

Discrimination

70 ■ Marianne Bertrand and Sendhil Mullainathan

Are Emily and Greg More Employable Than Lakisha and Jamal?

A Field Experiment on Labor Market Discrimination

Every measure of economic success reveals significant racial inequality in the U.S. labor market. Compared to Whites, African-Americans are twice as likely to be unemployed and earn nearly 25 percent less when they are employed (Council of Economic Advisers, 1998). This inequality has sparked a debate as to whether employers treat members of different races differentially. When faced with observably similar African-American and White applicants, do they favor the White one? Some argue yes, citing either employer prejudice or employer perception that race signals lower productivity. Others argue that differential treatment by race is a relic of the past, eliminated by some combination of employer enlightenment, affirmative action programs and the profit-maximization motive. In fact, many in this latter camp even feel that stringent enforcement of affirmative action programs has produced an environment of reverse discrimination. They would argue that faced with identical candidates, employers might favor the African-American one. Data limitations make it difficult to empirically test these views. Since researchers possess

far less data than employers do, White and African-American workers that appear similar to researchers may look very different to employers. So any racial difference in labor market outcomes could just as easily be attributed to differences that are observable to employers but unobservable to researchers.

To circumvent this difficulty, we conduct a field experiment that builds on the correspondence testing methodology that has been primarily used in the past to study minority outcomes in the United Kingdom. We send resumes in response to help-wanted ads in Chicago and Boston newspapers and measure callback for interview for each sent resume. We experimentally manipulate perception of race via the name of the fictitious job applicant. We randomly assign very White-sounding names (such as Emily Walsh or Greg Baker) to half the resumes and very African-American-sounding names (such as Lakisha Washington or Jamal Jones) to the other half. Because we are also interested in how credentials affect the racial gap in callback, we experimentally vary the quality of the resumes used in response to a given ad.

Marianne Bertrand and Sendhil Mullainathan. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," in *The American Economic Review* (September, 2004), pp. 991-992, 1006-1007, 1009-1013. Reprinted by permission of *The American Economic Review*.

Higher-quality applicants have on average a little more labor market experience and fewer holes in their employment history; they are also more likely to have an e-mail address, have completed some certification degree, possess foreign language skills, or have been awarded some honors. In practice, we typically send four resumes in response to each ad: two higher-quality and two lower-quality ones. We randomly assign to one of the higher- and one of the lower-quality resumes an African-American-sounding name. In total, we respond to over 1,300 employment ads in the sales, administrative support, clerical, and customer services job categories and send nearly 5,000 resumes. The ads we respond to cover a large spectrum of job quality, from cashier work at retail establishments and clerical work in a mail room, to office and sales management positions.

We find large racial differences in callback rates. Applicants with White names need to send about 10 resumes to get one callback whereas applicants with African-American names need to send about 15 resumes. This 50-percent gap in callback is statistically significant. A White name yields as many more callbacks as an additional eight years of experience on a resume. Since applicants' names are randomly assigned, this gap can only be attributed to the name manipulation.

Race also affects the reward to having a better resume. Whites with higher-quality resumes receive nearly 30-percent more callbacks than Whites with lower-quality resumes. On the other hand, having a higher-quality resume has a smaller effect for African-Americans. In other words, the gap between Whites and African-Americans widens with resume quality. While one may have expected improved credentials to alleviate employers' fear that African-American applicants are deficient in some unobservable skills, this is not the case in our data.

