

The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future

By
DEVAH PAGER

Have we conquered the problems of racial discrimination? Or have acts of discrimination become too subtle and covert for detection? This discussion serves to situate current debates about discrimination within the context of available measurement techniques. In this article, the author (1) considers the arguments from recent debates over the contemporary relevance of labor market discrimination; (2) provides a detailed introduction to experimental field methods for studying discrimination (also called audit studies), including an overview of the findings of recent audit studies of employment; (3) addresses the primary critiques of the audit methodology and the potential threats to the validity of studies of this kind; and (4) considers how we might reconcile evidence from field experiments with those from analyses of large-scale survey data, each of which points to markedly different conclusions. Only by gathering rigorous empirical evidence can we begin to understand the nature of race and racial discrimination in labor markets today.

Keywords: discrimination; race; employment; field experiment; audit study

In December 2002, the Equal Employment Opportunity Commission (EEOC) filed a lawsuit in the Wisconsin Federal Court against the Target Corporation, alleging discrimination against black job applicants at nearly a dozen Wisconsin stores. In depositions for the lawsuit, Target employees admitted to routinely

Devah Pager is an assistant professor of sociology and faculty associate of the Office of Population Research at Princeton University. Her research focuses on institutions affecting racial stratification, including education, labor markets, and the criminal justice system. Her current research has involved a series of field experiments studying discrimination against minorities and ex-offenders in the low-wage labor market. Recent publications include "Walking the Talk: What Employers Say versus What They Do" (American Sociological Review, 2005, with Lincoln Quillian) and "The Mark of a Criminal Record" (American Journal of Sociology, 2003).

NOTE: Support for this research comes from an NSF CAREER Award (NES-0547810). Views expressed in this document are those of the author and do not necessarily represent those of the granting agency.

DOI: 10.1177/0002716206294796

destroying the job applications of black individuals who attended job fairs held at several Milwaukee universities.¹

Examples of blatant forms of discrimination appear sporadically in a blitz of media attention. As much as these examples provide vivid illustration of lingering forms of racial bias, they simultaneously reinforce the notion that acts of discrimination in contemporary America are rare events committed by unusually malevolent actors. Under more typical circumstances, discrimination in America appears to have all but disappeared. Indeed, the presence of prominent black personalities, athletes, actors, and politicians provides an image of an open door to opportunity for blacks, one no longer conditioned by the stigma of skin color. In his book *Creating Equal: My Fight against Racial Preferences* (2000), Ward Connerly (sponsor of the successful proposition in California to end affirmative action) argued that liberals cling to a misguided belief in the persistence of racism, characterized by the “need to believe that Rosa Parks is still stuck in the back of the bus, even though we live in a time when Oprah is on a billboard on the side of the bus.”² Perhaps, then, periodic examples of discriminators “caught in the act” represent only extreme aberrations. Dramatic cases of discrimination may get extensive publicity even if they represent rare occurrences overall.

On the other hand, contemporary forms of discrimination may be simply more subtle and covert, leading to less frequent detection and awareness by the general public. In the contemporary United States, social and legal proscriptions against discrimination are strong, placing pressure on potential discriminators to conceal their motives in ways that are consistent with norms of color blindness. Employers (or other gatekeepers) who retain strong preferences for members of a particular race thus face clear incentives to mask their discriminatory actions behind nonracial justifications. It could be the case, then, that discrimination remains fairly routine in certain contexts, despite infrequent public exposure.³

Debates about the contemporary relevance of discrimination have been difficult to resolve, in part because of the challenges in identifying, measuring, and documenting its presence or absence in all but extreme cases. In this article, I consider the utility of one approach—the field experiment—as a tool for measuring employment discrimination. Though field experiments in this context are not without limitations, they offer certain unique advantages—namely, the opportunity to observe discrimination directly. The following discussion proceeds in four sections: First, I consider the arguments from recent debates over the contemporary relevance of discrimination in shaping employment opportunities. Second, I provide a detailed introduction to experimental field methods for studying discrimination (also called audit studies), an approach still relatively rare in studies of employment. I also present an overview of the findings of recent audit studies of employment. Third, I address the primary critiques of the audit methodology and examine potential threats to the validity of studies of this kind. Finally, I consider how we might reconcile evidence from field experiments with those from analyses of large-scale survey data, each of which points to markedly different conclusions about the relevance of discrimination in contemporary

labor markets. Overall, this discussion serves to situate current debates about discrimination within the context of available measurement techniques. It is only by gathering rigorous empirical evidence that we can make headway in understanding the nature of race and racial discrimination in labor markets today.

