

Discrimination in Low-Wage Labor Markets<sup>1</sup>

Devah Pager

Bruce Western

Bart Bonikowski

*Princeton University*

October, 2006

DRAFT:  
PLEASE DO NOT QUOTE OR CITE  
WITHOUT PERMISSION OF THE AUTHORS

---

<sup>1</sup> This research has been supported by grants from the National Science Foundation, the National Institute of Justice, the JEHT Foundation, the Princeton Research Institute on the Region and the Industrial Relations Section of Princeton University. We also gratefully acknowledge the support of the New York City Commission on Human Rights, and Commissioner Patricia Gatling. Direct all correspondence to Devah Pager, Department of Sociology, Princeton University, Princeton, NJ 08544, pager@princeton.edu

Despite a booming U.S. labor market through the late 1990s, racial differences in employment remain among the most intractable economic inequalities. While unemployment rates declined for all groups in the 1990s, young black men remained twice as likely to be unemployed relative to whites of their age. Further, racial inequality in measures of joblessness—which include those who have exited the formal labor market altogether—widened substantially over this period among young black and white men.<sup>2</sup> How can we explain these large and enduring inequalities in employment?

Historically, research on racial inequality in employment strongly emphasized the role of labor market discrimination. In *An American Dilemma*, Gunnar Myrdal argues that “the concentration of unemployment upon the Negro people is explainable only as the direct and indirect effects of discrimination” (p.998; see also Pettigrew, 1975; Allport, 1958). Continuing in this tradition, more recent research has investigated the preferences and attitudes of employers as possible sources of persistent racial disparities in the labor market. Relying on survey research and in-depth interviews, this line of work finds that firms are reluctant to hire young minority men—especially African Americans—because they are seen as unreliable, dishonest, or lacking in social or cognitive skills (Waldinger and Lichter 2003; Moss and Tilly 2000; Holzer 1999; Wilson 1996, chap. 5). Echoing earlier research on prejudice and discrimination, this work locates one important source of minority employment problems at the demand side of the labor market.

Recent research in both economics and sociology, however, offers a strong challenge to demand-side arguments. An influential paper by Derek Neal and William Johnson (1996), for example, uses data from the National Longitudinal Survey of Youth to estimate wage differentials between white, black, and Hispanic young men. After controlling for a measure of cognitive ability, these authors find that most or all of the racial wage gap among young male workers can be explained. This and other similar studies has reinforced the view that the vast majority of the employment problems of young minority men can be explained by skill or other individual deficiencies, rather than any direct effect of discrimination (Neal and Johnson 1996; Farkas and Vicknair 1996). Economist James Heckman puts the point most clearly in a 1998 paper: “...Most of the disparity in earnings between blacks and whites in the labor market of the 1990s is due to differences in skills they bring to the market, and not to discrimination within the labor market...” and he goes on to describe labor market discrimination “the problem of an earlier era” (Heckman, 1998:101-102).<sup>3</sup>

---

<sup>2</sup> Current Population Survey data show that in the early 1980s, only 14% of young white men with a high school diploma were not working compared to 25% of their black counterparts. By 2000, the jobless rate for young high-school educated white men had dropped below 10%, but joblessness among black men of the same age and education was around 22%. Racial inequality in joblessness had thus increased and employment rates for young noncollege blacks at the height of the economic boom in 2000 was little better than during the recession of the early 1980s.

<sup>3</sup> Note, Heckman’s first claim speaks to the *relative importance* of skill versus discrimination, which is not something we address here. Our goal is more modest in seeking to establish simply whether or not discrimination continues to represent a significant barrier to employment for minority workers, or whether it represents merely “the problem of an earlier era.”

Does employer discrimination continue to play an important role in shaping the labor market prospects of minority workers? Or are real skill differences between white and minority workers sufficient to explain both employers' perceptions of various groups and their differential employment outcomes? Answers to these questions remain elusive, in part because of notorious difficulties in measuring discrimination. Without observing actual hiring decisions, it is difficult to assess exactly how and when race shapes employer behavior. The present study addresses this question directly, using an innovative experimental field methodology that allows direct observation of employer hiring decisions. By presenting equally qualified applicants who differ only by race/ethnicity, it becomes possible to measure whether and to what extent racial considerations shape actual hiring decisions.

### **The Persistence of Race**

Economic models of discrimination predict that competitive markets will drive discriminatory employers out of business (Becker, 1967). Taste discrimination—the preference for one racial group over another—is economically costly because prejudice leads employers to over-pay for majority workers. Over time, then, discrimination based on racial animus or in-group preference should decline as economic forces pressure firms toward equal treatment.

Not all forms of discrimination, however, are as responsive to market pressures. Statistical discrimination—the attribution of group-level characteristics to an individual applicant—may persist under conditions of uncertainty and incomplete information. If accurate, group-level estimates of difficult-to-observe productivity characteristics can provide useful information in the screening of individual applicants (Aigner & Cain, 1977). In this case, while illegal, statistical discrimination may represent a rational and efficient strategy for employers. What happens when employers' estimates of group-level attributes are inaccurate or outdated? Perfect-market models assume that inefficiencies will be automatically eliminated over time, as relevant actors discover their practices to be suboptimal and correct for necessary modifications (Oettinger, 1996). Factors such as occupational segregation, imperfect information flows, and negative feedback effects, however, can impede awareness of changes and work to preserve existing outcomes (Tomaskovic-Devey & Skaggs, 1999; Whatley & Wright, 1994; Arrow, 1998; Loury, 2002).<sup>4</sup> These processes may lead employers to retain inaccurate beliefs far longer than optimal. Likewise, social psychological research suggests that stereotypes can display a stubborn resistance to change, as individuals unconsciously resist the integration of counterstereotypic information through biases in the gathering, processing, and recall of information (Fiske, 1998; Bodenhausen, 1988; Trope & Thomson, 1997; Fazio, 1986). Particularly in evaluating characteristics with some

---

<sup>4</sup> Labor economist Kenneth Arrow (1998) argues, “Each employer has a very limited range of experience, and so prior beliefs can remain relatively undisturbed. Indeed, to the extent that discrimination takes the form of segregation, then there will in fact be little experimentation to find out abilities.... The very fact of segregation will reinforce beliefs in racial differences” (p.97). Glenn Loury (2002) provides an elegant discussion of “the logic of self-confirming stereotypes” in which he articulates the negative feedback loop caused by statistical discrimination, whereby prior negative expectations lead to the emergence of real differences in job-relevant attributes, with the perceived link between the stigma (e.g., race) and productivity ultimately becoming realized (pp.26-33).

degree of ambiguity—such as the “soft skills” emphasized in many low-wage service jobs—stereotypes can filter information in ways that preserve expectations (Dovidio & Gaertner, 2000). Estimates of group-level characteristics, then, will depend on a range of influences which in some cases may delay or inhibit accurate updating.

### *Race Coding of Post-Industrial Skills*

Deindustrialization and the growth of service industries have changed the kinds of skills demanded by employers in urban labor markets. In the early 1970s more than one-third of non-college workers in urban areas were employed in heavy manual work in manufacturing, transport and utilities. By 2000, nearly three-quarters of those workers were in service sector jobs, about half in retail, hospitality, or personal services. While the old manual jobs demanded physical exertion in repetitious tasks, the new entry-level service sector jobs involve more communication with customers (Holzer 1996). In this case, employers are more likely to value interpersonal skills, and more intangibly, a positive attitude that projects a friendly and personable self. Moss and Tilly (2000), drawing on interviews with employers in four urban labor markets, thus argue for the importance of “soft skills” in contrast to the hard skills of numeracy, literacy, and other technical abilities. Soft skills consist of a capacity for effective communication, a “positive attitude,” and a high level of motivation. Although soft skills have been associated most closely by employers with service sector work, manufacturing sector employers also report a high value on interpersonal skills and motivation. A plurality of manufacturing sector employers interviewed by Waldinger and Lichter (2003) in Los Angeles reported that a “good attitude” was the most important quality sought in a job candidate.

