Are Emily and Brendan More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination

Marianne Bertrand and Sendhil Mullainathan *

November 18, 2002

Abstract

We perform a field experiment to measure racial discrimination in the labor market. We answer help-wanted ads in Boston and Chicago newspapers by sending resumes. To manipulate perception of race, each resume is randomly assigned either a very African American sounding name or a very White sounding name. This manipulation produces a significant gap in the rate of callbacks for interviews. White names elicit about 50% more callbacks than African American names. We also investigate how improvements in credentials affect discrimination. For each employment ad, we send resumes of higher and lower quality. For Whites, the higher quality resumes elicit 30 percent more callbacks. For African Americans, however, the higher quality resumes do not elicit significantly more callbacks. In other words, African Americans benefit little if at all from improving their credentials. The extent of discrimination is also remarkably uniform across occupations and industries. Similarly, Federal contractors (for whom affirmative action is better enforced) and employers who list "Equal Opportunity Employer" in their ad discriminate as much as other employers. In Chicago, we find that employers located in more African American neighborhoods discriminate less.

*University of Chicago Graduate School of Business, NBER and CEPR; MIT and NBER. Addresses: 1101 E. 58th Street, R0 229D, Chicago, IL 60637; 50 Memorial Drive, E52-380a, Cambridge, MA 02142. E-mail: marianne.bertrand@gsb.uchicago.edu; mullain@mit.edu. David Abrams, Victoria Bede, Simone Berkowitz, Hong Chung, Almudena Fernandez, Mary Anne Guediguian, Christine Jaw, Richa Maheswari, Beverley Martis, Alison Tisza, Grant Whitehorn, and Christine Yee provided excellent research assistance. We are also grateful to numerous colleagues and seminar participants for very helpful comments.

1 Introduction

Every measure of economic success reveals significant racial inequality in the US labor market. Compared to Whites, African Americans are more than twice as likely to be unemployed. And when employed, they earn nearly 25% less than their white counterparts (Council of Economic Advisers, 1998). We address two important questions raised by this inequality. First, do employers actively discriminate against African Americans in hiring? Faced with two observably similar applicants, do they favor the White one? Employers may discriminate for many reasons, from prejudice to rational beliefs that race correlates with unwanted traits. But whatever their behavioral motivation, discrimination would raise a hurdle for African Americans seeking jobs. Second, if there is discrimination, does improving credentials of African Americans affect it? For several reasons, discrimination may diminish for African Americans with more observable skills. For example, if discrimination is driven by employer worries that African Americans differ from Whites on unobservable dimensions, acquiring credentials may assuage these fears. But there are equally compelling reasons why discrimination may increase for African Americans with more observable skills. For example, employers may feel that these skills have less value because African Americans do not possess other traits (e.g. resourcefulness) to fully use them. The effectiveness of skills training in reducing race inequality depends on the direction of this effect.

To answer these two questions, we conduct a field experiment. We send resumes in response to help-wanted ads and use the rate of callback for interviews to measure the success of each resume. We experimentally manipulate perception of race. Since resumes rarely state race, we manipulate race via the name of the applicant. We randomly assign to half the resumes a very White sounding name (such as Emily Walsh or Brendan Baker) and to the other half a very African American sounding name (such as Lakisha Washington or Jamal Jones). To see how the effect of race differs by resume quality, we respond to most employment ads with both higher and lower quality resumes. Applicants with higher quality resumes have on average more labor market experience and fewer holes in their employment history; they are also more likely to have an email address, to have completed some certification degree and to have been awarded some honors. So we typically send four resumes in response to each ad: two high-quality and two low-quality resumes, with one of each randomly assigned an African American sounding name. In total, we respond to over 1300 ads in Boston and Chicago newspapers in the sales, administrative support, clerical and customer services job categories and send nearly 5000 resumes. The jobs we respond to cover a large spectrum of quality. The jobs range from cashier work at retail establishments and clerical work in a mailroom to a highly paid executive assistant for a CEO and management of sales at a large firm.

Despite random assignment of race, we find huge differences in call-back rates by race.[1] Applicants with White names need to send about 10 resumes to get one call-back whereas applicants with African American names need to send around 15 resumes to get one callback. This 50% gap in callback rates is statistically very significant. Since applicants' names are randomly assigned, this gap can only be attributed to the race-specific names, suggesting that employers discriminate (quite a bit) on the basis of race.

Race also affects the reward to having a better resume. Whites with higher quality resumes receive 30 percent more callbacks than Whites with lower quality resumes, a statistically significant difference. On the other hand, having a higher quality resume has a much smaller and statistically insignificant effect for African American applicants. This lower reward for African Americans suggests that, *in the current state of the labor market*, African Americans do not have strong individual incentives to build a stronger resume. Discrimination appears to bite twice, making it harder for African Americans to find a job and to improve their employability.

The experiment also reveals several other aspects of discrimination. First, since applicants' addresses are randomly assigned, we can study the effect of neighborhood of residence on the probability of callback. We find that living in a wealthier (or more educated or more White) neighborhood increases callback rates. But, interestingly, African Americans are not helped more than Whites by neighborhood. Second, the amount of discrimination in an industry is statistically uncorrelated with the race gaps (either wage or employment) observed for that industry in Census data. The same is true for the amount of discrimination we measure in different occupations. In fact, we find that statistically, discrimination levels are the same across all the occupations and industries covered in the experiment. We also find that federal contractors, who are more severely constrained by affirmative action laws, do not discriminate less; neither do larger employers or employers who explicitly state that they are an "Equal Opportunity Employer" in their employment ads. Finally, in the Chicago labor market, employers located in more African American neighborhoods are less likely to discriminate against African Americans.

[1]For ease of exposition, we refer to the effects uncovered in this experiment as "race" differences. Technically, however, these effects are about racial soundingness of names.

2 Evidence on Discrimination

Conventional data are limited in their ability to measure racial discrimination and analyze its mechanics.[2] With survey data, one usually compares the labor market performance of Whites and African Americans who have a similar set of skills. But such comparisons can be quite misleading. These data do not contain all the characteristics that employers observe when hiring, promoting or setting wages. So one can never be sure that the African American and White workers being compared are truly similar from the employer's perspective. As a consequence, any measured differences in outcomes can be attributed to these unobserved factors and not to discrimination per se.

This difficulty with conventional data has led some authors to use pseudo-experiments to study discrimination.[3] Goldin and Rouse (2000), for example, examine the effect of blind auditioning on the hiring process of orchestras. By looking at the treatment of female candidates before and after the introduction of blind auditions, they try to measure the amount of sex discrimination. When such "experiments" can be found, the resulting study can be very informative, but finding such valid experiments has proven extremely challenging.

A different set of studies, known as audit studies, attempt to place comparable minority and White subjects into actual social and economic settings and measure how each group fares in these settings. For example, one might examine whether similarly dressed African Americans and Whites have the same ease in getting a cab.[4] Labor market audit studies send comparable minority (African American or Hispanic) and White auditors in for interviews and measure whether one is more likely to get the job than the other.[5] While the results vary somewhat across studies, minority auditors tend to perform worse: they are less likely to get called back for a second interview and, conditional on getting called back, less likely to get hired.

These audit studies provide some of the cleanest non-laboratory evidence of labor market dis-

[2]See Altonji and Blank (1999) for a detailed review of the existing literature on racial discrimination in the labor market.

[3]Darity and Mason (1998) describe an interesting non-experimental study. Prior to the Civil Rights Act of 1964, employment ads would explicitly state racial biases, providing a direct measure of discrimination. Of course, as Arrow (1998) mentions, discrimination was at that time "a fact too evident for detection."

[4]Fix and Turner (1998) provide a survey of many such audit studies.

[5]Earlier hiring audit studies include Newman (1978) and McIntyre et al (1980). Three more recent studies are Cross et al (1990), Turner et al (1991) and James and DelCastillo (1991). Altonji and Blank (1999), Heckman and Siegelman (1992) and Heckman (1998) summarize these studies.

crimination. But they also have weaknesses, most of which have been highlighted in Heckman and Siegelman (1992) and Heckman (1998). First, these studies require that both members of the auditor pair are identical in all dimensions that might affect productivity in employers' eyes, except for race. Researchers usually match auditors on several characteristics (height, weight, age, dialect, dressing style, hairdo) and train them for several days to coordinate interview styles. Yet, critics note that this is unlikeley to erase the numerous differences that exist between auditors in a pair.

Another weakness of the audit studies is that they are not double-blind. Auditors know the purpose of the study. As Turner et al (1990) note: "The first day of training also included an introduction to employment discrimination, equal employment opportunity, and a review of project design and methodology." This may generate conscious or subconscious motives among auditors to generate data consistent or inconsistent with racial discrimination. As psychologists know very well, these demand effects can be strong. It is very difficult to insure that auditors will not want to do "a good job." Since they know the goal of the experiment, they can alter their behavior in front of employers to express (indirectly) their own views. Even a small belief by auditors that employers treat minorities differently can result in apparent discrimination. This effect is further magnified by the fact that unlike actual interviewees, the auditors are not in fact seeking jobs. They are therefore freer to let their beliefs about employers' attitude affect the interview.