The experiment also reveals several other aspects of the differential treatment by race. First, since we randomly assign applicants'

postal addresses to the resumes, we can study the effect of neighborhood of residence on the likelihood of callback. We find that living in a wealthier (or more educated or Whiter) neighborhood increases callback rates. But, interestingly, African-Americans are not helped more than Whites by living in a "better" neighborhood. Second, the racial gap we measure in different industries does not appear correlated to Census-based measures of the racial gap in wages. The same is true for the racial gap we measure in different occupations. In fact, we find that the racial gaps in callback are statistically indistinguishable across all the occupation and industry categories covered in the experiment. Federal contractors, who are thought to be more severely constrained by affirmative action laws, do not treat the African-American resumes more preferentially; neither do larger employers or employers who explicitly state that they are "Equal Opportunity Employers." In Chicago, we find a slightly smaller racial gap when employers are located in more African-American neighborhoods. . . .¹

Interpretation

Three main sets of questions arise when interpreting the results above. First, does a higher callback rate for White applicants imply that employers are discriminating against African-Americans? Second, does our design only isolate the effect of race or is the name manipulation conveying some other factors than race? Third, how do our results relate to different models of racial discrimination?

Interpreting Callback Rates

Our results indicate that for two identical individuals engaging in an identical job search, the one with an African-American name would receive fewer interviews. Does differential treatment within our experiment imply that employers are discriminating against African-Americans (whether it is ra-

tional, prejudice-based, or other form of discrimination)? In other words, could the lower callback rate we record for African-American resumes *within our experiment* be consistent with a racially neutral review of the *entire pool* of resumes the surveyed employers receive?

In a racially neutral review process, employers would rank order resumes based on their quality and call back all applicants that are above a certain threshold. Because names are randomized, the White and African-American resumes we send should rank similarly on average. So, irrespective of the skill and racial composition of the applicant pool, a race-blind selection rule would generate equal treatment of Whites and African-Americans. So our results must imply that employers use race as a factor when reviewing resumes, which matches the legal definition of discrimination.

But even rules where employers are not trying to interview as few African-American applicants as possible may generate observed differential treatment in our experiment. One such hiring rule would be employers trying to interview a target level of African-American candidates. For example, perhaps the average firm in our experiment aims to produce an interview pool that matches the population base rate. This rule could produce the observed differential treatment if the average firm receives a higher proportion of African-American resumes than the population base rate because African-Americans disproportionately apply to the jobs and industries in our sample.

Some of our other findings may be consistent with such a rule. For example, the fact that "Equal Opportunity Employers" or federal contractors do not appear to discriminate any less may reflect the fact that such employers receive more applications from African-Americans. On the other hand, other key findings run counter to this rule. As we discuss above, we find no systematic difference in the racial gap in callback across

occupational or industry categories, despite the large variation in the fraction of African-Americans looking for work in those categories. African-Americans are underrepresented in managerial occupations, for example. If employers matched base rates in the population, the few African-Americans who apply to these jobs should receive a higher callback rate than Whites. Yet, we find that the racial gap in managerial occupations is the same as in all the other job categories. This rule also runs counter to our findings on returns to skill. Suppose firms are struggling to find White applicants but overwhelmed with African-American ones. Then they should be less sensitive to the quality of White applicants (as they are trying to fill in their hiring quota for Whites) and much more sensitive to the quality of Black applicants (when they have so many to pick from). Thus, it is unlikely that the differential treatment we observe is generated by hiring rules such as these.

Potential Confounds

While the names we have used in this experiment strongly signal racial origin, they may also signal some other personal trait. More specifically, one might be concerned that employers are inferring social background from the personal name. When employers read a name like "Tyrone" or "Latoya," they may assume that the person comes from a disadvantaged background. In the extreme form of this social background interpretation, employers do not care at all about race but are discriminating only against the social background conveyed by the names we have chosen.

While plausible, we feel that some of our earlier results are hard to reconcile with this interpretation. For example, we found that while employers value "better" addresses, African-Americans are not helped more than Whites by living in Whiter or more educated neighborhoods. If the African-American names we have chosen mainly signal negative social background, one might have expected

the estimated name gap to be lower for better addresses. Also, if the names mainly signal social background, one might have expected the name gap to be higher for jobs that rely more on soft skills or require more interpersonal interactions. We found no such evidence.