Valued traits like attitude, motivation, and communication skills that have become associated with low-wage service sector work are often strongly race-coded by employers. Minority workers, especially young black men, are widely viewed by employers to be lacking these qualities compared to other low-skill workers. Wilson (1996, 116-117), for example, found that Chicago employers frequently doubted the communication skills of inner-city residents, expressing concern about their ability to write and speak in standard English. In addition to being singled out for poor language ability, black workers are often described as lacking motivation or work ethic. Kirschenman and Neckerman (1991, 227) report that among Chicago employers, “the widespread perception that black workers were unreliable or had a poor work ethic hurt them in the labor market.” Waldinger and Lichter (2003, 172-176) argue that employers’ perceptions of a poor work ethic among blacks also implied a “bad attitude” or a lack of deference. Because servility to employers and customers is desirable for service sector work, ascriptions of insubordination were damaging for low-skill black workers. Because these highly-valued traits of work ethic and attitude are measured more by subjective judgment than by objective tests or credentials, employers may be more likely to rely on stereotypes to assess the merit of job candidates (Moss and Tilly 2000). In short, the intangible skills that are highly valued in low-wage post-industrial work are vulnerable to race-coding, creating conditions for widespread discrimination.

Contemporary labor markets are characterized not just by the expansion of low-wage service work, but also by the increasing heterogeneity of the urban minority work force. Low-skill African American workers are now much more likely to compete with low-skill Latino workers. Employers' perceptions of insubordination disadvantages black workers in relation to Latinos. Waldinger and Lichter (2003) found that their sample of Los Angeles employers viewed Latino workers as more pliant and more reliable. The researchers related the perceived docility of Latino workers to immigration status, in which being an outsider to America society gave Latino workers fewer claims to equal treatment. Kirschenman and Neckerman (1991, 210) echo this idea, finding that Chicago employers favor Latino workers over blacks, but also favor non-native Mexicans over Puerto Rican-born U.S. citizens. In both field settings, in Los Angeles and Chicago, whites were viewed by employers as the most desirable workers, standing at the top of the racial hierarchy. Latinos occupied the middle ground, while blacks were the most disfavored, with young black men, last of all.

### *Race and Criminal Stigma*

Research on employer attitudes also indicates that black males are distinguished from other workers by their association with criminality. One of the most commonly endorsed stereotypes of African Americans relates to images of black violence or criminality (Eberhardt et al., 2004; Correll et al., 2002). Survey respondents rate blacks as more violent than any other racial or ethnic group, with the stereotype of aggressiveness and violence most frequently endorsed in ratings of African Americans (Sniderman and Piazza 1993, 45; see also Smith, 1991). The stereotype of blacks as criminals is deeply embedded in the collective consciousness of white Americans, irrespective of the perceiver's level of prejudice or personal beliefs (Devine and Elliot 1995; Graham and Lowery, 2004). In part, racial ascriptions of criminality may be related to a realistic understanding of the high levels of crime and incarceration among young black men. Pettit and Western (2004) describe the mass imprisonment of young black men in which nearly a third of those without college education will go to prison by their mid-thirties. The social fact of mass imprisonment is also dramatized in the popular culture, which tends to depict criminal episodes in a heavily racialized context. Media and political research shows that, even relative to their distribution among arrestees, blacks are disproportionately portrayed as criminal by local television news coverage. In political campaigns, voters' fears of street crime and violence have been stoked by images of young black men (Mendelberg 2001; Beckett 1997).

Racialized ascriptions of criminality extend beyond public opinion and popular culture into the workplace. Reporting on his Chicago employers, Wilson (1996) finds that black men are suspected of having criminal records, instilling fear in customers, of being "belligerent and dangerous," and of taking drugs. Other employer interviews suggest that black men are not perceived as literally criminal, but as excessively assertive. Thus Moss and Tilly (1999) report that employers see black workers as "having a chip on their shoulder." Waldinger and Lichter (2003) also report that Los Angeles employers view black workers as being more interpersonally aggressive than immigrant workers. More

than just unreliable or lazy, there is evidence that employers view young black men as threatening or confrontational.

Perceptions of the dangerousness of young black men indicate a prejudice in which the individualized stigma of criminality is projected on to young black men as a group. Pager's (2003) research in a Milwaukee audit study, for example, compares the magnitudes of racial and criminal stigma among pairs of job seekers who were randomly assigned a criminal record. Fielding teams of black and teams of white job applicants, she found that blacks without criminal records has roughly equal outcomes to whites with felony convictions. That the magnitude of the effect of being black is roughly equivalent to that of being a convicted felon underscores the significance of race in the eyes of Milwaukee employers. At the same time, the Milwaukee study sent blacks and whites to apply for different jobs so cross-racial comparisons may be confounded with employer effects which differ for the two groups. The comparisons of race and criminality thus remain tentative. We follow-up the Milwaukee study here by sending blacks and whites to apply for the same jobs, yielding a sharper test of the stigmatic equivalence of race and criminal background.

In sum, structural changes in urban economies that precipitated chronic unemployment among young black men also created conditions for renewed discrimination. In this account, the postindustrial economy created a demand for hard-to-observe skills that are often race-coded. In the current era of high incarceration rates, young black men, often stereotyped as lazy and unreliable, bear the added negative ascription of criminality. Together, these factors lead us to question the declining significance of race, and to initiate a contemporary investigation of the relevance of race in hiring decisions within the low-wage labor market.

### **Methods for Studying Labor Market Discrimination**

Racial discrimination in the labor market is typically studied by comparing the wages of whites and minorities. In this approach, the difference in wages is estimated while controlling for relevant human capital characteristics. Estimates from a variety of social surveys suggests that the black-white difference in hourly wages among men, controlling for schooling, work experience, and other indicators of human capital varies between about 10 and 20 percent (e.g., Cancio et al 1996; Neal and Johnson 1996; Darity and Meyers 1998). This residual method, in which discrimination is defined as the unexplained race difference in wages, is highly sensitive to the measurement of human capital. Where race differences in human capital are incompletely observed, the effect of discrimination may be over-estimated. Analyses of the effects of test scores thus argue that differences in cognitive ability accruing before the completion of schooling explains nearly all the black-white difference in wages (Neal and Johnson 1996; Farkas and Vicknair 1996). From the perspective of this research, standard studies over-estimate racial discrimination because cognitive differences between black and whites are inadequately controlled.<sup>5</sup>

---

<sup>5</sup> At the same time, misspecification of a statistical model can conversely lead to underestimation of the race effects. Subsequent analyses of the Neal & Johnson paper, for example, finds that adjusting for the years of schooling at the time the AFQT was taken leads the racial wage gap to re-enlarge (Carneiro et al.,

Residual estimates of discrimination infer employer behavior from data on workers' wages. Field experiments, by contrast, offer a more direct approach to the measurement of discrimination. This approach, referred to as an audit methodology, involves the use of matched pairs of job applicants—called testers—who apply to real job openings and record responses from employers. In studies of race discrimination, black and white testers are assigned equivalent resumes and are matched on a variety of characteristics like age, education, physical appearance, and interpersonal skills. Because black and white testers are sent to the same firms, and testers are matched on a wide variety of characteristics, much of the unexplained variation that confounds residual estimates of discrimination is experimentally controlled.

Audit studies conducted by the Urban Institute in the early 1990s provided evidence of racial preference among employers in several urban labor markets. In Chicago and Washington DC, white testers received more job offers than blacks. In Chicago, nearly 21 percent of white testers received job offers compared to 16 percent of successful black testers. In Washington, 36 percent of whites were successful compared to 22 percent of blacks. In a San Diego audit study, white testers were paired with Latinos, where they received job offers in 44 percent of applications, compared to 31 percent for Latinos. Only one of the Urban Institute studies, in Denver, did not reveal a clear preference for white job applicants over minorities (Heckman and Siegelman 1993 summarize the results).

Unfortunately, the use of the in-person audit studies remain rare, with the most recent studies of racial discrimination in employment conducted in the early 1990s. Moreover, the typical emphasis on a single comparison group leaves several significant features of contemporary urban labor markets unexplored. First, the racial and ethnic heterogeneity of contemporary labor markets in which ethnic minorities compete for jobs with whites and blacks has not been studied with a unified experimental design. In addition, under the novel conditions of mass imprisonment, no study has directly examined the hypothesis of the stigmatic equivalence of race and criminality.