Finally, these audit studies are extremely expensive, making it difficult to generate large enough samples to understand the nuances and mitigating factors of discrimination. In fact, these budgetary concerns worsen the problem of mismatched auditor pairs. Cost considerations force the use of a few pairs, meaning that any mismatch will easily drive the results. In fact, studies generally tend to find that outcomes significantly differ across pairs.

Our study circumvents these problems. First, because we only use resumes and not people, we can be sure to generate comparability. In fact, since race is randomly assigned to each resume, the same resume will sometimes be associated with a African American name and sometimes with a White name. This guarantees that any differences we find are due solely to the race manipulation. Second, the use of paper resumes also insulates us from demand effects. While the research assistants also know the purpose of the study, our protocol allows little room for conscious or subconscious deviations from the set procedures. Moreover, we can objectively measure whether the randomization occurred as expected. This kind of objective measurement is impossible in the case of the previous audit studies. Finally, the relatively lower cost of sending out resumes means that rather than a few auditor pairs, we can send out a large number of resumes. Besides giving us more precise estimates, it also allows us to examine the mechanics of discrimination from many more angles.[6]

3 Experimental Design

3.1 Creating a Set of Resumes to Send

The first step of the experimental design is to generate the resumes that are to be sent. The challenge is to produce a set of realistic and representative resumes without using ones that belong to actual job seekers since this could interfere with their job search. To achieve this goal, we start with resumes of actual job searchers but alter them sufficiently to create distinct resumes. The alterations maintain the structure and realism of the initial resume without compromising their owners.

We begin with resumes posted on two job search websites as the basis for our artificial resumes.[7] While the resumes posted on these websites may not be completely representative of the average job seeker, they provide a practical approximation.[8] We restrict ourselves to people seeking employment in our experimental cities (Boston and Chicago). We also restrict ourselves to four occupational categories: sales, administrative support, clerical services and customer services. Finally, we further restrict ourselves to resumes posted more than six months prior to the start of the experiment.

During this process, we classify the resumes within each occupational category into two groups: high and low quality. In judging resume quality, we use criteria such as labor market experience, existence of gaps in employment and skills listed. For example, resumes with large gaps in employment experience are more likely to be classified as low quality ones. Of course, such a classification is subjective but it is made independently of any race assignment (which occurs later). To further reinforce the quality gap between the two sets of resumes, we add to each high quality resume some

[6] A similar "correspondence" technique has been used in a few U.K. studies (Jowell and Precott-Clarke (1970), Brown and Gay (1985), Hubbock and Carter (1980)). These earlier studies had very limited sample size and focused mainly on documenting gap in call-backs between the minority and non-minority groups. Some of these studies also failed to fully match skills between minority and non-minority resumes by imposing differential education background by racial origin.

[7] The sites are (*www.careerbuilder.com* and *www.americasjobbank.com*).

[8] In practice, we found large variation in skill levels among people posting their resumes on these sites.

of these features: summer employment experience or while-at-school employment experience, volunteering experience, extra computer skills, certification for administrative positions, special honors or some military experience. This resume quality manipulation needs to be somewhat subtle to avoid making a higher quality job applicant overqualified for a given job. We tried to avoid this problem by making sure that these features were not all added at once and/or by limiting their strength. This leaves us with a pool of resumes, some of which are classified as "high" quality and some which are classified as "low" quality.

We then purge these resumes of the person's name and any contact information. To minimize similarity to actual job seekers, we use resumes from Boston job seekers to form templates for resumes to be sent out in Chicago and used the Chicago resumes to form templates for resumes to be sent out in Boston. To implement this migration, we alter the names of the schools and previous employers on the resumes. For each Boston resume, we use the Chicago resumes to replace a Boston school by a Chicago school. We use various sources to match high schools and colleges as well as possible on quality and demographic characteristics. We use a similar procedure to migrate Chicago resumes to Boston.[9] This procedure produces distinct but realistic looking resumes, similar in their education, experience and personal profiles to this sub-population of job searchers.[10]

3.2 Identities of Applicants

The next step is to generate identities for the fictitious job applicants: names, telephone numbers, postal addresses and (possibly) e-mail addresses. The choice of names is crucial to our experiment. To decide on which names are uniquely African American and which are uniquely White, we use name frequency data calculated from birth certificates of all babies born in Massachusetts between 1974 and 1979. We tabulate these data by race to determine which names are distinctively White and which are distinctively African American. Distinctive names are those that have the highest ratio of frequency in one racial group to frequency in the other racial group.

As a check of distinctiveness, we conducted a survey in various public areas in Chicago. Each respondent was given a personal name and asked to assess features of the person, one of which

[9]Note that for applicants with schooling or work experience outside of the Boston or Chicago areas, we leave the school or employer name unchanged. These are very rare events.

[10]We also generated a set of different fonts, layouts and cover letters to further differentiate the resumes. These are applied at the time the resumes are sent out.

being race. In general, the names led respondents to readily attribute the expected race for the person but there were a few exceptions and these names were disregarded.[11]

The final list of first names used for this study are reported in Appendix Table 1. The table also reports the frequency of these names in the Massachusetts birth certificates.[12] The African American first names used in the experiment are remarkably common in the population. This suggests that by using these names as an indicator of race, we are actually covering a large segment of the African American population.[13]

Applicant in each race/sex/city/resume quality share a phone number. This guarantees that we can precisely track employer callbacks in each cell, even if we were not able to match a callback to a specific resume. The phone lines we use are virtual ones with only a voice mail box attached to it.[14] Since people with different names use the same voice box, we cannot use the person name in the outgoing message. Each outgoing message is recorded by someone of the appropriate race and gender. Moreover, different people are used to record the messages on the high and low quality mailboxes.

While we do not expect positive feedback from an employer to take place by postal mail, resumes still need postal addresses. We therefore construct fictitious addresses based on real streets in Boston and Chicago using the White Pages. We select up to 3 addresses in each 5-digit zip code in Boston and Chicago. Within cities, we randomly assign addresses across all resumes.

Finally, we create 8 email addresses, 4 for Chicago and 4 for Boston. The e-mail addresses are registered on *Yahoo.com*, *Angelfire.com* or *Hotmail.com*. The specific addresses chosen are neutral with respect to both race and sex. Not all applicants are given an email address. As we explained above, the email addresses are used almost exclusively for the higher quality resumes.

This procedure leaves us with a bank of names, phone numbers, addresses and e-mail addresses which we can assign to the template resumes when responding to the employment ads.

[11]For example, Maurice and Jerome are distinctively African American names in a frequency sense yet are not perceived to be so by many people.

[12]We also tried to use more White sounding last names for White applicants and more African American sounding last names for African American applicants. The last names used for White applicants are: Baker, Kelly, McCarthy, Murphy, Murray, O'Brien, Ryan, Sullivan and Walsh. The last names used for African American applicants are: Jackson, Jones, Robinson, Washington and Williams.

[13]Is this a different segment of the African American population? We discuss in Section 5, whether our race effect could be interpreted as a social class effect.

[14]Each voice mail box had a variant of the message "I'm not in right now. Please leave a message."

3.3 Responding to Ads

The experiment was carried on between July 2001 and January 2002 in Boston and between July 2001 and May 2002 in Chicago. Over that period, we collected all employment ads in the Sunday editions of *The Boston Globe* and *The Chicago Tribune* in the sales, administrative support, and clerical and customer services sections. We eliminate any ad where applicants are asked to call or appear in person. This was in fact a rare event, as most ads we surveyed in these categories asked for applicants to fax in their resume. We log the name (when available) and contact information for each employer, along with any information on the specific position advertised and further requirements (such as education, experience, or computer skills). We also record whether or not the ad explicitly states that the employer is an equal opportunity employer.

For each ad, we use the bank of resumes to sample four resumes (two high-quality and two low-quality) that fit the job description and requirements as closely as possible.[15] In some cases, we slightly alter resumes to improve the quality of the match, such as adding the knowledge of a specific software program.

One of the high and one of the low quality resumes are drawn at random to receive African American names, the other high and low resumes receive White names.[16] We use male and female names for sales jobs, whereas we use nearly exclusively female names for administrative and clerical jobs to increase callback rates.[17] Based on sex, race, city and resume quality, we assign a resume the appropriate phone number. We also select an address at random from the pool of addresses. Finally, e-mail addresses were added to most high quality resumes.[18] The final resumes are then formatted, with fonts, layout and cover letter style chosen at random. The resumes are then faxed (or in a few cases mailed) to the employer.[19] All in all, we respond to more than 1300 employment ads over the entire sample period and send close to 5000 resumes.

[15] In some cases, our sample of resumes do not have four resumes that are appropriate for the ad. In these cases, we send two resumes to the ad.

[16] Though the names are repeatedly used, we guaranteed that no ad would receive multiple resumes with the same name.

[17] Male names were used for a few administrative jobs in the first month of the experiment.

[18] In the first month of the experiment, a few high quality resumes were sent without email addressed and a few low quality resumes were given email addresses. See Table 4 for details.

[19] As part of the fax process, we strip all identifiers from the outgoing fax to guarantee that employers could not see that all four faxes originate from the same locale.

3.4 Measuring Responses

We measure whether a given resume elicits a callback or e-mail back for an interview. For each call-back or email-back, we use the content of the message left by the employer (name of the applicant, company name, telephone number for contact) to match the call-back to the specific resume and ad.[20]

Any attempt by employers to contact applicants by postal mail cannot be measured in our experiment since the addresses are fictitious. Several human resource managers confirmed with us that employers rarely, if ever, contact applicants via postal mail to set up interviews.

