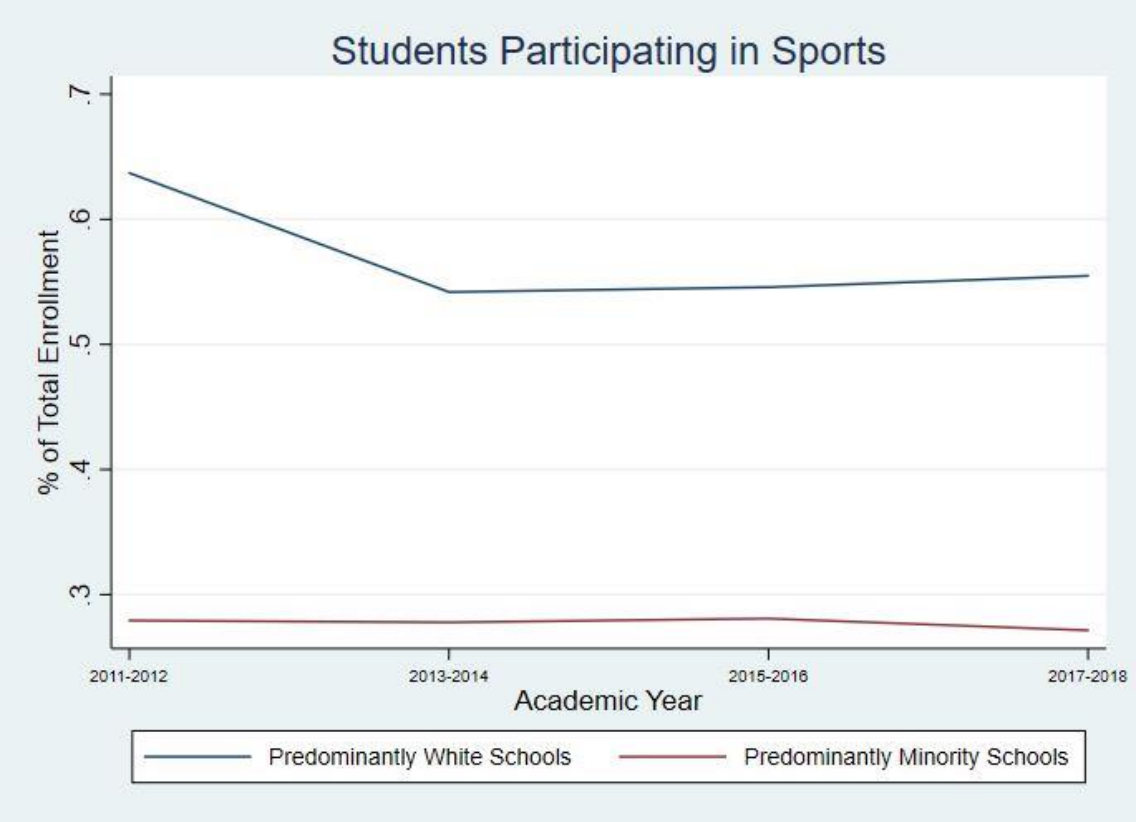
Graph 3. Employment-population ratio by race



Source: Bureau of Labor Statistics (2019)

"High schools with high concentrations of minority and low-income students generally have fewer resources for extracurricular activities" (NWLC and PRRAC, 2015). More than 40 percent of public high schools are either predominantly White or predominantly minority in the United States (NWLC and PRRAC, 2015). Athletic opportunities are significantly fewer in predominantly minority high schools compared to predominantly White high schools. On average, predominantly minority high schools provide less than half the spots on sports team than those offered in predominantly White high schools. This translates to predominantly White high schools having more than twice the rate of athletic participation than predominantly minority high schools.

Graph 4. Percent of students participating in sports in "predominantly White" and "predominantly minority" schools.



Source: Author's calculation using Civil Rights Data Collection (CRDC) years 2011-2018. Following NWLC and PRACC 2015 I define schools as predominantly White if 90% or more of enrolled students are White and predominantly minority if 10% or less of enrolled students are White. Participation is measured as the percent of enrolled students that participate in sports.

Graph 4 represents the gap in participation between predominantly White and predominantly minority schools. The percent of students participating in sports in predominantly White schools is roughly twice as large, around 60%, as the percent of student athletes in predominantly minority schools, around 30%. Data suggest that for the period analyzed, this participation gap has remained constant. Students who attend a predominantly White school are twice more likely to participate in high school sports.

Discrimination contributes to gaps in employment between Whites and Blacks and increases racial inequality. Young Blacks, ages 18 to 19, are twice more likely to be unemployed compared to Whites (Bureau of Labor Statistics, 2015). This project is the first study to rely on an experimental design to examine whether participation in high school sports reduces labor market discrimination for young, low-skilled Black males. A handful of studies have tried to find resume characteristics, quality, and socio-economic status, that reduce racial discrimination in the labor market, with little to no success. This study will use a correspondence study in which fictional resumes are sent to real job openings, callbacks are recorded, and any differences in callback rates are considered evidence of discrimination. All resume attributes are randomly assigned, and race is signaled through the applicants' names. The audit study will focus on 18–19-year-olds, who are more likely to report high school activities on resumes and will run in three large American cities: Chicago, Dallas, and New York. Callbacks will be recorded through email accounts and virtual voicemails. A pilot conducted in October 2016 finds preliminary evidence that sports participation reduces racial and gender-based discrimination. There is also evidence of heterogeneity in sports, different sports differing in their ability to reduce discrimination.

This research is directly focused on examining activities that may reduce discrimination in the labor market. Employment discrimination contributes to the longstanding gap in employment between Whites and Blacks, which in turn contributes to persistent racial inequality (Quillian and Hexel, 2016). Involvement in high school sports participation may reduce some of the discrimination experienced by young Black males in the labor market. Rigorously exploring the positive ramifications of athletic involvement has important policy implications. Schools in the

United States remain highly segregated along racial and economic lines, and schools with high concentrations of minority and low-income students generally have fewer resources for extracurricular activities (NWLC and PRRAC, 2015). This is the first experimental study to examine how sports participation may reduce racial and gender-based discrimination in the labor market. Results from this study will inform school officials about the importance of extracurricular activities, particularly for highly segregated schools where funding is already dire.

Disconnected Youth

This study will focus on a segment of disconnected youth, that is, 16–24-year-olds who have completed high school but not enrolled in college. In 2014 this population represented 6.1% of the youth population in the United States (Fernandes-Alcantara, 2015). Males are more likely to be disconnected compared to females (SSRC, 2019). This gender gaps widens for Blacks, 19.9% of Black males are disconnected compared to 14.8 percent of Black females. Policymakers and youth advocates have argued that this population lacks important social networks to aid in their transition to adulthood (Fernandes-Alcantara, 2015; The Annie E. Casey Foundation, 2012). More importantly, this population is likely to experience negative outcomes related to employment, housing, and health (The Annie E. Casey Foundation, 2012). These dire prospects make this a vulnerable and policy-relevant population. There are three reasons why this study will focus on disconnected youth: (ii) it is a vulnerable and policy-relevant population; (iii) young adults who have recently left high school are more likely to report the activities they participated in; (iii) this is a population that might be using job search engines as opposed to young adults still in school, who are more likely to approach potential employers in person (Holzer, 1996).

The remainder of this dissertation is structured as follows: Section 2 reviews the literature on labor market discrimination, sports participation, and labor market outcomes. Section 3 describes the audit methodology. Section 4 explains the experimental design. Section 5 reviews the main results. Section 6 discusses the results and limitations of the study.

Section 2. Theory & Literature Review

This work draws on two theoretical frameworks to establish the hypotheses I will test. First, it draws on the theories of labor market discrimination which explain differences in outcomes for Black and White job applicants as well as for male and female job applicants. Second, it draws from the human capital and signaling literatures to explain differences in outcomes for job applicants with and without extracurricular activities. These theoretical frameworks combined also help amplify the findings. In this section, I will discuss both theoretical frameworks and examine the existing literature on labor market discrimination.

Labor Market Discrimination Literature

Discrimination in the labor market is usually defined as when members of a minority (e.g., Blacks) are treated differently, usually less favorably, than members of a majority group (e.g., Whites) even when both exhibit identical productive characteristics (Autor, 2003). In the labor market, when two employees have the same level of productivity or offer employers the same levels of production, any difference in treatment between them can be attributed to discrimination. Models of discrimination broadly define this difference in treatment for a variety of group memberships. That is, the models explain discrimination in the labor market based on many characteristics, including gender, race, ethnicity, and age. Since discrimination models create a framework to understand differential treatment between two groups, under this framework, the minority will correspond to Blacks and males – the two main groups studied in this dissertation.

Models that explain discrimination in the labor market fall under two broad types: models of taste-based discrimination and models of statistical discrimination. Taste-based models begin with Becker (1971) and are usually divided into employer-based, employee-based, and customer-based discrimination. In Becker's classic model, which assumes perfect labor markets, White employers, workers, and consumers, experience disutility from employing, working with, or purchasing goods and services from Blacks.

Under the employer discrimination model, employers dislike hiring Blacks. When a Black applicant is hired, an employer considers the cost to be both the wage and the disutility from hiring the worker (Neumark, 2018). In this model, we think employers are utility maximizers with the following utility function:

$$U = U(\pi, B) = \pi - d \cdot B$$

where π is profits, B is the number of Black workers, and d is a constant larger than zero, which denotes discriminatory tastes against Blacks. We assume the production function is:

$$\gamma = f(W + B)$$

a simple production function where the only input is labor, and White (W) and Black (B) workers are perfect substitutes. "We also assume that labor supply is perfectly inelastic, so wages are determined solely by demand. Normalizing the price to one, and letting w denote the wage of group I, the first-order conditions are:

$$MP_L = w_W, MP_L = w_B + d"$$ (Neumark, 2018)

Because d is positive, this implies that $w_B < w_W$, and we can rewrite as follows

$$w_W = w_B + d$$

This means that if d is the same for all employers, then when $w_W = w_B + d$ employers are indifferent between hiring Black or White workers, so we do not expect discriminatory hiring practices and both types of workers are hired. That is, because the cost of both type workers is the same to employers. Blacks receive a lower wage; the cost of the disutility of hiring a Black employee plus the lower wages equals the cost of hiring a white employee. However, if $w_W \neq w_B + d$, employers will only hire whites if $w_W < w_B + d$. If the cost of hiring Whites (wages) is lower than the cost of hiring Black workers (wages + disutility from hiring Blacks), then employers will only hire Whites as they represent to them a cheaper input. Since the market hold employers with and without taste for discrimination, employers will only want to hire one race. This happens because employers with no taste for discrimination will capitalize on lower Black wages and only hire Blacks, while employers with a taste for discrimination will only hire Whites. This leads to an equilibrium where wages will adjust to the point where $w_W = w_B + d$. That is, when the cost of hiring differs between Blacks and Whites and there are enough firms in a competitive market with differing levels of state for discrimination, this model predicts that employers with no taste for discrimination will hire the lower cost Black workers at a lower wage. While employers with a taste for discrimination will hire White workers at a higher cost. Furthermore, in perfect competition the advantage exploited by the non-discriminators that hire at a lower cost means that discriminatory firms (with higher hiring costs) are driven out of the market. In equilibrium, wages will adjust until both types of employees are hired. In contrast, if *d* varies across employers, an equilibrium wage differential should be established such that $d > w_W - w_B$. That is, if employers are heterogeneous in their taste for discrimination a wage differential between White and Black workers arises. This is more likely the case, that not all employers have the same taste

for discrimination. Under this scenario and in the presence of competitive markets, hiring discrimination does not exist. Employers with no taste for discrimination experience a lower cost of production due to lower wages for Black employers, this would mean that under competition, employers with no taste for discrimination drive discriminating employers out of business. In summary, Becker predicts that in a competitive market there should be no taste-based discrimination as discriminators are driven out of the market by non-discriminators.

However, data on the United States labor market has found little evidence to support this classical model (Lang and Lehmann, 2012). Pager (2016) finds that firms that discriminate against Black applicants in her original audit studies are less likely to survive 6 years after, conforming the model of taste-based discrimination. Bertrand and Mullianathan (2008) find evidence that supports racial segregation in the labor market, trends in wage and earnings since the 1960s show evidence of a wage differential. If anything, wage and earnings differentials between Black and Whites are fairly large even after controlling for skills – around 10 percent (Lang and Lehmann, 2012).

Employee-based discrimination

Arrow (1972) expands the employee-based taste discrimination model, by exploring the implications of the model that arise when the employer does not experience disutility from hiring Black applicants, but coworkers do. Under this scenario, White employees may pressure employers in their hiring practices by demanding a premium for working alongside Blacks. Under

this model, the following expressions for profit (π) correspond to an all-White workforce, an all-Black workforce, and an integrated workforce:

$$\pi_W = f(W) - \omega_W W$$

$$\pi_B = f(B) - \omega_B B$$

$$\pi_{WB} = f(W + B) - (\omega_W + d)W - \omega_B B$$

Under this model, complete racial segregation would occur in the labor market due to the increased cost of hiring White workers to work alongside Black workers. Arrow (1972) expands this model to account for adjustment costs for moving from an all-White to an all-Black workforce and parts from the assumption that firms initially have a predominantly White workforce. Under these assumptions, Arrow demonstrates there is incomplete segregation that leads to lower wages and job opportunities for Black employees. That is, parting from equilibrium with equal wages, any decrease in the relative price of Black labor should in theory make an all-White workforce switch to an all-Black workforce when there is no taste for discrimination. This is due to the assumption that firms are profit maximizers then switching to an all-Black workforce represents an outward shift of their production function, can now produce that same amount of output at a lower cost. However, Arrow demonstrates this is only the case when there are no adjustment costs to switching. Once we account for adjustment costs, the result is incomplete segregation with lower wages for Blacks.

Frijters et al. (2006) use matched employer-employee data to test employee-based discrimination in Britain. In this paper, the authors explore two hypotheses that stem from

employee-based discrimination models: (i) that employees that work with a high proportion of coworkers from an ethnic minority report lower levels of job satisfaction (disutility from working with ethnic minorities) and (ii) that these employees perceive higher wages to compensate for lower job satisfaction due to an integrated workplace (wage-gap). The authors find evidence of lower job satisfaction for White employees. However, they find heterogenous effects on wage differential by gender. The largest effect on wage is perceived by White males, who are compensated, on average, 19% higher wages to work in integrated workplaces. This heterogeneous effect by gender, is likely due to the existing gender wage-gap.

Customer-based discrimination

Unlike the previous types of discrimination, which economic theory suggest would not prevail under a perfectly competitive market due to their profit-reducing nature (Becker, 1957), discrimination to satisfy consumer preferences is thought to increase profits (Leonard and Levine, 2006). Under this model, employers hire employees that better match the demographic of their customers when these have a taste for discrimination. When consumers have a taste for discrimination, they are willing to either pay a premium for shopping where employees match their demographics or at a discount where employees do not match their demographics. This would imply that sales will be higher at workplaces that match employee demographics with those of their customer base.

Leonard and Levine (2006) use employee data from more than 800 stores and match the demographics of the store location to test whether stores with similar demographics to those of

its community perform better. They find that increasing the share of Black employees in predominantly white communities, reduces sales. Increasing the share of Black employees by one standard deviation, or about 7%, decreases sales by about 2%.

Statistical Discrimination Models

Statistical models of labor market discrimination are divided into level-based statistical discrimination (Phelps, 1972) and variance-based statistical discrimination (Aigner and Cain, 1977). In models of level-based statistical discrimination, employers believe there are differences in the average quality of the different types of applicants, usually White and Black applicants. Statistical discrimination does not assume prejudice; it is viewed as a model of stereotyping based on assumed group averages. Decision-makers in statistical discrimination models have imperfect information about some important aspect of the individuals, such as their productivity, skills, qualifications, etc., this incentivizes the use of group statistics as proxies for these unobserved characteristics (Fang & Moro, 2011).

Under this model, an employer does not observe with certainty the skill level of applicants but observes group identity$j \in \{B, W\}$. Workers' skill q is assumed equal to their marginal productivity. Employers observe group identity and an imperfect or noisy signal of productivity $\theta = q + \varepsilon$, where ε is a normally distributed error term with mean zero.

In a competitive market, workers are paid their expected productivity conditional on the signal. "Each employer infers the expected value of q from the signal using available information,

including group identity. The skill and the signal are jointly normally distributed, and the conditional distribution of q is normal with mean equal to a weighted average of the signal and the unconditional group mean" (Fang & Moro, 2011).

$$E(q|\theta) = \frac{\sigma_j^2}{\sigma_j^2 + \sigma_{\varepsilon j}^2}\theta + \frac{\sigma_{\varepsilon j}^2}{\sigma_j^2 + \sigma_{\varepsilon j}^2}\mu_j$$

"If the signal is very noisy, i.e., if the variance of ε is very high, the expected conditional value of workers' productivity is close to the population average, regardless of the signal's value. At the other extreme, if the signal is very precise, variance is close to zero), then the signal provides a precise estimate of the worker's ability. Phelps suggested two cases that generate inequality" (Fang & Moro, 2011) under statistical discrimination. Case 1, levels-based discrimination, assumes that groups' signals are equally informative, but one group has lower average skills or human capital investment. In this scenario employers assign the group characteristic of lower skills to individual applicants, regardless of the individual's actual level of skill, resulting in disparate treatment between otherwise identical individuals. For example, when states established Ban the Box policies that prohibited employers' questions regarding criminal background, it inadvertently increased racial discrimination. In the absence of this information, employers assumed that Black applicants were more likely to have a criminal record, assigning their belief of a group characteristic to individual applicants. Under levels-based discrimination, if employers are given more information to show that the Black candidate is of higher quality than the expected average (or has a valuable skill employers consider rare in Black candidates compared to Whites), this information is expected to benefit more the Black candidates than the White candidates (Lahey & Oxley, 2021).

Case 2, variance-based statistical discrimination does not imply that on average Black applicants have lower skills than White applicants, but rather that the signal for such skills is not as strong for Blacks as compared to White. In other words, the unconditional distributions of skills are the same between the two groups, but the signals employers receive are differently informative. Under variance-based discrimination, additional positive information will not help Black applicants more than White applicants since the signal will be trusted for White applicants but not for similar Black applicants.

Discrimination Theory: Recent advancements

Although the two main models of discrimination in economics continue to be taste-based models of discrimination (Becker, 1957) and statistical discrimination models (Phelps, 1972; Arrow, 1972; Aigner and Cain, 1977), psychologists have made significant contributions to the understanding of the roots of discrimination (Bertrand and Duflo, 2016). This literature has looked for the foundations of taste-based discrimination by studying "personality development, socialization, social cognition, evolutionary psychology, and neuroscience" (Bertrand and Duflo, 2016). While in the past prejudiced was thought to be an unnatural way of thinking, Tajfel (1979) has shown that social identity plays an important role in prejudicial attitudes. At the same time, evolutionary psychology has demonstrated the importance of social identity and group membership to achieve cooperation. This perspective on discrimination, provides some complexity to the traditional model of 'taste-based' discrimination by exploring some of the roots that may cause 'distaste' between member of the majority and member of a minority group. In this vein, a

growing literature has explored the existence of unintentional or unconscious forms of bias. Implicit associations have been widely studied as a potential source of bias and discrimination in a variety of settings (Banaji and Greenwald, 1995; Bertrand et al., 2005; Dovidio et al., 2002). Researchers have used Implicit Association Tests (IAT), a method to indirectly measure the strength of association between two concepts (Oswald et al., 2013), to explain discriminatory behavior. However, evidence suggests that while IAT does seem to explain unconscious bias, this bias is usually moderated by other environmental factors. In other words, implicit bias does not always translate into discriminatory behavior.

Perhaps more interestingly, this literature provides some evidence that identifying the nature of the discrimination may be more complex than previously considered. Unconscious bias may influence our judgment when assessing expected productivity (Uhlmann and Cohen, 2005; Rudman and Glick, 2001). If this is the case, evidence of taste-based discrimination may be consistent with statistical discrimination in cases where our priors about minority applicants are inaccurate (Bohren et al., 2019). In other words, when we have inaccurate beliefs about signals of quality for a minority applicant, we may not account for this signal or greatly discount them when evaluating minority applicants. This in terms translates into signals of quality not having an impact on employment. From an empirical point of view, if we are to find such results, we would conclude that we have found evidence of taste-based discrimination, however, this would be incorrect. Employers are not necessarily acting based on a distaste for minority applicants, but rather based on inaccurate or incorrect priors about minority applicants. In other words, when distinguishing between taste-based and statistical discrimination we assume that statistical

discriminators have accurate beliefs about minority applicants, and assumption that has been disproved by the presence of unconscious biases.

Empirical work on discrimination in the labor market

Although the above theoretical framework has focused on racial discrimination, these models are also used to explain other forms of discrimination in the labor market, such as by sex, age, beauty, obesity, sexual orientation, and disability. Empirical work that studies discrimination in the labor market falls into two broad categories: non-experimental and experimental studies. I will briefly discuss some of the work done with decomposition methods, but focus mainly on audit studies, as they are more relevant for this study.

The first wave of non-experimental studies to test labor market discrimination used decomposition methods. This approach focuses on identifying the role of discrimination in generating wage gaps between groups. Economists have used wage regression estimations that attempt to control for differences related to productivity between the two groups and interpreting the remaining unexplained gap as an estimate of discrimination (Oaxaca, 1973; Neumark, 1988). There is a large wage gap between black and white men; much of this differential can be explained by differences in skills, at least for young men (O'Neill, 1990). Neal and Johnson (1996) find evidence that after controlling for age and performance on the Armed Forces Qualifying Test (AFQT), the Black-White wage gap is about seven percent and not statistically significant. However, work that is more recent has shown that controlling for other wage-determining characteristics increases the Black-White gap. Particularly, studies have found

that controlling for years of schooling reintroduces the Black-White gap because conditional on AFQT, Blacks tend to get more education than Whites (Rodgers and Spriggs, 1996; Carneiro et al., 2005; Lang and Manove, 2011).

Importantly, challenges arise from estimating discrimination using decomposition methods. First, the unexplained wage differential can always be attributed to unmeasured characteristics related to productivity that go unaccounted for in the regression analysis. Second, the rapid increase in the incarceration rates of Black men between the 1980s and 2000s generates serious problems with some of the data sources used for decomposition analysis. For example, Current Population Surveys exclude institutionalized individuals by design, which in turns removes from the sample frame many Black men. Although decomposition methods pose these challenges, they remain popular because they provide information about actual wages that experimental studies cannot.

