Studying Discrimination:
Fundamental Challenges and Recent Progress

Kerwin Kofi Charles

Edwin A. and Betty L. Bergman Distinguished Service Professor
Harris School of Public Policy
University of Chicago

Jonathan Guryan

Associate Professor of Human Development and Social Policy
Faculty Fellow, Institute for Policy Research
Northwestern University

Version: June, 2011

Abstract

This working paper discusses research on discrimination against blacks and other racial minorities in labor market outcomes, highlighting fundamental challenges faced by empirical work in this area. Specifically, for work devoted to measuring whether and how much discrimination exists, the authors discuss how the absence of relevant data, the potential noncomparability of blacks and whites, and various conceptual concerns peculiar to race may frustrate or render impossible the application of empirical methods used in other areas of study. For work seeking to arbitrate empirically between the two main alternative theoretical explanations for such discrimination as it exists, the paper distinguishes between indirect analyses, which do not directly study the variation in prejudice or the variation in information—the mechanisms at the heart of the two types of models reviewed in the paper—and direct analyses, which are more recent and much less common. The authors highlight problems with both approaches. They also discuss recent work, which, the various challenges notwithstanding, permits tentative conclusions about discrimination. They conclude by pointing to areas that might be fruitful avenues for future investigation.

1. INTRODUCTION

The U.S. Civil Rights movement of the 1950s and 1960s led to the eradication of various exclusionary policies, laws, and institutions that had long limited the life prospects of racial minorities. Differences between blacks and whites in many outcomes have since narrowed sharply, but some gaps remain stubbornly persistent, including large differences in labor market outcomes such as wages and employment probabilities. One possible explanation for these differences in labor market outcomes might be the continued presence of racial differences in various dimensions of skill; there remain, for example, disturbingly large racial differences in various measures of educational attainment and cognitive skill (Neal 2006). Another possible explanation for the remaining labor market outcome differences may be labor market discrimination, meaning that blacks may receive lower labor market returns relative to similarly productive whites.

This article identifies and discusses various challenges that, in our view, frustrate empirical research on racial labor market discrimination.[1] Much like the discrimination literature more generally, we focus on discrimination against blacks.[2] We also discuss some recent empirical discrimination studies that feature relatively novel approaches for dealing with the challenges to discrimination research that we outline. The article is not, and does not attempt to provide, a comprehensive review of the vast discrimination literature. We therefore do not reproduce a comprehensive summary of various theoretical models of discrimination, so ably provided by Lang & Lehmann (2010). Nor do we thoroughly list the huge set of empirical findings in this literature as such a summary would be too long for the purposes of the article and may readily be found elsewhere (see Cain 1986, Darity & Mason 1998, Altonji & Blank 1999).

Empirical work on discrimination generally applies microeconometric methods to individual-level data, eschewing methods such as calibration. The literature tries to answer two questions: First, how much discrimination in outcomes such as wages is there against blacks? Second, what theoretical model, or set of models, best accounts for such labor market discrimination as it exists? Research on either question confronts a number of major challenges, which we discuss in turn.

We focus first on the difficulty of measuring the degree of labor market discrimination against blacks. The main problem this line of inquiry confronts is that, in observational data, individuals of different races may systematically differ with respect to other determinants of labor market outcomes apart from race, including some that are unobserved. Individual cross-racial comparisons may thus involve persons who are not otherwise equal, rendering conclusions about discrimination invalid. The possible presence of omitted-variables bias bedevils most microeconometric work seeking to make causal inference, and economists have developed a variety of methods to address it. However, we argue that various concerns peculiar to the case of race render the use of these methods imperfect for measuring discrimination. In particular, we discuss the generally ignored conceptual problem of accurate racial classification, concerns about limited data availability, and the near-impossibility of experimental or quasi-experimental research designs in the context of studies of race. We present some strategies used in recent work to circumvent some of these problems and explain why, despite these interesting and innovative efforts, the question of how much discrimination there is in the labor market against blacks remains somewhat controversial.

We move next to a critical assessment of efforts to determine which theoretical framework best explains existing discrimination. Two classes of models dominate modern theoretical work on labor market discrimination. Prejudice models emphasize the presence of negative preference in the utility function of key labor market actors. The earliest and perhaps most influential versions of these models assume perfect labor markets, but subsequent models often assume the presence of frictions associated with either random or directed search. Statistical discrimination models emphasize the role of limited information in labor market transactions.[3] Papers purporting to test these two classes of models typically conduct what we consider to be indirect assessments—that is, they measure whether particular aspects of wage or employment-differentials distributions are consistent with the predictions of the model being tested.

Indirect tests, in the sense that we use the term, are ubiquitous in economics, so we raise no objection to them on this basis. The problem with indirect tests in the context of discrimination is that the two predominant types of discrimination models often yield similar predictions about labor market differentials. Indirect approaches to testing between discrimination models are thus ultimately only convincing to the degree that the particular

finding about labor market outcome differences is explicable exclusively by one of the models and not by some version of the other class of discrimination models. In our view, this challenge is rarely met, especially because of the variety of insightful versions of the two basic classes of models that have now appeared in the literature, examples of which we discuss below.

Rather than using an indirect approach of examining a model's ultimate predictions about labor market outcomes, an alternative strategy would be to test the specific mechanism—prejudice or information limitations—by which the model implies those outcomes are generated. Data to conduct this type of direct analysis are generally not readily available in the data sources most commonly used by economists, and available measures do not always perfectly line up with the theoretical construct outlined in a chosen model. These are serious challenges, but some recent work has nonetheless conducted exactly this kind of direct assessment, especially for prejudice models.[4] We describe results of this work and address some questions and concerns they raise.

In the final section of the paper, we selectively discuss other recent work that has appeared since the most recent comprehensive review of the empirical literature. We identify themes uniting these new analyses and some questions—both theoretical and empirical—that we believe have received too little attention in the literature, ending the discussion with ideas for future work.

2. ASSESSING HOW MUCH RACIAL DISCRIMINATION EXISTS IN THE LABOR MARKET

2.1. Basic Setup

Empirical work seeking to measure the extent of labor market racial discrimination usually starts with a simple statistical model in which an individual i with observable traits X_{it} and unobservable characteristics ε_{it} receives wages (or some other labor market outcome) y at time t given by

$$y_{it} = \alpha X_{it} + \delta B_i + \varepsilon_{it}. \tag{1}$$

Here the binary variable B_i indicates the person's race and equals 1 if he is black. The parameter δ answers the ceteris paribus question of how, given the statistical model, an individual's wages

would be different if he were black instead of white but otherwise remained exactly the same. Equivalently, δ measures how two different persons, one black and one white but otherwise identical, would differ in terms of their wages. By either view, δ represents labor market discrimination against blacks. Various modifications of this basic specification might be imagined. For example, there is no i subscript on the term δ, so Equation 1 ignores the possibility that discrimination may be heterogeneous across blacks (see Heckman et al. 1999 for a discussion of heterogeneous treatment effects).[5] And the formulation in Equation 1 assumes that the various determinants of wages enter the regression equation linearly and additively. Considerations such as heterogeneous effects or rich interaction effects are undoubtedly important, but we ignore discussion of them with the aim of focusing on concerns we think are more fundamental challenges for measuring discrimination.

The regression given in Equation 1 is identical to the standard specification in the program evaluation or policy evaluation literature. In the language of that literature, being black $\left(\text{or } B_i = 1\right)$ is the treatment in Equation 1, and the analysis aims to measure δ, the causal effect of that treatment. As in the program evaluation area more generally, the key challenge to estimating δ using data on a sample of blacks and whites is the econometric difficulty posed by the possible correlation between B_i and the unobservables ε_{it}. In particular, a regression performed on Equation 1 for a sample of blacks and whites will produce a biased estimate of δ unless race is independent of the unobservables ε, conditional on the observables X, or

$$Cov\left(B_i, \varepsilon_{it} \mid X_{it}\right) = 0. \tag{2}$$

The implication of the requirement of the race variable's conditional independence is that, in general, estimates of discrimination based on comparisons using observational data will be biased unless all determinants of the outcome that remain unobserved after conditioning on observable controls are on average equal for blacks and whites, or

$$E\left[\varepsilon_{it} \mid B_i = 1, X_{it}\right] = E\left[\varepsilon_{it} \mid B_i = 0, X_{it}\right]. \tag{3}$$

The conditional independence assumption is of course impossible to test, but it is easy to think of unobserved wage determinants, of which blacks and whites probably have different amounts because of U.S. racial history.[6] Empirical scholars have developed various tools to deal

with violations of the conditional independence assumption. Before discussing the application and limitations of these methods, we discuss challenges peculiar to the treatment variable "being black" in Equation 1 that simply do not arise in other program evaluation problems. These considerations impose what we call a taxonomical challenge on efforts to measure discrimination, affecting both the analyses that can be done and the inferences that may be drawn from them.

2.2. The Taxonomical Challenge

In coding the variable B_i, economists generally assign individuals to one of a few mutually exclusive racial categories, based either on the agent's self-report about race on a survey or on the race the individual is assigned by an outside observer such as a survey interviewer. What do designations such as black, white, or Latino mean? The view among biologists appears to be that they have no meaningful biological significance. And the prevailing convention in all the humanities and other social sciences besides economics is that race is entirely socially constructed, meaning that the racial bins into which people are sorted at a point in time are essentially merely arbitrary divisions, which can change over time or contexts (e.g., see Lopez 1994 for a discussion of the biological versus social construction of race).

To these arguments made by other types of scholars, economists might reasonably respond that the coarse definitions they use unarguably represent commonly understood and historically important groupings. People routinely report their race on surveys, suggesting that they understand perfectly well what is being asked of them. Persons commonly referred to and who call themselves black or white disproportionately possess particular readily identifiable phenotypic traits such as certain types of skin coloration and hair texture, common to people whose ancestry can be traced, at least in part, to particular regions in the world (e.g., Africa for blacks, Europe for whites).[7] And the different racial bins into which people are sorted on the basis of these (admittedly superficial) characteristics have each been subject to a particular array of influences and experiences. Given these points, racial designations as applied in empirical work in economics are not obviously controversial. These designations denote membership in particular groups, characterized by a distinct set of experiences, culture, and broadly shared physical traits. These classifications are taken to be permanent and ascriptive, in that they are

5

part of an individual's endowment and different from attributes such as years of schooling or place of residence, which agents partly choose.

But does this reasonable definition of racial classification satisfy the conditions generally understood about a treatment in the treatment effects literature? An unstated condition in that literature is that the analyst can characterize each observation as having received (or at least as having been assigned to) the treatment or not. The treatment may differ in degree, as in dosage-type studies, but it is presumed to not differ in type—meaning that each treated observation can be regarded as having received the same, widely understood, and agreed-upon thing. These conditions obviously hold for treatments such as a participant in a job-training program or a student in a small class, but matters are not as clear when the treatment is a person's race. Precisely because race incorporates both a superficial physical aspect and something having to do with background or life experiences, two persons who differ markedly in their physical traits might, because of the essentially identical backgrounds from which they come, both characterize themselves as black, even if some third person might consider this to be true for only one of them. Or persons who appear indistinguishable from each other to an observer might consider themselves to come from different racial groups because of their vastly different life experiences.

Built into the notion of race, therefore, is the problem that the analyst's presumption about whether an agent has received the treatment—whether the agent is or is not black—may differ from what others believe about that agent. A person calling herself black, who is so noted by the analyst, may not be regarded as black by the market actors with whom she interacts, and vice versa. Because discrimination research aims to determine whether people are treated differently by various market actors because of their race, knowing the race that observers ascribe to an individual would seem an important precondition for calling differential treatment discrimination. Having information about the race the individual calls herself is not a guarantee against this concern; it is at least theoretically possible that racial self-reports might differ from observer beliefs about race. It is, moreover, theoretically possible that either self- or observer racial classifications may change over time or contexts, further frustrating efforts to call observed differential racial outcomes racial discrimination.

An active research program outside of economics suggests that these issues may be more than mere theoretical concerns.[8] A literature in sociology suggests that racial self-classification

may be sensitive to a number of considerations: the way that the question is posed or the range of options presented to the respondent (Farley 2002, Snipp 2003); the location of interviews with racial questions, as shown in the analysis by Harris & Sim (2002) of teenagers interviewed at home and at school; and possible societal changes that augment racial pride, as has been suggested about the massive and unexplained increase in the number of Americans describing themselves as American Indian after 1960 (Nagel 1995).