3.5 Weaknesses of the Experiment

We have already highlighted the strengths of this experiment relative to previous audit studies. Now we discuss its weaknesses. First, our outcome measure is crude, even relative to the previous audit studies. Ultimately, one cares about whether applicants get a job or not and about the wage they are offered conditional on getting the job. Our procedure, however, simply measures callbacks for interviews. To the extent that the search process has even moderate frictions, one would expect that reduced interview rates would translate into reduced job offers. However, we are not able to translate our results into hiring rates or earnings gaps.

Another weakness is the resumes do not directly report race but instead suggest race through personal names. This leads to various sources of concern. First, while the names are chosen to make race salient, some employers may simply not notice the names or not recognize their racial content. As a result, our findings may under-estimate the extent of discrimination. In a related vein, because we are not assigning race but only race-specific name, our results are not representative of the average African American who may not have such a racially distinct name. However, as Appendix Table 1 indicates, the very African American sounding names we use are actually quite common, making this less of a concern.

Finally, and this is an issue pervasive in both our study and the pair-matching audit studies, newspaper ads represent only one channel for job search. As is well known from previous literature, social networks are the most common means through which people find jobs. This alternative job search channel can clearly not be studied here.

[20] Almost no employers used email to contact an applicant back.

4 Results

4.1 Is There Discrimination?

Table 1 presents our basic tabulations of the average callback rates by racial soundingness of names. Included in brackets under each rate is the number of applicants in that cell. Row 1 presents our results for the full data set (Chicago and Boston labor markets). Resumes with White names have a 10.08 percent chance of receiving a callback. Equivalent resumes with African American names, however, have only a 6.70 percent chance of being called back. This represents a difference in callback rates of 3.35 percentage points, or 50 percent, that can solely be attributed to the name manipulation. Column 4 shows that this difference is extremely significant. Put in other words, these results imply that a White applicant should expect on average one call back for every 10 ads she or he applies to; on the other hand, a African American applicant would need to apply to 15 different ads to achieve the same result.[21]

Rows 2 and 3 break down the full sample of sent resumes into the Boston and Chicago markets. About 20 percent more resumes were sent in Chicago (about 2700 resumes) than in Boston (about 2200 resumes). The average call back rate (across races) is lower in Chicago than in Boston. This might reflect differences in labor market conditions across both cities over the experimental period or maybe differences in the ability of the MIT and Chicago teams of undergraduates in selecting resumes that were good matches for a given help wanted ad. The percentage difference in call-back rates is however strikingly similar across both cities. White applicants are 48 percent more likely than African American applicants to receive a call-back in Chicago and 52 percent more likely in Boston. These racial differences are statistically significant in both cities.

Finally, rows 4 to 7 break down the full sample of sent resumes into female and male applicants. Row 4 displays the average results for all female names while rows 5 and 6 break the female sample into administrative (row 5) and sales jobs (row 6); row 7 displays the average results for all male names. As noted earlier, female names were used in both sales and administrative job openings whereas males names were used close to exclusively for sales openings.[22] Looking across occupations, we find a significant racial gap in callbacks for both males (49%) and females (50%).

[21] This exercise obviously assumes that African American applicants cannot assess a priori which firms are more likely to discriminate against them.

[22] Only about 6 percent of all male resumes were sent in response to an administrative job opening.

Comparing males to females in sales occupations, we find a somewhat larger racial gap among males (25% versus 49%). Interestingly, females in sales jobs appear to receive more call-backs than males; however, this (reverse) gender gap is statistically insignificant and economically smaller than any of the racial gaps discussed above.

The results in Table 1 pool together all the specific personal African American names and all the specific personal White names to compute an average racial gap. In Table 2, we examine whether there are name-specific effects. Because sample sizes are not large enough to separately consider each first and last name combination, we focus on first names. We report in Table 2 the mean call-back rate for each personal first name, ranking names in each sex/race group by ascending callback rate. Not surprisingly, we find variation in callback rates across names. Chance alone would produce such variation because of the rather small number of observations in each cell. We therefore formally test the hypothesis that the names within each sex-race category produce the same effect. We estimate a probit regression of the call back dummy on all the personal first names, allowing for clustering of the observations at the employment ad level. For all but one sex-race category, we cannot reject the hypothesis that all the first name effects are the same. Only for African American female names do we reject this null at a significant level. Five out of nine female African American names (Aisha, Keisha, Tamika, Lakisha and Tanisha) do worse than the worst female White name. The last four female names (Latoya, Kenya, Latonya and Ebony) perform only slightly below the average White female name.

We investigated two possible explanations for these name specific effects among African American females. First, we considered the possibility that employers might be relatively less familiar with some of the worst performing names and that it is this lack of familiarity that motivates their callback behavior. However, we found no obvious correlation between the name-specific call back rates and the relative frequency of each of the names, at least in the Massachusetts birth certificates. Second, we considered the possibility that the name fixed effects reflect differences in social class or economic background. To assess the relevance of this interpretation, we used some limited Massachusetts birth certificates data on mothers' education. More specifically, for each of the five most common African American female names in our sample (Aisha, Ebony, Keisha, Tamika and Tanisha), we were able to obtain information on the fraction of mothers having completed high school. We found no obvious relationship between the name-specific callback rates and mothers' education.[23]

Rather than studying the distribution of call backs at the applicant level, one can also tabulate the distribution of callbacks at the employment ad level. We do this in Table 3. More formally, we compute in that table the fraction of employers that treat White and African American applicants equally, the fraction of employers that favor White applicants and the fraction of employers that favor African American applicants. Because we send up to four resumes (2 Whites and 2 African Americans) in response to each sampled ad, the 3 categories above (equal treatment, Whites favored and African Americans favored) can each take 3 different forms. Equal treatment occurs when either: 1) no applicant gets call backed, 2) one White and one African American get called back or 3) two Whites and two African Americans get called back. Whites (African Americans) are favored when either: 1) only one White (African American) gets called back, 2) two Whites (African Americans) and no African American (White) get called back and 3) two Whites (African Americans) and one African American (White) get called back.

As Table 3 indicates, equal treatment occurs for about 87.5 % of the help-wanted ads. As expected, the major source of equal treatment comes from the high fraction of ads for which no callbacks are recorded (82.5 % of the ads). Whites are favored by nearly 9% of the employers, with a majority of these employers contacting exactly one White applicant. African Americans, on the other hand, are favored by only about 3.7% of employers. We formally test whether there is symmetry in the favoring of Whites over African Americans and African Americans over Whites by employers. We find that the difference between the fraction of employers favoring Whites and the fraction of employers favoring African Americans (5.11%) is statistically very significant ($p = .0000$).

In summary, our results so far suggest that there is a substantial amount of discrimination in the job recruiting process. Selecting from a pool of applicants with comparable credentials, employers tend to favor the applicants with White sounding names. The documented gap in call-back suggest that African Americans may need to send on average 50 percent more resumes than comparable Whites to achieve the same number of call backs. While the cost of sending additional resumes might not be large per se, this 50 percent gap could be quite substantial when compared to rate of arrival of new job openings.[24]

[23] Female names by ascending mother education are: Tamika, Keisha, Tanisha, Ebony and Aisha.

[24] Across the jobs that we sampled, for example, we would find only 5 to 15 new relevant ads in each city each week.

4.2 Do African Americans Receive Different Returns to Resume Quality?

The evidence we have put together so far is consistent with the presence of racial discrimination in the labor market. In the next step, we would like to learn more about how employers discriminate in the current state of the labor market. More specifically, we ask how employers respond to improvements in African American applicants' credentials. To answer this question, we examine how the racial gap in callback rates varies by resume quality.

As we mentioned in section 3, for almost every employment ad we respond to, we send four different resumes, two higher quality and two lower quality ones. Table 4 gives a better sense of which factors enter into this qualitative classification. In that table, we report mean and standard deviation for the most relevant resume characteristics. Column 1 focuses on the entire sample of sent resumes. Columns 2 and 3 break down the full sample into resumes with White names and resumes with African American names. Since applicants' names are randomized, there is no difference in resume characteristics across these two columns. Most relevant, columns 4 and 5 contrast higher and lower quality resumes. Higher quality applicants have more labor market experience and fewer employment holes (where an employment hole is defined as a period of at least 6 months without a reported job); they are more likely to have worked while at school and to report some military experience. Finally, higher quality applicants are more likely to have an email address, to have received some honors and to list some computer skills and special skills (such as a certification degree) on their resume.

The last 5 rows of Table 4 show some summary characteristics of the applicants' zip code address. Using 1990 Census data, we compute the fraction high-school dropouts, fraction college educated or more, fraction Whites, fraction African Americans and log(median per capita income) for each zip code used in the experiment. Since addresses were randomized within cities, these neighborhood quality measures are uncorrelated with race or resume quality.

The differences in callback rates between high and low resumes are presented in Table 5. The first thing to note is that the resume manipulation works: higher quality resumes receive higher call back rates. As row 1 indicates, we record a call back rate of more than 11 percent for White applicants with a higher quality resume, compared to 8.8 percent for White applicants with lower quality resumes. This is a statistically significant difference of 2.51 percentage points, or 30 percent. Most striking, however, is that African Americans experience much less of an increase in call-back rate for the same improvements in their credentials. African Americans with higher quality resumes receive call-backs 6.99 percent of the time, compared to 6.41 percent for African Americans with lower quality resumes. This is only a .58 percentage point, or 9 percent, increase and this difference is not statistically significant.