Experimental work on labor market discrimination falls into two categories: lab experiments and field experiments known as audit studies.

Lab experiments attempt to mimic personnel decisions by introducing vignettes. In such studies, researchers provide participants with a hypothetical scenario where they must select a job candidate for hiring. Rosen and Jerdee (1974) conducted one of the firsts vignette studies to estimate gender-based discrimination in a bank setting. Managers were provided with a hypothetical scenario that required them to select bank employees for promotion or development; the experimenters manipulated the gender of eligible employees. They found

significantly lower promotions and development for female employees. Correll et al. (2007) use a lab experiment to explore the labor market penalty for motherhood. Undergraduate subject evaluated same sex married applicants who differed on their status as parent. Mothers were rated lower on competence and commitment and were less likely to be recommended for the job. Fathers, on the other hand, were rated higher and more likely to be recommended for the job and a higher salary. Rudolph et al. (2009) conduct a metanalysis of vignettes to estimate weight-based discrimination. They find evidence of adverse effects of obesity on performance evaluations, promotion, and hiring. Büsch et al (2009) use vignettes to estimate age-based discrimination. They find evidence of age-based discrimination in both Germany and Norway. Lab experiments such as these have been useful in our understanding of mechanisms., There is growing concern that the samples used are usually not representative of actual managers and that the hypothetical nature of the study limit greatly limit the generalizability of the findings. As a respond to these challenges, field experiments have risen in popularity to study discrimination in the labor market.

Field Experiments

Field experiments on hiring discrimination are extensive, and they are divided between traditional audit studies and correspondence studies. Traditional audit studies use trained individuals of different races, ethnicities, etc., to pose as job candidates. These individuals are selected in ways that minimize physical differences between them, and they are trained to speak, act and dress similarly. Differences in hiring between the two candidates are treated as hiring discrimination. Correspondence studies, on the other hand, create fictitious job applicants by

randomly constructing résumés. Two of the earliest field experiments are Daniel (1968) and Jowell and Prescott-Clarke (1970) which find evidence of rational discrimination. However, Paiger (2003) is one of the most influential audit studies that estimated that employers were twice more likely to offer a job to a White applicant with a criminal record, than a Black applicant with no criminal record. One of the most influential correspondence studies is Bertrand and Mullianathan (2004) where job openings were identified on the Sunday newspaper in both Boston and Chicago. In this study, Bertrand and Mullianathan signaled the race of job applicants by assigning Black and White names to applicants. They relied on Social Security Administration (SSA) data to identify the most common names by race, and they test the perceived race of such names as way of ensuring that such names signal race. Theirs was the first correspondence audit to rely on names as a signal for race, and they found that applicants with Black sounding name are four times less likely to receive a callback compared to applicants with White sounding names. Moreover, this study showed that increasing the quality of the resume by reducing grammatical errors and increasing work experience, did close to nothing to reduce racial discrimination; these findings are consistent with a variance-based statistical discrimination where quality signals help White applicants but not Black applicants. After Bertrand and Mullianathan's (2004) study, a flurry of audit studies has been published studying hiring discrimination based on race, ethnicity, gender, age, sexual orientation, and nationality (Neumark, 2016). Findings consistently report hiring discrimination against the minority group in the cases of race and ethnicity (Blommaert et al., 2014). A handful of these studies have tried to identify ways to reduce discrimination and therefore, determining the nature of such discrimination, whether taste-based or statistical. Bertrand and Mullianathan (2004), do so by varying the quality of résumés and estimate whether

increased quality reduces racial discrimination. Their study finds that higher quality signals benefit whites more than blacks, a finding consistent with variance-based discrimination. Lahey (2008) estimates age discrimination in hiring and uses signals, such as phrases and hobbies and activities, that indicate increased flexibility and openness to change and activity to deter discrimination. She finds that these signals are less beneficial to younger workers and decrease the amount of discrimination experienced by older applicants, a finding that is consistent with level-based discrimination.

Sports Participation and Labor Market Outcomes Literature

The main theoretical premise through which sports participation may affect labor market outcomes are through the development and signaling of human capital. Human capital formation through sports participation can be through both cognitive skills (Mincer, 1958; Becker, 1964; Spence, 1978) and non-cognitive skills.

Mincer's model of human capital is based on rational theory. He formally explains that higher amounts of training (in length of time) represent higher costs for the individual (both direct cost of training/education and indirect costs of foregone earnings). Therefore, for an individual to choose to invest in higher levels of education/training it must be the case that he or she expects higher future earnings because of this choice. Formally stated:

l = the length of a working life without training

a_n= annual earnings for n years of training

V_n = the present value of their life-earnings at the start of training

r = the rate at which future earnings are discounted

t = 0,1,2,…, l – time, in years

d = difference in the amount of training, in years

e = base of natural logarithms

The present value of life-earnings for individuals with $(n - d)$ years of training is

$$V_{n-d} = \frac{a_n - d}{r}(e^{-r(n-d)} - e^{-rl})$$

The ratio, $k_{n,n-d}$, of annual earnings of individuals with different d years of training is found when $V_n = V_{n-d}$:

$$k_{n,n-d} = \frac{a_n}{a_{n-d}} = \frac{e^{-r(n-d)} - e^{-rl}}{e^{-rn} - e^{-rl}} = \frac{e^{r(l+d-n)} - 1}{e^{r(l-n)} - 1}$$

This implies that $k_{n,n-d}$ is a positive function of r and a negative function of l. That is, people with more training perceive higher earnings and the difference in earnings widens with larger differences in training.

Contributions to the foundations of human capital theory have emphasized the importance of non-cognitive skills in the development of human capital (Heckman and Rubinstein, 2001; Heckman and Carneiro, 2003; Heckman et al., 2006). Participation in sports might affect labor market outcomes through investments in a teamwork ethic, which is valued in the labor market (Rees and Sabia, 2010); or indirectly by providing athletes with access to a network of

employment (Jackson, 2011). Externalities could occur through the impact of sports participation on health, which is a signal of greater current and future productivity (Lechner, 2009; Rooth, 2011). Several studies have studied the relationship between sports participation and labor market outcomes. Initial studies were mainly correlational in nature due to their inability to adequately deal with the self-selection issues that arise when studying the effects of sports participation. I will first discuss these correlational studies and then, explore recent work that has tried to estimate the causal effect of sports participation on labor market outcomes.

In one of the first studies that explored the relationship between sports participation and labor market outcomes, Long and Caudill (2001) find that males who participated in intercollegiate athletics received a wage premium. However, some have questioned the reliability of these findings based on their use of simple OLS regression as an estimation strategy (Lechner and Downward, 2017). As previously mentioned, simple OLS regression does not solve the issue of self-selection into sports. Even when controlling for observable characteristics, we cannot rule out the presence of unobservable characteristics that are both correlated with selecting into sports and labor market outcomes. Ewing (1998, 2007) argues that high school athletes experience higher earnings compared to non-athletes and he shows that former high school athletes experience greater fringe benefits. However, none of these initial studies explicitly attempted to identify causal effects.

Later studies have tried to distinguish between selection and casual effects of sports participation on labor market outcomes through a variety of econometric strategies. Barron et al. (2000) uses

school characteristics, parental income, weight, and height of students as instruments for sports participation and finds positive effects of sports participation on wages. However, the identifying assumption is that the instruments have no direct effect on wages and can only affect wages through participation in sports. However, both school characteristics and parental health directly affect wages in the labor market (Card and Krueger, 1992). Moreover, the literature on "beauty economics" argues that weight and height also directly impact wages (Judge and Cable, 2004). Lipscomb (2007) uses a fixed effects model that exploits information on joining and quitting of clubs by individuals to control for selection effects and finds that high school athletes have higher educational attainment and are more likely to be employed than non-athletes are. Stevenson (2010) uses the enactment of Title IX as exogenous source of variation on female high school sport participation and finds that female high school athletes are more likely to be employed and earn higher wages once they leave high school. Rooth (2011) uses a correspondence study to estimate the effect of leisure sports on callback rates in Sweden and finds that candidates that signaled sports participation were 5 percent more likely to receive a callback compared to otherwise similar candidates. Finally, Lechner and Downward (2017) use a longitudinal administrative panel of German citizens and a matching approach to explore heterogeneous effects of sports participation on probability of being employed and wages. They find that, in general, team sports drive most of the effect on employment, compared to individual sports. Johnes (2018) uses matching methods and a longitudinal dataset to explore the effect of sports participation while in university on earnings in the UK. Through a variety of matching estimators, he finds that sports participation in university leads to an average wage premium of about 5 percent. Baker and Cornelson (2019) exploit the exogenous shock in female sports participation

created by Title IX to estimate the effect of sports participation on female labor market outcomes. They explore the hypothesis that increased female sports participation may lead to improved spatial skills that translates into higher labor market participation in traditionally male-dominated occupations. However, the authors find no evidence of increased spatial skills or higher female participation in male-dominated occupations due to increased female sports participation. Black et al. (2019) explore the effects of sports facilities on a host of outcomes including socioeconomic status. The authors use longitudinal data from the British National Child Development Study and exploit variation in school assignment and the quality of sports facilities. They find no evidence that quality of sports facility impacts adulthood measures of physical and mental health, or socioeconomic status. Except for Rooth (2011), studies that have explored the causal effect of sports participation on labor market outcome have relied on strong assumptions. The literature clearly points to an effect of sports on later-life outcome. However, challenges and further understanding of the mechanisms through which sports participation affects later-life outcomes still require further research.

The link between sports participation and labor market outcomes has been well documented in the literature, yet two main challenges remain: (1) convincingly addressing selection in to sports to estimate unbiased causal effects and (2) exploring other potential mechanisms through which sports participation translates to better labor market outcomes. This study addresses the former by relying on an experimental design to estimate causal effects, and the latter by introducing other treatment categories to assess whether employers may have a taste for athletes. This last point will be explained in more detail in the methods section of this narrative. More specifically,

this study will advance the current literature in several distinct ways: (1) it is the first experimental design to estimate the causal effect of high school sports participation on labor market outcomes in the United States, (2) it is the first study to convincingly explore differential effects of high school sports participation on labor market outcomes for Blacks males in the US, (3) it is the first study to experimentally explore possible heterogeneity in sports participation, team versus individual sports, in the US, and (4) it is the first study that tries to underpin whether the link between high school sports participation is driven by skill formation/signaling or whether employers may have a taste for athletes. In the next section, I explain the audit methodology and review some of the most influential studies to use this method.

Sports Participation and Discrimination

Drawing on the theories of discrimination explain at the beginning of this section, participation in sports can influence labor market discrimination in both models of discrimination: taste-based and statistical.

In a model of taste-based discrimination, employers may have a taste for athletes, making them more likely to hire athletes compared to otherwise similar applicants. Under this scenario, employers, coworkers and/or customers may experience increases in utility from interacting with athletes. This would predict higher labor market outcomes for athletes compared to non-athletes.

In a model of statistical discrimination, sports participation may act as a signal for certain skills valued in the labor market. This signal of quality is expected to decrease differential treatment between Black and White applicants. Moreover, sports participation is expected to be particularly beneficial to Black applicants in a model of levels-based statistical discrimination.

Section 3. Audit Methodology

Non-experimental research on labor market discrimination has relied heavily on decomposition methods to estimate racial differences in employment and wages after controlling for demographic characteristics and productivity determinants (Lang and Lehmann, 2012; Johnson and Neal, 1998; Kochhar, 2008). A classic critique of these techniques is that of potential omitted variable bias which arises when we are unable to account for all relevant characteristics related both to labor market outcome and race (Altonji and Blank, 1999; Quillian, 2006). Due to this shortcoming, researchers have started designing and implementing both lab and field experiments to measure discrimination. The audit methodology has become the standard field experiment for measuring discrimination in different markets: credit, housing, labor, and others.

Field experiments designed to measure discrimination are usually referred to as audit studies. "Audit studies allow researchers to make strong causal claims and explore questions that are difficult to answer with observational data" (Gaddis, 2018). These studies usually refer to field experiments where researchers randomize one or more characteristics of the individual to examine the effect of this characteristic on a particular outcome. Audit studies were first implemented in the 1970s by the Department of Housing and Urban Development to test for discrimination in the housing market (Yinger, 1995). These studies included 3,264 audits across 40 metro areas and found discrimination against Blacks in housing availability, treatment by real estate agents, reported terms and conditions, and the information requested by the real estate agents. During this time, other researchers started expanding the scope of audit studies to include housing discrimination based on parental and relationship status among women (Galster

and Constantine, 1991), racial and gender discrimination in new car prices (Ayres, 1991). Between the 1980s and 1990s, HUD conducted a second wave of audit studies (Yinger, 1991), and the Urban Institute conducted a series of employment audits (Mincy, 1993). The Urban Institute studies focused on discrimination against Hispanics in Chicago and San Diego in 1989. In 1990 the Urban Institute conducted another employment audit in Chicago and Washington DC to study discrimination against Blacks (Turner et al., 1991).

In the 2000s, two influential audit studies were conducted by researchers. Devah Pager's (2003) audit on race and criminal record in Milwaukee documented that employers are more likely to offer jobs to a White felon than to a Black applicant with no criminal record. In 2004, Bertrand and Mullainathan's correspondence study lead the way to an increase in correspondence audit as opposed to in-person audits, while incorporating the use of names and last names as signals for race (Gaddis, 2018). Currently, researchers tend to refer to in-person audits as audit studies, and virtual audits as correspondence audit. I will continue to use these terms when referring to these studies respectively.

In audit or correspondence studies, fictitious individuals who are identical except for race, sex, ethnicity, or any other study's characteristic, apply for jobs. Group differences in call-back rates are interpreted as measures of discrimination. The underlying assumption for this claim is that audit or correspondence studies create an artificial pool of job applicants, among whom there are no average differences by group, by construction. In other words, by removing differences in productivity across groups, differences in outcomes can be attributed to discrimination. Two

main types of audits study have been used to test for hiring discrimination, correspondence studies and in-person audits. I will briefly describe these two methodologies below while referencing some of the most influential studies to use them.

In-person audits

In-person audits, usually referred to as audit studies, require training matched pairs of individuals, usually referred to as testers, who apply to real job openings. Testers need to be carefully matched on a series of characteristics such as age, height, weight, physical appearance, communication style and other characteristics that may be relevant to employers. Resumes are also constructed for each tester to ensure equal levels of education and prior work experience. Testers' physical appearances provide a clear signal for race and gender. An added strength of in-person audit is that they allow researchers to collect qualitative data regarding the interactions between testers and potential employers. Under the assumption that testers have been successfully matched on all relevant characteristics, but race or gender, differences in job offers are attributed to discrimination. However, important challenges arise with this methodology that are worth mentioning. The validity of in-person studies rests on the assumption that testers are perceived as equally qualified applicants who differ only by race or gender. However, there is no way to test this assumption. If employers perceive differences in testers that are not attributed to their race, then differences in job offers would not represent discrimination (Heckman, 1998). Another important critique to in-person audits is that expectations and behavior of testers may systematically affect the results, also known as experimenter effects (Pager, 2007). If testers have certain expectations about their interactions with employers, this may influence their behavior

in ways that bias the results. That is, if Black testers expect to have poor interactions with potential employers, this expectation may negatively influence their behavior and interactions. In this scenario, expectations create a correlation between race and behavior/interaction during job applications and therefore the difference between white and black tester will no longer be race. This would result in a biased estimation of discrimination. An important advantage of audit studies in comparison to correspondence studies, is that researchers can observe actual job offers while correspondence studies only measure callbacks rates.

In-person audit studies have concluded that race affects hiring decisions. Although studies vary in their estimation of discrimination ranging from whites being 1.5 to 5 times more likely to receive a callback or job offer relative to similar Black applicants, they consistently provide evidence for racial discrimination in the labor market. In one of the most influential in-person audits, Pager (2003) shows that White applicants with a criminal record are 18 percent more likely to receive a callback or job offer relative to Black applicants with no criminal record.

Correspondence Studies (CS)

Correspondence studies use constructed resumes to apply to real job openings. Resumes are constructed in a way that matches fictitious applicants on all relevant characteristics, but the manipulation of interest. Correspondence studies only measure callback rates and not actual hiring. This is an important limitation that I will further explore in a later chapter. Correspondence studies have many advantages over traditional audit studies: lower costs per application allow for larger samples, researchers avoid experimenter effects by relying on paper applicants.

In correspondence studies an important aspect of the design is the construction of resumes to make them both seam realistic, and part of the group being studied.

Identification

Correspondence audits remove systematic differences between groups in observable characteristics. This means that for a sample of resumes with Black and White sounding names, other resume characteristics are identical other than group membership. Formally,

$$Y = \alpha + \beta B + \varepsilon$$

Where Y is callback rates and B is an indicator for Blacks, ε is uncorrelated with B by construction, so that OLS estimate $\hat{\beta}$ is an unbiased estimate of discrimination on callback rates. However, some researchers point that there can still be differences in employers' assumptions about unobservable differences between groups (Neumark, 2016). There is a theoretical debate around this issue that Heckman and Siegelman were the first to point out. I further discuss the Heckman critique in its own subsection. However, Yinger (1995) argues that this assumption about unobservable characteristics is indeed discrimination from employers. Given the same information about potential job applicants, any inference about unobservable differences between them is the result of discrimination. Following this reasoning, many researchers don't directly address the critique. However, Neumark (2012) argues that the critique is valid and proposes several corrections for estimating discrimination with correspondence study data. In particular, the underlying issue is that correspondence audit makes it hard for researchers to distinguish between statistical and taste-based discrimination. That is employers may have

different expectations about the mean productivity for Black and Whites, conditional on what they observe, or they simply may have a taste for White applicants. As discrimination theory poses, the more information employers have about applicants we would expect statistical discrimination to decrease. Even if this is true, the issue of distinguishing between statistical or taste-based discrimination with correspondence studies remains. This is one of the limitations of the methodology, the inability to parse out statistical from taste-based discrimination. However, Yinger's argument still holds, regardless of the type of discrimination, given the same information for two job applicants of different races, for example, any differential treatment based on employers' assumption of unobservable characteristic, must be discrimination.

Different set-ups: single vs. matched pair

In an audit or correspondence study, the researcher can decide whether an employer receives only a treatment or a control resume, or whether the employer receives both treatment and control resumes. That is, the researchers can choose to send one applicant to employers or send pairs (or triplets) to each employer to compare between each applicant. This last scenario is known as a matched pair design. The choice of how many resumes to send to a single employer comes with tradeoffs. Matched-pair designs increase study sample size, and it allows for the analysis of how a single employer differentiates different types of resumes through FE models. However, sending matched-pair resumes may make treatments more salient than they otherwise would be if employers are now more likely to compare resumes to each other than if they would only receive one treatment or control (Phillips, 2015). At first glance it may seem as if matched designs would be preferred over unmatched in scenarios where probability of detection is low.

This conclusion comes from the intuition that a matched design would be more efficient as early studies suggested (Kramer, 1991). Most audit studies use a matched-pair design for this reason (Neumark, 2018), and the added power of estimating how a single employer may discriminate among otherwise equal resumes. However, Vuole et al. (2018) argue that the gains in efficiency do not always outweigh the probability of detection. More precisely, whether a matched design will have significant gains in efficiency will depend on the share of employers that differentiate between matched pairs. That is, the percent of employers that call back only one of the profiles in a pair. If most employers treat resumes the same – and the caveat here is, both a positive and a negative response – then a matched pair is desired for efficiency purposes. Specifically, if more than 50% of employers treat pairs equally then a matched design is encouraged (Vuole et al., 2016). This threshold is the point at which each half of the resumes had either the same or different outcome per employer. When more than 50% of resumes experience the same outcome, then the unit of experimentation (an employer per job ad) influences the outcome and matching is preferred because estimation can rely on a smaller sample size. Conversely, when more than 50% of employers treat resumes differently an unmatched design is preferred for its efficiency gains. This seems straightforward at first glance, but the nuance lies in the fact that many job posts turn out to produce a negative response for both resumes in a pair. The overall callback rate of audit studies is not particularly high, which means that in many cases there will not be a differential treatment from an employer because each resume in a pair will not receive a callback.

Researchers have also argued that introducing fake resumes may have spillover effects on the employers' perception of applicant pool, decreasing external validity (Phillips, 2016). Phillips argues that introducing a pair of resumes changes the decision process and has estimated that correcting for this potential spill-over effect increases measured discrimination by about 20% on average. This seems to imply that in matched designs, we could be underestimating the degree of discrimination. This point is particularly problematic in cases where applicant pool is believed to be small to begin with.

There is also an ethical concern if many fake resumes are sent to a particular employer, which leads to hiring managers changing their hiring practices because of fake applicants. Overall, the disadvantages of matched pair designs tend to decrease when job openings receive a larger number of applications. However, most studies employ a matched pair design to control for heterogeneity across employers. In a matched pair design an employer views application from all groups being tested (e.g., Black and White applicants). This allows for estimation of discrimination at the firm-level with the introduction of fixed-effects estimates.

Basic characteristics

Resume information will provide employer with the basic characteristics of job seekers: names, education, job history, and skills. These resume characteristics are used to signal group membership: race, age, gender. Age and gender tend to be straight-forward to signal in resumes by avoiding gender-neutral names and including graduation year (Neumark et al., 2015). Race is traditionally not listed in resumes. Bertrand and Mullainathan use birth certificate data to signal

race through traditionally Black and White names. Since their study, CS have used names to signal race. However, a valid critique is that Black-sounding names may be associated with other characteristics that may influence hiring decision, other than race. Bertrand and Mullainathan try to address this issue by controlling for neighborhood quality, to control for socio-economic status. Addresses, usually included in resumes, are typically used to signal socio-economic status. I further discuss the use of names as a signal of SES in the next subsection.

Other signals have been used to signal group membership, Drydakis (2009) uses volunteering work with the gay community to signal sexual orientation. Many audit or correspondence studies have been conducted to explore discrimination in a variety of group membership: age, gender, race, ethnicity, sexual orientation, disability status, military service, and religion.

Names and race

Race in correspondence studies is signaled mainly through name, though some studies have used club membership instead (Bendick et al., 1994). Since Bertrand and Mullainathan (2004) names have become the main cue for signaling race in correspondence audits. Fryer and Levitt (2004) argue that distinctively Black sounding names are associated with mothers with lower Social-economic status (SES). They study birth certificate data from the state of California from 1961 to 2000, covering over sixteen million births. In their cross-sectional model they find that “Blacker names (defined as an index constructed by the authors – Black Name Index) are associated with lower-income zip codes, lower levels of parental education, not having private insurance, and having a mother who herself has a Blacker name” (p.786). However, their mother-fixed effects

results suggest that these association are weaker than they appear in the cross-sectional analysis. The authors conclude that after controlling for other SES characteristics, there is “no negative relationship between a distinctively Black name and later-life outcomes” (Fryer and Levitt, 2004).