Our study, conducted in 2004, extends earlier research in three ways. First, reflecting the racial heterogeneity of large urban labor markets, we record employer responses to matched black, white, and Latino job seekers. Second, to help calibrate the magnitude of racial preferences, we compare applicants affected by varying forms of stigma. Specifically, we compare minority applicants to white applicants just released from prison. Third, we draw from the extensive field notes taken by testers which describe their interactions with employers, providing qualitative evidence of the experience of racial discrimination in the contemporary American labor market.

### **Research Design and Methods**

The New York City Hiring Discrimination Study sent matched teams of testers to apply for 341 real entry-level jobs throughout New York City over nine months in 2004. The

---

2003). Likewise, the wage gap measured at later ages also shows much larger racial disparities (Tomaskovic-Devey et al., 2005).

testers were well-spoken, clean-shaven, young men, aged 22 to 26. Most were college-educated, between 5 feet 10 inches and 6 feet in height, recruited in and around New York City. They were matched on the basis of their verbal skills, interactional styles (level of eye-contact, demeanor, and verbosity), and physical attractiveness. The testers passed through a common training program to ensure uniform behavior in job interviews. While in the field, the testers dressed similarly and communicated with team-mates by cell phone to anticipate unusual interview situations. Testers were assigned fictitious resumes indicating identical educational attainment, and comparable quality of high school, work experience (quantity and kind), and neighborhood of residence. Testers presented themselves as high school graduates with steady work experience in entry-level jobs.

To study employer's treatment of whites compared to blacks and Latinos, we fielded two teams of three testers. The first team investigates racial queues within low-wage labor markets. In this team, a black, Latino, and white tester each applied for the same jobs. To help ensure comparability, the Latino testers spoke in unaccented English, were U.S. citizens of Puerto Rican descent, and like the other testers, claimed no Spanish language ability. This first team tests a standard racial hierarchy, with the white tester serving as a benchmark against which to measure variation in racial and ethnic discrimination. To test the stigmatic equivalence of race and criminality, the second team compares black and Latino testers to a white tester with a criminal record. The criminal record was typically disclosed in answer to the standard question on employment applications, "Have you ever been convicted of a crime? If yes, please explain." Testers were instructed to reveal that they had recently been released from prison after serving 18 months for a drug felony (possession with intent to distribute, cocaine). In addition, following Pager (2003), the white tester's criminal record was additionally signaled on the resume by listing work experience at a state prison, and by listing a parole officer as a reference.

For both teams, employers were sampled from job listings for entry-level positions, defined as jobs requiring no previous experience and no education greater than high school. Job listings were randomly drawn each week from the classified sections of *The New York Times*, *The Daily News*, *The New York Post*, and *The Village Voice*. The broad range of job listings allows for comprehensive coverage of the entry-level labor market in New York. From the available population of job listings, we took a simple random sample of advertisements each week. Testers in each team applied to each job within a 24-hour period, randomly varying the order of the applicants.

Our dependent variable recorded any positive response in which a tester was either offered a job or called back for a second interview. Callbacks were recorded by voicemail boxes set up for each tester. For employer  $i$  ( $i=1, \dots, N$ ) and tester  $t$  ( $t=W, B, \text{ or } L$  for white, blacks and Latinos), a positive response,  $y_{it}$ , is a binary variable that scores 1 for a job offer or callback, and 0 otherwise. Similar to Heckman and Seigelman's (1993) "net" measure of discrimination, we define the average level of differential treatment as the difference in mean positive response rates, say  $d_{WB} = \bar{y}_W - \bar{y}_B$ . Under the null hypothesis of equal treatment,  $d_{WB}=0$ , the number of positive responses received by each racial group is equal. For data on matched pairs, several statistical tests have been proposed that use



within-pair comparisons to account for the correlation of observations from the same pair (e.g., Heckman and Seigelman 1993; Agresti 1990). In our case, where three testers are sent to the same employer, we have a matched triplet and information from all three testers should ideally contribute to an inference about a contrast between any two. Ghosh, Chen, Ghosh and Agresti (2000) suggest that matched pairs can be fit with a hierarchical logistic regression with a random effect for each pair.

We generalize their approach to our matched triplets, fitting a random effect for each employer. If the probability of a positive response is given by  $E(y_{it})=p_{it}$ , the hierarchical model is written,

$$\text{Log} \left( \frac{p_{it}}{1-p_{it}} \right) = \alpha_i + \beta B_{it} + \gamma L_{it},$$

where  $B_{it}$  is a dummy variable for blacks,  $L_{it}$  is a dummy variable for Latinos, and the random effects for employers,  $\alpha_i$ , is given a normal distribution. The employer effects,  $\alpha_i$ , induce a correlation among observations from the same employers and reduces standard errors as in the usual matched-pair inference. The models are estimated with Markov Chain Monte Carlo methods. Intervals for the mean differences ( $d_{WB}$ ,  $d_{WL}$ , and  $d_{BL}$ ) are constructed by taking random draws from the posterior predictive distribution of  $y_{it}$ . Alternative methods that adjust for clustering by employer yield similar results to those reported below.

### *The Problems of Matching*

The quality of the audit results depends strongly on the comparability of the testers. Because race cannot be experimentally assigned, researchers must rely on effective selection and matching to construct audit pairs that represent comparable candidates with respect to all relevant characteristics—something that, according to critics, leaves substantial room for bias. Heckman and Seigelman (1993), for example, have argued that researchers know little about the hard-to-observe characteristics highly prized by employers.<sup>6</sup> If testers are poorly matched, evidence of discrimination may be merely an artifact of idiosyncratic tester characteristics.

Bertrand and Mullainathan (2005) remove tester effects in a “correspondence test” which sent resumes to employers with common white and African American names. Their design allows the random assignment of resume characteristics to white- and black-sounding names, largely removing concerns about unobserved characteristics. Sending matched resumes in Boston and Chicago, Bertrand, and Mullainathan (2005) find that those with white names received callbacks from employers 9.7 percent of the time, a significant difference from the 6.5 percent callback rate for black names.

---

<sup>6</sup> Although the work of Holzer (1996), Wilson (1996), Moss and Tilly (2000), and Waldinger and Lichter (2003), reviewed above, have since supplied additional information. Further, as an employer him/herself, the researcher must identify subtle cues about applicants that indicate their ability to perform. Whether or not these cues are explicit, conscious, or measurable, they are present in a researcher’s evaluation of tester candidates as they are for employers’ evaluations of entry-level job applicants. Like employers, researchers are affected by both objective and subjective/subconscious indicators of applicant quality in their selection and matching of testers in ways that should ultimately improve the nuanced calibration of test partners.

Because we prefer the use of in-person audits for our study of low-wage labor markets, the effective matching of testers becomes a key concern.<sup>7</sup> Instead of the usual summertime recruitment of college students, we matched 10 testers from over three hundred applicants from all over greater New York.<sup>8</sup> Successful applicants were subject to two lengthy screening interviews and a written test, a far more probing job selection process than the testers encountered in their fieldwork. Each tester passed through a standard training period, was required to dress uniformly, and was subject to periodic spot checks for quality control.

Despite these measures, uncontrolled tester effects remain a threat to inferences about discrimination. We assess the sensitivity of our results to testers in two ways. First, each tester may have a unique effect, but the average effect of the testers may be zero. In this case, the observations from each tester will be correlated and standard errors that ignore this clustering will tend to be too small. We allow for this possibility by fitting an additional random effect for each tester in our hierarchical logistic regression. Second, each tester may have a unique effect, but these effects may not average to zero. To assess the sensitivity of our results to each tester we perform a type of cross-validation in which the treatment effect is recalculated for a reduced data set, sequentially omitting those employers associated with each individual tester. Confidence intervals below are based on models including employer and tester random effects. These results are compared to cross-validation treatment effects based on subsets of the data in which individual testers are sequentially omitted.

Where critics of the audit methodology have often worried about imperfect matching of testers, the achievement of perfect matches may itself produce distortions in the hiring process. Because audit partners are matched on all characteristics that are most directly relevant to the hiring process (education, work experience, physical appearance, etc.), employers may be forced to privilege relatively minor characteristics simply out of necessity of breaking the tie. “By taking out the common components that are most easily measured, differences in hiring rates as monitored by audits arise from the idiosyncratic factors, and not the main factors, that drive actual labor markets” (Heckman, 1998: 111). If employers care only marginally about race, but are confronted with applicants equal on all other dimensions, this single characteristic may take on greater significance in that particular hiring decision than is true under normal circumstances, when evaluating real applicants who differ according to multiple dimensions.