A particularly noteworthy recent example in this area is the work by Saperstein & Penner (2010) studying racial classification in the National Longitudinal Study of Youth (NLSY). In this well-known panel data set, which spans more than two decades, individuals were asked to report their own race in 1979 and later in 2002. In addition, survey interviewers recorded what they (the interviewers) took the respondent's race to be after the interview was completed in 19 consecutive surveys. Saperstein & Penner show that for fully 20% of the sample, there was at least one change (and often many more) in interviewer racial classification of the same respondent across years. The authors also document many changes in racial self-reports between the 1979 and 2002 responses.

The data allow the authors to carefully examine the notion that these changes might be the result of random miscoding. Various pieces of evidence sharply undercut this possibility, including the remarkable stability of interviewer coding about respondent gender and the fact that particular types of changes in self-classification are associated with particular life episodes in the intervening years. For example, they show that among persons who said they were white in 1979, those who did not have a bout of incarceration at some point in the intervening 20 years again reported themselves as white in 2002 at a rate of 95%, while those who had been subsequently incarcerated described themselves as white at the later date at only a rate of 81%, with 8% calling themselves black in the later year.

Following on these results, we examine the consistency of racial self-reports in the Current Population Survey (CPS), another data source frequently used by economists to study racial differences. The CPS question used to code respondents' race was changed between December 2002 and January 2003.[9] The 2002 question offered respondents four possible answers: White, Black, American Indian/Aleut/Eskimo, or Asian/Pacific Islander. The 2003 question allowed respondents to choose as many options as they desired from a much longer list

of possible races. Because CPS respondents are surveyed eight times over the course of 16 months, we observe responses for some individuals who were asked both race questions. We examine individuals whose first outgoing rotation month was in 2002 and whose final outgoing rotation month was in 2003 and follow the methods of Madrian & Lefgren (1999) to match observations across CPS waves. Specifically, we first match by household identifier and person line number. We then discard observations if there is not a match on gender, if age increases by more than two years between 2002 and 2003, and if reported years of education are lower in 2003 than in 2002. Madrian & Lefgren also discard cases in which the race response does not match; given our desire to measure how this variable changes, we obviously do not discard these cases.

Table 1 presents cross-tabulations of responses to the two race questions. Responses to the 2002 question are summarized in the columns, and each row reports the response by what the approach takes to be the same individuals to the 2003 question. The bottom panel of the table shows the fraction giving the same or different reports in both years. Two patterns are immediately noticeable. First, a large majority of individuals listed the same single race in 2003, when there was the option of listing many races, as the race they reported in 2002, when they were limited to a single choice. Second, there was substantial change over time, especially among racial minority groups, in own racial classification. Among those describing themselves as black in 2002, 6.7% say they are white (and only white) a year later when given the choice to list multiple races. Furthermore, 8.6% of blacks report something other than being only black when offered the choice of multiple race reports. The variation is even greater among Asian/Pacific Islanders and American Indian/Aleut/Eskimos. For example, among those reporting being Asian/Pacific Islanders in 2002, 11.2% report being white in 2003, and nearly 25% report something other than being only Asian/Pacific Islander.

Unlike the Saperstein & Penner NLSY results mentioned above, which come from a large panel data set in which respondents are tracked carefully over many years, the changing racial designations in the CPS may actually merely reflect our inability to accurately match respondents across CPS waves. It is possible to assess the extent to which mismatching of individuals across years of the CPS might be responsible for these results by performing a benchmark analysis for gender. Specifically, we matched outgoing rotation observations in the same way as in the race exercise except that we did not drop observations that were mismatched

8

on gender. We find from this analysis that only 0.4% of the pairs reported a different gender in the two waves of the CPS. As we should expect that about half of any randomly matched pairs of individuals would be of the same gender, this suggests that about 0.8% of the sample might be matched to different individuals in the two waves of the CPS. This estimate of the mismatch rate is an order of magnitude smaller than the fraction of blacks in 2002 who report something other than black in 2003, strongly suggesting that the different race responses are by the same person.

At the very minimum, the results about changing racial classification in commonly used data sets raise some doubt about the notion that race is a fixed trait and should give pause to the scholar of discrimination. Conceptually, what does it mean to say that blacks are or are not discriminated against if individuals who say that they are black, or who are coded as black by observers, in one time period report or are given a different racial classification at another?

If the changing racial classification depicted in Table 1 or the likelihood of classification disagreements at a point in time were random, estimates of the extent of discrimination will probably be downward biased.[10] But there is some evidence that the type of racial classification a person gives in survey data may be related to observable characteristics—at least among blacks. Above we note the correlation between the propensity to call one's self black and incarceration found in the NLSY. Table 2 shows the results of ordinary least squares (OLS) regressions that relate the propensity to report a single race in 2003 to individual characteristics in the CPS. Columns 1 and 2 show results for those who report being black in 2002; the dependent variable is an indicator for reporting black only in 2003. Columns 3 and 4 show results for those who report being white in 2002; the dependent variable is an indicator for reporting white only in 2003. All regressions include only observations for which valid wage observations exist in both 2002 and 2003.

The results in column 1 indicate that among those reporting being black in 2002, conditional on log wages measured in 2002, educational attainment is positively related to the choice to identify as being only black in 2003. However, conditional on educational attainment, log wages are insignificantly negatively related to reporting being only black in 2003. The regression in column 2 adds a control for the increase in log wages between 2002 and 2003. Adding this control sharpens the difference in estimated coefficients on log wages and education. Both are significant, although opposite in sign, and the coefficient on wage growth is also

negative and significant. It is also interesting to compare the results for blacks with those for whites, shown in columns 3 and 4. Just as for blacks, education positively predicts reporting a single race for whites when given the choice of multiple race responses. However, whereas the opposite is true for blacks, conditional on education, both the level and growth in log wages positively predict reporting being only white in 2003.

Further analysis is necessary before any definitive conclusions can be drawn about the meaning of these estimates, especially given the possible mismatch problems in the CPS. However, these results hint at the possibility that the propensity to self-report a particular racial designation may be correlated with the level of growth or wages or with things (such as schooling) that are related to wages. If true, this means that the treatment variable race in Equation 1 may be endogenous with respect to labor market outcomes, with effects on the estimates of discrimination that are not obvious. An important task for future empirical work on discrimination will be to more thoroughly explore changing racial classification and to examine the theoretical issue of what determines racial identification. Here insights from the interesting body of theoretical work on identity and people's choice of racial and other putatively immutable ascriptive traits (Fang & Loury 2005, Akerlof & Kranton 2000) may be quite useful.

2.3. Limitations of Experimental Estimates

Leaving aside any concerns about the possibly fluid and selective nature of racial classification, and supposing henceforth that racial designation captures an immutable and perfectly understood construct, the key challenge for the empirical scholar measuring discrimination is how to estimate δ in the face of possible violations of Equation 2 or 3. Of course, the only way to guarantee conditional independence of the treatment variable in a framework such as Equation 1 is by running an experiment, in which there is random assignment to different values of the treatment variable. For this reason, an experiment with controlled random assignment is widely considered the gold-standard method for answering ex post questions about the causal effect of a treatment.[11] When the analyst is unable to conduct an experiment, it is often possible to at least conceive of a mechanism by which treatment could be experimentally manipulated or to imagine that there might exist situations in which the random assignment forthcoming from an actual experiment is effectively produced by some natural feature of the environment.

For studies of discrimination, however, there exists no realistic possibility of randomly varying individuals' race nor is there any likelihood that a natural experiment that does the same thing might be discovered. Even if it were possible to change the phenotypic correlates of race such as skin color or hair texture, no responsible social scientists would undertake this endeavor.[12]

Randomly varying a person's race is not even a well-formed idea in the first place. It is not clear that persons whose skin color or hair texture had been changed (were such manipulation possible or moral) would be black, given that racial designation indicates some notion of shared culture or life experiences with a particular group. Furthermore, even if it were possible to randomly manipulate race, it is not clear which characteristics it would be correct to manipulate and which should be held constant. To answer this question, it is necessary to define precisely what race is, something that has not been settled, as discussed above. But if it is impossible to say which individual characteristics should be held constant and which should be randomly varied in a hypothetical experiment designed to estimate the effect of race on wages, how are we to make the same decisions when running an OLS regression for the same purpose?

For both practical and conceptual considerations, the gold-standard method of measuring the causal effect of an individual being of a different race is ruled out to the scholar of discrimination. Moreover, that one cannot even conceive of an ideal experiment in the realm of the hypothetical implies that the causal question itself may not be well specified.

Whereas experimental modification of an individual's race is impossible, random manipulation of an observer's beliefs or knowledge about the racial classification of an agent with whom that observer interacts is something social scientists can do.[13] One could, for example, imagine researchers randomly telling some subset of market actors that a person of superficially indeterminate race is in fact black and telling another subset that the person is white. A comparison across these groups in their treatment of the individual would provide some evidence about whether those actors discriminated against people they thought to be black.[14]

Bertrand & Mullainathan (2004) essentially conduct exactly this type of analysis in their study of firms' treatment of résumés sent in for entry-level job applications. Bertrand & Mullainathan find that résumés that had been randomly given "black-sounding" names as opposed to "white-sounding" ones, but that were otherwise the same, were significantly less

likely to receive call backs. A fair reading of these results is that firms engage in discrimination against applicants thought to be black.

The experimental methods employed in this well-known and imaginative paper might not completely circumvent concerns about omitted-variables bias. In particular, conditional independence requires that the only thing that changes in an observer's mind upon examining a résumé with a distinctively black-sounding name is the belief that the person to whom the résumé refers is black. If instead the observer happens to believe that people with distinctively black names are also distinguished by other productivity-relevant traits, differential behavior toward résumés with black-sounding names may be the result of a change in beliefs about those other traits.[15] Or, because the most-black-sounding names are unorthodox in the distribution of all names encountered in the United States, one cannot be sure whether it is the (assumed) race of the agent that generates the differential treatment or the fact that the observer treats those with unorthodox names differently. Either behavior would represent a type of discrimination, but discrimination in the latter case would not be racial in nature.

Bertrand & Mullainathan's paper may be read as an extension of the traditional audit-study approach to measuring discrimination. In a typical audit study, the analyst measures the extent to which subjects of different races or genders are treated differently by randomly selected employers or landlords to whom these subjects are sent to seek jobs or to rent or buy homes (e.g., see Yinger 1986, Mincy 1993, Darity & Mason 1998, Neumark et al. 1996). Many of these studies document that black subjects in these transactions receive worse outcomes (e.g., lower wages, smaller likelihood of job offers). The use of audit-study approaches to measure discrimination has been sharply criticized, most notably by Heckman & Siegelman (1993) and Heckman (1998). One critique is the possibility that, despite the best efforts of the study designer, there is no way to be absolutely certain that the subjects (perhaps unwittingly) do not exhibit some behavior in their interactions with firms or landlords that has an independent effect on how those firms act toward them. This problem may be present whether audit pairs are randomized to various firms or not. Because Bertrand & Mullainathan never send an actual person to firms, but rather only the inanimate and perfectly controllable résumé, this is not a concern in their study.

Bertrand & Mullainathan's paper might be more affected by a second complaint made by critics of conventional audit studies. The criticism asserts that, even if selected firms discriminate against blacks sent to them randomly, this differential treatment need not measure the labor market discrimination blacks experience in equilibrium. The presence of discriminatory firms might cause blacks to adopt different job-search strategies than whites, leading them to perhaps send more résumés out than would otherwise identical whites. Alternatively, blacks might be expected in equilibrium to sort themselves across firms, minimizing their contact with especially discriminatory ones. As we discuss at greater length below, existing models suggest that market discrimination will be determined by the firm with which blacks interact at the margin. Because of considerations like this, audit results—even those based on random assignment as in Bertrand & Mullainathan's study—may either understate or overstate racial equilibrium outcome differences.