This critique has been unpacked from a variety of different lenses (Butler and Homola, 2017; Gaddis, 2017; Gaddis, 2018) and is to some extent like the Heckman and Siegelman critique. A first counterargument comes itself from Bertrand and Mullainathan (2004) who incorporate signals of SES in their construction of resumes. And while racial discrimination decreases when they control for SES as one would expect under statistical discrimination, a large Black-White callback remains, which the authors argue must be attributed to racial discrimination as signaled by distinctively Black sounding names.

An important distinction when crafting correspondence studies is that what really matters is the signal that evaluators are responding to (Butler and Homola, 2017). As researchers we are assuming that evaluators know these associations a priori and are responding to the signal of SES and not necessarily race. However, correspondence studies have shown that, especially in large cities, employers are unaware of these associations (Butler and Homola, 2017; Carbonaro and Schwarz, 2018). If anything, the argument about Blackness of names affecting the measure of discrimination in correspondence studies (Gaddis, 2017) seems more like a question of strength of the signal rather than confounding effects. When resumes are fully randomized, we are essentially comparing resumes that are identical to one another, on average, except for the manipulation of interest. That employers would assign a lower SES to a distinctively Black

sounding name residing at the same address (highly correlated with SES) as a White sounding name, but not so to a less sounding Black name seems like an issue with the strength of the signal. If we believe employers have a general idea of the demographics of certain neighborhoods and they use this information differently depending how distinctively Black the name is, this must mean that when the signal is strong for race they engage in discrimination.

A potentially more interesting criticism of signaling race through names relates to how generalizable the findings are. That is, not all African Americans have distinctively Black sounding names, one could argue that African Americans without distinctively Black sounding names probably do not experience the levels of discrimination measured in correspondence studies when applying to job postings online. However, if the argument regarding the strength of the signal is valid – it must mean that the measured levels of discrimination eventually catch up with Black applicants. They could avoid being racially discriminated against initially if they do not have a distinctively Black sounding name but may eventually experience similar levels of discrimination once their race becomes evident.

Heckman-Siegelman (HS) critique

As previously mentioned, the typical correspondence study uses a matching approach where the researcher randomly assigns treatment across two 'identical' resumes and sends out both resumes to a job opening (treatment and control) as previously explained. However, authors have criticized this approach arguing that there may be unobserved differences between the resumes that the researcher may not find important, but employers do (Heckman & Siegelman,

1993). Two strategies were developed to overcome this issue: 1) sending one resume per vacancy and pooling across treatment and control to identify the average treatment effect, 2) making the *distribution* of characteristics (components in a resume) identical between treatment and control (Heckman, 1998; Lahey, 2008). The first approach greatly reduces the sample size of audit studies, which would increase standard errors and make inference more difficult. The second approach relies on complete randomization of all resume components including the treatment. One of the strengths of the second approach is that by randomizing all components of the resumes, researchers can test for interaction effects between the treatment and other characteristics.

Despite these solutions to the HS critique, Neumark (2012) argues that the critique still holds because researchers may exclude important information in resumes that employers may value. He develops a formal framework to explain why this would result in biased measures of discrimination and subsequently develops a model to correct for this issue. Neumark's argument follows that when there is incomplete information correlated to productivity in constructed resumes, then employers will use their expectations to assign a productivity value to each resume and this expectation may differ between White and Black applicants. That is, employers view productivity as the sum of both observed measures of productivity (resume characteristics) and unobserved measured of productivity (information not readily available in a resume). For audit studies to produce unbiased estimates of discrimination must mean that the expected value of unobserved measured of productivity must be equal between White and Black applicants, as the observed measures are made equal by construction. Heckman (1998) and Heckman and

Siegelman (1993) argue that audit studies, and even correspondence studies, violate this assumption. Moreover, they argue that differences in what employers estimate to be the mean unobserved productivity-related characteristics between White and Black applicants would affect the results. This is surprising because it seems to directly refer to Becker's definition of statistical discrimination. If employers are assigning different levels of expected mean unobserved characteristics to resumes based on their group membership, race in this case, then by definition they are engaging in statistical discrimination (Phelps, 1972). While in-person audits still need to address the issue of potential unobserved, uncontrolled characteristics between testers there seems to be a consensus among researchers that the HS critique is not relevant in correspondence studies with full randomization or unmatched designs.

Limitations

Audit and correspondence studies have been deemed a powerful tool for researchers to study discrimination due to their high levels of internal validity and relatively straightforward design. However, there are important limitations to this methodology that must be explored to fully understand what can and cannot be learned from these studies.

One of the most evident limitations of audit and correspondence studies is that they are poorly suited to address questions regarding mechanisms. Some researchers sometimes refer to audits as black boxes, with which one can convincingly measure discrimination but not necessarily uncover the mechanisms that drive such differences (Pedulla, 2018). Some researchers have

come up with clever ways to uncover mechanisms with the methodology, but it is not always feasible (Lahey, 2008; Tilcsik, 2011; Kaas and Manger, 2011).

Another limitation that was been discussed throughout this chapter is the reliance on indirect signals. A lot of the assumptions of design rely on our expectation that decision-makers understand these signals as we understand them. However, that is not always the case as some studies have found that in certain contexts, employers may be unaware of characteristics like neighborhood demographics and school quality (Carbonaro and Schwarz, 2018). The consensus seems to move towards a need to test signals before fielding the study. However, this is not always straightforward. Eliciting honest responses for some of the more sensitive or controversial topics is not an easy task given issues of social desirability bias.

An important area for future development is that some recruitment processes have started to rely more heavily on AI (artificial intelligence) to make initial decisions. Future research needs to seriously engage with how this potentially changes the landscape for these types of studies. There is already evidence that algorithms can discriminate as well (Obermeyer et al., 2019). The question is how these processes affect the correspondence methodology and how will it need to be adjusted to incorporate these changes in hiring practice.

Another important limitation is the narrow scope of the outcome variables. Most studies look solely at differences in call-back rate, which limits the measure of discrimination to only the initial stages of job search. Some researchers have tried to complement their studies by looking at later

stages of the process or combine their results with readily available data (Pager and Quillian, 2005; Correll et al., 2007; Pedulla, 2016). However, the limitation remains that the measured variable in audit and correspondence studies is usually restricted to the initial stage of the hiring process. Although it is important for researchers to continue to think of creative ways to incorporate later stages of the hiring process in the design of audit and correspondence studies, researchers have argued and empirically demonstrated that the discrimination in early hiring stages translate into worse off outcomes in later stages. Bendick et al. (1999) and studies conducted by the International Labor Organization (ILO) estimate that around 90% of labor market discrimination occurs in the initial stages of interviews and job offer (Riach and Rich, 2002).

In the next section I will describe the research questions and explain the experimental design used in this correspondence study.

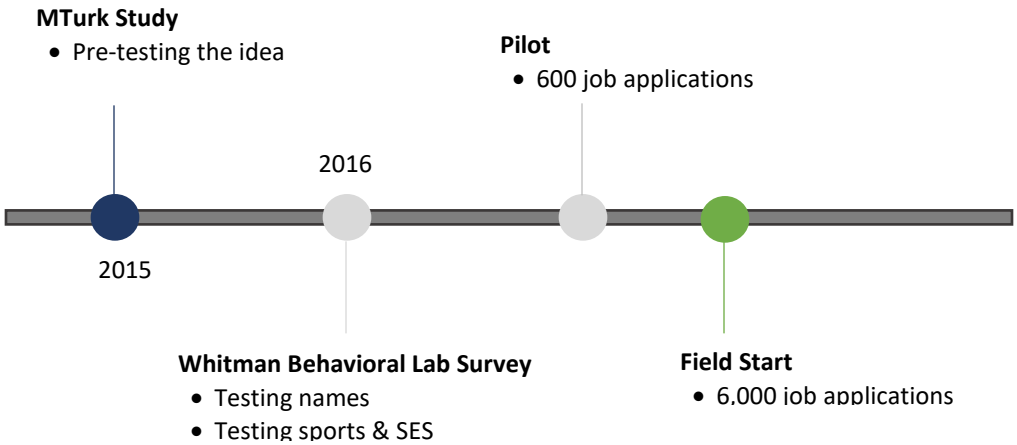
Section 4. Research Questions & Experimental Design

This dissertation attempts to answer whether recent high school graduates who participated in sports are more likely to receive a callback from a low-skill, entry-level job ad compared to otherwise similar applicants who did not participate in sports.

Second, whether high school sports participation is particularly beneficial for Black applicants, and therefore reduces discrimination in the labor market. This outcome would be consistent with statistical discrimination, which states that if employers are given more information to show that the Black candidate is of higher quality than the expected average (or has a valuable skill employers consider rare in Black candidates compared to Whites), this information is expected to benefit more the Black candidates than the White candidates (Lahey & Oxley, 2021).

The road to an experimental design

Timeline



Designing an experiment to answer specific research questions entails a long journey of sifting through choices and making unavoidable trade-offs. The perfect experiment doesn't truly exist, as we usually make trade-offs in almost all design decisions. Sometimes we need to forgo external validity to maintain internal validity, other times we must forgo answering all interesting questions to ensure we have enough power to answer those we have deemed most important. This experiment was in the field for 6-months, yet it went through a process of pre-testing, data collection, and piloting that took around 2 years. In this section, I will trace my steps through a pre-test of the idea, a perception analysis for race, sports and SES, and the piloting of the study, all of which were instrumental to the final design. With this tracing, I hope to explain the choices I made along the way, to better understand the final experimental design. Hopefully, this will leave the reader with a better understanding of the trade-offs faced and the reason I settled on the final design, as well as the implication of these choices for identification, external validity, and limitations of the experiment. In this section I explain in detail the following choices and walk through their trade-offs: target population, extracurricular activity status and which sports, names to signal race, addresses, high schools, and the geographical scope of the study. In this process I take the following steps:

1. MTurk Study to test the idea that sports participation matters for employment [Fall 2015]
2. Whitman Behavioral Lab Survey [Spring 2016]
3. Pilot of final experimental design (600 job applications) [Fall 2016]
4. Start of field work (6,000 job applications) [Winter/Spring 2016/17]

Setting up the idea – A choice Experiment with Conjoint Design

Prior to setting up an experiment with real employers, I conducted a survey experiment. I employed the now widely used online tool Amazon Mechanical Turk (MTurk) that allows researchers to efficiently identify and recruit participants for surveys and experiments.

For this survey experiment, I constructed profiles applying to an entry-level administrative position and randomized the profile characteristics. I then asked respondents to choose between the two profiles to fill a job opening described in the vignette. I randomly manipulate extracurricular activity and gender in each profile to estimate the effect of these characteristics on the probability of being chosen.

I do not find a statistically significant effect of participating in basketball while in high school on the probability of being chosen. Similarly, there's no statistically significant effect of participating in marching band on the probability of being chosen. However, gender seems to have a statistically significant effect on the probability of being chosen. Females are disproportionally chosen compared to males. Males seem to be about 50% less likely of being chosen for the job. Detailed information about the study design, estimation and results can be found in Annex A.

Creating profiles for the MTurk study introduced the first trade-off I faced as it is very unlikely that college graduates will indicate in their resumes whether they participated in high school sports, therefore the study was confined to high school graduates.

Target population

Focusing on high-school graduates meant limiting the study's external validity. This decision made sense to me for several reasons. First, work experience during and after high school matters for later-life outcomes (Carr et al., 1996). Using the National Longitudinal Survey of Youth, the authors find that work experience in high school positively affects labor force participation, employment status, and income even a decade after graduation. Moreover, this is a population of interest from a policy perspective since there is a growing number of people ages 16 to 24 that are both out of school and out of work, 6.5 million in 2011. Studies have shown that this vulnerable population, often referred to as disconnected youth, is burdened with future financial instability and grim employment prospects (The Annie E. Casey Foundation, 2012). Furthermore, young Blacks, ages 18 to 19, are twice more likely to be unemployed compared to Whites (Bureau of Labor Statistics, 2015). This implies that this is a particularly relevant population that stands to benefit from increased employment prospects and reductions in employment discrimination.

Second, if sports are particularly beneficial for the employment prospects of Blacks, the policy implication of the study was quite straightforward given the gap in athletic opportunities for high school students in predominantly minority schools (NWLC and PRRAC, 2015). That is, if sports participation signaled higher quality applicants, or developed skills valued in the labor market, theory would suggest that it would have a protective effect against discrimination. Blacks participating in sports would, in theory, perceive lower levels of discrimination than otherwise similar Black applicants who did not participate in sports. Third, choosing college graduates

presented two possibilities: listing college sports participation, which represents a much smaller portion of students compared to high school athletes or stating high school sports participation as a college student/graduate[1]. This last option did not seem in line with resume samples from the real world, where such activities were no longer listed, which further harmed the external validity of the study. For these reasons, I chose to limit the scope of the study to high school sports participation and low-skill, entry-level job opportunities.

Choosing Extracurricular Activities

The next manipulation, extracurricular activity, brings me to the second design choice. It became evident that comparing profiles with sports participation to profiles with no sports participation didn't quite answer the question of whether sports matter. The effect of sports is not necessarily obvious from a design that only incorporates 'sports' or 'no sports', it is unclear whether what matters is the sport or just having something 'extra' in your profile. Including manipulations helps researchers better understand the effect of certain characteristics on call-back rates (Pedulla, 2018). The extracurricular activity manipulation had three different *levels*: participation in the high school basketball team, participation in the high school marching band and no extracurricular activity. The reason for choosing basketball among the possible high school sports is twofold. First, choosing a gender-neutral sport allows for complete randomization without worrying about restricting cases due to a match between gender and sport that is highly unlikely (i.e., female playing American football), this greatly simplifies the analysis. Second, basketball is

[1] In the 2018-2019 academic year almost 8 million high school students participated in sports, compared to the roughly 500,000 college-athletes (NFHS, 2021; NCAA, 2021)

by far the most popular high school sport for both females and males (NFHS, 2015), which allows for complete randomization across gender

Participation in high school band is introduced for reasons previously mentioned. If I were to compare profiles with no extracurricular activities with profiles with high school sport participation, I wouldn't be able differentiate between the effect of sports and the effect of having an extracurricular activity listed in the profile, regardless of it being sports. Including another extracurricular activity, which has analogies to sports, increased my ability to estimate a causal effect on high school sport participation. Ideally, we want extracurricular activities that develop skills like those developed through sports participation. Feldman and Matjasko (2005) argue that certain extracurricular activities, like athletics, place students in a team environment where they develop important social and leadership skills that are valued in the labor market (e.g., leadership skills, being a team player, discipline). The literature has identified the following activities as meeting the characteristics mentioned above (Feldman and Matjasko, 2005): marching band and debate club. The introduction of these activities allows testing whether there is something particular about sports that is valued in the labor market. Finally, a third manipulation includes no extracurricular activities.

High School Sports

The choice of sport and extracurricular activity was one that required careful thought for several reasons. First, some sports are severely segregated by gender, played almost exclusively by boys or girls in high school (i.e., American Football for boys and Softball for girls). This gender segregation prevented a fully independent randomization of all resume characteristics, which

tends to be avoided in audit studies as it complicates the analysis. Second, some sports are highly associated to socio-economic status which meant confounding the effect of sports with a potential signal for SES. Table 1 shows how students from Syracuse University and Northwestern University perceived the SES of high school students participating in specific sports. Third, some sports have a marked cultural aspect to them, which may signal other attributes to employers. That is the case of Lacrosse and the lax bros (Randall, 2012). Fourth, some sports are highly dependent on geographical location which would reduce the possible scope of the field experiment and external validity of the findings. These nuances in sports and extracurricular activities led me to narrow activities to widely popular and gender-neutral sports and activities. These included: basketball, track and field, soccer, debate team, marching band, and instrument. Another important aspect of choosing extracurricular activities was the incorporation of both single and team activities as the literature suggest potential heterogeneity in their effects on labor market outcomes due to the development of different skills (Lechner and Downward, 2017).

Table 1. Extracurricular activities and SES

Activity	High Income	Low Income	Cannot tell
Basketball	15%	14%	71%
Marching Band	18%	11%	71%
Tennis	61%	0%	39%
Debate Club	19%	10%	71%
Soccer	23%	8%	69%
Student Council	53%	1%	46%
Track & Field	19%	7%	74%
Instrument	31%	6%	63%

After narrowing the population of interest and the manipulations of extracurricular activities (including the choice of sports), to continue crafting the design I needed to start focusing on the race signal. This meant choosing names that according to the Social Security Administration (SSA) were widely used for the cohort studied and perceived as distinctly Black or White.

Measuring perceived race and SES

In collaboration with the Whitman Behavioral Lab, I conducted a survey to measure perceived race of names the most common names by race from the SSA. In addition, I measured associations of extracurricular activities with socio-economic status. This was an important step, because most audit studies do not previously measure potential confounding characteristics, such as SES, with manipulations of interest (Pedulla, 2018). This step helps to build the argument that the chosen sports do not necessarily signal a particular SES, which means I can rule this out as a potential mechanism through which sports affects call-back rates.

Table 2. Names and perceived race

Names	Likely White	Likely African American	Cannot tell
Jamal Washington	1%	86%	14%
Daryl Williams	1%	81%	17%
Tyrone Lewis	2%	87%	11%
Aisha Jackson	1%	82%	17%
Latoya Wilson	1%	87%	12%
Keisha Jones	7%	64%	30%
Monique Lewis	1%	76%	23%
Greg Evans	69%	2%	29%
Jack Baker	69%	1%	31%
Todd Olson	69%	1%	30%
Emily Carter	68%	0%	31%
Hannah Myers	74%	0%	26%
Amy Price	70%	1%	29%
Laurie Fisher	69%	1%	30%
Terrell Johnson	12%	50%	38%
Luke Russell	69%	2%	30%

Source: Names from the Social Security Administration and perceived race is a calculation based on survey collected by the Author. Shaded names are those used in the experimental design.

Table 2 shows names drawn from the SSA and the percent of respondents that classified them as either: likely to be White, likely to be African American or cannot tell the race. Most of the names were perceived as belonging to the corresponding race (Table 3), however, about a fourth to a third were deemed as 'cannot tell'. It is plausible that this is a result driven by social desirability bias, however, I chose to omit these names in the experiment design and work strictly perceived as the corresponding race. A total of 12 names, 6 African American soundings names and 6 white sounding names – shaded in Table 6. Within each race group, half of the names are female names and the other half male.

Table 3. Perceived race of names

	Perceived Race			
Race SSA	White	African American	Cannot tell	Total
White	69.54%	1.02%	29.43%	100%
African American	3.14%	76.63%	20.23%	100%
Total	36.46%	38.69%	24.85%	100%

Source: Names from the Social Security Administration and perceived race is a calculation based on survey collected by the Author.

Table 4 shows the percent of respondents that classified a participant of popular extracurricular activities (NFHS, 2016) as being: likely high SES, likely low SES, cannot tell SES. From this table it was evident that certain sports and extracurricular activities were strongly associated with income, such as: tennis and student council. More importantly, it confirmed that the sports and extracurricular activities chosen (basketball, track and field, marching band, and debate club) did not seem to be strongly associated with socioeconomic status.

Table 4. Perceived socio-economic status of extracurricular activities

	SES			
Activity	High Income	Low Income	Cannot tell	Total
Basketball	15.12%	14.34%	70.54%	100%
Marching Band	18.39%	10.73%	70.88%	100%
Tennis	60.92%	0.38%	38.70%	100%
Debate Club	48.85%	3.46%	47.69%	100%
Soccer	23.46%	7.69%	68.85%	100%
Student Council	53.28%	0.77%	45.95%	100%
Track & Field	18.77%	6.90%	74.33%	100%
Instrument	30.65%	6.13%	63.22%	100%
Total	33.69%	6.30%	60.02%	100%

Location, High Schools & Addresses

After choosing names and extracurricular activities, I needed to determine the geographical scope of the field experiment. Boston and Chicago had been the most widely used sites for correspondence studies, as a lot of the studies were trying to build on the data collected by Bertrand and Mullainathan (Gaddis, 2018). However, I wanted to include more geographical variation to the study. I also needed a larger sample given that different manipulations in the design. These reasons led me to initially chose Chicago, Dallas, and NYC. However, as noted by Carbonaro and Schwarz (2018) sites need to be piloted to ensure adequate number of job postings to make an audit feasible. As I further discuss below, a pilot conducted prior to the field confirmed that these locations had large enough volumes of job postings to field the experiment in roughly 6-months. Timing and length of the audit is important as one needs to consider the business cycles, the job candidates created, and the jobs being applied to (Lahey and Beasley, 2009). In this setting, of high school graduates looking for low-skilled, entry level jobs, the corresponding cycle turned out to be mostly spring/summer. This will become relevant in the limitations section where I discuss if and how this may affect the generalizability of the findings.

The next steps included choosing high schools and addresses. Originally, I set out to choose addresses based on what researchers conducting audit studies traditionally do. Addresses were first identified based on demographic characteristics associated with the corresponding zip code that would ensure a neighborhood with both demographic and socio-economic diversity. However, zip codes with these characteristics were predominantly associated with highly segregated public high schools. I wanted to avoid using segregated schools because they could

clearly signal SES or school quality, which would create differences between Black and White resumes by construction. The alternative would be to allow complete randomization, but this also meant that even if race and school were independently randomized, profiles would not be otherwise 'equal'. As one profile could be assigned a predominantly white school associated with higher SES and quality, while another would be assigned a predominantly minority school associated with lower SES and quality. This seemed like a threat to the internal validity of the matched-pair experimental design and led me to work solely with integrated high school. Clearly, this decision has implications for the generalizability of the findings, and I discuss these in the limitations section.