Instead of relying on the subjective quality classification, Panel B directly uses resume characteristics to classify them. More specifically, we use a random subsample of one-third of the resumes to estimate a probit regression of the callback dummy on the resume characteristics listed in Table 4. We further control for a sex dummy, a city dummy, 6 occupation dummies and a vector of job requirements as listed in the employment ads.[25] We then use the estimated coefficients on the resume characteristics to rank the remaining two-thirds of the resumes by predicted callback. We classify as "high" resumes that have above median predicted callback; similarly, we classify as "low" resumes that have below median predicted callback. As one can see from Panel B, qualitatively similar results emerge from this analysis. While African Americans do appear to significantly benefit from higher quality resumes under this alternative classification, they benefit much less than Whites. The ratio of call-back rates for high versus low quality resumes is 1.66 for African Americans, compared to 2.81 for Whites.

In Table 6, we show the results of the probit estimation directly, but this time on the full sample. We start in column 1 with results for the full sample of sent resumes. As one can see, many of the resume characteristics have the expected effect on the likelihood of a callback. The addition of an email address, honors and special skills have a positive and significant effect on the likelihood of a callback.[26] Also, more experienced applicants are more likely to get called back: at the average number of years of experience in our sample (8 years), each extra year of experience increases the likelihood of a callback by about .4 percentage point. The most counterintuitive effects come from computer skills, which appear to come in negatively, and employment holes which appear to come in positively.

The same qualitative patterns hold in column 2 where we focus on White applicants. More importantly, the estimated returns to an email address, additional work experience, honors and special skills appear economically stronger for that racial group. For example, at the average

[25] See section 4.4 for more details on these occupation categories and job requirements.

[26] Note that the e-mail address dummy, because it is close to perfectly correlated with the subjective resume quality variable, may in part capture some other unmeasured characteristics that may have led us to categorize a given resume as higher or lower quality.

number of years of experience in our sample, each extra year of experience increases the likelihood of a call back by about .7 percentage point. Also, working while at school, while not a statistically significant determinant of callback in the full sample, yields positive returns for White applicants.

As might have been expected from the two previous columns, we find that the estimated returns on these resume characteristics are all economically and statistically weaker for African American applicants (column 3). In fact, all the estimated effects for African Americans are statistically insignificant, except for the return to special skills. Resume characteristics thus appear less predictive of callback rates for African Americans than they are for Whites. To see this more clearly, we predict call-back rates using either the regression in column 2 or the regression in column 3. The standard deviation of the predicted callback from column 2 is .064, whereas it is only 0.034 from column 3. In summary, employers simply seem to pay less attention or discount more the characteristics listed on the resumes with African American sounding names.

4.3 Effect of Applicant's Address on Call-Back

An incidental feature of our experimental design is the random assignment of address to the resumes. This allows us to examine whether and how an applicant's residential address, all else equal, affects the likelihood of a callback. In addition, and most importantly for our purpose, we can also ask whether African American applicants are helped more by residing in more affluent neighborhoods.

We perform this analysis in Table 7. We start (columns 1, 3 and 5) by discussing the effect of neighborhood of residence across all applicants. Each of these columns reports the results of a probit regression of the call back dummy on a specific zip code characteristic and a city dummy. Standard errors are corrected for clustering of the observations at the employment ad level. We find a positive and significant effect of neighborhood quality on the likelihood of a callback. Applicants living in Whiter (column 1), more educated (column 3) or higher income (column 5) neighborhoods have a higher probability of receiving a call back. For example, a 10-percentage point increase in the fraction of college-educated in a zip code of residence increases the likelihood of a callback by .6 percentage point.

In columns 2, 4 and 6, we further interact each of the zip code characteristic above with a dummy variable for whether the applicant is African American or not. Each of the probit regressions in these columns also includes an African American dummy, a city dummy and an interaction of the city dummy with the African American dummy. There is no evidence that African Americans benefit any more than Whites from living in a more White, more educated zip code. The estimated interactions between fraction White and fraction college educated with the African American dummy are economically very small and statistically insignificant. We do find an economically more meaningful effect of zip code average income level on the racial gap in callback; that effect, however, is statistically insignificant.

In summary, while the quality of the neighborhood of residence is a significant factor in employers' decision to contact a job applicant back, African Americans do not appear to benefit more than Whites from living in better neighborhoods. These findings are interesting. Indeed, if ghettos and bad neighborhoods are particularly stigmatizing for African Americans, one might have expected African Americans to be helped more by having a "good" address. Our results do not lend support to this argument.

4.4 Job Characteristics

In Section 4.2, we were interested in studying how the average (discriminating) employer responds to improvements in African American applicants' credentials. We showed that African Americans gained relatively less from improvements in skills and credentials, holding job type or employer identity constant. These results were consistent with the idea that, for example, an African American woman trained to be a secretary has little hope to reduce the discrimination that she faces by beefing up her resume with more experience, extra skills or other relevant credentials. Another question one might ask, though, is whether African Americans can avoid discrimination by simply moving to better jobs altogether. There might be several reasons for this to occur. First, the larger improvements in resume quality that this would entail may be perceived differently than the smaller improvements we discussed in Section 4.2. Second, recruiters for higher quality jobs may be less discriminatory. Finally, even if no less discriminatory, recruiters for higher quality jobs may fear more the consequence of a legal action from higher educated applicants. In order to investigate this question, we now turn to job characteristics and ask whether discrimination varies based on the specific job requirements or occupation listed in the employment ads.

Table 8 gives a description of the specific requirements stated in the sample of ads we respond to. About 80% of the ads state some form of requirement. About 43% of the ads require some minimum experience, of which roughly 50% simply ask for "some experience," 25% less than two years and 25% at least 3 years of experience. About 44% of ads mention some computer knowledge requirement, which can range from Excel or Word to more esoteric software programs. Good communication skills are explicitly required in about 12% of the ads. Organization skills are mentioned 7% of the time. Finally, only about 11% of the ads list an explicit education requirement. Of these, 8.8% require a high school degree, 49% some college (such as an associate degree) and the rest at least a 4-year college degree.[27]

How do these different job requirements affect the racial gap in callback? To answer this question, we show in column 2 the results of various probit regressions. Each entry in this column is the marginal effect of the requirement on the racial gap. Specifically, each entry comes from a separate probit regression where we regress a call-back dummy on an African American dummy, the requirement dummy and the interaction of the requirement dummy with the African American dummy. The reported coefficient is that on the interaction term. The point estimates show no consistent economic pattern and are all statistically insignificant. Other ways of estimating these effects produce a similar non-result. Among other things, we considered including a city dummy or estimating the effects separately by city; we also estimated one single probit regression including all requirements at once. In other words, these job requirements have no statistically discernible effect on the extent of discrimination.

Panel A of Table 9 investigates whether discrimination varies significantly across occupations. As we mention earlier, the specific subsections of the Sunday newspapers help-wanted sections that we sample for this study broadly relate to sales and administrative positions. This is reflected in the occupational distribution reported in column 1. To form this classification, we use the job description listed in the employment ad and map it into one of six broad Census categories: executives, administrators and managers; administrative supervisors; sales representatives; sales workers (retail and personal services), secretaries and clerical workers (administrative support).[28]

We report in the second to last column of Table 9 average earnings measures for each of these occupation categories. These measures are computed from the 1990 5 percent Census for Boston and Chicago. More specifically, we first estimate a micro wage regression of log(annual earnings) on

[27] Other requirements sometimes mentioned include typing skills for secretaries (with specific words per minute minimum thresholds) and, more rarely, foreign language skills.

[28] A large fraction of the clerical workers are information clerks, such as receptionists.

8 education level dummies, a quadratic in age, a sex dummy and a city dummy. We then compute the mean residual earnings in each occupation category. As expected, the sampled occupations are very different. The first two (about 22% of the data) correspond to better jobs on average that involve, among other things, the supervision of other workers; these two categories are associated with the highest earnings. Sales representatives and secretaries (about 50% of the data) involve less supervision but are still fairly well paying jobs. The last two occupations (sales workers and especially clerical workers) correspond to the lowest paying jobs in our sample.

The second and third columns of Panel A of Table 9 respectively report the call back rates for White and African American applicants in each of these occupation categories. Column 4 reports the ratio of callback rates while column 5 reports the difference. The average call-back rate across races seems to decrease with the level of job quality. Whites receive more callbacks than African American in all job categories. While the racial gap in callback rates varies somewhat across occupations, we cannot reject the null hypothesis that the gap is the same across all categories.

We however briefly discuss the point estimates by occupation and how they relate to Census measures of occupational earnings and racial gap in earnings. The smallest gap appears to be for executive job positions, where Whites only face a 33% higher chance than African American of being called back. Clerical jobs are a close second, with Whites only being called back 38% more often. Interestingly, these two job categories are at opposite extremes of the job quality spectrum, as captured by our earnings measure. This suggests no monotonous relationship between job quality and the extent of discrimination. The highest discrimination ratio happens for administrative supervisors, the second highest job quality category. For such job openings, Whites are 64% more likely to get a callback.