2.4. Limitations of Regression-Based (or Selection on Observables) Methods of Measuring Discrimination

Given the challenges associated with using experimental methods to measure equilibrium labor market racial discrimination, it is not surprising that the most common approach in the literature relies on some form on regression-based procedure. The discussion in the previous section highlights that it is conceptually and practically difficult to even define which wage determinants should be categorized as being a part of race and which should be things to control for. If one could clearly delineate between the two and it were somehow possible to observe and control for all the other determinants of labor market outcomes besides race, conditional independence of the treatment variable B_i in the regression in Equation 1 would be assured (see Barnow et al. 1980). Indeed, even if this condition is not met, so that the regression controls only for a large subset of additional determinants of market outcomes, there probably is still descriptive value in examining the estimates from regressions such as that in Equation 1, conditional on the variables (X) present in standard data sets.

The neoclassical model of wage determination posits that workers are paid their marginal product. The default position of most economists is thus that unexplained differences in wages reflect differences in dimensions of skill not accounted for by the scholar, rather than providing evidence of discrimination. Indeed, a series of papers (e.g., O'Neill 1990, Maxwell 1994,

Ferguson 1995) has shown that controlling for the most widely studied observable dimensions of skill, such as the amount or quality of formal schooling, reduces unexplained racial gaps in wages and other outcomes. In a well-known paper, Neal & Johnson (1996) relate racial wage gaps among young adults in the NLSY to a single measure of skill—the person's adolescent score on the Armed Forces Qualifying Test (AFQT). Neal & Johnson find that controlling for AFQT score dramatically lowers the unexplained gap in wages between blacks and whites and indeed causes it to disappear for some subgroups. Darity & Mason (1998) point out that there is no consensus about what the AFQT measures, with possibilities ranging from intelligence to school quality, and subsequent authors have found that when both years of schooling (measured at the time of the AFQT test or much later in life) and AFQT score are included in Neal & Johnson's regressions, substantial racial wage gaps re-emerge (see Rodgers & Spriggs 1996, Carneiro et al. 2005, Lang & Manove 2009).

Despite these caveats, it appears that controlling for dimensions of skill not generally accounted for in empirical exercises, such as whatever it is that the AFQT score captures, may affect the conclusions we draw about the presence of discrimination based on regression estimates. In particular, residual racial wage differences may simply reflect that available data simply do not allow us to control for all aspects of market-relevant skills carefully enough. At a minimum, Neal & Johnson's results and related papers show that it will be important for future work seeking to measure discrimination with regression methods to collect and incorporate new data on skills, including more of the noncognitive aspects of skill, the value of which for labor market and other outcomes has been the focus of much recent work by Heckman and others (e.g., see Heckman et al. 2006).

Another challenge confronted by regression-based estimates is best illustrated in the widely used Blinder-Oaxaca decomposition framework, which has a long history in studies of racial discrimination (see Blinder 1973; Corcoran & Duncan 1979; Neumark 1988; Oaxaca & Ransom 1994, 1998; Darity et al. 1996). Applications of this method relax the assumption in Equation 1 that the coefficient β on the control vector X is constant across races and assume instead that the labor market outcome received by a person of a given race B_i may be represented by a race-specific regression

$$y_{it}^B = \alpha_B X_{it}^B + \varepsilon_{it}^B, \qquad B_i = \{0,1\}, \tag{4}$$

14

where the constant term is subsumed in X to focus the equations on explained differences in labor market outcomes. Because the regression line for each race passes through the means of y and X, the difference in observed wages between blacks $(B_i = 1)$ and whites $(B_i = 0)$ is given by

$$E\left[\bar{y}_{it}^1 - \bar{y}_{it}^0\right] = \hat{\alpha}_1 \bar{X}_{it}^1 - \hat{\alpha}_0 \bar{X}_{it}^0 = \bar{X}_{it}^0\left(\hat{\alpha}_1 - \hat{\alpha}_0\right) + \hat{\alpha}_1\left(\bar{X}_{it}^1 - \bar{X}_{it}^0\right) \quad (5)$$

or

$$E\left[\bar{y}_{it}^1 - \bar{y}_{it}^0\right] = \hat{\alpha}_1 \bar{X}_{it}^1 - \hat{\alpha}_0 \bar{X}_{it}^0 = \bar{X}_{it}^1\left(\hat{\alpha}_1 - \hat{\alpha}_0\right) + \hat{\alpha}_0\left(\bar{X}_{it}^1 - \bar{X}_{it}^0\right), \quad (6)$$

where $\hat{\alpha}$ are the estimated regression coefficients. The expressions in Equations 5 and 6 show that mean explained labor market differentials between races can be decomposed into two pieces: one due to differences between blacks and whites in their observed characteristics $(\text{the } X's)$ and another due to differences in the prices that those observed traits receive in the labor market (the estimated $\alpha's$). Because the benchmark neoclassical account of labor market outcome determination presumes that skills are treated equally, with no regard given to traits such as race with which they are correlated, a natural measure of discrimination in the importance of cross-race differences is the first term in either Equation 5 or 6. The part of the differential attributable to differences in X is determined before agents enter the market and thus is strictly speaking not the object of interest in studies of labor market discrimination.[16]

Results from several studies using these decomposition methods have found evidence consistent with market discrimination, especially for black men. One reason for caution about these estimates is that they appear to depend critically on which of the decomposition expressions in Equation 5 or 6 is used. More fundamentally, however, whether decomposition-style methods provide a convincing answer to the thought experiment "How would blacks fare if the distribution of traits they brought to the market were exactly the same as the distribution of traits brought by whites?" depends both on the degree to which the traits of blacks and whites overlap in the sample at hand and what we are willing to assume about the functional form mapping traits X into the outcome y.

If we are willing to assume that the effect of available traits on y is linear, then even if no blacks in the sample are observed to have the same values of the relevant traits as do whites, the decomposition exercise described above will still be valid. If, however, the effect of observed

15

traits on the labor market outcome is highly nonlinear and of unknown functional form, then the decomposition exercise outlined above may not produce a valid answer to the thought experiment about equalizing the distribution of traits across races. Intuitively, this is because there will be values of X at which it is not possible to match blacks and whites.[17]

An illustration of the misleading inferences that might be drawn from a naive application of decomposition methods with a linearity assumption is provided by Barsky et al. (2002). These authors study how much of the racial wage gap can be attributed to differences in income and how much can be attributed to differences, across race, in how income is converted to wealth. They show that because incomes differ so dramatically across races, combined with the high nonlinearity of the effect of income on wealth, decompositions of the sort typically used in discrimination work may produce underestimates of the importance of differences in the distribution of income. Although Barsky et al. do not focus on labor market discrimination, the problems they highlight may well be relevant to the use of decompositions to infer the role of discrimination in racial wage gaps. Because the distributions of many measures of skill, such as test scores, differ quite substantially by race, and probably affect wages and other labor market outcomes in ways that are highly complicated and almost certainly nonlinear, estimates of discrimination that come from decomposition exercises may in fact be speculative.

2.5. Is There Consensus About Existing Discrimination?

Given these various concerns, it is reasonable to ask whether there are any facts about racial labor market discrimination on which the literature has converged. There is indeed a large, well-documented gap in wages and earnings between black and white men. Much of this gap remains after controlling for education and potential experience. This gap narrowed significantly in the years preceding and following the Civil Rights Act of 1964 (Smith & Welch 1989, Donohue & Heckman 1991, Chay 1995). This narrowing then stalled beginning in the early 1980s and has remained fairly stable in the years since (Bound & Freeman 1992, Altonji & Blank 1999).

Besides differences in completed education, there are also large differences by race in measures of ability on things such as test scores. The well-known paper by Neal & Johnson (1996) finds that controlling for the score on the AFQT taken prior to entry to the labor market eliminates two-thirds of the wage gap for racial minorities, but as noted above, recent work finds that the usual gaps reappear in regressions that control for both AFQT score and measures of

16

completed schooling. These findings point to the existence of nontrivial wage gaps that are not attributable to skill differences, suggesting the presence of significant racial discrimination in wages.

The vast majority of the racial discrimination literature focuses on comparisons of black and white men, largely because comparisons of black and white women are complicated by selection issues arising from differential marriage and fertility. Differences in wages between black and white women that do not control for selection are typically smaller than those observed for men, but Neal (2004) shows that once differential selection is accounted for, wage differences between black and white women are at least as large as for men.

The literature focuses on wages, but there have also historically been very large racial gaps in unemployment and labor force participation rates (Stratton 1993, Fairlie & Sundstrom 1997). Compared with the literature on wages, much less work has been done to investigate whether differences in these other outcomes are accounted for by differences in observable measures of skill or from discrimination. The much smaller focus on employment gaps in empirical work mimics the theoretical literature, which has generally concentrated on wages. We turn next to a discussion of the difficulties associated with empirically determining what theoretical models best account for such labor market discrimination as it exists.

3. TESTING MODELS OF DISCRIMINATION

There is a large theoretical literature on labor market discrimination, including a number of papers of considerable technical complexity and sophistication. Given the aims of this article, we do not review the details of these models, as is done in Lang & Lehmann's (2010) recent review. We concentrate instead on the fact that each of the various theoretical models in the literature features one of a small number of arguments. We briefly describe the essence of those arguments and discuss the key empirical predictions forthcoming from the models that employ them. We then discuss efforts to empirically test theoretical models of discrimination, pointing out both fundamental challenges and what we consider recent advances.

3.1. The Two Main Types of Discrimination Models

Broadly speaking, theoretical models of discrimination in economics can be split into two types. The point of departure of prejudice models is that some persons have a negative feeling—an

antipathy or aversion—toward persons of another race.[18] Of course, the idea that negative cross-racial feeling might exist and probably has something to do with the historically poor life circumstance of (especially) blacks did not originate with economists. However, in work that is widely credited with starting the discrimination literature in economics, Becker (1971 [1957]) formalizes this negative feeling and demonstrates how its presence among different types of white market actors might translate into discriminatory labor market outcomes for blacks in competitive settings.

Becker represents negative cross-racial sentiment as a disutility from cross-racial interaction.[19] He separately considers prejudice among three different types of market actors—employers, coworkers, and customers. Because it is modeled as disutility from interacting with blacks, prejudice operates effectively as sort of a psychic price, or cost, incurred by a prejudiced person when interacting with blacks. This psychic cost is added to the pecuniary cost or benefits that otherwise characterize the interaction. Thus a prejudiced employer incurs a total cost of employing a black worker equal to the wage paid to that person plus the employer's prejudice; a prejudiced coworker receives a total benefit from working alongside a black person at a particular job equal to the wage he receives at that job minus his prejudice; and a prejudiced customer pays a total price for a product bought from a black seller given by the money price of the product plus the customer's prejudice. Because cross-racial contact is likely most unpleasant to the most-prejudiced whites, they pay a higher psychic price in cross-racial interactions than do relatively unprejudiced whites. Indeed, whites who are totally unprejudiced pay no psychic price at all from cross-racial interactions.

This simple, intuitively appealing formulation produces a number of interesting implications, perhaps best illustrated with the employer-prejudice example. Assuming blacks and whites to be perfect substitutes in production, an employer with prejudice $d_j \geq 0$ chooses black and white labor $\left(L_b \text{ and } L_a \right)$ to maximize

$$U_j = f\left(K, L_a + L_b \right) - w_a L_a - w_b L_b - d_j L_b, \tag{7}$$

where f is a constant returns to scale production function, and w_b and w_a are the wages for black and white labor, respectively. The employer simply hires the type of labor that is least costly to her, meaning that she will hire only blacks if $w_b + d_j < w_a$ and will hire only whites if

18

the strict inequality is reversed. The employer's workforce will be racially segregated only if her prejudice and prevailing wages are such that $w_b + d_j = w_a$.

In the short run, when the number and size of firms are fixed, it is easy to see that some blacks will be employed first by the least-prejudiced employers, then others by the next least prejudiced, and so forth until all blacks are employed. If the distribution of prejudice among employers is smooth, the last employer to hire blacks will be just indifferent between hiring blacks and whites at prevailing wages. In equilibrium there is market clearance, and the wage gap observed in equilibrium between blacks and whites, $w_a^* - w_b^*$, is implicitly given by the prejudice d_j^* of the last employer—what Becker calls the marginal discriminator.