Table 5. High school characteristics

			Demographics					
Zip	City	School	White (%)	AA (%)	Hispanic (%)	ELL (%)	Low income (%)	Graduation rate (4-year)
	Chicago	Amundsen HS	27%	11%	66%	11%	89%	76%
	Chicago	Chicago Academy HS	26%	14%	58%	6%	82%	82%
	Dallas	Hillcrest HS	14%	18%	68%	23%	77%	89%
	Dallas	Seagoville HS	20%	19%	61%	24%	87%	84%
	NYC	Abraham Lincoln HS	35%	40%	24%	13%	61%*	66%
	NYC	Herbert H. Lehman HS	20%	24%	55%	10%	64%*	63%

*Free lunch eligible

Source: NCES, 2017

Instead of choosing addresses, I first chose integrated high school of similar qualities, and worked backwards to identify zip codes within the catchment area of those schools that exhibited demographic diversity. Using the National Center for Education Statistics (NCES) report cards, I sifted through the high school public system in Chicago, Dallas, and NYC to choose high schools in each city. I identified integrated high schools with comparable graduation rates, percent ESL (English as a Second Language) and percent free and reduced lunch in each city. I choose two high schools in each city and add a high school from a different city to complete the matched designed of three resumes per job posting. Table 5 shows the demographic characteristics of the chosen high schools in each city. Addresses were chosen based on the catchment area of the high school the applicant attended. I then use Zillow to ensure the addresses chosen across the catchment areas of the different high schools have comparable mean house prices. I do this to prevent differing signals of SES across addresses (Bertrand and Mullainathan, 2004).

Other resume characteristics

I created email addresses using Gmail.com, and each name has its own email address. This is different from previous correspondence studies, which tend to use generic emails that do not correspond to a particular name (Bertrand and Mullianathan, 2004). This is probably because Gmail has very strict policies about the number of accounts a single individual can open, which made the email creation process a tedious one. However, having individualized email accounts is widely common in the real world so one would worry about the signal that an email address that does not include the individual's name sends. In this sense, this study design more closely mimics real world resumes. For telephone numbers, I buy virtual numbers from Sonetel.com. Each name-

city combination has a telephone number[2]. I recruited volunteers to record outgoing messages; volunteers were matched to each name based on their race and gender. This was done to increase the realism of the resumes.

I chose work experiences by identifying real resumes in Indeed.com of people with a similar profile to my applicants (young high school graduates) applying to entry-level jobs[3]. Each work history was supplemented with a list of skills that was also randomly assigned to each resume. These skills included both general work ethics as well as specific skills relevant for the type of job[4].

The Pilot

During fall 2016, I conducted a pilot of the experimental design to ensure localities generated a large volume of job postings, estimate initial effects for power calculations, and test the overall implementation of the resume creation and data generating processes. For this pilot I applied to a total of 200 job postings and sent a total of 600 resumes – three per job posting. The following tables show the descriptive statistics for the pilot. Table 6 with descriptive statistics for the resumes sent during the pilot, shows that the randomization worked properly and that job openings seemed to behave similarly in all three cities.

[2] A total of 36 virtual numbers were bought, 12 phone numbers in each city.

[3] Work experience was chosen in a way that aligned with the job opening identified. For jobs in the food industry, resumes were given one year of part-time experience in a restaurant. Applicants to retail openings were given one year of part-time experience in a retail store and those applying to administrative positions were given one year of part-time experience as an administrative assistant.

[4] For example, applicants to administrative positions were given 'experience operating a multi-line phone system' as a skill. Applicants to retail position were given knowledge and experience with Point of Sale (POS) systems.

Table 6. Pilot descriptive statistics

Descriptive Statistics

	Resume
Sports	0.33
Other EC	0.33
No EC	0.33
Males	0.51
Females	0.49
Blacks	0.50
Whites	0.50
Chicago	0.38
Dallas	0.33
New York	0.30
Administrative	0.35
Food	0.33
Sales	0.33
Other city hs	0.33
N	600

Tables 7 shows the call-back rates and the percent of employers that treated triplets of resumes differently. The table 8 shows the same statistics for Bertrand and Mullainathan (2004) to put the findings in perspective. In general, response rate and the fraction of employers that engage in differential treatment was higher in the pilot than for Bertrand and Mullainathan (2004). A percent of employers who differentiate of 13.5 suggest that a matched design is preferred since it does not meet the threshold 50% established by Voule et al. (2018) for unmatched designs.

Table 7. Pilot average call-back

Average call-back

	Resume	Job
Call-backs	72	39
Employers who differentiated	-	27
Call-back rate	12%	19.5%
Percent of employers who differentiate (identification)	-	13.5%
N	600	200

Table 8. Putting table 7 in perspective (Bertrand and Mullainathan 2004)

Average call-back

	Resume	**Job**
Call-backs	410	154
Employers who differentiated	-	127
Call-back rate	8%	11.6%
Percent of employers who differentiate (identification)	-	9.6%
N	4890	1323

Table 9 explores the possibility that employers may detect the experiment through examining the effect of order in callback rates. Research use order to test whether the order in which resumes are received is correlated to callback rates. The intuition behind the test is that when employers detect a possible audit study, the probability of a callback reduces with the order in which the resumes are sent. That is, first resumes are more likely to receive a callback, when employers have not yet detected a possible audit study. Table 9 shows that order does not seem to affect the probability of callback.

Table 9. Testing detection in pilot

Resume order: testing for detection

	(1)
First resume	-0.025
	(0.022)
Second resume	-0.005
	(0.021)
Constant	0.13
	(0.024)
Observations	600

Notes: Results from a simple LPM. SE clustered at the ad.

Table 10. Effects of race and gender in the pilot

Race and gender

	(1)	(2)	(3)
Black	-0.046		-0.046
	(0.026)		(0.026)
Female		-0.023	-0.022
		(0.027)	(0.027)
Constant	0.143	0.131	0.15
	(0.026)	(0.024)	(0.031)
Observations	600	600	600

Notes: Results from a simple LPM. SE clustered at the ad.

Table 10 shows the effect of race and gender on likelihood of callback. Column (1) looks at the effect of race and estimates that resumes with a Black sounding names have 4.5 percentage points reduced likelihood of receiving a callback compared to resumes with a White sounding name. Column (2) shows that females have a 2.3 percentage points reduced likelihood of receiving a call back compared to otherwise similar male applicants. Column (3) shows the result in a single equation, the point estimates do not change in column (3) which is expected when the randomization works. Since both race and gender are independently randomized in each profile, the point estimate should not change when the estimation is done independently or combined.

Using the estimates in Table 10 I estimate power calculations with G*Power to estimate the required sample size given the point-estimate. Figures 1 and 2 show the results of the power calculations, which estimate a sample size of roughly 6,000 to detect a point estimate of -0.046 in the difference in callback rates between Black and white applicants.

Figure 1. Central and noncentral distributions

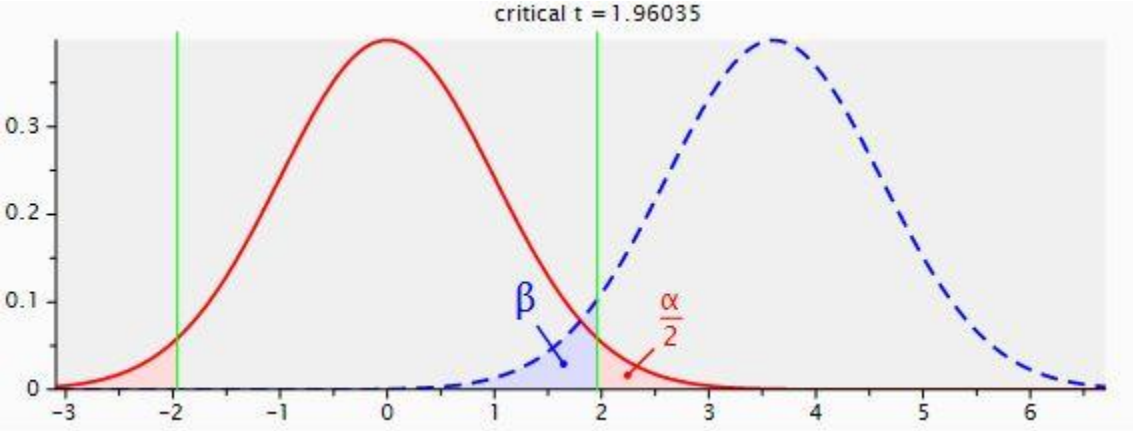


Figure 2. Protocol of power analyses

t tests – Means: Difference between two dependent means (matched pairs)

Analysis:	A priori: Compute required sample size		
Input:	Tail(s)	=	Two
	Effect size dz	=	0.046
	α err prob	=	0.05
	Power (1−β err prob)	=	0.95
Output:	Noncentrality parameter δ	=	3.6056489
	Critical t	=	1.9603502
	Df	=	6143
	Total sample size	=	6144
	Actual power	=	0.9500276

Table 11 explores the effect of sports participation on likelihood of response. Column (1) looks at the pooled effect of sports participation and other extracurricular activity on likelihood of response. Column (2) looks at the effect by sports and pooled extracurricular activities. Column (1) suggests that sports participation positively affects likelihood of response as well as extracurricular activities. Column (2) suggests that the effect of sports participation is being driven mainly by basketball. Soccer seems to negatively affect call back rates, as well as track and field.

Table 11. Effects of sports in the pilot

Effect of high school sports on likelihood of response

	(1)	(2)
Sports	0.005	
	(0.021)	
Basketball		0.060
		(0.046)
Soccer		-0.043
		(0.037)
Track & Field		-0.003
		(0.035)
Other Extracurricular	0.02	0.02
	(0.022)	(0.022)
Constant	0.115	0.105
	(0.019)	(0.022)
Observations	600	600

Notes: Results from a simple LPM. SE clustered at the ad.

Table 12 shows the effects of race, gender, and sports combined in the likelihood of response. Column (1) looks at the interaction between race and sports participation. The estimation suggests a protective effect of sports participation for profiles with Black sounding names. Column (2) looks at the interaction between gender and sports participation. The interaction effect also suggests an added benefit of sports participation for female profiles for likelihood of response.

Table 12. Effects of race, gender, and sports in the pilot

Race, gender and sports

	(1)	(2)
Black	-0.052	
	(0.034)	
Female		-0.036
		(0.03)
Sports	0.004	0.015
	(0.039)	(0.037)
Black*Sports	0.015	
	(0.061)	
Female*Sports		0.041
		(0.055)
Other Extracurricular	0.024	0.019
	(0.023)	(0.023)
Constant	0.14	0.14
	(0.026)	(0.029)
Observations	600	600

Notes: Results from a simple LPM. SE clustered at the ad.

Finally, the remaining tables show preliminary estimates of the effects of race, gender, and extracurricular activities on call-back rates. Overall, point estimates are in the expected direction.

Results from these tables had a direct impact on the final experimental design:

(1) The chosen cities generated a high enough value of job ads to warrant the experiment feasible in these locations and no large differences between cities where encountered.

(2) Racial discrimination estimates were like those found by Bertrand and Mullainathan (2004), which is always encouraging.

(3) Sports seemed to matter, even if the pilot did not produce enough statistical power to render the effect statistically significant.

(4) Estimates pointed to heterogeneity in sports but a clear trade-off with precision, as previously discussed throughout this document. These estimates led to the removal of soccer as a team sport for the final design.

(5) Only 13.5% of employers engaged in discriminatory behavior, clearly not making the threshold established by Voule et al. (2018) of 50%, suggesting that a matched design was to be used in the field.

(6) No evidence of detection as estimated by the effect of resume order (Bertrand and Mullainathan, 2004), further encouraging a matched design.

(7) Some encouraging results from the potential protective effect of sports for Black resumes.

(8) The process of randomization and data generation worked.

(9) Voicemails and email processes worked properly.

Bringing all the pieces together

To test for differential callback by high school sports participation, I sent paired resumes matched on all characteristics except for sports participation as indicated by their involvement in extracurricular activities, to entry-level job openings via Indeed.com. After creating the profiles, I sent resumes to job openings in Chicago, Dallas, and New York City. Over a period of six months, from March 2017 to August 2017, 30 entry-level job advertisements were identified across the three cities. To increase variation in the type of jobs applied, I chose job openings in retail, the food industry, and administrative positions. Once a job ad had been identified as belonging to one of the cities and job types, I verified that the ad was recently posted (not older than 5 days), that I hadn't previously applied to this job, and that I hadn't applied to a position from the same

employer in the past month. Recently posted jobs were chosen to increase chances of a callback, since older postings might have already been filled. As common practice with audit studies (Lahey 2008), I wait a month before applying to openings from repeated employers to minimize chances of detection. Once a job that meets these criteria is identified, I used Lahey and Beasley's randomizer to construct the resumes sent[5]. All the elements of the resume, including font-style are randomly assigned to each resume. The program chooses each line on a resume based on a set of choices that I feed into the program. I will walk the reader through an example, to illustrate how the randomization works.

Each randomization process constructs three resumes: one with sports participation, one with a non-sport extracurricular activity and one with no extracurricular activity. This allows for the comparison of "sports" and "other extracurricular activities" to a baseline of "no activity".

The first decision the program faces is what the style of the resume will be, which includes font-style, size, color and heading styles. The program randomly assigns each resume with a particular style. The next choice is the name on the resume. The program chooses from a total of 12 names, six Black soundings names and 6 white sounding names, chosen in the above detailed Whitman Survey. Within each race group, half of the names are female names and the other half male. Names are chosen without replacement. Once a name is chosen the program is instructed to assign the correspondent email address and telephone number associated with the name.

[5] Lahey and Beasley's randomizer program is straightforward and powerful; however, a basic understanding of html language is necessary in order to set up the randomizer to adequately structure a resume.

After pairing each name with the correspondent email address and telephone number, the randomizer chooses a home address for the resumes. Each city has six potential addresses from which the program chooses without replacement. These addresses were chosen based on the catchment area of the high school the applicant attended. All applicants have earned a high school diploma, and the school attended is chosen to match the address on the resume.

The last section of the resume was devoted to 'Other skills/Activities' this section always included knowledge of Microsoft Office programs (Word, Excel and PowerPoint). This section also included extracurricular activities. In each cluster of three resumes, one applicant was randomly assigned to a high school sport, one to another extracurricular activity and the third to no other activities[6]. Among those assigned to high school sports the program chose between basketball and track & field[7]. Basketball and track & field were identified as the most popular team and individual sports across the three study cities, with the largest number of both males and female athletes (NFHS, 2016). For other extracurricular activities, an applicant is either a member of the marching band or the debate team[8].

The resume randomizer created a Word file for each resume and an Excel file that summarized all the decisions that created the resume. After a cluster of resumes was sent to a job opening,

[6] Since this section always included information about MS Office, all three resumes in a cluster had the same number of subsections regardless of whether they were assigned to an extracurricular activity or not. This was done to ensure the length of the resume remained constant across treatments.

[7] A pilot conducted during October 2016 found evidence of heterogeneous effects of sports participation on callback rates. Team sports seemed to matter more for callbacks than individual sports. This insight led to the decision to have at least one of each sport. However, having more than two sports also implied that more power was needed to make any claims about heterogeneous effects of sports on callback rates.

[8] These activities were also chosen based on popularity and results from a pilot study that found evidence of heterogeneous effect of extracurricular activities on callback rates.

email addresses and voicemails were checked regularly for responses. All emails or voice messages that mentioned interest in the candidate such as an invitation to interview were coded as a positive response.

Samples for the different resume styles can be found on Figure 2A on the Appendix.

Estimation

The main equation looks at the effect of sports participation on responses:

(1) $\text{pr}[Response_i = 1] = B_0 + B_1 Sports + B_2 Other\ Extracurricular + B_3 Female + B_4 Black$

Because each job opening receives three somewhat similar resumes, I cluster standard errors at the job opening level.

Section 5. Main Results

Table 13 includes descriptive statistics of all résumés sent and of those that received a positive response. The first column of this table is evidence of randomization since characteristics should be roughly equal across applicants by design. The second column of this table shows the characteristics of applicants who received a positive response. Applicants who received a positive response are on average more likely to have participated in high school sports, be females and have a White sounding name. On average, applicants in Chicago and applying to job openings in the food industry are more likely to receive a positive response.

Table 13. Descriptive statistic of sent resumes

Descriptive Statistics

	All Resumes	**Positive Response**
Sports	0.33	0.37
Other EC	0.33	0.33
No EC	0.33	0.29
Males	0.51	0.45
Females	0.49	0.55
Blacks	0.50	0.35
Whites	0.50	0.65
Chicago	0.35	0.35
Dallas	0.34	0.34
New York	0.31	0.31
Administrative	0.34	0.29
Food	0.34	0.39
Sales	0.33	0.32
N	6000	659

Table 14. Main differences in call-backs

Average callback rates by race, gender and sports participation

	I. Overall callback rates		
	(percent callback)		
White	Blacks	Ratio	Difference (p-value)
11.01	6.00	1.84	5.00
[2988]	[3012]		(0.000)
	II. Callback rates for males		
	(percent callback)		
White	Blacks	Ratio	Difference (p-value)
10.51	4.30	2.44	6.20
[1541]	[1533]		(0.000)
	III. Callback rates for sports participation		
	(percent callback)		
White	Blacks	Ratio	Difference (p-value)
10.77	8.57	1.26	2.21
[891]	[1109]		(0.095)
	IV. Callback rates for basketball participation		
	(percent callback)		
White	Blacks	Ratio	Difference (p-value)
12.28	11.76	1.04	0.52
[464]	[476]		(0.095)
	V. Callback rates for male basketball participation		
	(percent callback)		
White	African American	Ratio	Difference (p-value)
11.28	8.65	1.30	2.60
[284]	[231]		(0.026)

Notes: Each section presents the average callback for White and African American, overall and by extracurricular participation and race. In brackets is the number of resumes sent to each group. The last column shows the difference across groups and the p-value from a test of proportion, testing equality of callback rates across groups

Table 14 shows the differences in callback rates between White and Black sounding names: overall, for resumes with extracurricular activities and for males and females. White sounding names consistently receive higher callbacks than Black sounding names, consistent with similar studies. Overall, resumes with White sounding names are almost two times more likely to receive a callback compared to resumes with Black sounding names. In other words, for every ten resumes a White applicant sends, they can expect to receive roughly one callback. However, Black applicants must send a total of 15 resumes to expect a similar callback rate.

The second panel explores the racial gap in callback rates between male and female applicants. Black male applicants perceive an average callback rate of 4.30. This finding is consistent with the characterization of disconnected youth being primarily black males. Of all the groups in this experimental design, black males face the lowest callback rate. This finding is also consistent with similar studies. Bertrand and Mullainathan (2004) find higher callback rates for female applicants in administrative, clerical and sales jobs. Similarly, Booth and Leigh (2010), Zhou et al. (2013), Berson (2012), and Jackson (2009) all find evidence of discrimination against males in their correspondence studies. A plausible explanation for these findings is that audit studies rely on a subset of sectors for job applications that tend to advertise employment online. The overwhelming majority of these job posting correspond to positions in the service industry. In these low-skilled, entry-level jobs, female applicants may be favored compared to male applicants. Most of the sales opening correspond to the fashion industry, where gender stereotypes may favor female applicants. Administrative openings tend to be for receptionist or office assistance where, again, gender stereotypes may favor female applicants. Food industry

openings include host and server positions, which can also favor women due to gender stereotypes.

The third panel in Table 14 estimates the average callback rates for sports participation (both track and basketball) for Black and White applicants. Having participated in sports increases the callback rates for Black applicants from 6 to 8.57, for a reduction in discrimination of about 40%. Exploring the average callback rates for those who participate in basketball, fourth panel in Table 14, shows that racial differences for applicants who participate in basketball are almost eliminated. The average callback of Black applicants who participate in basketball is 11.76, almost twice the average callback of overall Black applicants (6).

The last panel in Table 14 explores the racial differences in callback rates for males who participate in basketball. Participating in basketball seems to reduce discrimination for Blacks males by about 50%.

Next, I estimate a set of linear probability models to estimate the effect of race, gender, and extracurricular activity on the likelihood of response. The matched-triplets design allows for the estimation of within ad variation in responses. For all models I include ad fixed-effects, and robust standard-errors. I start by looking at racial differences in callback rates, I then look at how sports participation may affect these racial differences. Lastly, I focus on basketball which seems to drive the sports participation effect and how participating in basketball affects racial differences in callbacks for males. Focusing on Black males allows me to speak to the population of interest for

this dissertation, disconnected youth, and how high school sports participation may affect their employment prospects.

Table 15. Effects of gender and race

Effect of gender and race on likelihood of response

	(1)	(2)
Black	-0.056***	-0.045***
	(0.008)	(0.012)
Male	-0.028***	-0.018
	(0.008)	(0.012)
Black*Male		-0.021*
		(0.016)
Constant	0.127	0.122
	(0.006)	(0.008)
Observations	6000	6000

*significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Notes: Results from FE estimates. Standard errors in parenthesis. Robust SE

Table 15 starts with a simple estimation of the effect of race and gender on the likelihood of response. Column 1 indicates having a Black sounding name reduces the likelihood of response by 5.6 percentage points. This finding is statistically significant at less than 1 percent confidence level and it implies that resumes with white sounding names are almost twice as likely to receive a callback compared to resumes with Black sounding names. This estimate is consistent with racial discrimination and similar in magnitude to similar studies (Bertrand and Mullainathan, 2004; Nunley et al., 2015; Gaddis, 2015; Agan and Starr, 2016).

Being a male applicant reduces the likelihood by 2.2 percentage points, this effect is significant at a 5 percent confidence level. Although this effect may seem counterintuitive given the gender gap in employment, it is consistent with similar studies. As previously explained, this" reverse gender discrimination" against males could be explained by the types of jobs available to

correspondence audits. Another plausible explanation comes from the behavior of the gender gap in employment over the life cycle. Analysis of the life cycle of the gender gap for both employment and wage, suggests an inverse u-shape overtime (Goldin and Mitchell 2017; Joshi et al. 2019). These authors suggest that the gap starts to widen during child-bearing years. Although this study looks at young applicants, recent high school applicants, Goldin and Mitchell (2017) find evidence that the life cycle of female employment has been shifting with every subsequent generation from the 1950. Since the 1950, females have increased their labor force participation at younger ages and the dip in participation, potentially due to childbearing, is experienced later in life. This could help explain "reverse" gender discrimination even for older applicants, as is the case of Bertrand and Mullainathan (2004).

Column 2 examines the interaction between race and gender. Estimates show that males with Black sounding names are particularly at a disadvantage to receive a callback. The marginal effect of both a male and Black sounding name is -8.4 percentage points. The omitted group in this estimation is an applicant with a female, white sounding name which has an average likelihood of response of 12.2% – three times more likely to receive a callback compared to Black males, who have an average likelihood of response of 3.8%. The coefficient on the interaction between race and gender is significant at the ten percent level. These findings suggest that, in this study, Black females are discriminated against at a lower rate than Black males.