---

<sup>7</sup> In-person audits allow for the inclusion of a wide range of entry-level job types (which often require in-person applications); they provide a clear method for signaling race (through the physical presentation of job applicants), without concerns over the class-connotations of racially distinctive names (e.g., Fryer & Levitt, 2004); and they provide the opportunity to gather both quantitative and qualitative data, with information on whether or not the applicant receives the job as well as how he is treated during the interview process.

<sup>8</sup> Note: These 300 applicants had already been pre-screened for appropriate age, race, gender, and level of education.

Fortunately, the design of our study, which focuses on the early stages of the hiring process, avoids situations in which employers must choose only a single applicant. By using “callbacks” as one of our key dependent variables, we include cases which represent an employer’s first pass at applicant screening. Indeed, employers typically interview an average of eight applicants for each entry-level job they fill (see Pager, 2003). If race represents only a minor concern for employers, we would expect all members of our audit team to make it through the first cut. To the extent that race figures prominently even in these early rounds of review, we can infer that this characteristic has been invoked as more than a mere tie-breaker. In these cases, the evidence of race-based decision-making is quite strong.

### **Experimental Results**

The primary results from the audit study focus on the proportion of applications submitted by testers which elicited either a callback or job offer from employers, by race of the applicant. Our first team assesses the effects of race discrimination by comparing a black and Latino tester to a white tester. Positive response rates for each race-ethnicity group are reported in Figure 1. In applications to 171 employers, the white tester received a job offer or callback 31.0 percent of the time, compared to a positive response rate of 25.2 percent for Latinos and 15.2 percent for blacks.<sup>9</sup> Once we adjust for employer and tester effects, the confidence interval for the 5.9 point difference between whites and Latinos includes zero. However, the black-white difference in callback rates, 15.8 points, is substantively large and statistically significant. We can thus be highly confident that employers, on average, preferred white job seekers to equally qualified blacks. The positive response rate for blacks is also significantly lower than the rate for Latinos. These results thus yield clear evidence for a racial hierarchy in which white and Latino job applicants are significantly preferred by New York employers relative to equally qualified blacks. The results suggest that a black applicant would have to search twice as long for employment for as an equally qualified white applicant before receiving a positive response from an employer.

---

<sup>9</sup> Callbacks represent 36% of all positive responses (with the remainder job offers). Evidence of differential treatment among callbacks is strong relative to the overall rates of positive responses (callback rates for white, Latino, and black testers were 13, 10, 3%).

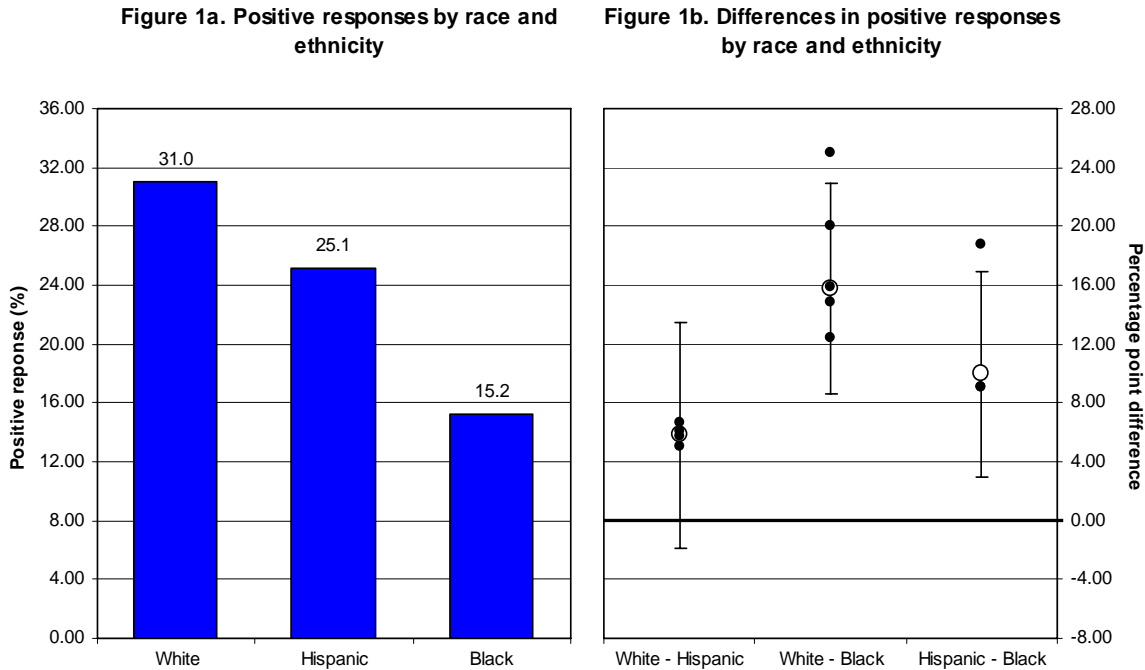


Figure 1. Left panel: Percentage of positive responses by race/ethnicity. Right panel: race/ethnic contrasts, error bars show a 95 percent confidence interval; points indicate cross-validation contrasts obtained by dropping tests associated with each tester.

The second panel of Figure 1 shows the contrasts between the three race groups. The points on the figure show the cross-validation results obtained by dropping employers associated with each tester. In each case, all contrasts remain consistently signed, indicating that employers treat blacks less positively regardless of which testers are applying for jobs.

To calibrate the effects of race against another stigmatized category, the ex-offender, we repeated the audit study this time assigning a criminal record to the white tester. Figure 2 shows the percentage of positive responses--job offers or callbacks--received by each tester. In this experiment, the preferential treatment of white applicants is largely eliminated. Whites with criminal records obtained positive responses in 17.1 percent of job applications, compared to 15.9 for Latinos, and 12.9 percent for blacks.<sup>10</sup>

<sup>10</sup> Note: Positive rates for all team members are lower in this experiment due to the staggered timing of the fieldwork by team. Callbacks represent 56% of positive responses for this team. Evidence of differential treatment among callbacks is strong relative to the overall rates of positive responses (callback rates for white, Latino, and black testers were 11, 9, 5%).

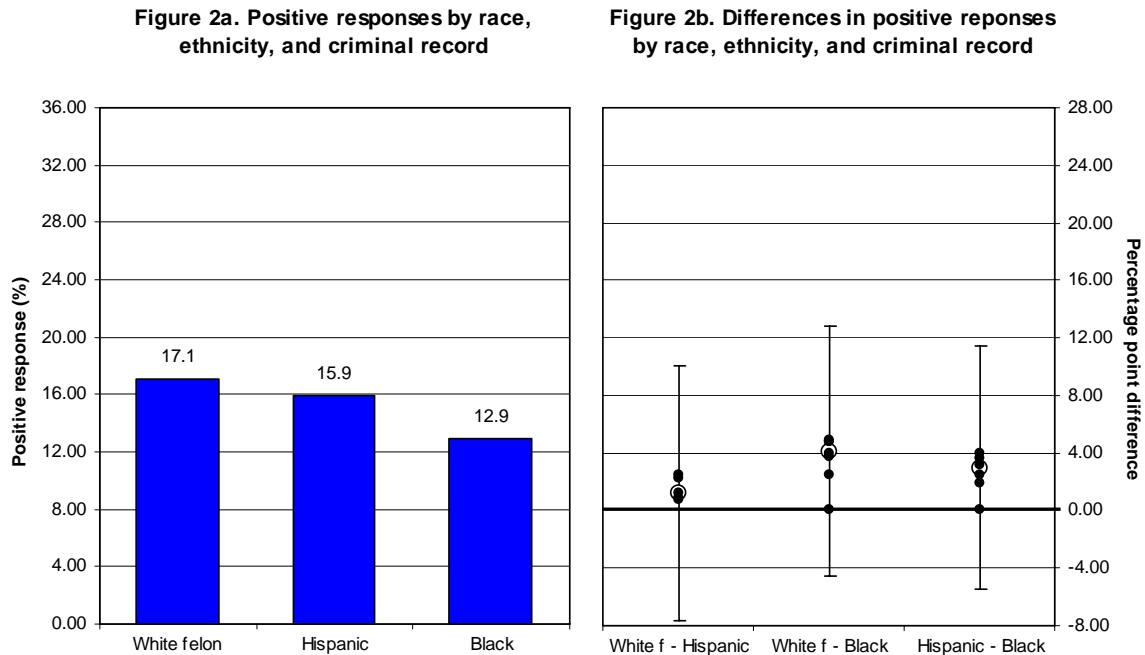


Figure 2. Left panel: Percentage of positive responses by race/ethnicity. The white tester is assigned a criminal record. Right panel: race/ethnic contrasts, error bars show a 95 percent confidence interval; points indicate cross-validation contrasts obtained by dropping tests associated with each tester.