Two key results follow from this reasoning. First, the equilibrium racial wage gap will not be determined by the mean level of prejudice among all employers in the market. Because market pressures sort blacks away from the most-prejudiced persons, wage gaps will be determined by the most-prejudiced employers with whom blacks interact. In fact, even if there are hugely prejudiced employers in a given market, if there are also enough totally unprejudiced employers $(d_j = 0)$ relative to the number of blacks, there will be no wage differential suffered by blacks in the short run. More generally, if blacks are a small minority in the labor market, only the prejudice in the left tail of the prejudice distribution matters for wage determination; prejudice at the 90th percentile of the prejudice distribution, for example, should be irrelevant as blacks will likely not interact with these persons in equilibrium.

The second result has to do with the ameliorative role of segregation on wage gaps. Because wage gaps ultimately derive from blacks having to interact with prejudiced whites, if the market operates in such a fashion that they could be isolated from such interactions—if the market were segregated by race—there will be no racial wage gap.

The basic intuition and key implications of the short-run version of the coworker version of Becker's model are quite similar: If the source of racial animus is coworkers rather than employers, and white and blacks workers can be perfectly segregated, there will be no wage difference among equally productive blacks and whites. If impediments to segregation are such that blacks and prejudiced whites must work together, the firm must pay whites forced to work with blacks a wage premium for doing so. Cost-conscious firms thus will tend to sort black

19

workers to ensure that they pay the smallest possible values of these premia—meaning that observed wage gaps will depend on the least-prejudiced whites with whom blacks are forced to interact as coworkers. In the customer prejudice case, whites who dislike buying goods made or sold by blacks will be willing to pay less for such goods, leading to a reduction in black relative wages. Acting in the other direction, whites who do not care about the race of the seller are willing to pay the same price for the good irrespective of the seller's race. Black producers have an incentive to sell to unprejudiced customers, and prejudiced customers will choose to purchase from white producers if prices are equal. If there are enough unprejudiced customers, black sellers need not sell at lower prices to attract prejudiced customers, and their wages are not correspondingly reduced. Equilibrium outcomes should again be determined by behavior at the margin.

Despite their interesting implications, two things about the neoclassical versions of prejudice (or taste-based) models of discrimination have long concerned economists. One is the long-standing unease among economists about theoretical arguments that hinge on changes to some part of the utility function. For example, the results from the standard employer discrimination models depend on firms maximizing something other than profit, because of the tastes of people running firms. The assumption that competitive firms maximize profit is central in neoclassical economics, so models in which it is seemingly assumed away—as in prejudice models—are in some sense unsatisfying, however interesting or intuitively appealing their implications.

The other concern about prejudice models (at least in their competitive form) is that, for some time, it was not obvious that the predictions of the model survive long-run competition. Employers who discriminate lose profits in the Becker model. Arrow (1973), extending a point made by Becker himself, stressed that if there were enough potential employers with no prejudice, or if existing employers with no prejudice extended their operations, wage differences would ultimately disappear in the long run as these employers hired available black workers. Later work has shown that, even in a competitive environment, what Arrow predicts need not arise if racial feeling takes the form of positive tastes for one group rather than animus toward another, or if attention were paid to the racial environment prejudiced employers would confront once they shut down (see Goldberg 1982 for a discussion of how taste discrimination based on

nepotism might survive long-run competition).[20] In addition, a variety of models have shown how wage discrimination arising from racial prejudice can survive when there is search.[21]

These various models reproduce many (although not always all) of the key insights of the basic Becker formulation. In particular, they tend to show that market forces sort the most racially prejudiced whites away from the objects of their prejudice and that how blacks fare in equilibrium should be a function of the most-prejudiced white with whom they are forced to interact in the relevant context.

3.2. Statistical Discrimination Models

The other types of models in the theoretical discrimination literature are statistical discrimination models. These models currently appear to be the most popular models among economists studying discrimination. Unlike prejudice models, statistical discrimination models do not presume up front that there is any negative racial sentiment on the part of employers. Rather, employers are typically represented as profit-maximizing agents who seek to hire the most able workers they can at prevailing wages. However, the nature of information is such that firms cannot accurately estimate workers' skills or productivity, given observable indicators of skill, and are especially bad at doing this for black compared with white workers. The firm faces a signal extraction problem and relies on race and individual-level indicators of productivity to solve that problem. The choices made by firms in their hiring or wage-setting decisions are fully rational: Firms maximize (expected) profit, the initial or induced levels of skill investment by race are fully consistent with firm decisions, and vice versa.

Early treatment of statistical discrimination models include Phelps (1972) and Arrow (1973), but perhaps the most complete early work is Aigner & Cain (1977). As with prejudice models, we quickly review this benchmark model as it represents the starting point for subsequent work in this area. Imagine that black (b) and white (a) workers apply for jobs. Workers are distinguished by their actual skill, or quality q, but firms only observe an error-ridden version of that skill, say $\tilde{q}$. Firms know the distribution of true skill by race; in particular, suppose that for persons of race $x = \{a,b\}$ it is known that

$$q_x \sim \Omega\left(\bar{q}_x, \sigma_q^2\right), \tag{8}$$

21

where Ω is the cumulative distribution function of the normal distribution. Finally, suppose that the error-ridden signal firms observe about a worker's ability is an unbiased indicator of ability. Thus

$$\tilde{q}_i = q_i + \eta_i \tag{9}$$

with the error of the signal, η, distributed

$$\eta_i \sim \Omega\left(0, \sigma_{\eta x}^2\right). \tag{10}$$

Firms pay wages equal to expected productivity, meaning that each person's wage y is given by

$$y_{ix} = E\left[q_{ix} \mid \tilde{q}_x, x\right] = \overline{q}_x\left(\frac{\sigma_\eta^2}{\sigma_q^2 + \sigma_\eta^2}\right) + \tilde{q}_x\left(\frac{\sigma_q^2}{\sigma_q^2 + \sigma_\eta^2}\right). \tag{11}$$

Each worker receives a wage that is a weighted average of the mean level of skill in his group and of the worker's level of the observed skill indicator, with weights that are functions of the variances in the unobserved skill and of the signal error. As the variance in the error of the signal falls relative to the variance in underlying skill (that is, as the signal becomes better), more weight is placed on the person's own indicator; as the signal becomes noisier, more weight is placed on the mean level of skill in the person's group.

Several interesting results follow from the benchmark result (Equation 11). Notice first that firms will pay both blacks and whites wages that are on average equal to their mean productivity. What about the payment received by a given black or white worker? Suppose that the quality of the signal is the same for blacks and whites but that blacks are known to have lower average skill $\overline{q}$, perhaps because of differences in educational opportunities. Then a black person will be paid less than a white person with the same value of the skill indicator, $\tilde{q}$. Notice that in this case, a black and a white worker with the same observed skills receive different wages, but this stems from no negative sentiment on the firm's part; it is the best decision the firm can make given its signal extraction problem. Notice also that the average difference in pay between blacks and whites is exactly equal to the racial difference in actual skills $\left(\overline{q}_b - \overline{q}_a\right)$.

Interestingly, the most common case studied in the literature assumes that mean skill $\overline{q}$ is the same across races, but the signal-to-noise ratio of the signal of skill differs between blacks

22

and whites. In this case, because mean wages paid to a group equal mean group productivity, blacks and whites are paid the same on average. However, the differences in variances across races mean that the weight placed on group mean skill versus the indicator of own skill will be different for blacks and whites. It is usually assumed that signal quality is worse for blacks.[22] This means that for blacks, compared with whites, more weight is given to the known mean productivity than to the individual skill indicator in Equation 11. Because wages are equal on average across races, blacks who are above the mean will be paid less than observationally similar whites, but for persons below the mean blacks should be paid more than observationally similar whites.

Using the benchmark model above as a starting point, subsequent theoretical work has concentrated on its various dynamic implications. One focus has been on how the presence of statistical discrimination affects the decision to invest in latent and observable dimensions of human capital. Lundberg & Startz (1983) consider a situation in which blacks and whites initially have the same underlying, unobserved productivity (q) and face the same opportunity for investments that can augment it. Using the insight from Equation 11 that greater measurement error in the accuracy with which firms can judge black productivity makes the return to such investment smaller for blacks, they show that blacks will choose not to undertake productivity-enhancing investment. Coate & Loury (1993) further explore this issue.

More recent work has raised questions about this well-known prediction of less human capital investment among blacks as a result of statistical discrimination. Lang & Manove (2009) observe that, unlike the assumption in previous work about skill investments being unobserved, key aspects of human capital, such as the amount or type of education, are observed by the market. Their model shows that, precisely because their productivity is more noisily assessed, blacks of any given ability have a greater incentive to engage in observable, verifiable investments such as schooling. In effect, education in the Lang & Manove framework lowers the error with which firms can judge black productivity (see also Arcidiacono et al. 2010 for closely related work).

The other, more dynamic direction taken in recent theoretical work on statistical discrimination builds on insights first presented by Farber & Gibbons (1996) about how the labor market resolves uncertainty about worker productivity. Farber & Gibbons argue that, in any

period, a worker's wage equals her expected output in that period given all available information about her, including her previous levels of production.[23] Firms effectively learn increasingly more about a worker's latent productivity over time by observing her various levels of output. Therefore, if there is a trait that is unobserved by the firm but observed by the analyst that is known to vary positively with ability, over time that trait should be an increasingly important determinant of wages.

Altonji & Pierret (2001) build on this key result from Farber & Gibbons (see also Pinkston 2006, Lange 2007). Their argument is as follows. Suppose (as is plainly the case) that race is correlated with some measure of productivity that is unobserved to the market. Firms either might statistically discriminate based on race or they might not. If firms learn workers' true productivity over time, as Farber & Gibbons suggest, the negative partial effect of race on wages should get smaller conditional on an interaction between race and time. And this interaction between race and time should itself have a negative partial effect on wages. If a worker's wages are determined increasingly more by her realized productivity the longer she spends in the labor market, and if some aspect of a worker's latent productivity is not accounted for in the process of initial wage setting (contrary to what would be true if the firm statistically discriminated using race), then conditional on an interaction between the latent (to the firm) measure of skill and time, the partial effect of race on wages should grow less negative over time.

3.3. Testing Between Two Types of Theories

The most fundamental challenge in empirically adjudicating between prejudice and statistical discrimination models is that they often predict the same outcomes. Our reading of the literature suggests that most tests in the literature, purporting to support or reject a particular type of model, are based on findings that are potentially explicable by some version of the other type of discrimination model.

A few examples from papers we admire illustrate this point. A prediction of Borjas & Bronars' (1989) model of customer prejudice with search is that the earnings of self-employed blacks will be lower than the earnings of comparable whites. And because the variance of returns to ability in self-employment is lower for high-ability black sellers than it is for high-ability whites, the variance in income among blacks across skill levels should be lower than for whites.

24

Borjas & Bronars find support for these predictions in the data. However, it is easy to see that this set of results could be the result of some other model, including one with the flavor of statistical discrimination. For example, if customers had no prejudice whatsoever, but knew that there was some difference in mean quality of goods made or sold by blacks, and offered prices for goods based on statistical discrimination in light of that knowledge or belief, we suspect that the predictions would be exactly the same as the ones Borjas & Bronars test and confirm. In our view, a similar sort of exercise can be done for most models that indirectly test prejudice models.

Taking an example in the other direction, we discuss above that an implication of statistical discrimination is that the greater error with which their true productivity is revealed should have an effect on blacks' willingness to invest in human capital. As noted, this insight is explored to different conclusions by Lundberg & Startz and by Lang & Manove. Of course, another reason that blacks might expect lower rewards for their human capital is simply that firms may discriminate against blacks because of prejudice. Nor does that prejudice have to be the same for all types of skill; firms may especially like or dislike uneducated blacks, which makes education either a protection against or an invitation to prejudiced labor market treatment.[24] The point is that evidence of the propensity of a larger or smaller human capital investment among blacks compared with whites is possibly explicable by either type of discrimination model.

As these examples demonstrate, most empirical exercises in the literature simply do not convincingly test between discrimination models. An approach that we believe offers some promise would be to analyze directly the two quite different mechanisms that the two types of models emphasize. Ideally this direct evidence would employ persuasive research designs, including exogenous variation in prejudice or in information. Even if this is not possible, it may still be possible to study associations involving prejudice or information that should be observed in the data if a particular model matters for racial labor market outcomes.