In the last column of Table 9, we report the race differences in earnings by occupation. These are again computed from the 1990 5 percent Census for Boston and Chicago. The racial gaps in earning are defined as the race differences in mean residual log(annual earnings) by occupation, where residual earnings are estimated as explained above. Interestingly, there does not seem to be a monotonous association between these estimated racial gaps in earnings and the racial gaps in callbacks. For example, executive and managerial positions have the lowest racial gap in call-backs but the second highest racial gap in earnings; similarly, administrative supervisors have the highest racial gap in callbacks but the second lowest racial gap in earnings.[29] On the other hand, secretarial positions are associated with the second-highest racial gap in callbacks and the highest racial gap in earnings. These results are interesting and potentially suggest that readily available measures of racial differences, such as the racial gap in earnings used above, may not give a reliable depiction of discrimination patterns in the economy. However, because the occupational differences in callback are not significant, these results should not be overstated.

In summary, our findings in this section provide no support for the view that, in the current state of the labor market, African Americans may succeed in reducing the discrimination that they face by moving to higher quality jobs. Indeed, we do not find any consistent evidence of a decrease in discrimination within higher quality occupations. These results, combined with our earlier findings in Section 4.2, suggest that African Americans may face relatively low individual incentives to invest in higher skills.

4.5 Employer Characteristics

The final important determinants of discrimination we need to consider are employer characteristics. Collecting such employer information is not obvious as most of this information is not readily available from the employment ads we respond to. In fact, the only piece of employer information we can directly collect from the employment ad is whether or not the employer explicitly states being an "Equal Opportunity Employer." In many cases, the name of the employer is not even mentioned in the ad and the only piece of information we can rely on is the fax number which applications must be submitted to.

We proceeded as follows. For employment ads that do not list a specific employer, we use the fax number to identify the company name via web reverse-lookup services. Based on the company names, we use three different data sources (*Onesource Business Browser*, *Thomas Register* and *Dun and Bradstreet Million Dollar Directory, 2001*) to track company information such as total employment, industry and ownership status. Using this same set of data sources, we also try to identify the specific zip code of the company (or company branch) that resumes were to be sent to. Finally, we use the Federal Procurement and Data Center website to find a list of companies that have federal contracts.[30] The race differences in call-back rates for the subsample where employer

[29] In fact, the racial gap in earnings is negative in this case.

[30] This website (www.fpdc.gov) is accurate up to and including March 21, 2000.

characteristics could be determined are very similar to those in the full sample.

Column 1 of Table 10 reports summary employer characteristics for the available data. Sample sizes for each variable are reported in parentheses. 29% of all employers explicitly state that they are "Equal Opportunity Employers". 11% are federal contractors and, therefore, might face greater scrutiny under affirmative action laws. The average company size is around 2000 employees but there is a lot of variation across firms. Finally, 73% of the firms are privately held, 15% are publicly traded and 12% are non-profits.

The second column of Table 10 presents the marginal effect of each of these characteristics on discrimination. As before, each entry corresponds to a separate probit regression where we regress a callback dummy on an African American dummy, the employer characteristic in that row and the interaction of the employer characteristic with the African American dummy. The reported coefficient is that on the interaction term. First, neither the "Equal Opportunity Employers" nor the federal contractors appear to discriminate less. In fact, each of these employer categories is associated with more discrimination, even though these effects are noisily estimated. Second, we find no effect of employer size on the degree of discrimination.[31] Point estimates indicate that publicly traded firms seem to discriminate more, while not-for-profit organizations seem to discriminate less; however, these effects are again noisily estimated.

Panel B of Table 9 documents how the racial gap in callbacks varies across broad industry categories. Our sample covers a variety of industries. Around 8% of the jobs are in manufacturing, 3% in transportation and communication, 22% are in wholesale and retail trade, 8% are in finance insurance and real estate, 27% are in business and personal services and 15% are in health, educational and social services. We are not able to determine a recruiter's industry in 16% of the cases.

Columns (2) through (4) show the callback rates for Whites and African Americans, and the differences and ratios of these two. In every industry except for transportation and communication (which corresponds to a very small subsample of jobs), African Americans fare worse than Whites. The biggest gap appears to be in finance, insurance and real estate (ratio of call-back is 2.44), while the smallest (outside of transportation) appears to be in health, educational and social services (ratio is 1.35). While the industrial differences in callbacks appear more pronounced than the

[31] Similar results hold when we measure employer size using a total sales measure rather than an employment measure.

occupational differences that we discussed above, we still cannot reject the null hypothesis that discrimination is the same across industries.

One however report how the point estimates by industry relate to Census measures of earnings and racial gap in earnings by industry. As we did for occupations, we compute such measures using the 1990 5 percent Census for Boston and Chicago. We follow a similar methodology to compute residual log(annual earnings) and mean residual log(annual earnings) by industry and industry-race cell. These Census measures are presented in the last two columns of the table.

Documenting the relationship between inter-industry wage differentials and discrimination is interesting in light of theories that predict that higher industry rents allow for more discrimination. If higher rents translate into higher wages through rent-sharing, this theory predicts high wage industries should discriminate more (Katz And Summers 1989). Such a prediction could evolve from a straightforward taste-based model of discrimination where employers are homogeneous in their dislike of minority workers but are differentially constrained by industry-specific product market pressures in their ability to engage in costly discrimination (Becker, 1961). If one abstracts from the small transportation and communication sector, it does appear that the two highest wage industries (manufacturing and finance, insurance and real estate) are also the two highest discrimination industries. However, this relationship is far from fully monotonous. For example, health, educational and social services have about average wage levels but also record the lowest level of discrimination.

Our final table, Table 11, focuses on the marginal effect of employer location on discrimination. As we mentioned above, we use as a measure of employer location the zip code of the company (or company branch) resumes were to be sent to. More specifically, we ask whether discrimination varies with the fraction African Americans in the employer's zip code. Each column in Table 11 corresponds to a different probit regression of a call-back dummy on an African American dummy, the fraction African Americans in the zip code and the interaction of the African American dummy with fraction African American in the zip code. Reported in the last row of that table is the mean fraction African Americans in the relevant sample. In columns 1 and 2, we use the full sample for which data are available. As we can see in column 1, the interaction term is positive, suggesting that employers located in more African American zip codes are more likely to call back African American applicants. However, the coefficient on this interaction term is not statistically significant.

Column 2 reestimates the same probit regression but allows for the degree of discrimination to vary by occupation categories, industries categories and city. This increases a bit the magnitude the of coefficient on the interaction term but that coefficient is still not significant.

Pooling Chicago and Boston in the regressions in columns 1 and 2 however hides an important difference in behavior across the two cities. The last four columns of Table 11 replicate the exercise above separately for Chicago (columns 3 and 4) and Boston (columns 5 and 6). We find very different patterns across the two cities. In Chicago, there is a larger and statistically significant effect of employer location on discrimination. Each 10-percentage point increases in the fraction African Americans in the employer's zip code reduces the racial gap in call back by close to 1 percentage point. Such a 10 percentage point increases corresponds to a little less than a move from the 25th percentile to the 75th percentile of the fraction African Americans in the Chicago sample. According to our estimates, equality of call-backs across races would occur in a zip code that is about 60 percent African American. This corresponds to about the 97th percentile in the Chicago sample.

On the other hand, we find no significant effect of the African Americanness of an employer's neighborhood in Boston. The estimated coefficient on the interaction term for Boston is both economically and statistically insignificant. One possible reason for this differential effect might be differences in both the level and variation of fraction African Americans across these two cities. The Chicago sample contains many more employers located in fairly African American neighborhoods. The average Chicago employer is located in zip code with 10.5 percent African Americans; the average Boston employer in zip code with 4.5 percent African Americans. The 25th and 75th percentiles of fraction African Americans in Chicago are respectively 1 percent and 12.5 percent; these are 0 percent and 5 percent in Boston.

Finally, in regressions not reported here, we also investigated whether other employers' zip code characteristics affect the level of discrimination in Chicago. Interestingly, we found qualitatively similar results to the ones above when we use fraction Whites, fraction college educated, fraction high school dropouts or median per capita income in the employer's zip code. These results all suggest that employers located in more affluent neighborhoods are more amenable to hiring African American applicants.

5 Interpretation

Could our results be driven by something other than racial discrimination? Two alternatives stand out. First, perhaps what appears as discrimination is actually the result of *reverse discrimination.* If African Americans are thought to be in high demand, then employers with average quality jobs might feel that an equally talented African American would never accept an offer from them and thereby never call her or him in for an interview. Such an argument might also explain why African Americans do not receive as strong of a return as Whites to better resumes since greater qualifications only worsen this problem.[32] But this argument would suggest that among the better jobs, we ought to see evidence of reverse, or at least less, discrimination. However, as we discussed in Section 4.4, we do not find any such evidence. Even among the better jobs in our sample, we find quite a bit of discrimination against African Americans.