There is one final potential confound to our results. Perhaps what appears as a bias against African-Americans is actually the result of *reverse discrimination*. If qualified 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 a return as Whites to better resumes, since higher qualification only strengthens this argument. But this interpretation would suggest that among the better jobs, we ought to see evidence of reverse discrimination, or at least a smaller racial gap. However, we do not find any such evidence. The racial gap does not vary across jobs with different skill requirements, nor does it vary across occupation categories. Even among the better jobs in our sample, we find that employers significantly favor applicants with White names.

Relation to Existing Theories

What do these results imply for existing models of discrimination? Economic theories of discrimination can be classified into two main categories: taste-based and statistical discrimination models. Both sets of models can obviously "explain" our average racial gap in callbacks. But can these models explain our other findings? More specifically, we discuss the relevance of these models with a focus on two of the facts that have been uncovered in this [chapter]: (i) the lower returns to credentials for African-Americans; (ii) the relative uniformity of the race gap across occupations, job requirements and, to a lesser extent, employer characteristics and industries.

Taste-based models (Gary S. Becker, 1961) differ in whose prejudiced "tastes" they emphasize: customers, coworkers, or employers. Customer and co-worker discrimination models seem at odds with the lack of significant variation of the racial gap by occupation and industry categories, as the amount of customer contact and the fraction of White employees vary quite a lot across these categories. We do not find a larger racial gap among jobs that explicitly require "communication skills" and jobs for which we expect either customer or coworker contacts to be higher (retail sales for example).

Because we do not know what drives employer tastes, employer discrimination models could be consistent with the lack of occupation and industry variation. Employer discrimination also matches the finding that employers located in more African-American neighborhoods appear to discriminate somewhat less. However, employer discrimination models would struggle to explain why African-Americans get relatively lower returns to their credentials. Indeed, the cost of indulging the discrimination taste should increase as the minority applicants' credentials increase.

Statistical discrimination models are the prominent alternative to the taste-based models in the economics literature. In one class of statistical discrimination models, employers use (observable) race to proxy for *unobservable* skills (e.g., Edmund S. Phelps, 1972; Kenneth J. Arrow, 1973). This class of models struggles to explain the credentials effect as well. Indeed, the added credentials should lead to a larger update for African-Americans and hence greater returns to skills for that group.

A second class of statistical discrimination models "emphasize[s] the precision of the information that employers have about individual productivity" (Altonji and Blank, 1999). Specifically, in these models, employers believe that the same observable signal is more precise for Whites than for African-Americans (Dennis J. Aigner and Glenn G. Cain, 1977; Shelly

J. Lundberg and Richard Startz, 1983; Bradford Cornell and Ivø Welch, 1996). Under such models, African-Americans receive lower returns to observable skills because employers place less weight on these skills. However, how reasonable is this interpretation for our experiment? First, it is important to note that we are using the same set of resume characteristics for both racial groups. So the lower precision of information for African-Americans cannot be that, for example, an employer does not know what a high school degree from a very African-American neighborhood means (as in Aigner and Cain, 1977). Second, many of the credentials on the resumes are in fact externally and easily verifiable, such as a certification for a specific software.

An alternative version of these models would rely on bias in the observable signal rather than differential variance or noise of these signals by race. Perhaps the skills of African-Americans are discounted because affirmative action makes it easier for African-Americans to get these skills. While this is plausible for credentials such as an employee-of-the-month honor, it is unclear why this would apply to more verifiable and harder skills. It is equally unclear why work experience would be less rewarded since our study suggests that getting a job is more, not less, difficult for African-Americans.

The uniformity of the racial gap across occupations is also troubling for a statistical discrimination interpretation. Numerous factors that should affect the level of statistical discrimination, such as the importance of unobservable skills, the observability of qualifications, the precision of observable skills and the ease of performance measurement, may vary quite a lot across occupations.