There are, of course, challenges to this direct approach. Measures of the negative racial sentiment that economic agents actually feel are not widely available, and available measures may not perfectly match the notion of prejudice in the model. How information affects people's expectations is a central part of statistical discrimination stories, but economists know little about how people form expectations in the real world and have only recently started to formally

explore how to measure them carefully (see Manski 2004 for a discussion about research on expectations). Finally, a given study might face the challenge of external validity—that the findings in the particular scenario studied do not generalize to the labor market more broadly. Despite these problems, examples of this direct approach have appeared in the recent literature. We discuss three below.

Using experimental methods, List (2004) tries to distinguish between taste-based and statistical discrimination in the baseball card market. The paper is described by Lang & Lehmann (2010), consistent with our characterization of this challenge in the literature, as "one of the few studies that attempt to identify the nature of discrimination rather than its mere presence." List recruited buyers and sellers of a particular, highly valued baseball card at a baseball card show. He finds that there was a significant price differential in the initial offer made for a card if the seller or buyer was a racial minority, with greater discrimination among sellers. Final transactions, after the buyer and seller had negotiated for a while, revealed much less discrimination. Among experienced buyers, there is little evidence of discrimination at all, although racial minorities had to expend more effort (that is, negotiate longer) to overcome discrimination.

List also runs a series of experiments. These include a dictator-game experiment in which he finds that, with the exception of white women, people were no more or less likely to share a transfer with someone of a different race. This suggests little racial animus among his subjects. He also ran a Chamberlain experiment in which buyers and sellers negotiated over the trade price. He finds that when sellers were told that reservation prices were randomly assigned, there was no difference in prices offered by race. By contrast, when they were not given this information and were thus unsure about sellers' reservation prices, sellers systematically offered lower prices to minorities.

This paper's results suggest that—at least as far as the baseball card market is concerned—transactions seem to be based on statistical inference about how reservation prices vary by race, rather than on something having to do with racial animus. The other merits of the paper aside, the obvious concern about these results is that the particular context in which they apply might have limited implications for the labor market as a whole.

A good example of a direct test of the central features of a statistical discrimination model is Autor & Scarborough (2008). Limited information is the key mechanism in statistical discrimination models: It is because they do not perfectly observe worker's abilities that firms rely on indicators such as race and place differential weight, by race, on the noisy indicators of productivity they do observe. Autor & Scarborough investigate what happens to the employment outcomes of blacks and whites when employers are provided with better information about the future productivity of potential workers, in the form of a new and better signal of worker ability. A model based on statistical discrimination suggests that this new information should have two effects on black outcomes. If the new signal is informative of productivity, highly productive black workers should benefit from the newly available information. However, if blacks score worse on average on the new employment screen, employers may shift hiring away from blacks and toward whites. Autor & Scarborough call this a trade-off between equity and efficiency. They show that the size of the latter effect depends not on the level of bias in the new information screen, but rather on the relative bias in the new information screen compared with the bias in employers' beliefs about productivity prior to the introduction of the new information.

They study the effect of new information using the example of the introduction of a computer-scored personality test into the hiring process at a firm. The scores on the personality test likely correlate well with productivity in this industry, and knowing these scores represents a marked improvement on the information previously available to firms about potential hires. Autor & Scarborough test their various predictions using a quasi-experimental research design. The firm they study has over 1,000 establishments, which introduce the test to the hiring process at different times over the course of a year. They take advantage of high-frequency personnel records and the differential timing of test introduction across establishments to estimate the effect of the test's introduction on the productivity of black hires and on the fraction of new hires that are black. Unfortunately, the only available measure of worker productivity is the length of the employment spell.

The results demonstrate that the introduction of the test leads to increases in employee tenure by 10%–12% among both whites and blacks. Better information (or a reduction in the error with which firms can project a worker's productivity) increases productivity among workers hired. Interestingly, despite the fact that black applicants score significantly worse on the test, its introduction into the hiring process does not have any noticeable effect on the

fraction of new hires that are black. This is a nice and rare example of a paper motivated by the important implications of a given discrimination model, which then tests the central mechanism of the model using a quasi-experimental research design.

A third example of this sort of empirical work is our own (Charles & Guryan 2008). We devote extra space to a discussion of some questions it raises both because we are the authors and because the paper's findings and its effort to directly test the key implication about the most simple prejudice model are not elsewhere present in the discrimination literature.

The test of the prejudice model in Charles & Guryan (2008) (henceforth CG) does not exploit exogenous variation in prejudice. Instead, the paper focuses on testing whether key associations between prejudice and wages implied by the Becker prejudice model are found in the data. The absence of these associations would tend to weaken support for the model, whereas their presence would point strongly toward an important role for the prejudice mechanism— especially if the particular associations tested are not obvious and are unlikely to be implied by other types of models. Recall that an implication of the Becker model is that because the market sorts blacks away from the most-prejudiced whites, outcomes for blacks should systematically depend on the left tail of the white prejudice distribution, as blacks are a small minority of the population. CG argue that this particular prediction (that it is the prejudice of relatively unprejudiced whites in a market that matters for wages rather than, say, the mean or some higher percentile point such as the 90th) is a subtle and initially counterintuitive prediction of a Becker-style prejudice model—the confirmation of which would lend credence to that model.

The study uses data from the General Social Survey (GSS), which over several years asks a representative sample of whites a series of questions having to do with racial sentiment. How the person feels about interracial marriage or whether they would vote for someone black are examples of the type of question asked. As the paper notes, none of these questions perfectly captures Becker's representation of prejudice as the disutility a person experiences from cross-racial interaction. However, a person's propensity to respond to them in a way consistent with having negative racial feelings probably correlates strongly with Becker's construct. Using responses to the separate racial-sentiment questions, CG create an individual prejudice index among whites in a given state and identify different percentile points in that prejudice distribution, differentially by state. Because of sample-size considerations, we pool GSS

28

observations over all years in the data to measure various percentiles of the distribution of prejudice in each state.

Using data from the 1977–2002 CPS, the paper then relates the average residual wage gap experienced by blacks in a state to the 10th, median, and 90th percentile of the white prejudice distribution in that state. Consistent with the Becker prediction, these results and other specifications consistently show that wage gaps are related to prejudice in the left tail of the distribution, but not to the level of median prejudice or prejudice at higher percentiles. Although not central to the analysis, given the percentile results, we also construct a measure of what the prejudice of the marginal discriminator is likely to be. Here too we follow the original Becker formulation, which argues that under a particular set of assumptions, this marginal person will be closely approximated by the p-th percentile of the white prejudice distribution if blacks constitute some percentage p of the workforce. We find that wage gaps are indeed systematically related to the prejudice of this marginal white, although the fit of these models varies across specifications.

Segregation features prominently in the Becker framework, and we test and confirm its main implication: that racial wage gaps should be negatively related to the fraction of blacks in the state. Using data from the National Longitudinal Education Study, we construct a measure of how much whites interact with blacks at work. Perhaps contrary to simple intuition, but consistent with the Becker formulation of prejudice, we find that greater racial integration in a state's workforce is associated with larger wage gaps. Moreover, we show that this estimated effect of greater integration is reduced when the regression controls for the prejudice of the marginal discriminator or the prejudice at the 10th percentile. This last finding is consistent with the reasoning in the Becker framework that the identity of the marginal discriminator depends on the extent to which the labor market can be segregated, and with Becker's idea that segregation and wage discrimination essentially substitute for each other.

We view these results as strongly supportive of the essential argument in the Becker prejudice model and consider them all the more important because these key predictions of that venerable model had not, to our knowledge, been previously tested. Several questions that might be raised about the results are directly addressed in the paper, but more recent results from the literature highlight two issues we use this opportunity to address.[25]

In a recent paper, Lang & Lehmann (2010) question whether the absence of a noticeable change over time in residual wage gaps is consistent with the decline in average prejudice over the past few decades.[26] Indeed, as Lang & Lehmann observe, evidence of a sharp decline over time in average prejudice can be found in the fact that the nation recently elected its first black president, overcoming in the process generations of bias.[27] The CG analysis does not address this important question; indeed, because of data constraints, the paper cannot definitely speak to how trends in residual wage gaps relate to prejudice changes within states. The paper mainly exploits cross-sectional variation, mostly because so few persons were asked about their prejudice by the GSS in a given state in a particular year that is difficult to accurately estimate different percentile points in a state's prejudice distribution in a particular year. Moreover, as the paper shows, there is almost no difference across regions in the decline in average prejudice over the time period studied.

The question Lang & Lehmann raise does suggest one exercise not done by CG. The key implication of the model CG test is that changes in mean prejudice among all whites in the country are not what should matter for national changes in black wages; changes in the prejudice of the marginal white are what should matter. We show here changes in the different percentile points in the distribution of white prejudice over the entire country, in which the numbers are created using the same method as in Charles & Guryan (2008). **Figure 1** shows that, although there has been a decline at all parts of the overall national prejudice distribution, the decline has been largest for the highest levels of prejudice: The persons at the 90th or 75th percentile of state distributions of prejudice among whites are on average much less prejudiced compared with 30 years ago. At lower percentile points in the distribution—that is, among whites with whom blacks might be expected to interact given the sorting implied by the prejudice model—the decline in prejudice over the past few decades has been much smaller. These patterns suggest that, even in the face of a decline in mean prejudice shown by both Lang & Lehmann and CG, persistent residual wage gaps are consistent with the fact that the prejudice of the marginal white person over the entire country has changed relatively little.

Another question prompted by recent results concerns the role of the South in CG's results. From the most recent reports from the Census Bureau, blacks have been moving to the South more than to any other region. And other recent work shows that there has been greater racial wage convergence in the South than in other regions (Vigdor 2006). Yet evidence for an

important role for prejudice in wage determination from CG comes, to a large degree, from comparisons between Southern states and the rest of the country. How can we reconcile these two sets of results—the CG results, on one hand, which appear to emphasize the greater prejudice of the South, and wage convergence and migration patterns, on the other hand, which suggest, at a minimum, sharply declining Southern prejudice relative to other areas?

The first response to this question is that although Southern/non-Southern differences do contribute importantly to the CG results, the paper also shows that the main results are qualitatively reproduced in a sample restricted to only Southern states. Also, CG show that although average prejudice has declined in the South, mirroring the decline in the rest of the country, the South remains the most-prejudiced region in the country. For consistency with the Becker model, the greater wage convergence over time in the South would have to coincide with greater convergence in the South of prejudice in the left tail of the distribution compared with other percentile points. **Figure 2** presents this analysis, which was not done by CG. The figure shows that most of the convergence in the South's prejudice did in fact occur in the left tail of the prejudice distribution. To account for the sample size issues discussed above, percentiles of the prejudice distribution are calculated by state and by year, and then the average is taken across all states in a Census region (i.e., South, Northeast, Midwest, West) in a year. The figures are noisy for the sample-size reasons noted above, but they seem to clearly show that convergence in the left tail of the prejudice distribution (the 10th and 25th percentiles) has been larger than that at the mean or in the right tail (the 90th percentile).

4. FINAL THOUGHTS

Despite the various problems outlined above, new research about racial discrimination continues to appear in the literature. Some of this work has not been about the labor market but has used reasoning in which statistical discrimination arguments feature prominently. For example, in a previous volume in this series, Persico (2009) describes the active body of work on racial profiling (see also Knowles et al. 2001, Antonovics et al. 2005, Anwar & Fang 2006, Sanga 2009). The striking thing about these papers, and work such as Altonji & Pierret (2001) on learning in the labor market, is that they use some variant of the statistical discrimination model to identify sharp and transparent predictions about conditional means or associations that ought to be observed in the data if the model is correct, and that should not be true otherwise. They

then test these predictions. As another example, Card et al. (2008) test a key prediction of Schelling's (1971) model about neighborhood racial makeup in the presence of preferences for being close to a neighbor of the same race that could not be true were the basic theory not valid. The style employed in these papers is similar to the CG approach for studying prejudice. We think many key predictions of the most basic versions of prejudice and statistical discrimination models about labor market discrimination remain to be tested using a similar approach and hope, and expect, to see papers attempting this work in the future.