The right hand panel of Figure 2 shows that the white-Latino difference is now close to zero and the confidence interval overlaps zero by a large margin. The white-black difference is now a statistically insignificant 4 percentage points, compared to a significant margin of nearly 16 points when the white tester had a clean record. As in the previous experiment, Latinos were preferred to blacks, but in this case the difference is not significant. As in the first experiment, none of these results are sensitive to the observations from any one tester. The cross-validation treatment effects, obtained by dropping employers associated with one particular tester, are all close to zero. These results indicate that regardless of which testers were sent into the field, the results were largely unchanged. Overall, the results suggest that New York employers view minority applicants as essentially equivalent to whites just out of prison.

### *Experimenter Effects*

An important critique of the audit methodology raises the problem of experimenter effects, or the possibility that the expectations or behaviors of testers can influence the audit results in nonrandom ways. For example, if a tester expects to be treated poorly by employers, he may appear more withdrawn, nervous, or defensive in interactions. The nature of the interaction may then create a self-fulfilling prophecy, in which the tester experiences poor outcomes, but for reasons unrelated to his race. Indeed, the possibility of experimenter effects represents one of the most serious threats to the validity of the audit experiment. As an indirect check on the problem of experimenter effects, we can compare those tests conducted with little or no in-person contact to those in which the

testers had extended interaction with the employer. If testers were acting in ways that fulfill their expectations of discrimination, we would expect outcomes for those tests conducted with interaction to show greater evidence of differential treatment than those without. If the results are consistent, or show weaker evidence of differential treatment, we can be more confident that experimenter effects are not driving the results.

The results of this test indicate that personal contact served to weaken the effect of race on hiring decisions (see Table 1).<sup>11</sup> When applicants has little chance to interact with the employer, for example, whites were 9.6 times more likely to receive a callback or job offer than their black partners (and 1.8 times more likely than their Latino partners). By contrast, when applicants had the opportunity to interact with employers, whites were just 1.9 times more likely to receive a callback or job offer than blacks, and equal to the outcomes of Latinos. Rather than enacting expectations of discrimination, then, interaction between testers and employers resulted in a substantial reduction in discrimination. The friendly, appealing qualities of the testers appear to mediate the effects of racial stereotypes, reducing the negative bias evident in more superficial reviews.

**Table 1. Percentage of positive responses and race differences, by level of personal contact**

Subsample	White	Latino	Black	Race Differences		
	(W)	(L)	(B)	W/L	W/B	L/B
Total	31.0	25.2	15.2	1.2	2.0	1.7
No personal contact	14.4	8.0	1.5	1.8	9.6	5.3
Personal contact	44.2	42.9	23.8	1.0	1.9	1.8

Subsample	White felon	Latino	Black	Race Differences		
	(Wf)	(L)	(B)	Wf/L	Wf/B	L/B
Total	17.1	15.9	12.9	1.1	1.3	1.2
No personal contact	9.4	10.6	3.4	0.9	2.8	3.1
Personal contact	27.0	22.4	34.0	1.2	0.8	0.7

Note: Personal contact varies across testers within teams. Tests involving personal contact represent 56% by white testers, 49% by Latino testers, and 61% by black testers in the first team (N=171); 44% of white testers, 45% of Latino testers, and 31% of black testers in the second team (N=170).

<sup>11</sup> Note: Because of the vastly different baseline positive response rates across groups, we calculate ratios as an indicator of disparate treatment rather than differences.

### **Qualitative Evidence of Discrimination**

Quantitative evidence from the audit experiment indicates the degree of racial preference among the sampled employers. The in-person design of the experiment allows us to further supplement these findings with qualitative evidence of differential treatment provided by the testers' field notes on their interactions in the job interviews. Research on employer preferences shows that that employers ascribe superior skills—soft and hard—to white workers. In job interviews described here, employers often discussed skills with each of the testers, but found the lack of skill to be disqualifying most for the minority job seekers. In these cases, employers appeared to filter information about applicant skills through the lens of race.

In the majority of job applications, contact between testers and employers was limited. Even where testers reported a formal interview or brief exchange, interactions that betrayed a clear racial preference were rare. Nevertheless, a number of cases reveal subtle ways in which the “objective” qualifications of testers are re-interpreted through the lens of race. In one case, Joe, an African American tester, was not allowed to apply for a sales position as a result of his lacking direct experience. He reports: “When [the employer] called me she handed me back my resume and told me they didn’t have any positions to offer me...She said... I needed a couple years of experience.” The employer voices similar concerns with Kevin, Joe’s white partner. Kevin writes: “[The employer] looked at my resume and said, ‘There is absolutely nothing here that qualifies you for this position.’” Despite the employer’s concerns, Kevin was then offered the sales job and asked to come back the next morning.

In another case, Josue, a Latino tester, applied for a job as a line cook at a mid-level Manhattan restaurant. He reports: “[The employer] then asked me if I had any prior kitchen or cooking experience. I told him that I did not really have any, but that I worked alongside cooks at [my prior job as a server]. He then asked me if I had any ‘knife’ experience and I told him no... He told me he would give me a try and wanted to know if I was available this coming Sunday at 2 p.m.” Simon, his white test partner, was also invited to come back for a trial period. By contrast, Joe, the black tester found, “they are only looking for experienced line cooks.” Joe continued, “I started to try and convince him to give me a chance but he cut me off and said I didn’t qualify.” Though none of the testers had direct experience with kitchen work, the white and Latino applicants were viewed as viable prospects, while the black applicant was rejected because he lacked experience.

In other cases, real skill or experience differences were perceived among applicants despite the fact that the testers’ resumes were designed to convey identical qualifications.<sup>12</sup> In one case, for example, the three testers applied for a job at a moving company. Joe, the African American applicant, spoke with the employer about his prior experience at a delivery company. Nevertheless, “[the employer] told me that he couldn’t use me because he is looking for someone with moving experience.” Josue, his Latino partner, presented his experience as a stocker at a delivery company and reports a similar

---

<sup>12</sup> Note: resume types were randomly varied across testers, with each resume used by testers from different race groups

reaction: “He then told me that since I have no experience... there is nothing he could do for me.” Simon, their white test partner, presented his identical qualifications to which the employer responds more favorably: “‘To be honest, we’re looking for someone with specific moving experience. But because you’ve worked for [a storage company], that has a little to do with moving.’ He wanted me to come in tomorrow between 10 and 11 for an interview.” The employer is consistent in his preference for workers with relevant prior experience, but he is willing to apply a more flexible, inclusive standard in evaluating the experience of the white applicant than in the case of the minority applicants. The shifting standards used by employers, offering more latitude to marginally skilled white applicants than similar qualified minorities, suggests that even the evaluation of “objective” information can be affected by underlying racial considerations.

These experiences are consistent with social psychological research on status characteristics, which demonstrates the process by which higher standards of ability are applied to low-status than high-status group members. According to Biernat and Kobrynowicz (1997), “a (high-status) man has to do less than a (low-status) woman to prove his ability, and he is allowed more latitude (more demonstrations of low ability) than a woman before lack of ability is inferred” (p.545-46; see also Yarkin et al., 1982). Similarly, our testers confronted signs of a double standard, where marginal work experience disqualifies low-status (minority) applicants but not the high status (white) applicant (see also Dovidio & Gaertner, 2000).