Table 16 summarizes the results of the effect of sports and extracurricular activities on likelihood of response. If we believe extracurricular activities are a signal for skills valued in the labor

market, then applicants who signal participation in extracurricular activities should be favored compared to applicants with no extracurricular activities. Kniffin et al. (2015) explore the relationship between leadership skills and participation in different extracurricular activity. They find that sports participation is associated with leadership skills and higher levels of employment and wages. However, they fail to find this association for participation in marching band and debate team, despite sociologists' argument that they develop similar skills (Feldman and Matjasko, 2005). Kniffin et al. (2015) suggests that the effect of extracurricular activity on the likelihood of response will vary by extracurricular activity.

Table 16. Effects of sports and extracurricular activities

Sports and other extracurricular activities

	(1)	(2)
Sports	0.024**	
	(0.009)	
Other Extracurricular	0.010	
	(0.009)	
Basketball		0.048**
		(0.012)
Track		0.001
		(0.010)
Band		0.022*
		(0.12)
Debate		-0.013
		(0.009)
Constant	0.075	0.075
	(0.006)	(0.006)
Observations	6000	6000

*significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Notes: Results from a simple LPM. Standard errors in parenthesis. SE clustered at the ad.

Column 1 shows that participating in high school sports increases the likelihood of response by 2.4 percentage points compared to applicants who do not have any activities in their resumes. This effect is significant at the five percent level. Applicants with no extracurricular activities have,

on average, a likelihood of callback of 7.5%. Whereas applicants with sports have, on average, a likelihood of response of 9.5%. Participating in non-athletic extracurricular activities (marching band and debate) does not seem to affect callback rates. Although the point estimate is positive, and half the size of the coefficient on sports participation, it is not statistically significant. These finding are consistent with Kniffin et al. (2015).

In column 2, we explore potential heterogeneity in team versus individual sports and different non-athletic extracurricular activities. Coleman (1961), Blau (1968), and Feldman and Matjasko (2005), argue that team sports develop different sets of skills compared to individual sports. Specifically, team sports setting tends to be associated with more prosocial behavior. Researchers argue that this behavior is valued in the labor market as it translates into employees that tend to be more loyal to the company and work well with others (Organ et al., 2006).

Participation in basketball drives the effect of sports participation. Participating in basketball increases the likelihood of response by 4.8 percentage points compared to resumes with no other extracurricular activities. In other words, applicants who participate in basketballs are roughly twice more likely to receive a callback compared to applicants who do not report any extracurricular activity. This effect is statistically significant at the 5 percent confidence level[9]. Participating in Track and Field does not seem to affect the likelihood of response.

[9] One of the high schools chosen (Lincoln High School) in New York City, is known for their strong basketball program. In the Appendix, I show that even though the effect of basketball on response is larger in New York City, the effect is also present in Chicago and Dallas.

Table 17 looks at how participating in basketball, the driver of the sports effect, affects Black male applicants. To explore this difference, I look solely at the subsample of males in the experiment. For comparison reasons, I present the same results for the female sample in Table 18. If we believe sports participation is a signal for skills valued in the labor market, such as leadership, and prosocial behavior, then we would Black applicants with sports participation would perceive higher callback rates than Black applicants with no sports participation if we think discrimination is statistical in nature. On the other hand, if Black applicants with sports participation perceive similar levels of discrimination compared to Black applicants with no sports participation discrimination, this could be evidence of taste-based discrimination. However, it can also mean that signals of quality are not perceived differently between Black and Whites (Uhlmann and Cohen, 2005), in which case, statistical discrimination cannot be ruled out completely.

Table 17. Effects of race and basketball for males

Race and basketball for male sample

	(1)	(2)
Black	-0.061***	-0.068***
	(0.009)	(0.010)
Basketball	0.029**	0.009
	(0.014)	(0.020)
Black*Basketball		0.042*
		(0.025)
Constant	0.10	0.10
	(0.008)	(0.008)
Observations	3074	3074

*significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Notes: Results from a simple LPM. Standard errors in parenthesis. SE clustered at the ad.

Column 2 of Table 17 estimates that Black males who did not participate in sports receive an average callback rate of 3.2 percent compared to white males who did not participate in sports who received an average callback rate of 10 percent However, if the Black male played basketball, their callback rate was 7.4 percent. Playing basketball seems to reduce the discrimination effect against Black makes by over 60 percent. This finding is consistent with statistical discrimination. That sports participation seems to benefit Blacks more than it benefits white applicants is consistent with employers using additional information to better assess a potentially 'noisy' estimate of expected productivity. This reduction in discrimination, by roughly 60%, is consistent with educational interventions aimed at closing the achievement gap (between Blacks and Whites). Interventions such as smaller classes, and social and psychological supports reduce the achievement gap by a range of 40-60% (Krueger and Whitmore, 2001; Cohen-Zada et al., 2009).

Table 18 looks at the effects of race and basketball participation for the female sample, but finds no evidence that basketball is particularly beneficial for black females. This suggests that participating in basketball reduces discrimination for Black males, the most disadvantaged group in the sample. In other words, participating in basketball may positively affect the employment prospects of disconnected youth.

Table 18. Effects of race and basketball for females

Race and basketball for female sample

	(1)	(2)
Black	-0.041**	-0.033**
	(0.011)	(0.011)
Basketball	0.026	0.060
	(0.017)	(0.040)
Black*Basketball		-0.061
		(0.040)
Constant	0.11	0.11
	(0.008)	(0.008)
Observations	2926	2926

*significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Notes: Results from a simple LPM. Standard errors in parenthesis. SE clustered at the ad.

Exploring heterogeneity by job type and cities

I further analyze the data to explore whether main results are being driven by certain characteristics, or whether there is evidence of heterogeneity in the effects by both job types and geographical location.

Table 19 explores the likelihood of callback rates for African American resumes by job type. Differing levels of discrimination by job type is sometimes considered evidence of distinct types of discrimination in specific settings. In other words, when different job types have differing levels of interaction with customers, and job type is associated with different levels of discrimination it can be considered evidence of customer-based discrimination. In this design, sales and administrative openings have a higher degree of customer interaction compared to food jobs.

Table 19 shows slightly lower levels of discrimination based on race for food openings. However, testing for coefficient equality in Table 20 shoes that these estimates are not statistically significantly different from one another. In other words, I do not find evidence that the estimated levels of discrimination vary by job type.

Table 19. Effects of race by job

Effect of race on likelihood of response by job type

	Admin	Food	Sales
	(1)	(2)	(3)
Black	-0.052***	-0.044***	-0.053***
	(0.011)	(0.013)	(0.012)
Constant	0.098	0.119	0.112
	(0.009)	(0.010)	(0.010)
Observations	2,025	2,025	1,950

*significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Notes: Results from a simple LPM. Standard errors in parenthesis. SE clustered at the ad.

Table 20. Test of equality by job

Testing coefficient equality

I. Admin = Sales
chi2(1) = 0.01
Prob>chi2 = 0.9411
II. Admin = Food
chi2(1) = 0.20
Prob>chi2 = 0.6561
III. Sales = Food
chi2(1) = 0.25
Prob>chi2 = 0.6189

Notes: Each section shows a pair-wise test of the hypothesis that the estimate of discrimination is the same for the different job types.

Tables 21 and 22 explore heterogeneity in effects based on location. Table 21 explores whether racial discrimination varies based on the city. Column (1) reflects a lower measure of discrimination for Black applicants in New York City. However, testing for coefficient equality in Table 22 shows that I cannot rule out that estimates of discrimination are the same for all three cities.

Table 21. Effects of race by geography

Effect of race on likelihood of response by city

	New York City	Chicago	Dallas
	(1)	(2)	(3)
Black	-0.035***	-0.058***	-0.054***
	(0.013)	(0.012)	(0.012)
Constant	0.103	0.116	0.11
	(0.009)	(0.010)	(0.009)
Observations	1,845	2,115	2,040

*significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

Notes: Results from a simple LPM. Standard errors in parenthesis. SE clustered at the ad.

Table 22. Testing equality by geography

Testing coefficient equality

I. NYC = Chicago
chi2(1) = 1.66
Prob>chi2 = 0.1980
II. NYC = Dallas
chi2(1) = 1.08
Prob>chi2 = 0.2986
III. Chicago = Dallas
chi2(1) = 0.07
Prob>chi2 = 0.7867

Notes: Each section shows a pair-wise test of the hyphothesis that the estimate of discrimination is the same for the different cities.

The results from this study suggest that participation in high school sports, particularly in basketball, has a positive effect on likelihood of response for young low-skill applicants applying to entry-level positions in Chicago, Dallas, and New York City. The effect of sports participation seems to be particularly beneficial to applicants with Black sounding name, reducing some of the discrimination these applicants experience in the labor market. This result is consistent with statistical discrimination models, suggesting that at least some of the discrimination experienced by young Blacks in the labor market is statistical in nature. In the next section, I discuss some to the potential mechanisms that might be at play and the limitations of the study.

Section 7. Discussion and Limitations

Potential Mechanisms

An important disadvantage of correspondence audits is their limited ability to explore the mechanisms that may lead to discrimination. One can think of at least two mechanisms by which sports participation reduces racial discrimination in the labor market. First, sports participation can be thought as a quality signal. If this is true, then the reduction in differential treatment implies that at least part of the observed racial discrimination is statistical in nature. However, prior studies have found little evidence to support this mechanism. Bertrand and Mullianathan (2004) found that higher quality résumés did not lead to a reduction in discrimination for applicants with Black sounding names. This may suggest that sports participation may not necessarily be a quality signal, and therefore the reduction in differential treatment must be operating through a different mechanism.

Several studies in the field of behavioral sciences have found that employers are more likely to hire candidates who appear like them (Rivera, 2012). More specifically, employers are more likely to hire applicants with similar leisure pursuits and experiences. In a country where sports are a multimillion-dollar industry and where more than 50 percent of high school students participate in sports, it is plausible that sports may qualify as a shared experience for many Americans (NFHS, 2017). This mechanism may also explain why sports are particularly beneficial to Blacks. If most employers are white, then white applicants are more likely to appear like employers. Sports participation may not add much to the employer's perception of similarity between them and a white applicant. However, for a Black applicant, who may be perceived as dissimilar to their

potential employer, a shared experience (such as sports) may have a large effect on the employer's perception of similarity.

Limitations

As with most correspondence audits, there are very specific ways in which the study design limits the generalizability of the findings presented in this paper. Broadly speaking, study design choices limit external validity through geography, hiring process, population of interest, type of jobs, race signals, and extracurricular activities.

Geography

This study took place in three large US cities (Chicago, Dallas, and New York); effects of sports and gender may vary depending on the specific cities used. Results suggest that discrimination seems to be lower in NYC compared to Chicago and Dallas. This heterogeneity in the effects by geography could be evidence that the study results are not generalizable to all large cities in the United States. Moreover, it is unclear how the results compare to measures of discrimination in smaller cities. This is an important limitation of correspondence audits, their inability to generate sufficient sample to run experiments in smaller localities (Carbonaro and Schwarz, 2018). Carbonaro and Schwarz (2018) argue that results from correspondence audits in larger cities may differ significantly from those conducted in smaller cities. They argue that resume signals related to addresses and high school quality may be more salient in smaller localities in comparison to larger cities. In their own correspondence audit, they find that racial composition of high schools has no effect on response rate in larger cities. However, they are unable to compare these results

with results from a correspondence audit in a smaller locality because the hiring process in smaller localities is incompatible with correspondence audits. Smaller localities rely less on online job posting for entry-level jobs, and more on walk-ins. This leads to an inability to identify large enough number of job openings in a small locality to conduct a correspondence audit. This difference in hiring process is, in and of itself, a threat to the external validity of correspondence audits.

Hiring process

Similarly, correspondence audits rely on employers that recruit online. This characteristic of the study removes from the potential sample of employers, those that do not recruit online and those that recruit online through their own internal platforms. The latter are also removed from correspondence audit because in most cases they tend to require some form of personal identification such as Social Security Number (SSN). The degree to which these types of employers differ in their screening processes will determine how potentially biased results from correspondence studies are. That is, if employers who recruit online are more likely to engage in racial screening through automation, then results from correspondence audit will tend to overestimate measures of discrimination in the labor market. However, it is unclear whether employers who rely on online recruiting react differently to resume signals than employers who do not, and researchers have argued that further work is needed to better understand this limitation of the study design (Gaddis, 2018).

Population of Interest

As previously mentioned, this study focuses on recent high school graduates looking for low-skilled, entry-level job opportunities. Although, I have argued why this is a population of interest from a policy perspective, it is still the case that the results from this study may not be generalizable to a wider population. Related to the previous limitation, the hiring process for low-skilled, entry-level positions may differ in important ways from higher-skilled job openings. In the former, the expectation is that many applicants apply to a single position, a situation with a high labor supply. While as skill level increases one would expect labor supply to decrease. The importance of certain signals in a scenario with many competitors (high labor supply) may differ significantly from a scenario with few competitors (low labor supply) (Phillips, 2016). If we take this statement as true, correspondence audits that focus on low-skilled, entry-level openings may provide little information about measures of discrimination for high-skilled openings.

Race signal

Race in correspondence studies is signaled mainly through name, though some studies have used club membership instead (Bendick et al., 1994). Since Bertrand and Mullainathan (2004) names have become the main cue for signaling race in correspondence audits. Fryer and Levitt (2004) argue that distinctively Black sounding names are associated to mothers with lower SES. Although I have previously explained the ways in which this critique has been unpacked and addressed in the literature, it remains that not all African Americans have distinctively Black sounding names, one could argue that African Americans without distinctively Black sounding names probably do not experience the levels of discrimination measured in correspondence studies when applying

to job postings online. However, if the argument regarding the strength of the signal is valid – it must mean that the measured levels of discrimination eventually catch up with Black applicants. They could avoid being racially discriminated against initially if they do not have a distinctively Black sounding name but may eventually experience similar levels of discrimination once their race becomes evident.

Extracurricular activities

This study only focuses on two popular sports, basketball and track and field. It is plausible that other sports may have different effects on callback rates. I have explained that an important reason for choosing these sports is that they did not seem to send strong signals of SES. If this is the case, it may mean that sports that are associated with higher SES (golf, tennis) may have a different effect on likelihood of response than the sports chosen in this design. However, given the population of interest the choice of sports seems coherent. The average high school students may not have the means to access sports that signal higher SES, since they tend to be highly concentrated in predominantly White high schools (NWLC and PRRAC, 2015).

Call back rate as outcome

Another important limitation is the narrow scope of the outcome variable, call backs. This limits the measure of discrimination to only the initial stages of job search. Some researchers have tried to complement their studies by looking at later stages of the process or combine their results with readily available data (Pager and Quillian, 2005; Correll et al., 2007; Pedulla, 2016). However, the limitation remains that the measured variable in audit and correspondence studies is usually

restricted to the initial stage of the hiring process. Although it is important for researchers to continue to think of creative ways to incorporate later stages of the hiring process in the design of audit and correspondence studies, researchers have argued and empirically demonstrated that the discrimination in early hiring stages translate into worse off outcomes in later stages. Bendick et al. (1999) and studies conducted by the International Labor Organization (ILO) estimate that around 90% of labor market discrimination occurs in the initial stages of interviews and job offer (Riach and Rich, 2002).

Finally, this study focuses on young adults who recently finished high school, attended a high school with racial diversity and live in neighborhoods that correspond to the catchment area of these high schools. All this to say that the results from this study may not be generalizable to a broader population. However, they seem to suggest that sports participation may reduce labor market discrimination at least for a subset of the population.

Conclusions and policy implications

High school sports participation seems to matter for the employment prospect of young, low-skilled workers looking for entry-level positions. Black males, who are more likely to become disconnected youth, seem to benefit the most from participating in basketball when applying to these positions. High school basketball participation seems to reduce the experience of discrimination in the labor market for Black applicants, consistent with statistical discrimination.

The estimated effect is consistent with educational interventions aimed at reducing the achievement gap between Black and White students.

Policy implications of these findings must consider the limited scope of the research design and the complex dynamics of school resources. The findings from this dissertation focus on a particular set of high schools and a particular set of sports. More research is needed to understand how these effects may differ when looking at highly segregated schools and a wider set of sports. Further understanding under what circumstances high school sports participation affects labor market outcomes, will shed light into possible policy recommendations. Expanding athletic programs seems to be one of the many policy decisions that can affect later life outcomes for Blacks, in particular Black males. As with other interventions, we need more research to understand the cost-effectiveness of these interventions and how they compare to one another in terms their potential impacts and costs.

Schools have budgetary constraints which means that expanding certain programs comes at the expense of other programs and/or resources. Future research should also focus on understanding how the effect of sports participation on labor market outcomes compares to the effect other school programs may have on labor market outcomes. However, effects on labor market outcomes may not be the only outcomes of interest when understanding the value of certain school programs. Outcomes such as school performance, graduation rates, health and wellbeing are also important. More research is needed to understand how athletic participation

may affect these outcomes and how it compares to other interventions with the similar objectives.

To make specific policy recommendations that may advocate for more access to sports programs in high schools, further research is needed to understand how school demographics and particular sports affect labor market outcomes and discrimination against Black applicants. Furthermore, research is needed to understand how sports participation affects other important outcomes for high school students such as school performance, graduation rates, health, and wellbeing, and how these effects compare to other school programs. However, this study shows that expanding athletic programs could reduce Black males' likelihood of becoming disconnected youth, and therefore should be considered by policy makers as a policy that can reduce racial inequality for young Black males.

Concurrently, predominantly minority schools have less athletic opportunities than predominantly white schools (NWLC and PRRAC, 2015). Findings from this study suggest that it is worthwhile exploring the ways in which more disadvantaged school district could create and/or expand athletic opportunities for students who will likely benefit from sports participation as demonstrated by this study. Title IV of the Civil Rights Act of 1964 states that states and school district have an obligation "to provide students with equal access to [educational] resources without regard to race, color, or national origin." Extracurricular programs such as sports are included in this definition of educational resources.

APPENDIX

Setting up the idea – A choice Experiment with Conjoint Design

Prior to setting up an experiment with real employers, I conducted a choice experiment. I used the now widely used online tool Amazon Mechanical Turk (MTurk) that allows researchers to efficiently identify and recruit participants for 'lab-on-the-field' surveys and experiments. MTurk is an online tool that gives employers and employees access to both a workforce and labor. Businesses and developers can post tasks, which are usually computational in nature, and employees can accept. It is something like a virtual labor market. One of the tasks that are frequently used is responding to surveys. Using this task granted me access to many respondents for a relatively low cost. Respondents were paid $0.25 for participation, well within the acceptable payment for short research surveys (Mason and Suri 2012).

This setup allowed me to pretest the idea that sports participation matters for hiring decisions, the initial question I set out to answer. Using a vignette and conjoint design, I wanted to test if sports participation affected respondents' probability of choosing a constructed profile for an entry-level job. The idea of the conjoint experiment was to explore potential interactions between sports participation and other resume characteristics. The first step for this experiment was creating comparable profiles that seemed realistic. Since it is very unlikely that college graduates will indicate in their resumes whether they participated in high school sports, the study was confined to high school graduates. This was the first trade-off I faced – focusing on high-school graduates meant limiting the study's external validity. This decision made sense to me for several reasons. First, work experience during and after high school matters for later-life

outcomes (Carr, Wright and Brody 1996). Using the National Longitudinal Survey of Youth, the authors find that work experience in high school positively affects labor force participation, employment status, and income even a decade after graduation. Moreover, this is a population of interest from a policy perspective since there is a growing number of people ages 16 to 24 that are both out of school and out of work, 6.5 million in 2011. Studies have shown that this vulnerable population, often referred to as disconnected youth, is burdened with future financial instability and grim employment prospects (The Annie E. Casey Foundation 2012). Furthermore, young Blacks, ages 18 to 19, are twice more likely to be unemployed compared to whites (Bureau of Labor Statistics 2015). This implies that this is a particularly relevant population that stands to benefit from increased employment prospects and reductions in employment discrimination.

Second, if sports happen to be protective for Blacks, the policy implication of the study was quite straightforward given the gap in athletic opportunities for high school students in predominantly minority schools (NWLC and PRRAC 2015). That is, if sports participation signaled higher quality applicants, or developed skills valued in the labor market, theory would suggest that it would have a protective effect against discrimination. African Americans participating in sports would, in theory, perceive lower levels of discrimination than otherwise similar Black applicants who did not participate in sports. Third, choosing college graduates presented two possibilities: listing college sports participation, which represents a much smaller portion of students compared to high school athletes or stating high school sports participation as a college student/graduate[10].

[10] In the 2018-2019 academic year almost 8 million high school students participated in sports, compared to the roughly 500,000 college-athletes (NFHS 2021, NCAA 2021)

This last option did not seem in line with resume samples from the real world, where such activities were no longer listed, which further harmed the external validity of the study. For these reasons, I chose to limit the scope of the study to high school sports participation and low-skill, entry-level job opportunities.

To ensure the realistic component of the profiles, I used online databases that collect real applicants' resumes (i.e., CareerBuilder and Americas Job Bank) and modify them to ensure that applicants are not overqualified for low-skill, market entry jobs. Figure 1A shows a screenshot of a possible combination.

Figure 1A. Applicant's profiles

Name	Parker	Alex
Professional Experience	Cashier at a local ice cream parlor 2014-2015	Cashier at a local bakery 2014-2015
Qualifications, Skills and Qualities	• Good leadership skills and ability to work well in a group setting. • Professional work ethic. • Great communications and customer service skills. • Excellent performance review - past employment. • Fluent in French	• Great communications and customer service skills. • Good leadership skills and ability to work well in a group setting. • Ability to perform several tasks simultaneously. • Great work ethic. • Fluent in German
Education	High School Diploma, 2014 GPA - 3.45	High School Diploma, 2014 GPA - 3.5
Other Skills/Activities	Proficient in MS Office and MS Excel Athlete for the high school Basketball Team 2012-2014	Proficient in MS Office and MS Excel .
Gender	Female	Male

An important caveat of conjoint experiments is that the number of manipulations must remain manageable to ensure statistical power (Hainmueller, Hopkins and Yamamoto, 2013). This restriction led me to focus only on a couple manipulations: gender and extracurricular activities. Race was completely omitted from this pre-test. The reason for this omission was twofold: the issues that arise from directly listing race and not having conducted a prior perception analysis of names signaling race. Listing race directly in the profile generated a couple of issues. First, race is not typically listed directly in a profile which may lead respondents to believe that the study had a race component. If this is the case, one would be worried about respondents: a) answering in accordance with their beliefs about the study objectives or b) engaging in social desirability bias when choosing between profiles. Both these issues would generate bias in the estimates. As I will explain later, I believe some of these issues partly explain the gender results of this study. Second, as previously discussed, signaling race with names is a nuanced exercise that requires careful selection and testing. This exercise came after the MTurk pre-test, and I explain it in more detail further below.