In empirical economics more generally, there has been a notable increase in the number of papers employing experimental methods, whether in a lab or in the field. Above we discuss papers using this approach both to identify discrimination in the labor market (Bertrand and Mullainathan 2004) and to test between explanations for its sources (List 2004). A closely related set of papers looks at discrimination in the quasi-experimental settings provided by game shows (Levitt 2004, Antonovics et al. 2005). What these studies sacrifice in terms of possible external validity is made up for by the transparency and persuasiveness of their research designs. We suspect that many other similar papers, hopefully focusing on labor market outcomes, will be forthcoming in the literature.

There is growing output from the theoretical side of the discrimination literature. Recent models, centering on beliefs and information as statistical discrimination models do, have yielded interesting results about subjects such as belief flipping, evolutionary statistical discrimination, and stereotypes (e.g., Coate & Loury 1993; Loury 2002, 2006; Fryer 2007; Levin 2009). All these theoretical efforts represent a long-overdue effort to think deeply about the source of erroneous racial stereotypes, and how beliefs evolve. However, because few, if any, of their predictions have been tested in the data, there are opportunities for new empirical work on these questions in the future.

There are also new ideas about the nature of prejudice in the literature. Some recent work studies the difference between explicit and implicit prejudice (Greenwald et al. 1998, Bertrand et al. 2005). In a similar vein, Price & Wolfers (2010) estimate the differential treatment of white and black NBA basketball players by referees of different race. Price & Wolfers attribute the differences they observe to subconscious split-second decision making, which suggests a more nuanced type of prejudice than Becker's notion of a willingness to pay a price to avoid

interaction. Other work examining the different ways that explicit and implicit prejudice affects markets is likely to be forthcoming in the literature.

Economists thus far have spent little time exploring the sources of prejudice, such as whether it derives from battles over economic resources, whether it indicates a preference toward in-group members or distaste directed at out-group members, or how and why it is affected by things such as education. Although Becker's formulation of prejudice as distaste for cross-racial interaction has offered great insights into the way that racial feelings may create discriminatory outcomes for minorities, there may be fresh insights to be gleaned from the large body of work in social psychology that studies both the form and sources of racial prejudice (see Charles & Guryan 2009 for discussion). As economic theorists engage with work such as Fiske (1998), the classic studies of Allport (1954) and Tajfel (1981), and the famous robber's cave study by Sherif et al. (1988 [1961]), there may be fresh insights into ways to test for the importance of prejudice in labor markets. As all these examples show, despite the difficulty of establishing definitive evidence about whether discrimination exists at all, or why it does, there are many avenues for creative work in the future.

LITERATURE CITED

Aigner DJ, Cain GG. 1977. Statistical theories of discrimination in labor markets. *Ind. Labor Relat. Rev.* 30:175–87

Akerlof GA, Kranton RE. 2000. Economics and identity. *Q. J. Econ.* 115:715–53

Allport GW. 1954. *The Nature of Prejudice*. Reading, MA: Addison-Wesley

Altonji JG, Blank RM. 1999. Race and gender in the labor market. In *Handbook of Labor Economics,* Vol. 3C, ed. OC Ashenfelter, D Card, pp. 3143–259. New York: Elsevier

Altonji JG, Pierret CR. 2001. Employer learning and statistical discrimination. *Q. J. Econ.* 464:313–50

Antonovics K, Arcidiacono P, Walsh R. 2005. Games and discrimination: lessons from 'The Weakest Link'. *J. Hum. Resour.* 40:918–47

Anwar S, Fang H. 2006. An alternative test of racial prejudice in motor vehicle searches: theory and evidence. *Am. Econ. Rev.* 96:127–51

Arcidiacono P, Bayer P, Hizmo A. 2010. Beyond signaling and human capital: education and the revelation of ability. *Am. Econ. J. Appl. Microecon.* 2:76–104

Arrow K. 1973. The theory of discrimination. In *Discrimination in Labor Markets*, ed. OC Ashenfelter, A Rees, pp. 3–33. Princeton, NJ: Princeton Univ. Press

Autor DH, Scarborough D. 2008. Does job testing harm minority workers? Evidence from retail establishments. *Q. J. Econ.* 123:219–77

Barnow B, Cain G, Goldberger A. 1980. Issues in the analysis of selectivity bias. *Eval. Stud.* 5:42–59

Barsky R, Bound J, Charles KK, Lupton JP. 2002. Accounting for the black-white wealth gap: a nonparametric approach. *J. Am. Stat. Assoc.* 97:663–73

Becker GS. 1971 (1957). *The Economics of Discrimination*. Chicago: Univ. Chicago Press. 2nd ed.

Bertrand M, Chugh D, Mullainathan S. 2005. Implicit discrimination. *Am. Econ. Rev.* 95:94–98

Bertrand M, Mullainathan S. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *Am. Econ. Rev.* 94:991–1013

Black D. 1995. Discrimination in an equilibrium search model. *J. Labor Econ.* 13:309–34

Blinder AS. 1973. Wage discrimination: reduced form and structural estimates. *J. Hum. Resour.* 8:436–55

Borjas GJ, Bronars SG. 1989. Consumer discrimination and self-employment. *J. Polit. Econ.* 97:581–605

Bound J, Freeman RB. 1992. What went wrong? The erosion of relative earnings and employment among young black men in the 1980s. *Q. J. Econ.* 107:201–32

Bowlus AJ, Eckstein Z. 2002. Discrimination and skill differences in an equilibrium search model. *Int. Econ. Rev.* 43:1309–45

Cain GG. 1986. The economic analysis of labor market discrimination: a survey. In *Handbook of Labor Economics*, Vol. 1, ed. O Ashenfelter, R Layard, pp. 693–785. Amsterdam: North-Holland

Card D, Krueger AB. 1992. School quality and black-white relative earnings: a direct assessment. *Q. J. Econ.* 107:151–200

Card D, Mas A, Rothstein J. 2008. Tipping and the dynamics of segregation. *Q. J. Econ.* 123:177–218

Carneiro P, Heckman JJ, Masterov D. 2005. Labor market discrimination and racial diferences in pre-market factors. *J. Law Econ.* 48(1):1–39

Charles KK, Guryan J. 2007. *Prejudice and the economics of discrimination*. NBER Work. Pap. 13661

Charles KK, Guryan J. 2008. Prejudice and wages: an empirical assessment of Becker's *The Economics of Discrimination. J. Polit. Econ.* 116:773–809

Charles KK, Guryan J. 2009. Taste-based discrimination. In *The New Palgrave Dictionary of Economics*, ed. SN Durlauf, LE Blume. New York: Palgrave Macmillan; http://www.dictionaryofeconomics.com/article?id=pde2009_T000251

Chay KY. 1995. *The impact of federal civil rights policy on black economic progress: evidence from the Equal Opportunity Act of 1972*. Ind. Relat. Sect. Work. Pap. 346, Princeton Univ.

Coate S, Loury GC. 1993. Will affirmative-action policies eliminate negative stereotypes? *Am. Econ. Rev.* 83:1220–40

Corcoran M, Duncan GJ. 1979. Work history, labor force attachment, and earnings differences between the races and the sexes. *J. Hum. Resour.* 14:3–20

Darity W Jr, Guilkey D, Winfrey W. 1996. Explaining differences in economic performance among racial and ethnic groups in the USA: the data examined. *Am. J. Econ. Sociol.* 55:411–26

Darity WA Jr, Mason PL. 1998. Evidence on discrimination in employment: codes of color, codes of gender. *J. Econ. Perspect.* 12:63–90

DiNardo J. 2002. *Propensity score reweighting and changes in wage distributions*. Unpublished manuscript, Univ. Michigan

DiNardo J, Fortin N, Lemieux T. 1996. Labor market institutions and the distribution of wages, 1973–1993: a semi-parametric approach. *Econometrica* 64:1001–45

Donohue JJ III, Heckman J. 1991. Continuous versus episodic change: the impact of Civil Rights policy on the economic status of blacks. *J. Econ. Lit.* 29:1603–43

Fairlie RW, Sundstrom WA. 1997. The emergence, persistence and recent widening of the racial unemployment gap. *Ind. Labor Relat. Rev.* 52:252–70

Fang H, Loury GC. 2005. Toward an economic theory of dysfunctional identity. In *The Social Economics of Poverty: On Identities, Groups, Communities and Networks*, ed. CB Barrett, pp. 12–55. London: Routledge

Farber HS, Gibbons R. 1996. Learning and wage dynamics. *Q. J. Econ.* 111:1007–47

Farley R. 2002. Racial identities in 2000: the response to the multiple-race response options. In *The New Race Question: How the Census Counts Multiracial Individuals*, ed. J Perlmann, MC Waters, pp. 33–61. New York: Russell Sage

Ferguson R. 1995. Shifting challenges: fifty years of economic change toward black-white earnings equality. *Daedalus* 124:37–76

Fiske ST. 1998. Stereotyping, prejudice and discrimination. In *Handbook of Social Psychology,* Vol. 4, ed. DT Gilbert, ST Fiske, pp. 357–414. New York: Oxford Univ. Press

Fryer RG Jr. 2007. Belief flipping in a dynamic model of statistical discrimination. *J. Public Econ.* 91:1151–66

Fryer RG Jr, Levitt SD. 2004. The causes and consequences of distinctively black names. *Q. J. Econ.* 119:767–805

Goldberg MS. 1982. Discrimination, nepotism, and long-run wage differentials. *Q. J. Econ.* 97:307–19

Goldin C, Rouse C. 2000. Orchestrating impartiality: the impact of blind auditions on the sex composition of orchestras. *Am. Econ. Rev.* 90:715–41

Goldsmith AH, Hamilton D, Darity W Jr. 2007. From dark to light: skin color and wages among African-Americans. *J. Hum. Res.* 42:701–38

Greenwald AG, McGhee DE, Schwartz JLK. 1998. Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74:1464–80

Grogger J. 2008. *Speech patterns and racial wage inequality*. Harris Sch. Work. Pap. Ser. 008-13, Univ. Chicago

Harris DR, Sim JJ. 2002. Who is multiracial? Assessing the complexity of lived race. *Am. Sociol. Rev.* 67:614–27

Heckman JJ. 1998. Detecting discrimination. *J. Econ. Perspect.* 12:101–16

Heckman JJ, LaLonde R, Smith J. 1999. The economics and econometrics of active labor market programs. In *Handbook of Labor Economics,* Vol. 3C, ed. OC Ashenfelter, D Card, pp. 1865–208. New York: Elsevier

Heckman JJ, Siegelman P. 1993. The urban institute audit studies: their methods. In *Clear and Convincing Evidence: Measurement of Discrimination in America*, ed. M Fix, R Struyk, pp. 187–258. Washington, DC: Urban Inst.

Heckman JJ, Stixrud J, Urzua S. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *J. Labor Econ.* 24:411–82

Knowles J, Persico N, Todd P. 2001. Racial bias in motor-vehicle searches: theory and evidence. *J. Polit. Econ.* 109:203–29

Lang K. 1986. A language theory of discrimination. *Q. J. Econ.* 101:363–82

Lang K, Lehmannn J-YK. 2010. *Racial discrimination in the labor market*. Unpublished manuscript, Boston Univ.

Lang K, Manove M. 2009. *Education and labor market discrimination*. Unpublished manuscript, Boston Univ.

Lang K, Manove M, Dickens WT. 2005. Racial discrimination in markets with announced wages. *Am. Econ. Rev.* 95:1327–40

Lange F. 2007. The speed of employer learning. *J. Labor Econ.* 25:1–35

Levin J. 2009. The dynamics of collective reputation. *B.E. J. Theor. Econ.* 9(1):27

Levitt SD. 2004. Testing theories of discrimination: evidence from Weakest Link. *J. Law Econ.* 47:431–52

List JA. 2004. The nature and extent of discrimination in the marketplace: evidence from the field. *Q. J. Econ.* 119:49–89

Lopez IFH. 1994. The social construction of race: some observations on illusion, fabrication, and choice. *Harv. Civ. Rights-Civ. Lib. Law Rev*. 29:1–62

Loury G. 2006. Racial stigma: toward a new paradigm for discrimination theory. In *Understanding Poverty*, ed. A Banerjee, R Benabou, D Mookherjee, pp. 401–8. New York: Oxford Univ. Press

Loury GC. 2002. *The Anatomy of Racial Inequality*. Cambridge, MA: Harvard Univ. Press

Lundberg SJ, Startz R. 1983. Private discrimination and social intervention in competitive labor markets. *Am. Econ. Rev.* 73(3):340–47

Madrian B, Lefgren LJ. 1999. *A note on longitudinally matching current population survey respondents*. NBER Work. Pap. T0247

Manski C. 2004. Measuring expectations. *Econometrica* 72:1329–76

Maxwell N. 1994. The effect on black-white wage differences of differences in the quantity and quality of education. *Ind. Labor Relat. Rev.* 47:249–64

Mincy RB. 1993. The Urban Institute audit studies: their research and policy context. In *Clear and Convincing Evidence: Measurement of Discrimination in America*, ed. M Fix, RJ Struyk, pp. 165–86. Washington, DC: Urban Inst.