Though the issue of skills and experience did come up frequently in conversations with job applicants, many of these low-wage jobs required little in the way of concrete qualifications. As in other research on employers in low-wage labor markets, we found that employers often emphasized personality factors above objective skills or experience, expressing the desire for someone that could “fit in.” Particularly in sales or customer service positions, employers wanted someone who could represent the corporate image or fit in with the customers and other employees. On one occasion, a black tester, Joe, was told by a gallery owner that he was “looking for somebody that ‘spoke his language,’ in other words, someone that fit into the culture of the store and the position.” Joe reports the conclusion of the visit: “He said to come tomorrow at 9am. He said that this is only a trial period to see if I am ‘the right fit.’” Soon after Joe’s visit, his white test partner, Kevin, spoke with the same employer: “He read over my resume and asked if I am ready to start. I said ‘right now?’ He said, ‘Yes.’ He said he needed someone right now because they were having an exhibit Saturday. He said they will hire anyone that walks in.” As Kevin was leaving, he reports hearing the interviewer comment to another (white) employee, “I like him. He makes a good impression.”

In cases like this, it appears that employers are actively constructing white testers as more appropriate even though their resumes and styles of self-presentation are similar to those of minority testers. Though the employer in this case tells Kevin that his need for workers is so great that he will hire anyone who walks in the door, his response to the black applicant who arrived earlier does not convey the same urgency. Joe is also offered the job, but it is on a provisional basis, and the employer has doubts he will fit in. Kevin, on



other hand, has no trouble convincing the employer that he would make a good fit. In a job application without a long or probing interview, that involved little more than the presentation of a resume, the job criteria of “fitting in” seems like an oblique expression of racial preference.

In addition to employers’ concerns about “fit” for black applicants, a few interactions revealed employers’ racial stereotypes. In one case, for example, three testers were interviewed for a marketing position that involved contact with teenagers and young adults. According to Simon, a white tester, the employer’s “only question was ‘Are you a friendly person?’” Josue, a Latino tester, was asked a similar question, though the employer also warned that “no one can be late more than two times or they will have to be fired.” When the employer interviewed Joe, their black test partner, “He asked if I thought I would be able to approach 14 to 18 year old girls without intimidating them.” Concerns about discipline, work ethic, and potentially threatening demeanor emerged here in interviews with minority testers, while the white applicant was simply questioned about his level of sociability.

Testers’ reports of their interactions with employers offer clues about hiring decisions in entry-level labor markets. A pattern in these interactions, when compared side by side, is the use of double standards--seeking higher qualifications from blacks than non-blacks, or viewing whites as more qualified than minorities presenting equivalent resumes. Recent research emphasizes employers’ use of statistical discrimination to estimate difficult-to-observe productivity characteristics (Moss and Tilly 1999; Waldinger and Lichter 2003). Where we have detailed field notes on job interviews, the interactions we observe suggest that employers also use race in interpreting and weighting observable skill characteristics.<sup>13</sup> Standards appeared to shift as employers evaluated the qualifications of various applicants differently depending on their race or ethnicity.

### *Race-Coded Job Channeling*

The testers’ narratives, reporting their experiences at the conclusion of each audit, provide vivid illustrations of the kinds of the channeling that takes place. For example, our fieldnotes record the experience of Josue, one of our Hispanic testers, in an audit of a retail clothing company: *Josue describes the various young white 20-something women running the place. One of the women interviews him and asks about past work experience. She asks him what job he’s applying for—“I told her ‘sales associate.’”* [Note: the last serious job listed on Josue’s resume was as a sales assistant at a sporting goods store]. *She then told me that there was a stock position and asked if I would be interested in that.”* Josue ended up getting the stocker job, and was asked to start the next day.

In another case, Zuri, a black testers, applied for a sales position at a lighting store. A sign on the glass in front of the store indicated, “Salesperson Wanted.” Zuri describes the

---

<sup>13</sup> The different weighting of applicant skills or experience is also consistent with the view that employers view information about black applicants as less reliable or precise than information about white applicants (Aigner & Cain, 1977).

following interaction: *When she asked what position I was looking for I said I was open, but that since they were looking for a salesperson I would be interested in that. She smiled, put her head in her hand and her elbow on the table and said, "I need a stock boy. Can you do stock boy?"* Zuri's white and Hispanic test partners, by contrast, were each able to apply for the advertised sales position.

In many cases, these instances of channeling are coded as "positive responses" in the initial analyses. Indeed, our key concern is about access to employment of any kind. But this general focus masks some of the racial biases at play. Indeed, the experience of channeling was not limited to a handful of cases. A more systematic analysis of the testers' experiences provides support for these anecdotal experiences. We coded all instances of job channeling and, across both our teams—when the white had a clean record, and when he presented a criminal record—we counted 53 cases compared to 172 positive responses. By comparing the original job title to the suggested job type, these 53 cases were then categorized as downward channeling, upward channeling, lateral channeling, or unknown. Downward channeling is defined as (1) a move from a job involving contact with customers to a job without, say from server to busboy; (2) a move from a white collar position to a manual position, say from dispatcher to driver; or (3) a move in which hierarchy is clear, say from manager to server. Upward channeling is defined as a move in the opposite direction. We focus on these two types of channeling for our current analysis. After eliminating all cases in which a team was similarly channeled, we have 23 additional cases of differential treatment unrecorded by our initial measurement of job offers and callbacks (Table 2).

Like hiring criteria, job placement is also patterned by race. Black applicants were channeled into lower positions in 9 cases, Latinos were channeled down in 5 cases, whereas whites experienced downward channeling in only 1 case. Many of these cases were restaurant jobs in which the tester applied for a position as server but was steered to a job as busboy or dishwasher. Almost all were cases in which the original position required extensive customer contact while the suggested position did not (e.g., salesperson to stocker). Sometimes, testers were guided into lower positions because their resumes indicated limited work experience, but race differences in channeling suggests insufficient work experience was more penalizing for minorities than whites.

**Table 2. Job Channeling by Race**

<b>original job title</b>	<b>suggested job</b>
<b>Blacks channeled down</b>	
Server	Busser (324)
Counter person	Dishwasher/porter (102)
Server	Busboy (189)
Assistant manager	Entry fast food position (258)
Server	Busboy/runner (269)
Retail sales	Maintenance (399)
Counter person	Delivery (176)
Sales	Stockboy (831)
Sales	Not specified <sup>(a)</sup>
<b>Hispanics channeled down</b>	
Server	Runner (199)
Sales	Stock (2)
Steam cleaning	Exterminator (79)
Counter person	Delivery (176)
Sales	Stock person (503)
<b>Whites channeled down</b>	
Server	Busboy (192)
<b>Hispanics channeled up</b>	
Carwash attendant	Manager (1058)
Warehouse worker	Computer/office (1001)
<b>Whites channeled up</b>	
Line Cook	Waistaff (254)
Mover	Office / Telesales (784)
Dishwasher	Waistaff (858)
Driver	Auto detailing (948)
Kitchen job	“Front of the house” job (5)
Receptionist	Company supervisor (347)

(a) employer told tester “sales might not be right for you...”

Note: numbers in parentheses refer to employer ID codes

Indeed, whites were more often channeled up than down. In at least six cases, white testers were encouraged to apply for jobs that were of a higher-level or required more customer contact than the initial position they inquired about. In one case, the white tester was even encouraged to apply for a supervisory position, despite limited work experience. The white tester, Kevin, reports, “[The employer] then asked me if I had any experience in construction. I told him I did not. He asked if I would be okay working with people that have thick accents like his. I told him that was fine. He then told me that he wanted me to be his new company supervisor.”

The data on channeling offer two kinds of evidence. First, as in the testers' field notes, there is evidence that employers apply more stringent hiring criteria to minority workers. Second, minorities are disproportionately channeled out of customer service positions. Employers' reluctance to place young minority men in customer service positions is consistent with other research in which employers view minority applicants as lacking communication skills or as otherwise discomfiting for customers. In addition to whether or not the tester gets the job then, the type of job also reveals a racialization of employment decisions.

### **Discussion**

Sending trained testers with equivalent resumes to apply for entry-level jobs revealed a strong racial preference among employers in New York City. We find evidence that New York employers were about twice as likely to prefer whites over blacks. These results are robust to a variety of tester and experimental effects. The testers' field notes also showed that New York employers did not obviously view black job applicants as less skilled, as other researchers have found. Instead, New York employers tended to rule out black applicants by holding them to a higher standard. New York employers were also more likely to steer black workers into lower jobs, and out of jobs involving customer service.