The second characteristic, gender, is randomly assigned to profiles independently from extracurricular activity. When we put all the possible combinations together, we end up with a 6-conditions experiment where comparisons happen across all the possible combinations and randomization occurs simultaneously in both profiles. Other profile characteristics included: name, experience, skills and qualifications, and education.

I use a type of stated preference experiment that uses conjoint design, an experimental technique for handling situations in which a decision maker must deal with options that simultaneously vary across two or more attributes. Conjoint design enables researchers to estimate the causal effect of multiple treatment components and assess several hypotheses simultaneously. Respondents choose among a set of alternatives, where each alternative has randomly varied attributes (Hainmueller et al., 2013). This allows for the estimation of the marginal effect of each component on the probability of being chosen. This allows me to randomize more than one component at the same time, and therefore test multiple hypotheses using one experiment (Hainmueller et al., 2013). This design imposes randomization on both profiles, allowing for the comparison of the full set of possibilities, which is where the strength of the design resides. That is, results are estimated from pooling across all possible comparisons. With conjoint design, the estimate of interest is the Average Marginal Component Effect (AMCE). AMCE is the marginal effect of an attribute averaged over the joint distribution of the remaining attributes. It estimates the increase in the population probability that a profile would be chosen if the value of its 1^{st} attribute (A) were changed from A_0 to A_1, averaged over all the possible values of the other attributes.

I recruited participants based in the US and with some managerial experience or experience hiring. Once they qualified and accepted the task, I put them in the position of a manager that must hire for a vacancy at their company, I present them with the job description followed by the profiles, and I make them choose who they would hire. I then have them repeat the exercise

a second time and the survey conclude with a set of socio-demographic questions and manipulation checks.

Drawing from the literature and given the experimental design my expectations for the results are summarized by the following hypotheses:

H_1: Basketball players will have a higher probability of being chosen compared to profiles with no extracurricular activity.

H_2: Band members will have a higher probability of being chosen compared to profiles with no extracurricular activity.

H_3: Basketball and band will differentially affect the probability of being chosen.

H_4: The effect of extracurricular activity will vary by gender.

Due to the experimental manipulation, the resulting empirical strategy is rather simple and straightforward. My main estimating equations are the following:

$$(1)\ Chosen_i = \beta_0 + \beta_1 Basketball_i + \beta_2 Band_i + \varepsilon_i$$

$$(2)\ Chosen_i = \beta_0 + \beta_1 Basketball_i + \beta_2 Band_i + \beta_3 Male_i + \beta_4 (Basketball_i * Male_i) + \varepsilon_i$$

Where β_1 is the difference in probability of a basketball player being chosen relative to a profile with no extracurricular activity. β_2 is the difference in probability of a band member being chosen relative to a profile with no extracurricular activity. β_3 is the difference in probability of a male being chosen relative to a female profile. And β_4 is the coefficient that captures the interaction,

and therefore difference in probability of a male basketball player being chosen relative to a female with no extracurricular activity.

Table 1A. Descriptive statistics MTurk

Descriptive Statistics	
Statistic	**Mean**
Age	35
Gender	
Male	53.85%
Female	46.15%
Age	35.00%
Race/Ethnicity	
White	71.04%
African American	12.67%
Asian	7.24%
Latino	4.98%
Native American	3.17%
Sport Participation	
Played sports in high school	45%
Education	
Some high school	0.90%
High school diploma	5.43%
Some college	38.91%
College degree	45.25%
Graduate Degree	9.50%
Employment	
Currently employed	83.26%
Managers	
Experience as managers	69.68%
N	221

Table 1A presents descriptive statistics for the survey participants. The sample of participants is slightly more male than female (~54%), predominantly White (71%), and educated (84% high school diploma or more). Tables 4 & 5 present the experimental results. The point estimate for

AMCE on sports is in the expected direction, positive, I cannot rule out that there is no effect of sports participation on being chosen.

Table 2A. Estimation of the Average Marginal Component Effects

	Chosen	**Chosen**	**Chosen**
	(1)	**(2)**	**(3)**
Basketball		**0.0229**	**0.0228**
		(0.04)	(0.04)
Band		**0.0650**	**0.0661**
		(0.04)	(0.04)
Constant	**0.5*****	**0.471*****	**0.468*****
	(0.02)	(0.03)	(0.03)
Respondent's Characteristics	No	No	Yes
R^2	0	0.0027	0.0028
Observations	884	884	884

Table 2A looks at the effect of basketball participation and participation in marching band. Column (2) looks at the effect without controlling for respondents' characteristics. High school basketball participation increases the probability of being chosen by 2.29 percentage points, and participation in marching band increases the probability of being chosen by 6.5 percentage points. Results of Column (3) which include controls for respondents' characteristics are similar to the results without controls. None of the estimates are statistically significant.

Table 3A looks at the interaction between gender and extracurricular activity. The interactions between gender and extracurricular activities are both positive, however, are not statistically significant.

Table 3A. Estimation of the Average Marginal Component Effects including gender

	Chosen	Chosen	Chosen
	(1)	(2)	(3)
Basketball	**0.0150**	**0.0149**	**0.0024**
	(0.04)	(0.04)	(0.06)
Band	**0.0616**	**0.0623**	**0.0614**
	(0.04)	(0.04)	(0.05)
Male	**-0.133*** **	**-0.134*** **	**-0.1654*** **
	(0.03)	(0.03)	(0.06)
Basketball*Male			**0.0235**
			(0.08)
Band*Male			**0.0739**
			(0.08)
Constant	**0.5*** **	**0.52*** **	**0.555*** **
	(0.02)	(0.04)	(0.03)
Respondent's Characteristics	No	Yes	Yes
R^2	0.020	0.020	0.0214
Observations	884	884	884

I do not find a statistically significant effect of participating in basketball while in high school on the probability of being chosen. Similarly, there's no statistically significant effect of participating in marching band on the probability of being chosen. However, gender seems to have a statistically significant effect on the probability of being chosen. Females are disproportionally chosen compared to males. Males seem to be about 50% less likely of being chosen for the job. This discrimination against males appears to contradict our understanding regarding gender-based discrimination in the labor market (Neumark, 1995), which raises some concerns that I will address under the limitations of the findings.

Limitations MTurk Study

Like in any study, there are important limitations to the findings, these are mainly some variations of threats to the external validity of the findings.

The first of these threats relates to the sample of respondents. Amazon Mechanical Turk respondents may not be representative of the population of employers. Even when the survey asked for respondents to have had some experience as an employer, employers in Amazon Mechanical Turk may not be representative of the population of employers. This in turn may be one explanation for the type of gender-based discrimination described in the findings. In this sense, our findings may not be generalizable to all US employers, which significantly limits my ability to make causal statements regarding the effect of high school sport participation on employability.

The second threat is related to the type of sport participation. In this experimental setup, I use a very narrow representation of high school sport participation, athlete in the high school basketball team. If we think there is heterogeneity in sport participation and that certain sports may signal different skills or background, then my study is unable to speak to this. In this sense, I can only estimate the effect of participating in basketball while in high school, which also limits the generalizability of the findings.

Third, this study uses a very specific job description (customer service representative). Again, if we think sport participation may matter more in particular positions and/or industries the results

of this study are unable to speak to this. The findings may only be relevant for customer service representatives. Other issues that arise from this threat relates to gender roles. It is possible that the reason we see such a gender bias in chosen applicants could be due to the notion that customer service positions are thought to be female dominated positions (Blau, 1984).

Another threat to the generalizability of the findings comes from the types of profiles represented in the experiment. These are newly high school graduates with very limited labor market experience and very particular sets of skills, which also limits to whom the findings are generalizable to.

Finally, as with most experimental design, this is a hypothetical scenario where respondents have very little or close to nothing at stake. If employers in the real world, make decisions in a way that significantly differs from how respondents in a hypothetical scenario make such decisions this becomes an important threat to the external validity of the findings.

Although the MTurk study did not find statistically significant results to support my hypotheses, it provided useful information for the final experimental design. Specifically, it forced me to carefully consider narrowing the scope of the study, as it became clear that I would not be able to answer all interesting questions with one experiment. First, narrowing the population of interest to high school graduates searching for low-skill, entry-level jobs – as previously explained. Second, incorporating an extracurricular activity manipulation to distinguish between

effect of ‘any extracurricular activity’ and sports. Third, choosing the appropriate sports and extracurricular activities.

Figure 2A. Model resumes

Jamal Washington

2016 N Prairie Ave Dallas, TX 75204 ; phone: 469-532-0730 email: jmllwshngtn@gmail.com

Education

2012-2016 High School Diploma, Herbert H. Lehman High School - New York, NY

Work Experience

2014-2017

Kitchen Team
Shake Shack - New York, NY
Part-time
Responsibilities:

- Communicate with other crew members and managers effectively
- Prepare and pack food according to policies and procedures
- Make sure there are enough supplies available and order food supplies before they run out
- Manage transfers of supplies from the store to the kitchen, clean and sanitize customer sitting areas, kitchen and counters
- Prepare and pack food according to policies and procedures

Skills

- Very motivated person
- Organized and responsible
- Hard-working
- Superb communications skills

Other Skills

- Food Handler Certificate
- Athlete in the High School Track Team

Monique Lewis

3301 N Menard Ave. Chicago, IL 60634 ; phone: 312-757-6141 email: lewis.monique07@gmail.com

Education

2012-2016 High School Diploma, Chicago Academy High School - Chicago, IL

Work Experience

2014-2017

Sales Representative
Bloomindale's - Chicago, IL
Part-time
Responsibilities:

- Identify and assess customers' needs to achieve satisfaction, meet personal and team sales targets
- Handle complaints and provide appropriate solutions and alternatives
- Report and provide feedback to management
- Follow communication procedures, guidelines and policies when addressing clients

Skills

- Highly motivated individual
- Excellent interpersonal skills
- Responsible and well-organized
- Eager to learn and work hard

Other Skills/Activities

- Advanced abilities in most Microsoft Office programs (Word, Excel and PowerPoint) and POS System
- Member of the High School Marching Band

Latoya Wilson

90 La Salle St New York, NY 10027 | phone: 646-491-6377 email: latoyawilson03@gmail.com

Education

2012-2016 High School Diploma, Hillcrest High School - Dallas, TX

Previous Work

2014-2017

Sales Associate
Dillards - Dallas, TX
Part-time
Responsibilities included:

- In charge of stocking and organizing merchandise in my assigned section
- Attend to customers' needs
- Manage any complaints from customers or communicate to the appropriate personnel
- Contribute to team efforts by accomplishing related tasks as needed

Skills

- Willing to work hard
- Very responsible and organized
- Great people skills
- Extremely motivated

Others

- Advanced abilities in most Microsoft Office programs (Word, Excel and PowerPoint) and POS System
- Member of the High School Marching Band

Bibliography

Adida, C. L., Laitin, D. D., & Valfort, M. A. (2010). Identifying barriers to Muslim integration in France. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 22384–22390.

Agan, A., & Starr, S. (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics*, *133*(1), 191-235.

Agerström, J., & Rooth, D. O. (2011). The role of automatic obesity stereotypes in real hiring discrimination. *Journal of Applied Psychology*, *96*(4), 790.

Agerström, J., Björklund, F., Carlsson, R., & Rooth, D. O. (2012). Warm and competent Hassan = Cold and incompetent eric: A harsh equation of real-life hiring discrimination. *Basic and Applied Social Psychology*, *34*, 359–366.

Ahmed, A. M., Andersson, L., & Hammarstedt, M. (2012). Does age matter for employability? A field experiment on ageism in the Swedish labour market. *Applied Economics Letters*, *19*, 403–406.

Ahmed, A. M., Andersson, L., & Hammarstedt, M. (2013). Are gay men and lesbians discriminated against in the hiring process?. *Southern Economic Journal*, *79*(3), 565-585.

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Ilr Review*, *30*(2), 175-187.

Albert, R., Escot, L., & Fernández-Cornejo, J. A. (2011). A field experiment to study sex and age discrimination in the Madrid labour market. *The International Journal of Human Resource Management*, *22*(02), 351-375.

Altonji, J. G., & Blank, R. M. (1999). Race and gender in the labor market. *Handbook of Labor Economics, 3*, 3143–3259.

Altonji, J. G., & Pierret, C. R. (2001). Employer learning and statistical discrimination. *The quarterly journal of economics*, *116*(1), 313-350.

Ameri, M., Schur, L., Adya, M., Bentley, F. S., McKay, P., & Kruse, D. (2018). The disability employment puzzle: A field experiment on employer hiring behavior. *ILR Review*, *71*(2), 329-364.

Anderson, D. (2001). 'If You Let Me Play': The Effects of Participation in High School Athletics on Students' Educational and Labor Market Success. *Available at SSRN 258751*.

Anderson, L., Fryer, R., & Holt, C. (2006). Discrimination: experimental evidence from psychology and economics. *Handbook on the Economics of Discrimination*, 97-118.

Andriessen, I., Nievers, E., Dagevos, J., & Faulk, L. (2012). Ethnic discrimination in the Dutch Labor Market: Its relationship with job characteristics and multiple group membership. *Work and Occupations, 39*, 237–239.

Andreoni, J., & Petrie, R. (2008). Beauty, gender and stereotypes: Evidence from laboratory experiments. *Journal of Economic Psychology*, *29*(1), 73-93.

Arceo-Gomez, E. O., & Campos-Vazquez, R. M. (2014). Race and marriage in the labor market: A discrimination correspondence study in a developing country. *American Economic Review, 104*, 376–380.

Arrow, K. J. (1998). What has economics to say about racial discrimination?. *Journal of economic perspectives*, *12*(2), 91-100.

Arrow, K. J. (1972). Some mathematical models of race discrimination in the labor market. *Racial discrimination in economic life*, 187-204.

Åslund, O., & Skans, O. N. (2012). Do anonymous job application procedures level the playing field?. *ILR Review*, *65*(1), 82-107.

Attström, K., & Attström, K. (2007). Discrimination against native Swedes of immigrant origin in access to employment. In *International Labor Organization, International Migration Papers 86E*.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics, 21*(1), 1-42.

Ayres, I. (1991). Fair driving: Gender and race discrimination in retail car negotiations. *Harvard Law Review*, 817-872.

Ayres, I., & Siegelman, P. (1995). Race and gender discrimination in bargaining for a new car. *The American Economic Review*, 304-321.

Baert, S. (2014). Career lesbians. Getting hired for not having kids? *Industrial Relations Journal, 45*, 543–561.

Baert, S. (2015). Field experimental evidence on gender discrimination in hiring: Biased as Heckman and Siegelman predicted? *Economics: The Open-Access, Open-Assessment E-Journal, 9*, 25.

Baert, S. (2016). Wage subsidies and hiring chances for the disabled: Some causal evidence. *European Journal of Health Economics, 17*, 71–86.

Baert, S. (2018). Facebook profile picture appearance affects recruiters' first hiring decisions. *New Media & Society*, *20*(3), 1220-1239.

Baert, S., & Balcaen, P. (2013). The impact of military work experience on later hiring chances in the civilian labour market. Evidence from a field experiment. *Economics: The Open-Access, Open-Assessment E-Journal, 7*, 37.

Baert, S., & Omey, E. (2015). Hiring discrimination against pro-union applicants: The role of union density and firm size. *Economist, 163*, 263–280.

Baert, S., & Verhofstadt, E. (2015). Labour market discrimination against former juvenile delinquents: Evidence from a field experiment. *Applied Economics, 47,* 1061–1072.

Baert, S., & Vujić, S. (2016). Immigrant volunteering: A way out of labour market discrimination? *Economics Letters, 146*, 95–98.

Baert, S., Jong, A. Pin, R., De Freyne, L., & Parmentier, S. (2014). *Political ideology and labour market discrimination*. Conference presentation at the spring meeting of Young Economists 2014.

Baert, S., Cockx, B., Gheyle, N., & Vandamme, C. (2015). Is there less discrimination in occupations where recruitment is difficult? *ILR Review, 68*, 467–500.

Baert, S., De Pauw, A. S., & Deschacht, N. (2016a). Do employer preferences contribute to s ticky floors? *ILR Review, 69*, 714–736.

Baert, S., De Visschere, S., Schoors, K., Vandenberghe, D., & Omey, E. (2016b). First depressed, then discriminated against? *Social Science & Medicine, 170*, 247–254.

Baert, S., Norga, J., Thuy, Y., & Van Hecke, M. (2016c). Getting grey hairs in the labour market. A realistic experiment on age discrimination. *Journal of Economic Psychology, 57*, 86–101.

Baert, S., Rotsaert, O., Verhaest, D., & Omey, E. (2016d). Student employment and later labour market success: No evidence for higher employment chances. *Kyklos, 69*, 401–425.

Baert, S., Albanese, A., du Gardein, S., Ovaere, J., & Stappers, J. (2017). Does work experience mitigate discrimination? *Economics Letters, 155*, 35–38.

Bailey, J., Wallace, M., & Wright, B. (2013). Are gay men and lesbians discriminated against when applying for jobs? A four-city, internet-based field experiment. *Journal of homosexuality*, *60*(6), 873-894.

Baker, M., & Cornelson, K. (2019). IX and the spatial content of female employment—Out of the lab and into the labor market. *Labour Economics*, *58*, 128-144.

Banaji, M. R., & Greenwald, A. G. (1995). Implicit gender stereotyping in judgments of fame. *Journal of personality and social psychology, 68*(2), 181.

Banerjee, A., Bertrand, M., Datta, S., & Mullainathan, S. (2009). Labor market discrimination in Delhi: Evidence from a field experiment. *Journal of comparative Economics*, *37*(1), 14-27.

Barron, J. M., Ewing, B. T., & Waddell, G. R. (2000). The effects of high school athletic participation on education and labor market outcomes. *Review of Economics and Statistics*, *82*(3), 409-421.

Bartolucci, C. (2013). Gender wage gaps reconsidered a structural approach using matched employer-employee data. *Journal of Human Resources*, *48*(4), 998-1034.

Bartolucci, C. (2014). Understanding the native–immigrant wage gap using matched employer-employee data: evidence from Germany. *ILR Review*, *67*(4), 1166-1202.

Bartoš, V., Bauer, M., Chytilová, J., & Matějka, F. (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review, 106*(6), 1437-75.

Baum, C. L., & Ford, W. F. (2004). The wage effects of obesity: a longitudinal study. *Health economics, 13*(9), 885-899.

Becker, Gary S. (1957). The Economics of Discrimination. Chicago, IL: The University of Chicago Press.

Becker, Gary, (1964), Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, First Edition, *National Bureau of Economic Research, Inc*, https://EconPapers.repec.org/RePEc:nbr:nberbk:beck-5.

Behaghel, L., Crépon, B., & Le Barbanchon, T. (2015). Unintended effects of anonymous resumes. *American Economic Journal: Applied Economics*, *7*(3), 1-27.

Bendick Jr, M., Brown, L. E., & Wall, K. (1999). No foot in the door: An experimental study of employment discrimination against older workers. *Journal of Aging & social policy*, *10*(4), 5-23.

Bendick Jr, M., Jackson, C. W., & Reinoso, V. A. (1994). Measuring employment discrimination through controlled experiments. *The review of black political economy*, *23*(1), 25-48.

Bendick Jr, M., Jackson, C. W., & Romero, J. H. (1997). Employment discrimination against older workers: An experimental study of hiring practices. *Journal of Aging & Social Policy*, *8*(4), 25-46.

Bendick Jr, M., & Nunes, A. P. (2012). Developing the research basis for controlling bias in hiring. *Journal of Social Issues*, *68*(2), 238-262.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*(3), 351–368.

Berman, P. W., O'Nan, B. A., & Floyd, W. (1981). The double standard of aging and the social situation: Judgments of attractiveness of the middle-aged woman. *Sex Roles*, *7*(2), 87-96.

Berson, B. (2012). *Does competition induce hiring equity? Documents de travail du Centre d'Economie de la Sorbonne 12019*.

Bertrand, M., & Duflo, E. (2016). Review on field experiments on discrimination. In A. Banerjee

& E. Duflo (Eds.), *Handbook of field experiments*. Cambridge: Abdul Latif Jameel Poverty Action Lab.

Bertrand, M., Chugh, D., & Mullainathan, S. (2005). Implicit discrimination. *American Economic Review*, *95*(2), 94-98.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American economic review*, *94*(4), 991-1013.

Biddle, J. E., & Hamermesh, D. S. (2013). Wage discrimination over the business cycle. *IZA Journal of Labor Policy*, *2*(1), 1-19.

Black, D. A. (1995). Discrimination in an equilibrium search model. *Journal of labor Economics*, *13*(2), 309-334.

Black, N., Johnston, D. W., Propper, C., & Shields, M. A. (2019). The effect of school sports facilities on physical activity, health and socioeconomic status in adulthood. *Social Science & Medicine*, *220*, 120-128.

Blau, F. D. (1984). Occupational segregation and labor market discrimination. *Sex segregation in the workplace, 117*, 122-23.

Blau, P. M. (1968). Social exchange. *International encyclopedia of the social sciences*, *7*(4), 452-457.

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of economic literature*, *55*(3), 789-865.

Blinder, A. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources, 8*, 436–455.

Blommaert, L., Coenders, M., & Van Tubergen, F. (2014). Ethnic discrimination in recruitment and decision makers' features: Evidence from laboratory experiment and survey data using a student sample. *Social indicators research*, *116*(3), 731-754.

BLS. 2019. Unemployment Status of the civilian noninstitutional population by age, sex, and race. Bureau of Labor Statistics. https://www.bls.gov/cps/cpsaat03.pdf

Bohren, J. A., Haggag, K., Imas, A., & Pope, D. G. (2019). *Inaccurate statistical discrimination: An identification problem* (No. w25935). National Bureau of Economic Research.

Bóo, F. L., Rossi, M. A., & Urzúa, S. S. (2013). The labor market return to an attractive face: Evidence from a field experiment. *Economics Letters*, *118*(1), 170-172.

Booth, A. L., & Leigh, A. (2010). Do employers discriminate by gender? A field experiment in female-dominated occupations. *Economics Letters, 107*, 236–238.