Second, perhaps employers are inferring more than just race from applicants' names. More specifically, maybe employers are inferring social class. When employers read a name like "Tyrone" or "Latoya," they may associate that name with the ghetto or other disadvantaged social background. Of course, because African Americans on average do in fact come from poorer backgrounds than Whites, this argument would have to be sharpened. These names would need to be more reflective of economic background than being African American already is. While plausible, several of our results are inconsistent with this interpretation. First, recall that the African American sounding names we use are not as atypical as they may seem. In fact, as Appendix 1 shows, they are quite common among African Americans. Second, for the subset of African American female names where we had access to data on social background (mother's education to be precise), we found no correlation between social background and callback rates. Finally, and perhaps most telling, in Table 7, we found that African Americans are not helped more than Whites by living in more White or more-educated neighborhoods. If the African American names were mostly to signal negative social background, one might have expected a better address to yield greater returns for the African American names than for White names.

What do the results in this paper imply for existing models of discrimination? They suggest more caution in using these models to predict basic patterns of discrimination. The simplest statis-

[32] Another possible interpretation is that employers do not believe the higher quality resumes for African American applicants, finding them "too good to be true." This is of course another form of discrimination. However, given the rather minor resume improvements we consider, we find this explanation hard to believe.

tical discrimination models, for example, would intuitively suggest that racial gaps in employment or earnings should decrease when African Americans improve their observable skills. Since it is the fear of bad unobservables that leads employer to discriminate, reducing this fear should reduce discrimination. But this is counter to what we find. In fact, we find that African Americans have little to gain (in terms of reduced discrimination) from improving their credentials, either within jobs or across jobs. Of course, more complicated models of statistical discrimination could be made to fit the facts reported here. But then these are the models that should be used in the future.

This evidence provides mixed support for taste-based models of discrimination. On the one hand, simple taste-based models may counterfactually suggest that, as observable skills rise, the racial gap should diminish since the opportunity cost of discriminating rises. But taste-based models could be made to fit that fact if one were to assume that employers are especially prejudiced against more skilled African Americans. On the other hand, we do find some evidence suggesting that discrimination is lower when recruiters are located in more African American neighborhoods. This evidence seems more in line with a taste-based model of discrimination, where employers located in more African American neighborhoods are less discriminating, maybe because they are themselves more likely to be African American. This employer location evidence is consistent with customer-based discrimination models. On the other hand, recall that we find that jobs requiring good "communication skills" show no more discrimination yet these are the jobs were customer discrimination may be most important.

As a whole, these results underline the ambiguity of existing models of discrimination. Statistical and taste-based models encompass too large a set of alternatives to be meaningful distinctions. With suitable tweaking of assumptions (either information sets or preferences), each can be made to fit the facts ex post. Assessing the realism of these alternative assumptions may be an important step towards a better understanding of the theoretical nature of racial discrimination.

6 Conclusion

Our evidence suggests that discrimination is an important factor in why African Americans do poorly in the labor market. An African American seeking work gets far fewer callbacks for each resume she or he sends. Equally importantly, African Americans also find it hard to fight discrimination by improving their observable skills or credentials. Our evidence shows that employers reward African Americans far less for observable skills and credentials, possibly dulling their incentives to acquire such skills.

Our results on differential returns to skill also has important policy implications. Training alone may not alleviate the barriers raised by discrimination. For training to work, some force outside the context of our experiment would have to be strong. For example, massive training at the national level may change the structure of discrimination. But small training programs likely will not help. In fact, if African Americans recognize how employers view these skills, they may be rationally reluctant to even participate in them. In general, these results provide one reason for the persistence of racial inequality over time.

References

Altonji, Joseph G and Rebecca M Blank, "Race and Gender in the Labor Market," in Orley Ashenfelter and David Card eds, *Handbook of Labor Economics*, vol. 3, Elsevier Science B.V. (1999).

Arrow, Kenneth J, "What Has Economics to Say About Racial Discrimination?" *The Journal of Economic Perspectives*, 1998, 12 (2): 91-100.

Becker, Gary S., *The Economics of Discrimination*, 2nd Edition, University of Chicago Press: Chicago, IL (1961).

Brown, C. and P. Gay, *Racial Discrimination 17 Years After the Act*, Policy Studies Institute: London, UK (1985).

Council of Economic Advisers, *Changing America: Indicators of Social and Economic Well-Being by Race and Hispanic Origin*, September 1998. *http://w3.access.gpo.gov/eop/ca/pdfs/ca.pdf.*

Cross, H, G Kennedy, J Mell and W Zimmermann, *Employer Hiring Practices: Differential Treatment of Hispanic and Anglo Job Applicants*, Urban Institute Press, Washington, DC (1990).

Darity, William A Jr and Patrick L Mason, "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender," *The Journal of Economic Perspectives*, 1998, 12 (2): 63-90.

Fix, Michael and Margery Austin Turner (eds.), *A National Report Card on Discrimination in America: The Role of Testing*, 1998.

Goldin, Claudia and Cecilia Rouse, "Orchestrating Impartiality: The Impact of Blind Auditions on Female Musicians," *The American Economic Review*, 2000.

Heckman, James J, "Detecting Discrimination," *The Journal of Economic Perspectives,* 1998, 12 (2): 101-116.

Heckman, James J and Peter Siegelman, "The Urban Institute Audit Studies: Their Methods and Findings," in Michael Fix and Raymond J Struyk, *Clear and Convincing Evidence: Measurement of Discrimination in America*, Urban Institute Press, Lanham, MD (1992).

Hubbock, J. and S. Carter, *Half a Chance? A Report on Job Discrimination against Young Blacks in Nottingham*, Commission for Racial Equality: London, UK (1980).

James F and S W DelCastillo, "Measuring Job Discrimination by Private Employers Against Young Black and Hispanic Seeking Entry Level Work in the Denver Metropolitan Area," Mimeo, University of Colorado at Denver, Denver, 1991.

Jowell, R abd Prescott-Clark P, "Racial Discrimination and White-Collar Workers in Britain," *Race*, 1970, 11: 397-417.

Katz, Lawrence F. and Lawrence H. Summers, 1989, Industry Rents: Evidence and Implications, Brookings Papers on Economic Activity, Microeconomics, 209-275.

McIntyre, Shelby J, Dennis J Moberg and Barry Z Posner, "Discrimination in Recruitment: An Empirical Analysis: Comment,"*Industrial and Labor Relations Review*, 1980, 33 (4): 543-547.

Newman, Jerry M, "Discrimination in Recruitment: An Empirical Analysis," *Industrial and Labor Relations Review*, 1980, 32 (1): 15-23.

Turner, Margery A, Michael Fix and Raymond J Struyk, *Opportunities Denied, Opportunities Diminished: Racial Discrimination in Hiring,* Urban Institute Press: Washington, DC (1991).

Turner, Margery A, Raymond J Struyk and J Yinger, *Housing Discrimination Study Synthesis*, Urban Institute Press: Washington, DC (1991).

Wienk, R E, J C Simonson and F J Eggers, *Measuring Discrimination in American Housing Markets: The Housing Market Practices Survey,* US Department of Housing and Urban Development: Washington, DC (1979).

Table 1
Mean Call-back Rates By Racial Soundingness of Names [a]

	Call-back Rate for White Names	*Call-back Rate for Black Names*	*Ratio*	*Difference (p-value)*
Sample:				
All sent resumes	**10.06%** [2445]	**6.70%** [2445]	**1.50**	**3.35%** (.0000)
Chicago	**8.61%** [1359]	**5.81%** [1359]	**1.48**	**2.80%** (.0024)
Boston	**11.88%** [1086]	**7.83%** [1086]	**1.52**	**4.05%** (.0008)
Females	**10.33%** [1868]	**6.87%** [1893]	**1.50**	**3.46%** (.0001)
Females in administrative jobs	**10.93%** [1363]	**6.81%** [1364]	**1.60**	**4.12%** (.0001)
Females in sales jobs	**8.71%** [505]	**6.99%** [529]	**1.25**	**1.72%** (.1520)
Males	**9.19%** [577]	**6.16%** [552]	**1.49**	**3.03%** (.0283)

[a] Notes:

1. The table reports, for the entire sample and different subsamples, the call-back rates for applicants with a white-sounding name (column 1) and a black-sounding name (column 2), as well as the ratio (column 3) and difference (column 4) of these call-back rates. In brackets in each cell is the number of resumes sent in that cell.
2. Column 4 also reports the p-value for a test of proportion testing the null hypothesis that the call-back rates are equal across racial groups.

Table 2
Mean Call-Back Rates By First Name[a]

White Female

Name	Mean Call-back
Emily	**8.3%** [228]
Anne	**9.0%** [244]
Jill	**9.3%** [204]
Allison	**9.4%** [233]
Sarah	**9.8%** [193]
Meredith	**10.6%** [189]
Laurie	**10.8%** [195]
Carrie	**13.1%** [168]
Kristen	**13.6%** [214]
Ho: All white female first name effects are the same	p-value: .7574

Black Female

Name	Mean Call-back
Aisha	**2.2%** [180]
Keisha	**3.8%** [183]
Tamika	**5.4%** [258]
Lakisha	**5.5%** [200]
Tanisha	**6.3%** [207]
Latoya	**8.8%** [227]
Kenya	**9.1%** [198]
Latonya	**9.1%** [231]
Ebony	**10.5%** [209]
Ho: All black female first name effects are the same	p-value: .0150

White Male

Name	Mean Call-back
Neil	**6.6%** [76]
Geoffrey	**6.8%** [59]
Brett	**6.8%** [59]
Brendan	**7.7%** [65]
Greg	**7.8%** [51]
Todd	**8.7%** [69]
Matthew	**9.0%** [67]
Jay	**13.2%** [68]
Brad	**15.9%** [63]
Ho: All white male first name effects are the same	p-value: .6427

Black Male

Name	Mean Call-back
Rasheed	**3.0%** [67]
Tremayne	**4.3%** [70]
Kareem	**4.7%** [64]
Darnell	**4.8%** [42]
Tyrone	**5.3%** [76]
Jamal	**6.6%** [61]
Hakim	**7.3%** [55]
Leroy	**9.4%** [64]
Jermaine	**11.3%** [53]
Ho: All black male first name effects are the same	p-value: .7067

[a]Notes:

1. Sample: All sent resumes.
2. This table reports call-back rates by first name. Within each sex/race group, first names are ranked by increasing call-back rate. In brackets in each cell is the number of resumes sent in that cell.
3. The four tests reported in the table are log-likelihood tests obtained from a single probit regression of a dummy variable for whether a call-back was recorded or not on all the first name fixed effects.