This discussion suggests that perhaps other models may do a better job at explaining our findings. One simple alternative model is lexicographic search by employers. Employers receive so many resumes that they may use quick heuristics in reading these resumes. One such heuristic could be to simply read

no further when they see an African-American name. Thus they may never see the skills of African-American candidates and this could explain why these skills are not rewarded. This might also to some extent explain the uniformity of the race gap since the screening process (i.e., looking through a large set of resumes) may be quite similar across the variety of jobs we study.

Conclusion

This paper suggests that African-Americans face differential treatment when searching for jobs and this may still be a factor in why they do poorly in the labor market. Job applicants with African-American names get far fewer callbacks for each resume they send out. Equally importantly, applicants with African-American names find it hard to overcome this hurdle in callbacks by improving their observable skills or credentials.

Taken at face value, our results on differential returns to skill have possibly important policy implications. They suggest that training programs alone may not be enough to alleviate the racial gap in labor market outcomes. For training to work, some general-equilibrium force outside the context of our experiment would have to be at play. In fact, if African-Americans recognize how employers reward their skills, they may rationally be less willing than Whites to even participate in these programs.

NOTE

1. For further details on this experiment, see Bertrand and Mullainathan (2004).

REFERENCES

- Aigner, Dennis J. and Cain, Glenn G. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review*, January 1977, 30(1), pp. 175-87.
- Altonji, Joseph G. and Blank, Rebecca M. "Race and Gender in the Labor Market," in Orley Ashenfelter

- and David Card, eds., *Handbook of labor economics*, Vol. 30. Amsterdam: North-Holland, 1999, pp. 3143–259.
- Arrow, Kenneth J. "The Theory of Discrimination," in Orley Ashenfelter and Albert Rees, eds., *Discrimination in labor markets*. Princeton, NJ: Princeton University Press, 1973, pp. 3–33.
- Becker, Gary S. *The economics of discrimination*, 2nd Ed. Chicago: University of Chicago Press, 1961.
- Bertrand, Marianne and Sendhil Mullainathan. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94 (September 2004): 991–1013.
- Cornell, Bradford and Welch, Ivo. "Culture, Information, and Screening Discrimination." *Journal of Political Economy*, June 1996, 104(3), pp. 542–71.
- 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>.
- Lundberg, Shelly J. and Starz, Richard. "Private Discrimination and Social Intervention in Competitive Labor Market." *American Economic Review*, June 1983, 73(3), pp. 340–47.
- Phelps, Edmund S. "The Statistical Theory of Racism and Sexism." *American Economic Review*, September 1972, 62(4), pp. 659–61.

71 ■ Claude Steele

Stereotype Threat and African-American Student Achievement

Over the past four decades African-American college students have been more in the spotlight than any other American students. This is because they aren't just college students; they are a cutting edge in America's effort to integrate itself in the nearly forty years since the passage of the Civil Rights Act. These students have borne much of the burden for our national experiment in racial integration. And to a significant degree the success of the experiment will be determined by their success.

Nonetheless, throughout the 1990s the national college dropout rate for African Americans has been 20 to 25 percent higher than that for whites. Among those who finish college, the grade point average of Black

students is two-thirds of a grade below that of whites. . . .

Virtually all aspects of underperformance—lower standardized test scores, lower college grades, lower graduation rates—persist among students from the African-American middle class. This situation forces on us an uncomfortable recognition: that beyond class, something racial is depressing the academic performance of these students.

Some time ago two of my colleagues, Joshua Aronson and Steven Spencer, and I tried to see the world from the standpoint of African-American students, concerning ourselves less with features of theirs that might explain their troubles than with features of the world they see. A story I was told recently

Claude Steele. "Stereotype Threat and African-American Student Achievement," in *Young, Gifted, and Black: Promoting High Achievement Among African-American Students*, Theresa Perry, Claude Steele, and Asa Hilliard III, pp. 109–116, 120–123, 172–174, 176–177. Copyright © 2003 by Theresa Perry, Claude Steele, and Asa Hilliard III. Reprinted by permission of Beacon Press, Boston.