We also found that employers were more open to Latinos than blacks. Latinos received fewer callbacks and job offers than whites with clean records, though the margin was not statistically significant. In both our teams, the positive response rate for Latino applicants was about 50 percent higher than for blacks and in one case the difference was statistically significant. The generality of this result certainly deserves more study. The Puerto Ricans of New York that our Latino testers represented are a longstanding community of U.S. citizens. In other local labor markets where markers of citizenship and accent are more prominent sources of difference, evidence of ethnic discrimination may well be stronger. Still, the New York results offer clear evidence of a racial hierarchy in which whites and Latinos are at the top, and blacks follow some distance behind.

The magnitude of the penalty of blackness is underscored in its comparison with the penalty of a criminal record. A black applicant with a clean record fares no better than a white applicant recently released from prison. These results confirm Pager's (2003) finding in Milwaukee, though the inference is stronger here where whites with criminal records applied for the same jobs as blacks with clean records. A criminal record is indeed a significant barrier to employment; but the stigma of race poses a barrier equally as large.

While we find robust evidence of a large racial preference, we should be careful not to interpret these results as showing the level of discrimination actively experienced by minority job seekers in the New York labor market. If black workers can identify and apply to non-discriminating employers, actual experiences of discrimination will be limited (Becker 1970). Likewise, our sampling design, based on employers not workers, over-represents small firms relative to their share of employment. Our sample thus

includes many restaurants and independent retailers, for whom hiring is less bureaucratic, and whom lack the human resource departments that manage the equal employment opportunity obligations of large firms (Dobbin et al. 1993). Still, our sampled employers are offering the kinds of low-skill service work that dominates low-wage urban labor markets. The audited employers also use newspaper advertising to find job applicants, indicating an openness to a diverse range of job applicants relative to those who rely on job referrals from their own employees and social contacts (Petersen et al. 2000; Elliott 2000). In sum, we find significant evidence of labor market discrimination against young African American men, and this is a key precondition for the enduring influence of discrimination. However, our research design does not allow us to estimate the contribution of discrimination to observed racial inequalities, or to compare the effects of discrimination to other causes of inequality.

The significant evidence of discrimination found in this study contrasts sharply with recent research showing that racial inequality in wages is largely explained by differences in cognitive skill (Farkas and Vicknair 1996; Neal and Johnson 1996). How might these findings be reconciled? Labor market discrimination can be thought of as having both direct and indirect effects. Direct effects represent the kind of differential treatment observed by the audit study, where one applicant is treated differently than another on the basis of race. But the presence of discriminatory employers can have consequences even beyond those applicants are directly affected. First, as noted above, the presence of discrimination in the labor market may lead workers to differentially sort across employers, such that minority job seekers queue for jobs offered by employers less likely to discriminate. These dynamics would lead to longer search or wait times for minority job seekers, even if not reflected in ultimate wage offers. Second, the experience of discrimination may also add to the psychic costs of the job search process. If discouragement and poor treatment in the job search screen out all but the most motivated and most able black workers, black wage earners would represent an increasingly select group. We would thus expect to see both a large employment gap and a relatively small wage gap between blacks and whites. This is, indeed, what we observe. In this scenario, discrimination has not been eliminated in the post-civil rights period as some contend, but remains a vital component of a complex pattern of racial inequality.

## **Appendix A. Robustness Checks**

We examine the robustness of our primary results by examining racial and ethnic contrasts for different subsets of the data (Table 1A). To account for learning or adaptation by the testers we estimate effects for the first and second halves of the experimental period. In each period, whites and Latinos receive significantly more positive responses than blacks, and whites receive slightly more positive responses than Latinos. To examine whether our results depend strongly on any particular area within New York, we separate the experimental effects by location. Over half the audited employers were located in Manhattan. The pattern of black disadvantage was found throughout Manhattan and in the outer boroughs. Finally, we studied whether the tester first sent to an employer was more likely to be successful. The order in which testers were sent was randomized, and the experimental effects are similar regardless of which tester interviewed first. In short, these results indicate a large racial preference among New York employers for white job applicants over black, and smaller preference for whites over Latinos. We also find that Latino job seekers are preferred significantly more than blacks. All these results are robust to tester effects, experimental effects, and appear to be roughly uniform across New York City.

Table 1A. Percentage of positive responses and race differences, by date, employer address, and race of first tester, New York Hiring Discrimination Study.

Subsample (N)	White (W)	Latino (L)	Black (B)	Race Differences (SE)		
				W-L	W-B	L-B
Total (171)	31.0	25.2	15.2	5.9 (3.0)	15.8 (2.8)	9.9 (3.0)
<i>Date</i>						
Feb 23 - Apr 7 (84)	29.8	23.8	9.5	6.0 (4.6)	20.2 (4.2)	14.3 (4.6)
Apr 8 - Jul 16 (84)	33.3	27.4	21.4	6.0 (3.9)	11.9 (3.9)	6.0 (3.9)
<i>Location<sup>a</sup></i>						
Below 34th St. (56)	23.2	21.4	12.5	1.8 (5.3)	10.7 (5.2)	8.9 (5.3)
34-72nd St. (46)	30.4	21.7	17.4	8.7 (5.3)	13.0 (5.2)	4.4 (5.3)
Above 72nd St. (18)	33.3	22.2	5.6	11.1 (7.4)	27.8 (8.8)	16.7 (7.4)
Other (50)	40.0	34.0	20.0	6.0 (5.9)	20.0 (4.9)	14.0 (5.9)
<i>Race of first tester</i>						
White (69)	27.5	23.2	10.1	4.4 (4.4)	17.4 (4.4)	13.0 (4.3)
Black (45)	40.0	31.1	20.0	8.9 (6.8)	20.0 (5.6)	11.1 (6.8)
Latino (54)	29.6	24.1	18.5	5.6 (4.9)	11.1 (4.9)	5.6 (4.9)

<sup>a</sup> Street addresses are for Manhattan

Note: Standard errors are calculated with the bootstrap. Observations in subgroups may not sum to total number of observations due to missing data.

These robustness checks also support the finding that employers did not distinguish strongly between whites with criminal records and minority job seekers (Table 2). When the audit study is divided into two periods from March into mid-April, and from mid-April to early August, we find that treatment effects over the entire study are close to zero. Treatment effects are close to zero all throughout New York City, although there is some evidence of a preference for whites with criminal records outside of Manhattan. Still, standard errors for treatment effects in the outer boroughs imply little certainty in the direction of employer preference. Finally, we obtain the same result of zero treatment effects regardless of which tester is sent first to apply for the job. In sum, the finding of uniform treatment of whites with criminal records and minorities with clean records is supported over the whole of the experimental period, throughout New York City, and regardless of which tester first makes contact with the employer.

Table 2A. Percentage of positive responses and race differences, by date, employer address, and race of first tester. The white tester is assigned a criminal record

	White felon (Wf)	Latino (L)	Black (B)	Race Differences (SE)		
				Wf-L	Wf-B	L-B
Total (170)	17.1	15.9	12.9	1.2 (3.3)	4.1 (2.9)	2.9 (3.3)
<i>Date</i>						
Mar 2 - Apr 13 (84)	16.7	14.3	10.7	2.4 (4.5)	6.0 (3.9)	3.6 (4.5)
Apr 14 - Aug 6 (82)	17.1	17.1	15.9	0.0 (5.2)	1.2 (4.4)	1.2 (5.2)
<i>Location<sup>a</sup></i>						
Below 34th St. (52)	9.6	9.6	3.9	0.0 (5.4)	5.8 (3.2)	5.8 (5.4)
34-72nd St. (46)	13.0	17.4	13.0	-4.4 (5.3)	0.0 (5.3)	4.4 (5.3)
Above 72nd St. (7)	0.0	0.0	0.0	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
Other	29.0	21.0	21.0	8.1 (6.6)	8.1 (6.5)	0.0 (6.6)
<i>Race of first tester</i>						
White (56)	19.6	17.9	12.5	1.8 (6.5)	7.1 (5.8)	5.4 (6.5)
Black (58)	19.0	15.5	15.5	3.5 (4.8)	3.5 (3.5)	0.0 (4.8)
Latino (52)	11.5	13.5	11.5	-1.9 (6.4)	0.0 (5.7)	1.9 (6.4)

<sup>a</sup> Street addresses are for Manhattan

Note: Standard errors are calculated with the bootstrap. Observations in subgroups may not sum to total number of observations due to missing data.