Table 3
Distribution of Call-Backs By Employment Ad [a]

Equal Treatment:	*No Call-back*	*1W+1B*	*2W+2B*
87.37%	**82.56%**	**3.46%**	**1.35%**
[1162]	[1098]	[46]	[18]
Whites Favored (WF):	*1W+0B*	*2W+0B*	*2W+1B*
8.87%	**5.93%**	**1.50%**	**1.43%**
[118]	[79]	[20]	[19]
Blacks Favored (BF):	*1B+0W*	*2B+0W*	*2B+1W*
3.76%	**2.78%**	**.45%**	**.53%**
[50]	[37]	[6]	[7]
Ho: WF=BF			
(p-value)			
(.0000)			

[a] Notes:

1. Sample: all sent resumes.
2. This table documents the distribution of call-backs at the employment ad level. "No Call-Back" is the fraction of ads for which none of the experimental applicant received a call-back. "1W+1B" is the fraction of ads for which exactly one white and one black applicant received a call-back. "2W+2B" is the fraction of ads for which exactly two white applicants and two black applicants received a call-back. "Equal Treatment" is defined as the sum of "No Call-Back," "1W+1B," "2W+2B." "1W+0B" is the fraction of ads for which exactly one white applicant and no black applicant received a call back. "2W+0B" is the fraction of ads for which exactly two white applicants and no black applicant received a call-back. "2W+1B" is the fraction of ads for which exactly two white applicants and one black applicant received a call-back. "White Favored" is defined as the sum of "1W+0B," "2W+0B," and "2W+1B." "1B+0W" is the fraction of ads for which exactly one black applicant and no white applicant received a call-back. "2B+0W" is the fraction of ads for which exactly two blacks applicants and no white applicant received a call-back. "2B+1W" is the fraction of ads for which exactly two black applicants and one white applicant received a call-back. "Black Favored" is defined as the sum of "1B+0W," "2B+0W," and "2B+1W."
3. In brackets in each cell is the number of employment ads in that cell.

Table 4
Resume Characteristics: Summary Statistics [a]

	Sample:				
	All Resumes	**White Names**	**Black Names**	**Higher Quality**	**Lower Quality**
Characteristic:					
Years of experience	7.82	7.84	7.81	8.27	7.38
	(5.04)	(5.07)	(5.00)	(5.28)	(4.75)
Volunteering experience?	.42	.41	.41	.79	.03
(Y=1)	(.49)	(.49)	(.49)	(.41)	(.16)
Military experience?	.10	.09	.10	.18	.00
(Y=1)	(.30)	(.29)	(.30)	(.39)	(.06)
Email address?	.48	.48	.48	.92	.03
(Y=1)	(.50)	(.50)	(.50)	(.27)	(.17)
Employment holes?	.45	.45	.45	.34	.56
(Y=1)	(.50)	(.50)	(.50)	(.47)	(.50)
Work in school?	.56	.56	.56	.72	.40
(Y=1)	(.50)	(.50)	(.50)	(.45)	(.49)
Honors?	.05	.05	.05	.07	.03
(Y=1)	(.22)	(.23)	(.22)	(.25)	(.18)
Computer skills?	.82	.81	.83	.91	.73
(Y=1)	(.38)	(.39)	(.37)	(.29)	(.44)
Special skills?	.33	.33	.33	.36	.30
(Y=1)	(.47)	(.47)	(.47)	(.48)	(.46)
Fraction high school dropouts	.19	.19	.19	.19	.18
in applicant's zip code	(.08)	(.08)	(.08)	(.08)	(.08)
Fraction college or more	.21	.21	.21	.21	.21
in applicant's zip code	(.17)	(.17)	(.17)	(.17)	(.17)
Fraction whites	.54	.54	.54	.54	.55
in applicant's zip code	(.33)	(.33)	(.33)	(.33)	(.33)
Fraction blacks	.31	.31	.31	.32	.31
in applicant's zip code	(.33)	(.33)	(.33)	(.33)	(.33)
Log(per capita income)	9.55	9.55	9.55	9.54	9.56
in applicant's zip code	(.56)	(.56)	(.55)	(.54)	(.57)
Sample Size	4890	2445	2445	2458	2432

[a] Notes:

1. The table reports means and standard deviations for the resume characteristics. Column (1) refers to all resumes sent; column (2) refers to resumes with white-sounding names; column (3) refers to resumes with black-sounding names; column (4) refers to higher quality resumes; column (5) refers to lower quality resumes.

Table 5
Mean Call-Back Rates By Race and Resume Quality [a]

Panel A: Subjective Measure of Quality

	Low	High	Ratio	Difference (p-value)
White Names	**8.80%**	**11.31%**	**1.29**	**2.51%**
	[1216]	[1229]		(.0391)
Black Names	**6.41%**	**6.99%**	**1.09**	**0.58%**
	[1216]	[1229]		(.5644)

Panel B: Predicted Measure of Quality

	Low	High	Ratio	Difference (p-value)
White Names	**5.04%**	**14.18%**	**2.81**	**9.14%**
	[834]	[804]		(.0000)
Black Names	**5.14%**	**8.58%**	**1.66**	**3.44%**
	[817]	[816]		(.0060)

[a]Notes:

1. Panel A reports the mean call-back rates for applicants with a white-sounding name (raw 1) and black-sounding name (raw 2) depending on whether the resume was qualified as a lower quality (column 1) or higher quality (column 2). In brackets is the number of resumes sent for each race/quality group. Column 4 reports the p-value of a test of proportion testing the null hypothesis that the call-back rates are equal across quality groups within each racial group.

2. For Panel B, we use a third of the sample to estimate a probit regression of the call-back dummy on the set of resume characteristics as displayed in Table 4. We further control for a sex dummy, a city dummy, 6 occupation dummies and a vector of dummy variables for job requirements as listed in the employment ad (see Section 4.4 for details). We then use the estimated coefficients on the set of resume characteristics to estimate a predicted call-back for the remaining resumes (2/3 of the sample). We call "high quality" resumes the resumes that rank above the median predicted call-back and "low quality" resumes the resumes that rank below the median predicted call-back. In brackets is the number of resumes sent for each race/quality group. Column 4 reports the p-value of a test of proportion testing the null hypothesis that the call-back rates are equal across quality groups within each racial group.

Table 6
Effect of Resume Characteristics on Likelihood of Call-Back [a]

Dependent Variable: Call-Back Dummy

Sample:	**All Resumes**	**White Names**	**Black Names**
Years of experience (*10)	.07	.13	.02
	(.03)	(.04)	(.03)
Years of experience^2 (*100)	-.02	-.04	-.00
	(.01)	(.02)	(.01)
Volunteering? (Y=1)	-.01	-.01	.00
	(.01)	(.02)	(.01)
Military experience? (Y=1)	-.00	.01	-.01
	(.02)	(.02)	(.02)
Email? (Y=1)	.02	.03	.00
	(.01)	(.01)	(.01)
Employment holes? (Y=1)	.02	.03	.01
	(.01)	(.02)	(.01)
Work in school? (Y=1)	.01	.02	-.00
	(.01)	(.01)	(.01)
Honors? (Y=1)	.05	.07	.02
	(.02)	(.03)	(.02)
Computer skills? (Y=1)	-.02	-.03	-.00
	(.01)	(.02)	(.01)
Special skills? (Y=1)	.05	.07	.04
	(.01)	(.02)	(.01)
Ho: Resume characteristics effects are all zero (p-value)	55.73 (.0000)	59.83 (.0000)	20.78 (.0227)
Standard deviation of predicted call-back	.047	.064	.034
Sample size	4890	2445	2445

[a] Notes:

1. Each column gives the results of a probit regression where the dependent variable is a call-back dummy. Reported in the table are the estimated marginal change in probability for the continuous variables and the estimated discrete change for the dummy variables. Also included in each regression are a city dummy, a sex dummy, 6 occupation dummies and a vector of dummy variables for job requirements as listed in the employment ad (see Section 4.4 for details).
2. Sample in column (1) is the entire set of sent resumes; sample in column (2) is the set of resumes with white-sounding names; sample in column (3) is the set of resumes with black-sounding names.
3. Standard errors are corrected for clustering of the observations at the employment ad level.
4. Reported in the second to last row are the p-value for a χ^2 testing that the effects on the resume characteristics are all zero.
5. Reported in the last row are the standard deviations of the predicted call-back rate.