Debates about the contemporary relevance of discrimination have been difficult to resolve, in part because of the challenges in identifying, measuring, and documenting its presence or absence in all but extreme cases.

The Declining Significance of Race (as an Explanatory Variable)

Much has changed in American race relations since the middle of the twentieth century. The civil rights movement brought with it a wave of reform, undermining the previously entrenched system rooted in a racialized allocation of opportunity. Historic legislative and court decisions banning segregated schools, prohibiting discrimination in employment and public accommodations, and extending the franchise created a new horizon of opportunities previously unavailable to African Americans. Antidiscrimination law and affirmative action provided twin vehicles for the enforcement and promotion of equal opportunity for America's racial minorities, spurring an unprecedented growth of black upward mobility (Wilson 1978; Harrison and Bennett 1995). With the shifting legal context, the social context of discrimination was transformed dramatically as well. Whereas in 1940 fewer than half of Americans believed that blacks should attend the same schools as whites or have the same chances of getting a job, by 1970 the balance had shifted toward an endorsement of the principles of racial equality; by 1995 more than 95 percent of Americans would support the ideals of racial integration and equality of opportunity.⁴ In the wake of this historic transformation, many grew confident that American society had moved beyond the fault lines of race. Lingering signs of racial inequality could be viewed as the eroding vestiges of the previous era rather than as the continuing product of contemporary racial injustice.

Prominent intellectuals of the post-civil rights era were quick to document this transformation. Most notably, in 1978, William Julius Wilson published the now classic treatise on black America, titled *The Declining Significance of Race*, in

which he skillfully argued that the problems facing African Americans in the modern industrial period had more to do with class than race. Discrimination, Wilson argued, was no longer paramount in shaping the outcomes of blacks; rather, issues of poverty, education, and employment opportunities were far more important to improving the well-being of individuals from all race groups. Discrimination was not the problem, but rather a lack of jobs, caused by structural changes in the economy, which underlay black—and other groups’—disadvantage. Wilson’s book clearly picked up the developing zeitgeist of the time. Indeed, in the thirty years since its publication, we have seen a notable decline in attention to the problems of racial discrimination in academic and policy discussions.⁵

Consistent with notions of a “declining significance of race,” racial disparities on a number of key indicators have diminished or disappeared since the 1960s. Rates of high school graduation have narrowed to just a few percentage points difference, and the black-white test score gap appears to be following a similar trajectory (Jencks and Phillips 1998).⁶ These improvements in the human capital attainment of blacks, along with a liberalization of opportunity, have facilitated greater performance in the labor market, with blacks becoming increasingly well represented in occupational sectors previously dominated by whites, and a shrinking of the wage gap through 1980 (Farley 1997; Mare 1995; Harrison and Bennett 1995).

Despite visible improvements, however, blacks continue to lag behind whites on key dimensions of inequality. Particularly among those at the bottom half of the distribution, rapid gains beginning in the 1960s slowed, and in some cases reversed, during the 1980s and 1990s. Even at the high point of economic expansion in the late 1990s when unemployment rates were dropping steadily for all groups, black men were still more than twice as likely to be unemployed relative to their white counterparts. Over time blacks, and young black men in particular, have become increasingly likely to drop out of the labor market altogether when faced with the prospect of long-term unemployment or marginal employment opportunities (Holzer, Offner, and Sorensen 2005).⁷

How can we explain these persistent disparities? The truth of the matter is, the employment problems of blacks are vastly overdetermined. Far more factors contribute to black employment problems than would be necessary to produce the trends we observe: the manufacturing sector declined, jobs moved from the central city, black test scores have lagged behind those of whites as the returns to skill have increased, blacks have less effective social networks for finding work, blacks face increasing competition from women and immigrants (Wilson 1987; Freeman and Holzer 1986; Murnane, Willett, and Levy 1995; Waldinger 1999). Interestingly, in this litany of possible explanations, rarely nowadays do we hear mention of the oldest and most basic interpretation. Does discrimination continue to contribute to the employment problems of African Americans?