Nagel J. 1995. American Indian ethnic renewal: politics and the resurgence of identity. *Am. Sociol. Rev.* 60:947–65

Neal D. 2004. The measured black-white wage gap among women is too small. *J. Polit. Econ.* 112:S1–28

Neal D. 2006. Why has black-white skill convergence stopped? In *Handbook of the Economics of Education*, ed. E Hanushek, F Welch, pp. 511–634. New York: Elsevier

Neal DA, Johnson WR. 1996. The role of premarket factors in black-white wage differences. *J. Polit. Econ.* 104:869–95

Neumark D. 1988. Employer's discriminatory behavior and the estimation of wage discrimination. *J. Hum. Resour.* 23:279–95

Neumark D, Bank RJ, Van Nort KD. 1996. Sex discrimination in restaurant hiring: an audit study. *Q. J. Econ.* 111:915–42

Oaxaca RL, Ransom MR. 1994. On discrimination and the decomposition of wage differentials. *J. Econom.* 61:5–21

Oaxaca RL, Ransom MR. 1998. Calculation of approximate variances for wage decomposition differential. *J. Econ. Soc. Meas.* 24:55–61

O'Neill J. 1990. The role of human capital in earnings differences between black and white men. *J. Econ. Perspect.* 4:25–45

Persico N. 2009. Racial profiling? Detecting bias using statistical evidence. *Annu. Rev. Econ.* 1:229–53

Phelps ES. 1972. The statistical theory of racism and sexism. *Am. Econ. Rev.* 62:659–61

Pinkston JC. 2006. A test of screening discrimination with employer learning. *Ind. Labor Relat. Rev.* 59:267–84

Price J, Wolfers J. 2010. Racial discrimination among NBA referees. *Q. J. Econ.* 125:1859–87

Rodgers WM III, Spriggs WE. 1996. The effect of federal contractor status on racial differences in establishment-level employment shares: 1979–1992. *Am. Econ. Rev.* 86:290–93

Rubinstein Y, Brenner D. 2011. *Pride and prejudice: using ethnic-sounding names and inter-ethnic marriages to identify labor market discrimination*. Unpublished manuscript, Brown Univ.

Sanga S. 2009. Reconsidering racial bias in motor vehicle searches: theory and evidence. *J. Polit. Econ.* 117:1155–59

Saperstein A, Penner A. 2010. The race of a criminal record: how incarceration colors racial perception. *Soc. Forces* 57:99–113

Sasaki M. 1999. An equilibrium search model with coworker discrimination. *J. Labor Econ.* 17:377–407

Schelling TC. 1971. Dynamic models of segregation. *J. Math. Sociol.* 1:143–86

Sherif M, Harvey OJ, White BJ, Hood WR, Sherif CW. 1988 (1961). *The Robbers Cave Experiment: Intergroup Conflict and Cooperation*. Middletown, CT.: Wesleyan Univ. Press

Smith JP, Welch FR. 1989. Black economic progress after Myrdal. *J. Econ. Lit.* 27:519–64

Snipp CM. 2003. Racial measurement in the American census: past practices and implications for the future. *Annu. Rev. Sociol.* 29:563–88

Stratton LS. 1993. Racial differences in men's unemployment. *Ind. Labor Relat. Rev.* 46:451–63

Tajfel H. 1981. *Human Groups and Social Categories: Studies in Social Psychology*. Cambridge, UK: Cambridge Univ. Press

Vigdor JL. 2006. *The new promised land: black-white convergence in the American South, 1960–2000*. NBER Work. Pap. 12143

Yinger J. 1986. Measuring racial discrimination with fair housing audits: caught in the act. *Am. Econ. Rev.* 76:881–93

[1]Because of our focus on racial discrimination in the labor market, we do not discuss the large literature on discrimination in other markets (such as the market for housing) or labor market discrimination against other disadvantaged groups (such as women). Nonetheless, many arguments we make about the challenges of doing convincing work about racial discrimination carry over closely to work in those other contexts.

[2]The relative thinness of empirical research on discrimination against other types of racial minorities has been long lamented (see Altonji & Blank 1999). One reason for the paucity of research on discrimination against other minority groups (or by one minority group toward another) is that, unlike blacks and whites, for whom such data have always been collected, disaggregated data for other racial groups have not been consistently coded in the large data sets typically used by economists. This situation will hopefully change in the future, as scholars make use of the new, richer racial information now available on surveys such as the U.S. Census.

[3]Other types of discrimination models that have appeared in the literature often incorporate one of the insights central to these two classes of models. For example, in Black's (1995) model, discriminatory treatment derives from local monopsony power enjoyed by incumbent employers. In this framework, the fact that some firms enjoy differential monopsony power over black and white workers arises because of the presence of employers whose racial prejudice is such that they do not hire black workers at all (see also Sasaki 1999 and Bowlus & Eckstein 2002 for models with search and prejudice).

[4]Altonji & Blank (1999) use the expression direct evidence to mean something different than we do here. Although it is not defined in their paper, from the context it is clear that they use the expression to refer to persuasive empirical evidence about the existence of discrimination. Here we use the term direct to mean a focus on the mechanism implied by the particular model.

[5]Even if there are heterogeneous discrimination effects, the parameter δ from the formulation in Equation 1 measures the average labor market discrimination experienced by blacks in the sample.

[6]The historical differences in the quality of the schooling received by blacks and whites (Card & Krueger 1992) represent a likely source of unmeasured differences in productivity.

[7]The mapping from physical traits to race is not perfect. Many people who are commonly understood to belong to different racial groups share similar physical features. Likewise, many blacks in the United States differ so widely with respect to these same markers that their hair texture or skin color would assure that they would not be considered black in other parts of the world or perhaps even in the United States at other points in time.

[8]In studies of racial discrimination based on superficial identification, such as whether the police are more likely to pull over "black" motorists on the freeway, it obviously matters a great deal for our conclusions about the extent of discriminatory treatment if the persons describing themselves as black in the sample would not have been thought so by a third party based on casual inspection.

[9]This change in the way that race is coded in the CPS carries over from a fundamental change in how the U.S. Census asked people about their racial identification beginning with the 2000 Census.

[10]This is the familiar argument about there being attenuation bias if the regressor of interest is randomly mismeasured.

[11]Some treatments might be conceptually amenable to experimental manipulation but would still never be conducted because of moral considerations. The case of the health effects of smoking is a famous example in which the current consensus about the adverse causal effects of smoking has been established through the careful gathering of observational data rather than from purely experimental evidence.

[12]Although not experimental in nature, there is a literature that measures wage and other labor market differences by skin tone (see, e.g., Goldsmith et al. 2007). These results document penalties for darker-skinned individuals, but because the results are observational in nature, there are the usual concerns about possible unmeasured differences in productivity-relevant traits, including differences in family background.

[13]Goldin & Rouse (2000) exploit an exogenous change in knowledge to measure the effects of discrimination against women. They study a change in symphony audition procedures, whereby candidates performed behind a screen and thus were not seen by jurors, rather than in the open as had previously been done. They find that female musicians are more likely to be hired when

auditions were blind, when knowledge of gender was unknown to the relevant market actor. This study concerns gender discrimination, but one can easily imagine a similar methodology for identifying the effects of race on hiring or promotion.

[14]In a different cultural context, Rubinstein & Brenner (2011) explore whether wage differences between Ashkenazi and Sephardic Jews in Israel result from discrimination. Comparing males born to Sephardic mothers and Ashkenazi fathers with those born to Ashkenazi mothers and Sephardic fathers, they show that wages are lower in the latter group despite the fact that they have more observable skills. The authors attribute this difference to the former primarily having Ashkenazi surnames and the latter primarily having Sephardic surnames. They also show that married women's wages are related more strongly to the ethnicity of married women's father-in-law's ethnicity than to their father's ethnicity.

[15]Bertrand & Mullainathan address these concerns by controlling on their résumés for standard measures of skill for the jobs in question and present some supportive data on birth-certificate data from Massachusetts to show that maternal background traits do not drive their results. However, the idea that certain types of names are correlated with traits that might otherwise affect labor market treatment is demonstrated by Fryer & Levitt (2004) in their work on black naming conventions among infants born in California.

[16]This reasoning abstracts from the question of whether anticipated future discrimination affects people's current human capital choices. Given the importance of skill differences, the interplay between expectations of future market discrimination and human capital investment decisions is an important area for future work.

[17]Some recent applied econometrics work, following the important paper of DiNardo et al. (1996), uses some version of propensity-score-reweighting techniques to create counterfactual distributions of traits for use in decomposition exercises. This does not necessarily totally get around the issue of fundamental noncomparability for wildly differing distributions. As DiNardo (2002) notes, "Extremely low (or high) values of the propensity score are a potential problem. Intuitively, if the propensity score for having received the treatment is very small, this means that there are none (or few) treatment observations that look like the corresponding 'control' observation."

[18]Most economic models are silent about the sources of racially prejudicial sentiments.

[19]The particular representation adopted by Becker for prejudice, which also has been followed by later writers, is merely one of several plausible alternatives. For example, a review of racial history in the United States and elsewhere indicates that even whites with the most racist sentiments were often comfortable interacting with blacks—so long as a particular hierarchical relationship governed the interaction. We are unaware of any model that incorporates this notion of hierarchy into its representation of prejudice.

[20]Charles & Guryan (2007, 2008) point out that many employers who shut down would have to become employees elsewhere. Presuming that their negative racial sentiments are carried with them to this new role, interaction with black coworkers (over whose hiring they have no control) could bring disutility as large as the profit foregone by the refusal to hire blacks when they could control such hiring at their own firms.

[21]Black (1995) shows in a simple search model that the presence of very prejudiced employers in the market, who would not hire blacks under any circumstances, generates monopsony power in those relatively unprejudiced firms who do hire blacks. Emphasizing potential differences in fixed costs of operation, Black shows how prejudiced firms can survive long-run competition. Bowlus & Eckstein (2002) extend Black's model in a variety of important ways. Lang et al. (2005) present an interesting model of directed search in which firms announce wages and hire the most productive person among persons applying for the job. Because of prejudice, no black person gets a job to which a white has applied. In the equilibrium of their model, firms either (*a*) post a higher wage and hire only the white workers who apply or (*b*) post a lower wage and choose from among black applicants. Their model suggests that durable wage gaps can derive from the existence of relatively low levels of prejudice.

[22]Greater variance in the quality of black schooling is one justification for this assumption, but it may also be the case that differences in cultural characteristics such as language may also make it more difficult for employers to understand blacks than whites (see Lang 1986 and Grogger 2008 for work on race, language, and communication problems).

[23]Farber & Gibbons make the strong but important assumption that the full vector of worker productivity is publicly observed.

[24]The idea that prejudiced persons may feel differently toward different types of blacks is captured in the old expression "uppity black" and the more durable "lower-class black." If the

prejudiced persons especially dislike uneducated blacks, blacks may invest more in education compared with whites of similar skill, not because education lowers the error with which a statistically discriminating firm can judge a person's productivity, but because blacks wish to avoid prejudice-based negative treatment.

[25]For example, in robustness exercises, Charles & Guryan (2008) address concerns about differences in observed ability across regions by directly controlling for the difference in school quality, by race, across regions. They also instrument for the share of the current workforce that is black in a state in two-stage least squares analyses to address concerns about the possible endogeneity of this variable.