Table 7
Effect of Applicant's Address on Likelihood of Call-Back [a]

Dependent Variable: Call-Back Dummy

Zip code characteristic:	**Fraction whites**		**Fraction college or more**		**Log(per capita income)**	
Zip code characteristic	.021 (.012)	.023 (.016)	.057 (.023)	.055 (.031)	.019 (.007)	.014 (.010)
Zip code characteristic* Black	—	-.005 (.025)	—	.002 (.050)	—	.010 (.015)

[a] Notes:

1. Each column gives the results of a probit regression where the dependent variable is the call-back dummy. Reported in the table are the estimated marginal change in probability. Also included in columns (1), (3) and (5) is a city dummy; also included in columns (2), (4) and (6) is a city dummy and a city dummy interacted with a race dummy.
2. Sample in all regressions is the entire set of sent resumes ($N = 4890$).
3. Standard errors are corrected for clustering of the observations at the employment ad level.

Table 8
Job Requirements [a]

Requirement:	**Sample Mean (st. dev.)**	**Marginal Effect on African-American Call Backs**
Any requirement? (Y=1)	.80	.024
	(.41)	(.015)
Experience? (Y=1)	.43	-.019
	(.49)	(.012)
of which:		
some	50.2%	
two years or less	24.8%	
three years or more	25.0%	
Computer skills? (Y=1)	.44	-.004
	(.50)	(.013)
Communication skills? (Y=1)	.12	.000
	(.33)	(.018)
Organization skills? (Y=1)	.07	.019
	(.26)	(.026)
Education? (Y=1)	.11	-.025
	(.31)	(.019)
of which:		
high school degree	8.8%	
some college	48.5%	
4-year college degree	42.7%	
Total number of requirements	1.18	.001
	(.93)	(.006)

[a]Notes:

1. Column (2) reports means and standard deviations (in parentheses) for the job requirements.
2. Column (3) reports the marginal effect of the job requirement listed in that row on discrimination. Specifically, each cell in column (3) corresponds to a different probit regression of the call-back dummy on a black dummy, a dummy for the requirement listed in that row and the interaction of the requirement dummy with the black dummy. Reported in each cell is the estimated discrete change for the interaction term. Standard errors are corrected for clustering of the observations at the employment ad level.
3. Sample is all sent resumes ($N = 4890$).

Table 9
Racial Gap in Call-back by Occupation and Industry [a]

Panel A: Occupation Break-Down

	% of Ads	*Call-back Rates for* *White Names*	*Black Names*	*Ratio*	*Difference*	*Log(W)*	*Race Gap*
Executive and managerial	14.5%	7.91%	5.95%	1.33	1.96%	.28	.16
Administrative supervisors	7.7%	9.57%	5.85%	1.64	3.72%	.23	-.02
Sales representatives	15.2%	8.04%	5.09%	1.58	2.95%	.08	-.08
Sales workers, retail and personal services	16.8%	10.46%	7.05%	1.48	3.41%	-.07	.01
Secretaries	33.9%	10.49%	6.63%	1.58	3.86%	.16	.31
Clerical workers, admin. support	11.9%	13.75%	9.96%	1.38	3.79%	-.45	.08

H_0: Racial gap is the same across occupations
p-value=.975

Panel B: Industry Break-Down

	% of Ads	*Call-back Rates for* *White Names*	*Black Names*	*Ratio*	*Difference*	*1990 Census Statistics* *Log(W)*	*Racial Gap in Log(W)*
Manufacturing	8.3%	6.93%	3.96%	1.75	2.97%	.14	.15
Transportation and communication	3.0%	12.16%	14.86%	.82	-2.70%	.21	.11
Wholesale and retail trade	21.5%	8.76%	5.71%	1.53	3.05%	-.17	.19
Finance, insurance and real estate	8.5%	10.63%	4.35%	2.44	6.28%	.17	.11
Business and personal services	26.8%	11.30%	6.71%	1.68	4.59%	-.15	.21
Health, educational and social services	15.5%	12.14%	9.50%	1.28	2.64%	-.04	.13
Other/unknown	16.4%	8.71%	6.47%	1.35	2.24%	—	—

H_0: Racial gap is the same across industries
p-value=.1923

[a]Notes:

1. This table reports call-back rates by race and occupation (Panel A) and by race and industry (Panel B). Sample is all sent resumes ($N = 4890$).
2. The two tests reported in the table are log-likelihood tests obtained from two separate probit regressions. In Panel A, we regress the call-back dummy on 6 occupation dummies, a black dummy and the interaction of the black dummy with the six occupation dummies. In Panel B, we regress the call-back dummy on 7 industry dummies, a black dummy and the interactions of the black dummy with the 7 industry dummies. In each case, the null hypothesis tested is that the interaction term effects are all the same.
3. The last two columns of the table report earnings statistics from the 1990 5 percent Census for Boston and Chicago. The first of these two columns reports mean log(residual annual earnings) in the occupation or industry category. The second column reports mean white-black gap in log(residual annual earnings) in the occupation or industry category. Log(residual annual earnings) are obtained from a micro wage regression of log(annual earnings) on 10 education dummies, a quadratic in age, a sex dummy and a city dummy.

Table 10
Employer Characteristics [a]

Characteristic:	Sample Mean (st. dev.)	Marginal Effect on African-American Call Backs
Equal opportunity employer? (Y=1)	.29	-.010
(N=4890)	(.45)	(.012)
Federal contractor? (Y=1)	.11	-.027
(N=3118)	(.31)	(.018)
Log(employment)	5.74	-.000
(N=1702)	(1.74)	(.032)
Ownership status:		
(N=2894)		
Privately held	73.0%	.003
		(.019)
Publicly traded	15.5%	-.023
		(.016)
Not-for-profit	11.5%	.040
		(.045)

[a] Notes:

1. Column (2) reports means and standard deviations (in parentheses) for the employer characteristics. Sample sizes for each characteristic are reported in column (1).
2. Column (3) reports the marginal effect of the employer characteristic listed in that row on discrimination. Specifically, each cell in column (3) corresponds to a different probit regression of the call-back dummy on a black dummy, a dummy for the employer characteristic listed in that row and the interaction of the employer characteristic with the black dummy. Reported in each cell is the estimated coefficient on the interaction term. Standard errors are corrected for clustering of the observations at the employment ad level.

Table 11
Effect of Employer's Address on Likelihood of Call-Back [a]

Dependent Variable: Call-Back Dummy

Sample:	**Both Cities**		**Chicago**		**Boston**	
Black	-.039	—	-.055	—	-.019	—
	(.010)		(.013)		(.019)	
% blacks in employer's zip code	-.008	-.012	-.044	-.058	.205	.171
	(.054)	(.048)	(.048)	(.047)	(.172)	(.110)
Black*	.059	.075	.087	.086	.016	-.032
%blacks in employer's zip code	(.071)	(.060)	(.044)	(.046)	(.307)	(.201)
Industry dummies	No	Yes	No	Yes	No	Yes
Black*industry dummies	No	Yes	No	Yes	No	Yes
Occupation dummies	No	Yes	No	Yes	No	Yes
Black*occupation dummies	No	Yes	No	Yes	No	Yes
City dummy	No	Yes	No	Yes	No	Yes
Black*city dummy	No	Yes	No	Yes	No	Yes
Mean % blacks in employer's zip code	.082		.106		.047	
	(.154)		(.185)		(.085)	
Sample size:	1930		1142		788	

[a]Notes:

1. Each column gives the results of a probit regression where the dependent variable is a call-back dummy. Reported in the table are the estimated marginal change in probability for the continous variable and the estimated discrete change for the dummy variables.
2. Sample in all regressions is the set of sent resumes for which we could determine the employer's zip code.
3. Standard errors are corrected for clustering of the observations at the employment ad level.

Appendix Table 1
First Names[a]

White-Sounding		Black-Sounding	
Name	Frequency	Name	Frequency
Females			
Allison	4.7%	Aisha	3.6%
Anne	5.0%	Ebony	4.3%
Carrie	3.5%	Keisha	3.7%
Emily	4.7%	Kenya	4.0%
Jill	4.2%	Latonya	4.7%
Laurie	4.0%	Lakisha	4.1%
Kristen	4.4%	Latoya	4.6%
Meredith	3.9%	Tamika	5.3%
Sarah	3.9%	Tanisha	4.2%
Fraction of All Births:		Fraction of All Births	
3.8%		7.1%	
Males			
Brad	1.3%	Darnell	0.9%
Brendan	1.3%	Hakim	1.1%
Geoffrey	1.2%	Jermaine	1.1%
Greg	1.0%	Kareem	1.3%
Brett	1.2%	Jamal	1.2%
Jay	1.4%	Leroy	1.3%
Matthew	1.4%	Rasheed	1.4%
Neil	1.6%	Tremayne	1.4%
Todd	1.4%	Tyrone	1.6%
Fraction of All Births:		Fraction of All Births	
1.7%		3.1%	

[a]Notes:

1. This table tabulates the different first names used in the experiment and the frequencies with which each of these names was used. Also reported for each race-sex category is the fraction of all births in that race-sex category with these first names (from the Massachusetts birth certificates, 1974 to 1979).