Booth, A. L., Leigh, A., & Varganova, E. (2012). Does ethnic discrimination vary across minority groups? Evidence from a field experiment. *Oxford Bulletin of Economics and Statistics, 74*, 547–573.

Borghans, L., Duckworth, A., Heckman, J., & ter Weel, B. (2008). The economics and psychology of personality traits. *The Journal of Human Resources, 4* (43), 972-1059.

Bowles, S., Gintis, H., & Osborne, M. (2001). The determinants of earnings: a behavioral approach. *Journal of Economic Literature, 4* (39), 137-1177.

Burn, I. (2015). Legal differences in state non-discrimination laws and the effect of employment protections for gay men. *Available at SSRN 2631898*.

Bursell, M. (2014). The multiple burdens of foreign-named men—evidence from a field experiment on gendered ethnic hiring discrimination in Sweden. *European Sociological Review, 30*, 399–409.

Büsch, V., Dahl, S. A., & Dittrich, D. A. (2009). An empirical study of age discrimination in Norway and Germany. *Applied Economics, 41*(5), 633-651.

Butler, D. M., & Homola, J. (2017). An empirical justification for the use of racially distinctive names to signal race in experiments. *Political Analysis, 25*(1), 122-130.

Caliendo, M., & Gehrsitz, M. (2016). Obesity and the labor market: A fresh look at the weight penalty. *Economics & Human Biology*, *23*, 209-225.

Capéau, B., Eeman, L., Groenez, S., & Lamberts, M. (2012). *Two concepts of discrimination: Inequality of opportunity versus unequal treatment of equals* (Ecore Discussion Paper Series 2012–58).

Carbonaro, W., & Schwarz, J. (2018). Opportunities and challenges in designing and conducting a labor market resume study. *Audit studies: Behind the scenes with theory, method, and nuance*, 143-158.

Card, D., & Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of political Economy*, *100*(1), 1-40.

Card, D., & Payne, A. A. (2002). School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of public economics, 83*(1), 49-82.

Carlsson, M. (2010). Experimental evidence of discrimination in the hiring of first- and second-generation immigrants. *Labour*, *24*, 263–278.

Carlsson, M. (2011). Does hiring discrimination cause gender segregation in the Swedish labor market? *Feminist Economics*, *17*, 71–102.

Carlsson, M., & Eriksson, S. (2017). Do attitudes expressed in surveys predict ethnic discrimination?. *Ethnic and Racial Studies, 40*(10), 1739-1757.

Carlsson, M., & Rooth, D. O. (2007). Evidence of ethnic discrimination in the Swedish labor market using experimental data. *Labour Economics, 14*, 716–729.

Carlsson, M., & Rooth, D. O. (2012). Revealing taste-based discrimination in hiring: A correspondence testing experiment with geographic variation. *Applied Economics Letters, 19*, 1861–1864.

Carneiro, P., Heckman, J. J., & Masterov, D. V. (2005). Labor market discrimination and racial differences in premarket factors. *The Journal of Law and Economics, 48*(1), 1-39.

Case, A., & Paxson, C. (2008). Stature and status: height, ability, and labor market outcomes. *Journal of Political Economy, 3*(116), 499-532.

Castillo, M., & Petrie, R. (2010). Discrimination in the lab: Does information trump appearance?. *Games and Economic Behavior*, *68*(1), 50-59.

Castillo, M., Petrie, R., Torero, M., & Vesterlund, L. (2013). Gender differences in bargaining outcomes: A field experiment on discrimination. *Journal of Public Economics*, *99*, 35-48.

Cawley, J., Meyerhoefer, C., & Newhouse, D. (2007). The impact of state physical education requirements on youth physical activity and overweight. *Health economics*, *16*(12), 1287-1301.

Cediey, E., & Foroni, F. (2008). *Discrimination in access to employment on grounds of foreign origin in France: A national survey of discrimination based on the testing methodology of the International Labour Office*. ILO.

Census Bureau. (2021). Poverty data by race. https://www.census.gov/topics/income-poverty/poverty.html

Charles, K. K., & Guryan, J. (2008). Prejudice and wages: an empirical assessment of Becker's The Economics of Discrimination. *Journal of political economy*, *116*(5), 773-809.

Guryan, J., & Charles, K. K. (2013). Taste-based or statistical discrimination: The economics of discrimination returns to its roots. *The Economic Journal*, *123*(572), F417-F432.

Coate, S., & Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes?. *The American Economic Review*, 1220-1240.

Cohen-Zada, D., Gradstein, M., & Reuven, E. (2009). Class size and the regression discontinuity design: the case of public schools. *Available at SSRN 1533972*.

Colcombe, S., & Kramer, A. (2010). Fitness effects on the cognitive function of older adults. *Psychological Science, 2*(14), 125-130.

Coleman, J. S. (1961). Athletics in high school. *The Annals of the American Academy of Political and Social Science*, *338*(1), 33-43.

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty?. *American journal of sociology*, *112*(5), 1297-1338.

Council of Economic Advisers (US), Council of Economic Advisers, Washington, DC., United States Advisory Board to the President's Initiative on Race, United States. Government Printing Office, & President's Initiative on Race (US). (1998). *Changing America: Indicators of social and economic well-being by race and Hispanic origin*. US Government Printing Office.

Cross, H. (1990). Employer hiring practices: differential treatment of Hispanic and Anglo job seekers.

Currie, J., & Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of labor economics*, *3*, 3309-3416.

Daniel, W. W. (1968). *Racial discrimination in England: based on the PEP report* (Vol. 257). Penguin.

Darity, W. A., & Mason, P. L. (1998). Evidence on discrimination in employment: Codes of color, codes of gender. *Journal of Economic Perspectives*, *12*(2), 63-90.

Darolia, R., Koedel, C., Martorell, P., Wilson, K., & Perez-Arce, F. (2016). Race and gender effects on employer interest in job applicants: New evidence from a resume field experiment. *Applied Economics Letters, 23*, 853–856.

Davison, H. K., & Burke, M. J. (2000). Sex discrimination in simulated employment contexts: A meta-analytic investigation. *Journal of Vocational Behavior*, *56*(2), 225-248.

Decker, S. H., Ortiz, N., Cassia, S., & Hedberg, E. (2015). Criminal stigma, race, and ethnicity: The consequences of imprisonment for employment. *Journal of Criminal Justice, 43*, 108–121.

Derous, E., Ryan, A. M., & Nguyen, H. H. (2012). Multiple categorization in resume screening: Examining effects on hiring discrimination against Arab applicants in field and lab settings. *Journal of Organizational Behavior, 33*, 544–570.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology*, *56*(1), 5.

Dickinson, D. L., & Oaxaca, R. L. (2009). Statistical discrimination in labor markets: An experimental analysis. *Southern Economic Journal*, *76*(1), 16-31.

Dittrich, M., Knabe, A., & Leipold, K. (2014). Gender differences in experimental wage negotiations. *Economic Inquiry*, *52*(2), 862-873.

Doleac, J. L., & Stein, L. C. (2013). The visible hand: Race and online market outcomes. *The Economic Journal*, *123*(572), F469-F492.

Donohue, J. J., & Heckman, J. J. (1991). Continuous versus episodic change: The impact of civil rights policy on the economic status of blacks.

Dovidio, J. F., Gaertner, S. E., Kawakami, K., & Hodson, G. (2002). Why can't we just get along? Interpersonal biases and interracial distrust. *Cultural diversity and ethnic minority psychology*, *8*(2), 88.

Drydakis, N. (2009). Sexual orientation discrimination in the labour market. *Labour Economics, 16*, 364–372.

Drydakis, N. (2010a). Labour discrimination as a symptom of HIV: Experimental evaluation: The greek case. *Journal of Industrial Relations, 52*, 201–217.

Drydakis, N. (2010b). Religious affiliation and labour bias. *Journal for the Scientific Study of Religion, 49*, 472–488.

Drydakis, N. (2011). Women's sexual orientation and labor market outcomes in Greece. *Feminist Economics, 11*, 89–117.

Drydakis, N. (2012a). Estimating ethnic discrimination in the labour market using experimental data. *Southeast European and Black Sea Studies, 12*, 335–355.

Drydakis, N. (2012b). Sexual orientation and labour relations: New evidence from Athens, Greece. *Applied Economics, 44*, 2653–2665.

Drydakis, N. (2014). Sexual orientation discrimination in the Cypriot labour market. Distastes or uncertainty? *International Journal of Manpower, 35*, 720–744.

Drydakis, N. (2015). Measuring sexual orientation discrimination in the UK's labour market; a field experiment. *Human Relations, 68*, 1769–1796.

Drydakis, N., & Vlassis, M. (2010). Ethnic discrimination in the Greek labour market: Occupational access, insurance coverage and wage offers. *Manchester School, 78,* 201–218.

Duguet, E., Petit, P., & Petit, P. (2005). Hiring discrimination in the French financial sector: an econometric analysis on field experiment data. *Annales d'Economie et de Statistique*, 79-102.

Duguet, E., Leandri, N., L'Horty, Y., & Petit, P. (2010). Are young french jobseekers of ethnic immigrant origin discriminated against? A controlled experiment in the Paris area. *Annals of Economics and Statistics / Annales d'Économie et de Statistique, 99*(100), 187–215.

Eckel, C. (2008). The gender gap: Using the lab as a window on the market. In *Frontiers in the Economics of Gender* (pp. 239-258). Routledge.

Edo, A., Jacquemet, N., & Yannelis, C. (2019). Language skills and homophilous hiring discrimination: Evidence from gender and racially differentiated applications. *Review of Economics of the Household*, *17*(1), 349-376.

Eide, E., & Ronan, N. (2001). is participation in high school athletics an investment or a consumption good?: Evidence from high school and beyond. *Economics of Education Review, 20*(5), 431-442.

Eriksson, S., & Rooth, D. O. (2014). Do employers use unemployment as a sorting criterion when hiring? Evidence from a field experiment. *American economic review*, *104*(3), 1014-39.

Ewing, B. T. (1998). Athletes and work. *Economics Letters*, *59*(1), 113-117.

Ewing, B. T. (2007). The labor market effects of high school athletic participation: Evidence from wage and fringe benefit differentials. *Journal of Sports Economics*, *8*(3), 255-265.

Exley, C. L., Niederle, M., & Vesterlund, L. (2020). Knowing when to ask: The cost of leaning in. *Journal of Political Economy*, *128*(3), 816-854.

Fang, H., & Moro, A. (2011). Theories of statistical discrimination and affirmative action: A survey. *Handbook of social economics*, *1*, 133-200.

Farber, H. S., Silverman, D., & Von Wachter, T. M. (2017). Factors determining callbacks to job applications by the unemployed: An audit study. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *3*(3), 168-201.

Feldman, A. F., & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of educational research*, *75*(2), 159-210.

Fernandes, A. L., & Gabe, T. (2009). *Disconnected youth: A look at 16-to 24-year olds who are not working or in school*. DIANE Publishing.

Fershtman, C., & Gneezy, U. (2001). Discrimination in a segmented society: An experimental approach. *The Quarterly Journal of Economics*, *116*(1), 351-377.

Figinski, T. F. (2017). The effect of potential activations on the employment of military reservists: Evidence from a field experiment. *ILR Review*, *70*(4), 1037-1056

Finlay, K. (2008). *Effect of employer access to criminal history data on the labor market outcomes of ex-offenders and non-offenders* (No. w13935). National Bureau of Economic Research.

Fix, Michael, & Struyk, R. J. (1993). Clear and Convincing Evidence: Measurement of Discrimination in America. The Urban Institute Press.

Flory, J. A., Leibbrandt, A., & List, J. A. (2015). Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions. *The Review of Economic Studies*, *82*(1), 122-155.

Foster, A. D., & Rosenzweig, M. R. (1993). Information, learning, and wage rates in low-income rural areas. *Journal of Human resources*, 759-790.

Frijters, P., Shields, M., Theodoropoulos, N., & Price, W. (2006). Testing for employee discrimination in Britain using matched employer-employee data. *Department of Economics University of Cyprus Discussion Paper*, *8*, 40.

Fryer Jr, R. G., & Levitt, S. D. (2004). The causes and consequences of distinctively black names. *The Quarterly Journal of Economics*, *119*(3), 767-805.

Gaddis, S. M. (2015). Discrimination in the credential society: An audit study of race and college selectivity in the labor market. *Social Forces, 93*, 1451–1479.

Gaddis, S. M. (2018). An introduction to audit studies in the social sciences. In *Audit studies: Behind the scenes with theory, method, and nuance* (pp. 3-44). Springer, Cham.

Galarza, F. B., & Yamada, G. (2014). Labor market discrimination in Lima, Peru: Evidence from a field experiment. *World Development*, *58*, 83-94.

Gallo, E., Grund, T., & James Reade, J. (2013). Punishing the foreigner: implicit discrimination in the Premier League based on oppositional identity. *Oxford Bulletin of Economics and Statistics*, *75*(1), 136-156.

Galster, G., & Constantine, P. (1991). Discrimination against female-headed households in rental housing: Theory and exploratory evidence. *Review of Social Economy*, *49*(1), 76-100.

Ghayad, R. (2013). Escaping the Unem ployment Trap: Does Industry-Specific Human Capital Matter? Evidence from a Field Experiment.

Glover, D., Pallais, A., & Pariente, W. (2017). Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores. *The Quarterly Journal of Economics*, *132*(3), 1219-1260.

Gneezy, U., List, J., & Price, M. K. (2012). *Toward an understanding of why people discriminate: Evidence from a series of natural field experiments* (No. w17855). National Bureau of Economic Research.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The quarterly journal of economics*, *118*(3), 1049-1074.

Goldberg, M. S. (1982). Discrimination, nepotism, and long-run wage differentials. *The quarterly journal of economics*, *97*(2), 307-319.

Goldin, C., & Mitchell, J. (2017). The new life cycle of women's employment: Disappearing humps, sagging middles, expanding tops. *Journal of Economic Perspectives*, *31*(1), 161-82.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of" blind" auditions on female musicians. *American economic review*, *90*(4), 715-741.

Gomez-Pinilla, F. (2008). The influence of diet and exercise on mental health through hormensis. *Aging Research Review, 7*, 49-62.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, *102*(1), 4.

Gronau, R. (1988). Sex-related wage differentials and women's interrupted labor careers-The chicken or the egg. *Journal of labor Economics*, *6*(3), 277-301.

Groshen, E. L. (1991). The structure of the female/male wage differential: Is it who you are, what you do, or where you work?. *Journal of Human Resources*, 457-472.

Grossman, M. (2017). 1. On the Concept of Health Capital and the Demand for Health. In *Determinants of Health* (pp. 6-41). Columbia University Press.

Hainmueller, J., Hopkins, D., Yamamoto, T. (2013). Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments. *Political Analysis, 22*(1), 1-30.

Hamermesh, D. S. (2011). Beauty pays. In *Beauty Pays*. Princeton University Press.

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic literature*, *42*(4), 1009-1055.

Heckman, J. J. (1998). Detecting discrimination. *Journal of economic perspectives*, *12*(2), 101-116.

Heckman, J. J., & Carneiro, P. (2003). Human capital policy.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review*, *91*(2), 145-149.

Heckman, J. J., & Siegelman, P. (1993). The Urban Institute audit studies: Their methods and findings. *Clear and convincing evidence: Measurement of discrimination in America*, 187-258.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, *24*(3), 411-482.

Hedegaard, M. S., & Tyran, J. R. (2018). The price of prejudice. *American Economic Journal: Applied Economics*, *10*(1), 40-63.

Heilman, M. E. (1984). Information as a deterrent against sex discrimination: The effects of applicant sex and information type on preliminary employment decisions. *Organizational Behavior and Human Performance*, *33*(2), 174-186.

Hellerstein, J. K., & Neumark, D. (1999). Sex, wages, and productivity: An empirical analysis of Israeli firm-level data. *International Economic Review*, *40*(1), 95-123.

Hellerstein, J. K., & Neumark, D. (2008). Workplace segregation in the United States: Race, ethnicity, and skill. *The review of economics and statistics*, *90*(3), 459-477.

Hellerstein, J. K., Neumark, D., & Troske, K. R. (1999). Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of labor economics*, *17*(3), 409-446.

Hellester, M. D., Kuhn, P., & Shen, K. (2014). Employers' Age and Gender Preferences: Direct Evidence from Four Job Boards. *Unpublished manuscript, University of California, Santa Barbara*.

Hipes, C., Lucas, J., Phelan, J. C., & White, R. C. (2016). The stigma of mental illness in the labor market. *Social Science Research*, *56*, 16–25.

Holzer, H. J. (1996). Employer hiring decisions and antidiscrimination policy. Institute for Research on Poverty, University of Wisconsin-Madison.

Holzer, H. J. (1998). Why do small establishments hire fewer blacks than large ones?. *Journal of Human Resources*, 896-914.

Holzer, H. J., & Neumark, D. (2000). What does affirmative action do?. *ILr review*, *53*(2), 240-271.

Hosoda, M., Stone-Romero, E. F., & Coats, G. (2003). The effects of physical attractiveness on job-related outcomes: A meta-analysis of experimental studies. *Personnel psychology*, *56*(2), 431-462.

Hutchens, R. (1986). Delayed payment contracts and a firm's propensity to hire older workers. *Journal of labor economics*, *4*(4), 439-457.

Ioannides, Y. M., & Datcher Loury, L. (2004). Job information networks, neighborhood effects, and inequality. *Journal of economic literature*, *42*(4), 1056-1093.

Issacharoff, S., & Harris, E. W. (1997). Is age discrimination really age discrimination: the ADEA's unnatural solution. *NYUL Rev.*, *72*, 780.

Jackson, L. A. (1992). *Physical appearance and gender: Sociobiological and sociocultural perspectives*. SUNY Press.

Jackson, M. (2009). Disadvantaged through discrimination? The role of employers in social stratification. *British Journal of Sociology, 60*, 669–692.

Jackson, M. O. (2011). An overview of social networks and economic applications. *Handbook of social economics*, *1*, 511-585.

Jacquemet, N., & Yannelis, C. (2012). Indiscriminate discrimination: A correspondence test for ethnic homophily in the Chicago labor market. *Labour Economics, 19*, 824–832.

Johnes, G. (2018). A sporting chance: on the impact of sports participation on subsequent earnings. *Economics Bulletin*, *38*(1), 146-151.

Johnson, W. R., & Neal, D. A. (1997). *Basic skills and the black-white earnings gap*. University of Virginia, Thomas Jefferson Center for Political Economy.

Jowell, R., & Prescott-Clarke, P. (1970). Racial discrimination and white-collar workers in Britain. *Race*, *11*(4), 397-417.

Judge, T. A., & Cable, D. M. (2004). The effect of physical height on workplace success and income: preliminary test of a theoretical model. *Journal of Applied Psychology*, *89*(3), 428.

Kaas, L., & Manger, C. (2012). Ethnic discrimination in Germany's labour market: A field experiment. *German economic review*, *13*(1), 1-20.

Kahn, L. M. (2000). The sports business as a labor market laboratory. *Journal of economic perspectives*, *14*(3), 75-94.

Katz, D., & Braly, K. W. (1935). Racial prejudice and racial stereotypes. *The Journal of abnormal and social psychology*, *30*(2), 175.

Kite, M. E., Stockdale, G. D., Whitley Jr, B. E., & Johnson, B. T. (2005). Attitudes toward younger and older adults: An updated meta-analytic review. *Journal of social issues*, *61*(2), 241-266.

Kleykamp, M. A. (2009). Great place to start? The effect of prior military service on hiring. *Armed Forces & Society*, *35*, 266–285.

Kniffin, K. M., Wansink, B., & Shimizu, M. (2015). Sports at work: Anticipated and persistent correlates of participation in high school athletics. *Journal of Leadership & Organizational Studies*, *22*(2), 217-230.

Kochhar, R. (2008). *Latino labor report, 2008: Construction reverses job growth for Latinos*. Washington, DC: Pew Hispanic Center.

Krause, A., Rinne, U., & Zimmermann, K. F. (2012). Anonymous job applications of fresh Ph. D. economists. *Economics Letters*, *117*(2), 441-444.

Krings, F., Sczesny, S., & Kluge, A. (2011). Stereotypical inferences as mediators of age discrimination: The role of competence and warmth. *British Journal of Management*, *22*(2), 187-201.

Kroft, K., Lange, F., & Notowidigdo, M. J. (2013). Duration dependence and labor market conditions: Evidence from a field experiment. *The Quarterly Journal of Economics*, *128*(3), 1123-1167.

Krueger, A. B., & Whitmore, D. M. (2001). The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR. *The Economic Journal*, *111*(468), 1-28.

Kruse, D., & Schur, L. (2003). Employment of people with disabilities following the ADA. *Industrial Relations: A Journal of Economy and Society*, *42*(1), 31-66.

Kuhn, P., & Shen, K. (2013). Gender discrimination in job ads: Evidence from china. *The Quarterly Journal of Economics*, *128*(1), 287-336.

Kuhn, P., & Weinberger, C. (2005). Leadership skills and wages. *Journal of Labor Economics*, *23*(3), 395-436.

Lahey, J. N. (2008). Age, women, and hiring an experimental study. *Journal of Human resources*, *43*(1), 30-56.

Lahey, J. N., & Beasley, R. A. (2009). Computerizing audit studies. *Journal of Economic Behavior & Organization*, *70*(3), 508-514.

Lahey, J. N., & Oxley, D. R. (2021). Discrimination at the Intersection of Age, Race, and Gender: Evidence from an Eye-Tracking Experiment. *Journal of Policy Analysis and Management*, *40*(4), 1083-1119.

Lang, K., & Lehmann, J. Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, *50*(4), 959-1006.

Lang, K., & Manove, M. (2011). Education and labor market discrimination. *American Economic Review*, *101*(4), 1467-96.

Lang, K., Manove, M., & Dickens, W. T. (2005). Racial discrimination in labor markets with posted wage offers. *American Economic Review*, *95*(4), 1327-1340.

Laouenan, M., & Rathelot, R. (2017). Ethnic discrimination on an online marketplace of vacation rental.

Lazear, E. P. (1979). Why is there mandatory retirement?. *Journal of political economy*, *87*(6), 1261-1284.

Lechner, M. (2009). Long-run labour market and health effects of individual sports activities. *Journal of health economics*, *28*(4), 839-854.

Lechner, M., & Downward, P. (2017). Heterogeneous sports participation and labour market outcomes in England. *Applied Economics*, *49*(4), 335-348.