According to most Americans, the answer is no. The majority of white Americans believe that a black person today has the same chance at getting a job as an equally qualified white person, and only a third believe that discrimination is an important explanation for why blacks do worse than whites in income, housing, and jobs.⁸ Public opinion thus favors the idea that discrimination is of vanishing importance, at least as a direct cause of present-day inequalities.

Scholarly opinion likewise remains divided on the question of discrimination. Social psychologists have extensively documented the subtle distortions that take place when race is involved in the course of reasoned evaluations. Despite the widespread conscious endorsement of racial equality, deep-seated stereotypes about the intelligence, work ethic, criminality, and cultural dispositions of various groups continue to frame our evaluations and decision making in social situations.⁹

In contrast to social psychological research that shows a strong persistence of racial stereotypes and discrimination, a growing body of research in sociology and economics has challenged the notion that contemporary labor market outcomes are influenced by race. Wilson's work, mentioned earlier, highlights the importance of structural changes in the economy that, although disproportionately affecting poor and working-class blacks, are race-neutral in origin. More recent work has moved from structural to individual explanations, emphasizing the growing importance of skill in today's economy and the persistent black-white skill gap as a key source of contemporary racial disparities. A series of influential studies, for example, indicate that when relevant individual characteristics—in particular, cognitive ability—have been accounted for, racial disparities in wages among young men narrow substantially or disappear (Farkas 2003; Farkas and Vicknair 1996; Murnane, Willett, and Levy 1995; O'Neill 1990). Neal and Johnson (1996), for example, used data from the National Longitudinal Study of Youth to analyze the black-white wage gap for a cohort born between 1958 and 1963. They found that a measure of cognitive ability (AFQT) explains fully three-quarters of the wage gap for young men. This line of research has reinforced the view that the vast majority of the employment problems of young minority men can be explained by skill or other individual deficiencies, rather than any direct effect of discrimination. Economist James Heckman (1998, 101) summarized this position most clearly: "Most of the disparity in earnings between blacks and whites in the labor market of the 1990s is due to differences in skills they bring to the market, and not to discrimination within the labor market." He went on to describe discrimination as "the problem of an earlier era."

Have we conquered the problems of racial discrimination? Or have acts of discrimination become too subtle and covert for detection? These questions are difficult to answer using standard techniques of observation and analysis. In the following discussion, I consider the use of field experiments for studying discrimination in low-wage labor markets. Because field experiments remain relatively uncommon in studies of employment (and thus fewer resources exist for those interested in pursuing this methodology), there is some value to providing a fairly detailed description of both the conceptual and practical dimensions of this approach.

The Methodology of Field Experiments for Studies of Discrimination

Experimental methods provide a powerful means of isolating causal mechanisms. Traditional experiments typically begin with clearly defined "treatment"

and “control” conditions, to which subjects are randomly assigned.¹⁰ All other environmental influences are carefully controlled. A specific outcome variable is then recorded to test for differences between groups. Often subjects are not told the purpose of the experiment to ensure a naive or “natural” reaction to the experimental condition. *Field experiments* blend experimental methods with field-based research, relaxing certain controls over environmental influences to better simulate real-world interactions. While retaining the key experimental features of matching and random assignment important for inferences of causality, this approach relies on real contexts (e.g., actual employment searches, real estate markets, consumer transactions, etc.) for its staged measurement techniques. For example, rather than asking undergraduate subjects to rate hypothetical job applicants in a lab experiment, a field experiment would present two equally qualified job applicants to real employers in the context of real job searches.

*Field experiments blend experimental methods
with field-based research, relaxing certain
controls over environmental influences to better
simulate real-world interactions.*

Field experiments designed specifically for the measurement of discrimination are typically referred to as audit studies. The audit methodology was first pioneered in the 1970s with a series of audits conducted by the Department of Housing and Urban Development to test for racial discrimination in real estate markets (Yinger 1995; Wienk et al. 1979; Hakken 1979). The approach has since been applied to numerous settings, including mortgage applications, negotiations at a car dealership, and hailing a taxi (Turner and Skidmore 1999; Ayres and Siegelman 1995; Ridley, Bayton, and Outtz 1989; Yinger 1995; Massey and Lundy 2001; Cross et al. 1990; Turner and Skidmore 1991; Bendick, Jackson, and Reinoso 1994; Neumark 1996).¹¹ In the case of employment discrimination, two main types of audit studies offer useful approaches: correspondence tests and in-person audits.