[26]Using the responses to the some of the individual racial-sentiment questions in the GSS, Lang & Lehmann present a graph that reproduces the basic pattern CG show in a graph with their unidimensional index: that average racial prejudice has fallen substantially in the United States over the past 30 years. Both CG and Lang & Lehmann presume that responses to these questions are accurate reflections of people's feelings. It is possible, of course, that all that has declined is people's willingness to report their true racial sentiment. This would not mean that the measures would be without meaning: If it is indecorous to say one is racist, someone who gives what would widely be viewed as racist answers to these questions is likely quite prejudiced.

[27]It bears mentioning here that the taxonomical challenge mentioned above in which race is concerned applies to our statements about Barack Obama. Based purely on his ancestry, there is as much logical reason for our saying that the nation recently elected its first black president as there is for saying that it elected its 44th white one.
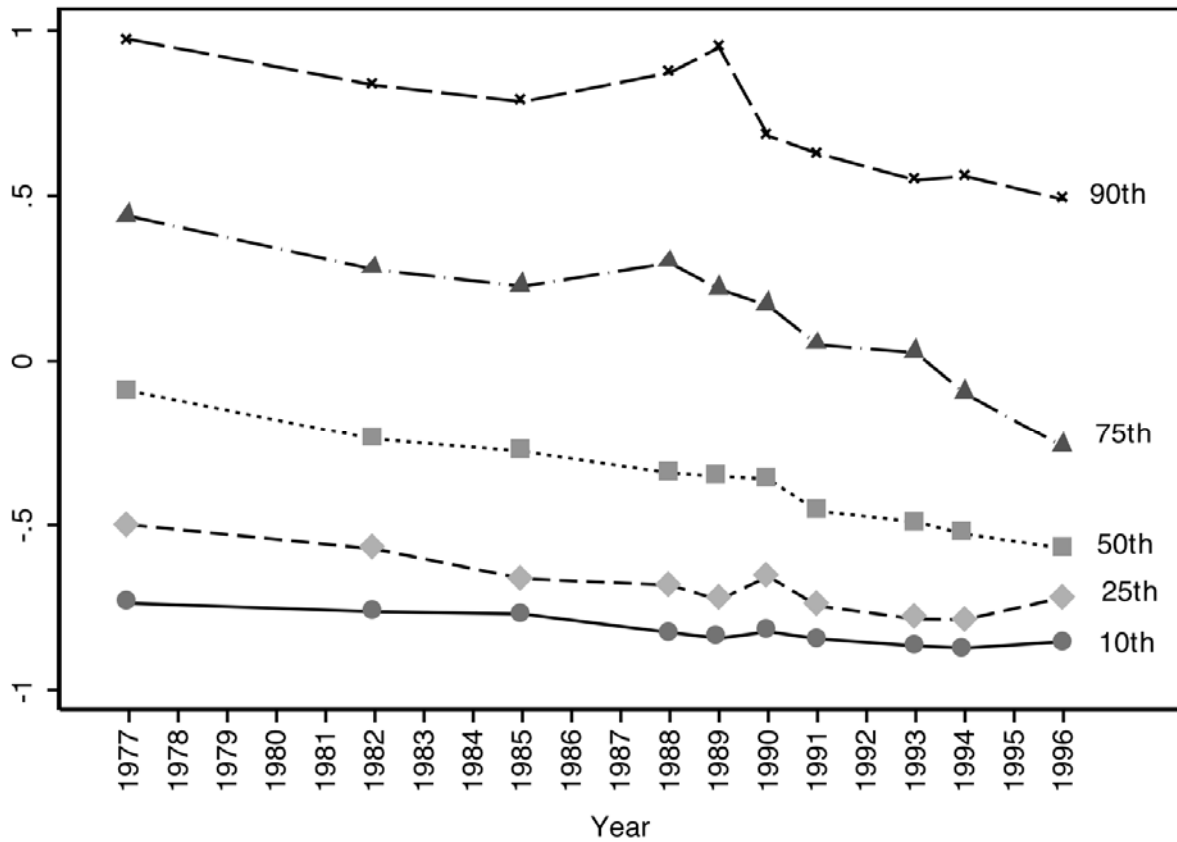
Figure 1: Trends in various percentiles of racial prejudice in the United States measured using data from the General Social Survey in certain years between 1977 and 1996.

To calculate the 10th percentile of prejudice in the United States in a given year, the 10th percentile is first calculated for each state in that year. The weighted average is then taken for that year across all states, where the weight is the state population from the 1990 census. Each of the other measures (25th percentile, median, 75th percentile, and 90th percentile) is calculated analogously.
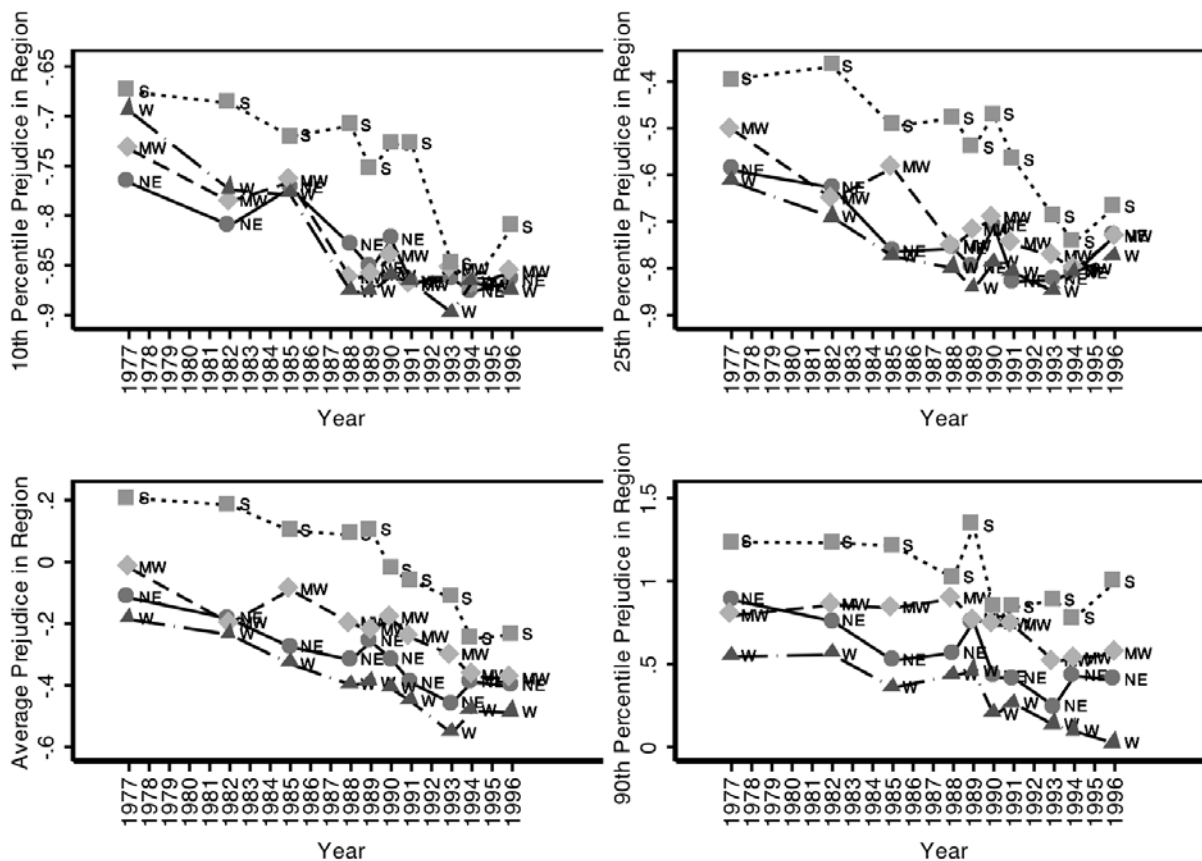
Figure 2 Trends in various measures of racial prejudice calculated from General Social Survey data for each of the four census regions in various years between 1977 and 1996.

To calculate the 10th percentile of prejudice in a region in a given year, the 10th percentile is first calculated for each state in that year. The weighted average is then taken for that year across states within a census region, where the weight is the state population from the 1990 census. Each of the other measures (25th percentile, average, and 90th percentile) is calculated analogously. Abbreviations: MW, Midwest; NE, Northeast; S, South; W, West.

Table 1:
Responses to multiple- and single-race questions in the Current Population Survey, 2002 and 2003

Race report in 2003	Race report in 2002 (percent)			
	White	Black	American Indian/Aleut/Eskimo	Asian/Pacific Islander
White	98.14	6.66	28.77	11.20
Black	0.48	91.36	1.83	0.84
American Indian/Aleut/Eskimo	0.23	0.27	54.55	0.20
Asian/Pacific Islander	0.28	0.35	1.42	75.22
Hawaiian/Pacific Islander	0.05	0.08	0.25	5.94
White-Black	0.09	0.35	0.00	0.05
White–American Indian/Aleut/Eskimo	0.58	0.01	13.01	0.08
White-Asian	0.09	0.02	0.00	1.57
White-Hawaiian	0.03	0.00	0.00	1.12
Black–American Indian/Aleut/Eskimo	0.01	0.59	0.08	0.00
Black-Asian	0.00	0.05	0.00	0.10
Black-Hawaiian	0.00	0.00	0.00	0.03
American Indian/Aleut/Eskimo-Asian	0.00	0.01	0.00	0.00
Asian-Hawaiian	0.00	0.00	0.00	1.80
White-Black– American Indian/Aleut/Eskimo	0.01	0.23	0.00	0.05
White–American Indian/Aleut/Eskimo–Asian	0.00	0.00	0.00	0.05
White-Asian–Hawaiian Pacific Islander	0.00	0.00	0.00	0.03
White-Black–American Indian/Aleut/Eskimo –Asian	0.00	0.00	0.00	0.00
Two or three races	0.01	0.00	0.08	1.60
Four or five races	0.00	0.01	0.00	0.13
Same as in 2002	98.14	91.36	54.55	75.22
Different from 2002	1.86	8.64	45.45	24.78

Note: Shown is the percent of respondents who gave each of four race answers in 2002 who then gave each of the possible race answers in 2003. Percentages sum to 100 in each column excluding the bottom two rows. The bottom panel shows the fraction of respondents, by 2002 race, who gave the same and different answers in 2003 as in 2002. Calculations are from the 2002 and 2003 Current Population Survey (CPS) outgoing rotation groups. The sample includes all respondents who were in their fourth month in sample in 2002 and their eighth month in sample in 2003. Observations are matched across CPS months using household identifier and person line number. Observations are dropped from the analysis if they do not match gender and if age responses in the adjacent years do not differ by zero, one, or two years.

Table 2:
Relationship between indication of single race in 2003 and individual characteristics

Dependent variable:	(1)	(2)	(3)	(4)
	Blacks		Whites	
	Just black in 2003		Just white in 2003	
Log(wage)_{02}	−0.008	−0.021	0.004	0.005
	(0.009)	(0.011)	(0.001)	(0.001)
$\text{Log(wage)}_{03} - \text{Log(wage)}_{02}$		−0.023		0.002
		(0.010)		(0.001)
Education/10	0.040	0.056	0.012	0.011
	(0.022)	(0.023)	(0.003)	(0.003)
Pot. experience/100	0.077	0.103	−0.004	−0.008
	(0.134)	(0.134)	(0.019)	(0.019)
Pot. experience^2/100	0.002	0.002	0.000	0.000
	(0.003)	(0.003)	(0.000)	(0.000)
R^2	0.005	0.006	0.002	0.002
N	4,777	4,777	48,279	48,279

Note: Shown are the results from four regressions, all based on data from the Current Population Survey outgoing rotation groups from 2002 and 2003. Regressions reported in columns 1 and 2 are restricted to those who self-indentified as being black in 2002. The dependent variable for the regressions reported in columns 1 and 2 is an indicator for whether the respondent reported his race as being black and black only in 2003. Regressions reported in columns 3 and 4 are restricted to those who self-identified as being white in 2002. The dependent variable for the regressions reported in columns 3 and 4 is an indicator for whether the respondent reported his race as being white and white only in 2003. In addition to the sample restrictions described in the text, regressions are further restricted to include only those observations with a valid wage in both 2002 and 2003.