Lee, H. A., & Khalid, M. A. (2016). Discrimination of high degrees: Race and graduate hiring in Malaysia. *Journal of the Asia Pacific Economy, 21*, 53–76.

Leonard, J. S., & Levine, D. I. (2006). The effect of diversity on turnover: A large case study. *ILR Review*, 59(4), 547-572.

Lipscomb, S. (2007). Secondary school extracurricular involvement and academic achievement: A fixed effects approach. *Economics of Education Review*, *26*(4), 463-472.

List, J. A. (2004). The nature and extent of discrimination in the marketplace: Evidence from the field. *The Quarterly Journal of Economics*, *119*(1), 49-89.

Long, J. E., & Caudill, S. B. (1991). The impact of participation in intercollegiate athletics on income and graduation. *The Review of Economics and Statistics*, 525-531.

Lopez Bóo, F., Rossi, M., & Urzúa, S. (2013). The labor market return to an attractive face: Evidence from a field experiment. *Economics Letters, 118*, 170–172.

Lundberg, S. J., & Startz, R. (1983). Private discrimination and social intervention in competitive labor market. *The American economic review*, *73*(3), 340-347.

Lundborg, P., Nystedt, P., & Rooth, D. O. (2014). Body size, skills, and income: evidence from 150,000 teenage siblings. *Demography*, *51*(5), 1573-1596.

Make the Road New York. (2010). Transgender need not apply: Gender identity job discrimination in New York City's retail sector.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*(1), 1–23.

Maurer-Fazio, M. (2012). Ethnic discrimination in China's internet job board labor market. *IZA Journal of Migration, 1*, 1–24.

Maurer-Fazio, M., & Lei, L. (2015). As rare as a panda. How facial attractiveness, gender, and occupation affect interview callbacks at Chinese firms. *International Journal of Manpower, 36*, 68–85.

McGinnity, F., & Lunn, P. D. (2011). Measuring discrimination facing ethnic minority job applicants: An Irish experiment. *Work, Employment and Society, 25*, 693–708.

Midtbøen, A. H. (2013). The invisible second generation? Statistical discrimination and immigrant stereotypes in employment processes in Norway. *Journal of Ethnic and Migration Studies, 40*, 1657–1675.

Midtbøen, A. H. (2016). Discrimination of the second generation: Evidence from a field experiment in Norway. *Journal of International Migration and Integration, 17*, 253–272.

Miller, C. (2017). The persistent effect of temporary affirmative action. *American Economic Journal: Applied Economics*, *9*(3), 152-90.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of political economy*, *66*(4), 281-302.

Mincer, J. (1974). Schooling, Experience, and Earnings. Columbia University Press.

Mincy, R. B. (1993). The Urban Institute audit studies: their research and policy context. *Clear and Convincing Evidence: Measurement of Discrimination in America*, 165-86.

Mishel, E. (2016). Discrimination against queer women in the US workforce: A résumé audit study. *Socius*, *2*, 2378023115621316.

Mobius, M. M., & Rosenblat, T. S. (2006). Why beauty matters. *American Economic Review*, *96*(1), 222-235.

Moreno, M., Ñopo, H., Saavedra, J., & Torero, M. (2012). Detecting gender and racial discrimination in hiring through monitoring intermediation services: the case of selected occupations in Metropolitan Lima, Peru. *World Development*, *40*(2), 315-328.

Mujcic, R., & Frijters, P. (2013). Still not allowed on the bus: It matters if you're black or white!.

National Center for Education Statistics. (2021). Digest of Education Statistics. https://nces.ed.gov/programs/digest/current_tables.asp

Neal, D. A., & Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of political Economy*, *104*(5), 869-895.

Neumark, D. (2004). Employers' discriminatory behavior and the estimation of wage discrimination. In *Sex Differences in Labor Markets* (pp. 163-177). Routledge.

Neumark, D. (2012). Detecting discrimination in audit and correspondence studies. *Journal of Human Resources, 47*(4), 1128-1157.

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature, 56*(3), 799-866.

Neumark, D., Bank, R. J., & Van Nort, K. D. (1996). Sex discrimination in restaurant hiring: An audit study. *The Quarterly journal of economics, 111*(3), 915-941.

Neumark, D., Burn, I., & Button, P. (2019). Is it harder for older workers to find jobs? New and improved evidence from a field experiment. *Journal of Political Economy, 127*(2), 922-970.

Neumark, D., & Button, P. (2014). Did age discrimination protections help older workers weather the Great Recession?. *Journal of Policy Analysis and Management, 33*(3), 566-601.

Neumark, D., & Rich, J. (2019). Do field experiments on labor and housing markets overstate discrimination? A re-examination of the evidence. *ILR Review, 72*(1), 223-252.

Neumark, D., Song, J., & Button, P. (2017). Does protecting older workers from discrimination make it harder to get hired? Evidence from disability discrimination laws. *Research on Aging, 39*(1), 29-63.

Neumark, D., & Stock, W. A. (1999). Age discrimination laws and labor market efficiency. *Journal of Political Economy, 107*(5), 1081-1125.

Neumark, D., & Stock, W. A. (2006). The labor market effects of sex and race discrimination laws. *Economic Inquiry, 44*(3), 385-419.

NFHS. (2014). 2013–14 High School Athletics Participation Survey. http://www.nfhs.org.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much?. *The quarterly journal of economics*, *122*(3), 1067-1101.

Norton, M. I., Sommers, S. R., Apfelbaum, E. P., Pura, N., & Ariely, D. (2006). Color blindness and interracial interaction: Playing the political correctness game. *Psychological Science*, *17*(11), 949-953.

Norton, M. I., Vandello, J. A., & Darley, J. M. (2004). Casuistry and social category bias. *Journal of personality and social psychology*, *87*(6), 817.

Nunley, J. M., Pugh, A., Romero, N., & Seals, R. A. (2015). Racial discrimination in the labor market for recent college graduates: Evidence from a field experiment. *B.E. Journal of Economic Analysis & Policy*, *15*, 1093–1125.

NWLC & PRRAC. (2015). Finishing Last: Girls of Color and School Sports Opportunities. National Women's Law Center & Poverty & Race Research Action Council

O'Neill, J. (1990). The role of human capital in earnings differences between black and white men. *Journal of economic Perspectives*, *4*(4), 25-45.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693-709.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447-453.

OECD. 2013. *International Migration Outlook 2013*, Paris: OECD.

Oh, S. J., & Yinger, J. (2015). What have we learned from paired testing in housing markets?. *Cityscape*, *17*(3), 15-60.

Olian, J. D., Schwab, D. P., & Haberfeld, Y. (1988). The impact of applicant gender compared to qualifications on hiring recommendations: A meta-analysis of experimental studies. *Organizational Behavior and Human Decision Processes*, *41*(2), 180-195.

Oreopoulos, P. (2011). Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes. *American Economic Journal: Economic Policy*, *3*(4), 148-71.

Orfield, G., Kucsera, J., & Siegel-Hawley, G. (2012). E pluribus... separation: Deepening double segregation for more students.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, *105*(2), 171.

Pager, D. (2003). The mark of a criminal record. *American journal of sociology*, *108*(5), 937-975.

Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *The Annals of the American Academy of Political and Social Science*, *609*(1), 104-133.

Pager, D. (2016). Are firms that discriminate more likely to go out of business?. *Sociological Science*, *3*, 849-859.

Pager, D., Bonikowski, B., & Western, B. (2009). Discrimination in a low-wage labor market: A field experiment. *American sociological review*, *74*(5), 777-799.

Pager, D., & Quillian, L. (2005). Walking the talk? What employers say versus what they do. *American Sociological Review*, *70*(3), 355-380.

Parsons, C. A., Sulaeman, J., Yates, M. C., & Hamermesh, D. S. (2011). Strike three: Discrimination, incentives, and evaluation. *American Economic Review, 101*(4), 1410-35.

Patacchini, E., Ragusa, G., & Zenou, Y. (2015). Unexplored dimensions of discrimination in Europe: Homosexuality and physical appearance. *Journal of Population Economics, 28*, 1045–1073.

Pedulla, D. S. (2016). Penalized or protected? Gender and the consequences of nonstandard and mismatched employment histories. *American sociological review, 81*(2), 262-289.

Pedulla, D. S. (2018). Emerging frontiers in audit study research: Mechanisms, variation, and representativeness. In *Audit studies: behind the scenes with theory, method, and nuance* (pp. 179-195). Springer, Cham.

Petit, P. (2007). The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector. *Labour Economics*, *14*(3), 371-391.

Persico, N., Postlewaite, A., & Silverman, D. (2004). The effect of adolescent experience on labor market outcomes: The case of height. *Journal of political Economy, 112*(5), 1019-1053.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review*, *62*(4), 659-661.

Phillips, D. C. (2019). Do comparisons of fictional applicants measure discrimination when search externalities are present? Evidence from existing experiments. *The Economic Journal*, *129*(621), 2240-2264.

Pierné, G. (2013). Hiring discrimination based on national origin and religious closeness: Results from a field experiment in the Paris area. *IZA Journal of Labor Economics, 2*, 4.

Pingitore, R., Dugoni, B. L., Tindale, R. S., & Spring, B. (1994). Bias against overweight job applicants in a simulated employment interview. *Journal of applied psychology, 79*(6), 909.

Polinko, N. K., & Popovich, P. M. (2001). Evil thoughts but angelic actions: Responses to overweight job applicants. *Journal of Applied Social Psychology*, *31*(5), 905-924.

Price, J., & Wolfers, J. (2010). Racial discrimination among NBA referees. *The Quarterly journal of economics*, *125*(4), 1859-1887.

Quillian, L. (2006). New approaches to understanding racial prejudice and discrimination. *Annual review of sociology*, 299-328.

Quillian, L., & Hexel, O. (2016, April). Trends and Patterns in Racial Discrimination in Hiring in America, 1974–2015. In *Meetings of the Population Association of America, Washington, DC*.

Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences*, *114*(41), 10870-10875.

Randall, E. (2021, June 5). *What's a Lax Bro? Don't Ask a Lax Bro*. The Atlantic. https://www.theatlantic.com/national/archive/2012/06/whats-lax-bro-dont-ask-lax-bro/327290/

Ravaud, J. F., Madiot, B., & Ville, I. (1992). Discrimination towards disabled people seeking employment. *Social Science & Medicine*, *35*(8), 951-958.

Rees, D. I., & Sabia, J. J. (2010). Sports participation and academic performance: Evidence from the National Longitudinal Study of Adolescent Health. *Economics of Education Review*, *29*(5), 751-759.

Reuben, E., Sapienza, P., & Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, *111*(12), 4403-4408.

Riach, P. A. (2015). A field experiment investigating age discrimination in four European labour markets. *International Review of Applied Economics*, *29*(5), 608-619.

Riach, P. A., & Rich, J. (2002). Field experiments of discrimination in the market place. *The economic journal*, *112*(483), F480-F518.

Riach, P. A., & Rich, J. (2010). An experimental investigation of age discrimination in the English labor market. *Annals of Economics and Statistics/Annales d'économie et de Statistique*, 169-185.

Rich, J. (2014). What do field experiments of discrimination in markets tell us? A meta analysis of studies conducted since 2000.

Riggio, R. E., & Throckmorton, B. (1988). The relative effects of verbal and nonverbal behavior, appearance, and social skills on evaluations made in hiring interviews 1. *Journal of Applied Social Psychology*, *18*(4), 331-348.

Rivera, L. A. (2012). Hiring as cultural matching: The case of elite professional service firms. *American sociological review*, *77*(6), 999-1022.

Rodgers III, W. M., & Spriggs, W. E. (1996). What does the AFQT really measure: Race, wages, schooling and the AFQT score. *The Review of Black Political Economy*, *24*(4), 13-46.

Rooth, D. O. (2009). Obesity, attractiveness, and differential treatment in hiring a field experiment. *Journal of human resources*, *44*(3), 710-735.

Rooth, D. O. (2010). Automatic associations and discrimination in hiring: Real world evidence. *Labour Economics*, *17*(3), 523-534.

Rooth, D. O. (2011). Work out or out of work—the labor market return to physical fitness and leisure sports activities. *Labour Economics*, *18*(3), 399-409.

Rosen, B., & Jerdee, T. H. (1974). Influence of sex role stereotypes on personnel decisions. *Journal of Applied Psychology*, *59*(1), 9.

Rosen, Benson, and Thomas H. Jerdee. 1976. "The Influence of Age Stereotypes on Managerial Decisions." *Journal of Applied Psychology* 61(4): 428-32.

Rosen, B., & Jerdee, T. H. (1976). The influence of age stereotypes on managerial decisions. *Journal of applied psychology*, *61*(4), 428.

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of social issues*, *57*(4), 743-762.

Rudolph, C. W., Wells, C. L., Weller, M. D., & Baltes, B. B. (2009). A meta-analysis of empirical studies of weight-based bias in the workplace. *Journal of Vocational Behavior*, *74*(1), 1-10.

Ruffle, B. J., & Shtudiner, Z. E. (2015). Are good-looking people more employable?. *Management Science*, *61*(8), 1760-1776.

Schneider, S., & Becker, S. (2005). Prevalence of physical activity among the working population and correlation with work-related factors: results from the first German National Health Survey. *Journal of occupational health*, *47*(5), 414-423.

Schwartz, R. D., & Skolnick, J. H. (1962). Two studies of legal stigma. *Social problems*, *10*(2), 133-142.

Seippel, Ø. (2006). Sport and social capital. *Acta sociologica*, *49*(2), 169-183.

Siddique, Z. (2011). Evidence on caste based discrimination. *Labour Economics*, *18*, S146-S159.

Singer, M. S., & Sewell, C. (1989). Applicant age and selection interview decisions: Effect of information exposure on age discrimination in personnel selection. *Personnel Psychology*, *42*(1), 135-154.

Social Science Research Council. (2019). *Youth Disconnection.* http://measureofamerica.org/youth-disconnection-2019/

Solnick, S. J. (2001). Gender differences in the ultimatum game. *Economic Inquiry*, *39*(2), 189-200.

Spence, M. (1978). Job market signaling. In *Uncertainty in economics* (pp. 281-306). Academic Press.

Steffens, M. C. (2004). Is the implicit association test immune to faking?. *Experimental psychology*, *51*(3), 165.

Stempel, C. (2005). Adult participation sports as cultural capital: A test of Bourdieu's theory of the field of sports. *International Review for the Sociology of Sport*, *40*(4), 411-432.

Stevenson, B. (2010). Beyond the classroom: using title IX to measure the return to high school sports. *The Review of Economics and Statistics, 2*(2), 284-301.

Stone, A., & Wright, T. (2013). When your face doesn't fit: Employment discrimination against people with facial disfigurements. *Journal of Applied Social Psychology*, *43*, 515–526.

Tajfel, H., & Turner, J. (1979). Social identity theory. *dikutip dari*, 20.

The Annie E. Casey Foundation (2012). Youth and Work: Restoring Teen and Young Adult Connections to Opportunity. https://www.aecf.org/resources/youth-and-work

Tilcsik, A. (2011). Pride and prejudice: Employment discrimination against openly gay men in the United States. *American Journal of Sociology*, *117*(2), 586-626.

Tinsley, M. (2012). Too much to lose: Understanding and supporting Britain's older workers'. Policy Exchange.

Tosi, H. L., & Einbender, S. W. (1985). The effects of the type and amount of information in sex discrimination research: A meta-analysis. *Academy of Management Journal*, *28*(3), 712-723.

Tunstall, R., Green, A., Lupton, R., Watmough, S., & Bates, K. (2014). Does poor neighbourhood reputation create a neighbourhood effect on employment? The results of a field experiment in the UK. *Urban Studies, 51*, 763–780.

Turner, M. A., Fix, M., & Struyk, R. J. (1991). *Opportunities denied, opportunities diminished: Racial discrimination in hiring*. The Urban Insitute.

Uhlmann, E. L., & Cohen, G. L. (2005). Constructed criteria: Redefining merit to justify discrimination. *Psychological Science, 16*(6), 474-480.

USA Facts. (2021). Data visualization. https://usafacts.org/data/

Vuolo, M., Uggen, C., & Lageson, S. (2018). To match or not to match? Statistical and substantive considerations in audit design and analysis. In *Audit studies: Behind the scenes with theory, method, and nuance* (pp. 119-140). Springer, Cham.

Weichselbaumer, D. (2003). Sexual orientation discrimination in hiring. *Labour economics*, *10*(6), 629-642.

Weichselbaumer, D. (2004). Is it sex or personality? The impact of sex stereotypes on discrimination in applicant selection. *Eastern Economic Journal*, *30*(2), 159-186.

Weichselbaumer, D. (2015). Testing for discrimination against lesbians of different marital status: A field experiment. *Industrial Relations: A Journal of Economy and Society*, *54*(1), 131-161.

Weiss, E. M., & Maurer, T. J. (2004). Age discrimination in personnel decisions: A reexamination 1. *Journal of applied social psychology*, *34*(8), 1551-1562.

Widner, D., & Chicoine, S. (2011). It's all in the name: Employment discrimination against Arab Americans. *Sociological Forum*, *26*, 806–823.

Wood, M., Hales, J., Purdon, S., Sejersen, T., & Hayllar, O. (2009). A test for racial discrimination in recruitment practice in British cities. *Department for Work and Pensions Research Report*, *607*.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of experimental social psychology*, *10*(2), 109-120.

Wright, B. R. E., Wallace, M., Bailey, J., & Hyde, A. (2013). Religious affiliation and hiring discrimination in New England: A field experiment. *Research in Social Stratification and Mobility*, *34*, 111–126.

Yinger, J. (1991). Acts of discrimination: Evidence from the 1989 housing discrimination study. *Journal of Housing Economics*, *1*(4), 318-346.

Yinger, J. (1995). *Closed doors, opportunities lost: The continuing costs of housing discrimination*. Russell Sage Foundation.

Zhou, X., Zhang, J., & Song, X. (2013). Gender discrimination in hiring: Evidence from 19,130 resumes in China. *Available at SSRN 2195840*.

Zschirnt, E., & Ruedin, D. (2016). Ethnic discrimination in hiring decisions: a meta-analysis of correspondence tests 1990–2015. *Journal of Ethnic and Migration Studies*, *42*(7), 1115-1134.

Zussman, A. (2013). Ethnic discrimination: Lessons from the Israeli online market for used cars. *The Economic Journal*, *123*(572), F433-F468.

LAURA L. RODRIGUEZ ORTIZ
República Dominicana
Calle San Martin de Porres no. 5, Ens. Naco
Tel. 849-919-9585
laural.rodriguezortiz@gmail.com
August 2022

PERSONAL INFORMATION

Full Name Laura Lorena Rodriguez Ortiz

EDUCATION

SYRACUSE UNIVERSITY, Maxwell School of Citizenship and Public Affairs — Syracuse, NY
PhD Candidate in Public Administration and International Affairs. (*expected august 2022)*
Fields of specialization: social policy and public finance
Dissertation: "Sports Participation and Labor Market Discrimination: A Correspondence Study"

SYRACUSE UNIVERSITY, Maxwell School of Citizenship and Public Affairs — Syracuse, NY
M.A. Economics, 2013

SYRACUSE UNIVERSITY, Maxwell School of Citizenship and Public Affairs — Syracuse, NY
Master in Public Administration (MPA), 2012

PONTIFICIA UNIVERSIDAD CATOLICA MADRE Y MAESTRA — Sto. Dgo., Dom. Rep.
B.A. in Economics, summa cum laude

GRANTS, HONORS & AWARDS

Roscoe-Martin Graduate Award, Maxwell School, December 2016	$730
Maxwell School Doctoral Dissertation Fellowship, 2016-2017	$20,000
Summer Research Grant, Maxwell School, Summer 2016	$1,500
Summer Research Grant, Maxwell School, Summer 2015	$1,750
Summer Research Grant, Maxwell School, Summer 2014	$1,500
Fulbright Scholarship, 2011-2012	

EXPERIENCE

INNOVATIONS FOR POVERTY ACTION (IPA) — Santo Doming, DR

2022 – **Senior Research Manager – LAC**

Managing projects in Latin America and the Caribbean on climate change, financial inclusion, gender, teen pregnancy, refugee populations, social protection, and agriculture

2018 – 2022 **Research Manager, Dominican Republic**

Managing projects in the Dominican Republic on climate change, financial inclusion, and social protection.

2017 – 2018 **CENTER FOR POLICY RESEARCH, SYRACUSE UNIVERSITY** — Syracuse, NY
Research Assistant
Colleen M. Heflin

2013 – 2017 **Graduate Assistant**
Leonard M. Lopoo

Summer 2014 **SYRACUSE UNIVERSITY/PRINCETON UNIVERSITY** — Bogotá, Colombia
Research Assistant

2012 – 2013 **SAY YES TO EDUCATION/SYRACUSE UNIVERSITY** — Syracuse, NY
Research Assistant

2010 – 2011 **UNITED NATIONS DEVELOPMENT PROGRAMME (UNDP)** — Sto. Dgo., Dom. Rep.
Junior Economist

2008 – 2010	**UNITED NATIONS DEVELOPMENT PROGRAMME (UNDP)** **Administrative Assistant**	Sto. Dgo., Dom. Rep.

PUBLISHED WORK

Heflin, Colleen M., Claire E. Altman, and ***Laura L. Rodriguez.*** "Food insecurity and disability in the United States." *Disability and health journal* (2018).

OTHER EXPERIENCES

SYRACUSE UNIVERSITY

Teaching Assistant

Public Budgeting (PAI 734) – Master course	Summer 2017
Introduction to Statistics (PAI 721) – Master course	Fall 2016
Public Budgeting (PAI 734) – Master course	Summer 2016
Quantitative Analysis: Program Evaluation (PAI 722) – Master course	Spring 2016

Substitute Lecturer

Fundamentals of Policy Analysis (PAI 723) – Master course	Sept. 16, 2015

Invited Lecturer

Research Methods for Public Administration (PAI 803) – Doctoral course	Oct. 5, 2015, Oct. 10, 2016 & Oct. 4, 2017

FIRE DEPARTMENT NATIONAL DISTRICT 2022-
Cuerpo de Bomberos del Distrito Nacional
Volunteer Firefighter

SKILLS/ RESEARCH INTERESTS

Language skills: Spanish (native). English (fluent). German (advanced). French & Italian (basic)
Computer skills: STATA, SPSS, Qualtrics, Sawtooth, Python y Microsoft Office
Research Interests: Social policy, program evaluation, behavioral insights and experimental