## Appendix B. Results by Tester Teams.

In the course of fielding two three-person teams of testers we used ten different testers: two Latinos, four African Americans, and four whites. In each three-person team consisting of a white, black, and Latino, the ten testers were combined into 6 different unique combinations. Before pooling the data across combinations of testers, Heckman and Seigelman (1993) recommend testing for the homogeneity of responses across combinations. A chi-square test within each team fails to reject the null hypothesis of homogeneity across combinations. With this evidence of homogeneity, we report treatment effects pooled across testers. Table B reports the detailed experimental results for each unique combination of testers.

**Table B. Detailed Experimental Results, by Unique Combination of Testers.**

Group	Who Gets Positive Response (%):						N	
	None	W + L	W + B	L + B	W	L		B
<i>White without criminal record</i>								
1	69.2	4.4	3.3	11.0	7.7	4.4	.0	91
2	67.9	11.3	0.0	7.5	9.4	3.8	.0	53
3	33.3	33.3	0.0	33.3	.0	.0	.0	6
4	18.2	.0	0.0	36.4	18.2	18.2	9.1	11
5	66.7	.0	0.0	.0	33.3	.0	.0	3
6	57.1	14.3	0.0	28.6	.0	.0	.0	7
Total	63.7	7.6	1.8	12.9	8.8	4.7	.6	171
<i>White with criminal record</i>								
1	77.8	8.3	2.8	5.6	2.8	.0	2.8	36
2	74.4	2.4	2.4	4.9	7.3	6.1	2.4	82
3	100.0	.0	.0	.0	.0	.0	.0	2
4	60.0	.0	.0	20.0	.0	20.0	.0	5
5	56.1	2.4	2.4	12.2	14.6	7.3	4.9	41
6	75.0	.0	.0	.0	.0	.0	25.0	4
Total	70.6	3.5	2.4	7.1	7.6	5.3	3.5	170

Note: W = white; L = Latino; B = black

## References

- Agresti, Alan. 1990. *Categorical Data Analysis*. New York, NY: Wiley.
- Aigner, Dennis J., and Glen G. Cain. 1977. "Statistical theories of discrimination in labor market." *Industrial and Labor Relations Review* 30:749-776.
- Altonji, Joseph G., and Charles R. Pierret. 1997. "Employer learning and statistical discrimination." NBER Working Paper 6279.
- Becker, Gary S. 1971. *The Economics of Discrimination*. Second Edition. Chicago, IL: University of Chicago Press.
- Biernat, Monica and Diane D. Kobrynowicz. 1997. "Gender and race-based standards of competence: lower minimum standards but higher ability standards for devalued groups." *Journal of Personality and Social Psychology* 72: 544-557.
- Cancio, A. Silvia, T. David Evans, and David J. Maume. 1996. "Reconsidering the Declining Significance of Race: Racial Differences in Early Career Wages." *American Sociological Review* 61:541-56.
- Carneiro, Pedro, James J. Heckman, and Dimitriy V. Masterov. 2003. "Labor Market Discrimination and Racial Differences in Premarket Factors," National Bureau of Economics Research Working Paper 10068.
- Tomaskovic-Devey, Donald, Melvin Thomas and Kecia Johnson. 2005. "Race and the Accumulation of Human Capital across the Career: A Theoretical Model and Fixed-Effects Application," *American Journal of Sociology* 111 (1): 58-89.
- Cawley, John, Karen Conneely, James Heckman, and Edward Vytlačil. 1997. "Cognitive Ability, Wages, and Meritocracy," in *Intelligence, Genes, and Success: Scientists Respond to the Bell Curve*, eds. Bernie Devlin, Stephen Fienberg, Daniel Resnick, and Kathryn Roeder. New York: Springer Verlag.
- Correll, Joshua, Bernd Wittenbrink, and Charles M. Judd. 2002. "The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals." *Journal of Personality & Social Psychology* 83:1314-1329.
- Daniel, W.W. 1968. *Racial Discrimination in England*. Harmondsworth: Penguin.
- Dovidio, John F., and Samuel L. Gaertner. 2000. "Aversive racism and selection decisions." *Psychological Science* 11:315-319.
- Eberhardt, Jennifer L., Phillip Atiba Goff, Valerie J. Purdie, and Paul G. Davies. 2004. "Seeing black: Race, crime, and visual processing." *Journal of Personality & Social Psychology* 87:876-893.

- Elliott, J.R. 2000. "Class, Race, and Job Matching in Contemporary Urban Labor Markets." *Social Science Quarterly* 81:1036-1052.
- Farkas, George. and K. Vicknair. 1996. "Appropriate Tests of Racial Wage Discrimination Require Controls for Cognitive Skill: Comment on Cancio, Evans, and Maume." *American Sociological Review* 61:557-60.
- Fiske, Susan. 1998. "Stereotyping, prejudice, and discrimination." Pp. 357-411 in *The handbook of social psychology*, edited by Daniel Gilbert, Susan Fiske, and Gardner Lindzey. Boston: McGraw Hill.
- Fiske, Susan, and S. L. Neuberg. 1990. "A continuum model of impression formation: From category-based to individuating processes: Influence of information and motivation on attention and interpretation." Pp. 1-74 in *Advances in experimental psychology*, edited by M. P. Zanna. New York: Academic Press.
- Fryer, Ronald G. Jr., and Steven D. Levitt. 2004. "The causes and consequences of distinctively black names." *The Quarterly Journal of Economics* 119 767-805.
- Ghosh, Malay, Ming-Hui Chen, Atalanta Ghosh, and Alan Agresti. 2000. "Hierarchical Bayesian Analysis of Binary Matched Pairs Data." *Statistica Sinica* 10:647-57.
- Heckman, James and Peter Siegelman. 1993. "The Urban Institute Audit Studies: Their Methods and Findings." Pp. 187-258 in *Clear and Convincing Evidence: Measurement of Discrimination in America*, edited by Michael Fix and Raymond Struyk. Lanham, MD: University Press of America.
- Holzer, Harry J. 1996. *What Employers Want: Job Prospects of Less-Educated Workers*. New York: Russell Sage Foundation.
- Kirschenman, Joleen and Katherine Neckerman. 1991. "'We'd Love to Hire Them, but...': The Meaning of Race for Employers." Pp. 203-234 in *The Urban Underclass*, edited by Christopher Jencks and P. Peterson. Washington, D.C.: Brookings Institute.
- Kobrynowicz, Diane D. and Monica Biernat. 1997. "Decoding subjective evaluations: How stereotypes provide shifting standards." *Journal of Experimental Social Psychology* 33: 579-601.
- Loury, Glenn C. 2002. *The Anatomy of Racial Inequality*. Cambridge MA: Harvard University Press.
- Moss, Philip and Chris Tilly. 2001. *Stories Employers Tell: Race, Skill, and Hiring in America*. New York: Russell Sage.

- Neal, Derek, and William Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104:869-895.
- Oaxaca, Ronald. 1973. "Male-Female Wage differentials in Urban Labor Markets." *International Economic Review* 14:693-709.
- Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108: 937-975.
- Petersen, Trond, Ishak Saporta, and Marc-David L. Seidel. 2000. "Offering a Job: Meritocracy and Social Networks." *American Journal of Sociology* 106:763-816.
- Pettit, Becky and Bruce Western. 2004. "Mass Imprisonment and the Life Course: Race and Class Inequality in U.S. Incarceration." *American Sociological Review* 59:151-169.
- Skogan, Wesley G. 1995. "Crime and the Racial Fears of White Americans." *Annals of the American Academy of Political and Social Science* 539:59-71.
- Yarkin, K.L., J.P. Town, and B.S. Wallston. 1982. "Blacks and Women Must Try Harder: Stimulus Person's Race and Sex Attributions of Causality." *Personality and Social Psychology Bulletin* 8: 21-24.
- Waldinger, Roger and Michael I. Lichter. 2003. *How the Other Half Works: Immigration and the Social Organization of Labor*. Berkeley, CA: University of California Press.
- Wilson, William Julius. 1978. *The Declining Significance of Race: Blacks and Changing American Institutions*. Chicago, IL: University of Chicago Press.
- Wilson, William Julius. 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Vintage Books.