Correspondence tests

The correspondence test approach, so named for its simulation of the communication (correspondence) between job applicants and employers, relies on fictitious matched resumes submitted to employers by mail or fax. In these studies,

two or more resumes are prepared reflecting equal levels of education and experience. The race (or other group characteristic) of the fictitious applicant is then signaled through one or more cues, with race randomly assigned to resume type across employers (i.e., minority status is assigned to one resume for half the employers, the other resume for the other half; this is to ensure that any unobserved differences between resumes will not be correlated with the measured effects of race).¹² Reactions from employers are then typically measured by written responses (to staged mailing addresses) or callbacks (to voice mail boxes) for each applicant. An exemplary study of this kind was recently conducted by Marian Bertrand and Sendhil Mullainathan (2004).¹³ In this study, the researchers prepared two sets of matched resumes reflecting applicant pools of two skill levels. Using racially distinctive names to signal the race of applicants, the researchers mailed out resumes to more than thirteen hundred employers in Chicago and Boston, targeting job ads for sales, administrative support, and clerical and customer services positions. The results of their study indicate that white-sounding names were 50 percent more likely to elicit positive responses from employers relative to equally qualified applicants with “black” names.¹⁴ Moreover, applicants with white names received a significant payoff to additional qualifications, while those with black names did not. The racial gap among job applicants was thus higher among the more highly skilled applicant pairs than among those with fewer qualifications.

The advantage of the correspondence test approach is that it requires no actual job applicants (only fictitious paper applicants). This is desirable for both methodological and practical reasons. Methodologically, the use of fictitious paper applicants allows researchers to create carefully matched applicant pairs without needing to accommodate the complexities of real people. The researcher thus has far more control over the precise content of “treatment” and “control” conditions. Practically, the reliance on paper applicants is also desirable in terms of the logistical ease with which the application process can be carried out. Rather than coordinating job visits by real people (creating opportunities for applicants to get lost, to contact the employer under differing circumstances [e.g., when the employer is out to lunch, busy with a customer, etc.], and so on), the correspondence test approach simply requires that resumes are sent out at specified intervals. Additionally, the small cost of postage or fax charges is trivial relative to the cost involved in hiring individuals to pose as job applicants.

While correspondence tests do have many attractive features, certain limitations of this design have led other researchers to prefer the in-person audit approach.

Problems signaling key applicant characteristics. Because correspondence tests rely on paper applications only, all relevant target information must be conveyed without the visual cues of in-person contact. In the case of gender or ethnicity, identifiable names can easily convey the necessary information using gender-specific or ethnically identifiable names (see Riach and Rich 2002; Lahey 2005). In the case of age discrimination, several studies have relied on high school graduation dates to convey the applicants’ age difference (Bendick,

Brown, and Wall 1999; Lahey 2005). Researchers who wish to study black-white differences, on the other hand, face a somewhat more challenging task. The Bertrand and Mullainathan (2004) study discussed above, for example, used racially distinctive names to signal the race of applicants. Names like “Jamal” and “Lakisha” signaled African Americans, while “Brad” and “Emily” were associated with whites. While these names are reliably associated with their intended race groups, some critics have argued that the more distinctive African American names are also associated with lower socioeconomic status, thus confounding the effects of race and class. Indeed, mother’s education is a significant (negative) predictor of a child having a distinctively African American name.¹⁵ The use of names to test for black-white differences, then, is complicated by the social context in which racially distinctive names are situated.

Other correspondence test studies have used the “extracurricular activities” or “voluntary memberships” section of the resume to bolster the signal of the applicant’s race.¹⁶ Membership in the student league of the NAACP, for example, would strongly signal an African American applicant. The matched “white” applicant would then be given a race-neutral activity (e.g., Student Democratic Alliance), which, in the absence of any racial identifiers, is typically (by default) associated with whites.¹⁷ Whatever strategy is used, it is important that resumes are pretested carefully before using them in the field. Names, extracurricular activities, neighborhoods, and high schools may each have connotations that are not readily apparent to the researcher. Directly assessing these connotations/associations is an important first step in developing the materials necessary for a strong test of discrimination.

Limited sample of jobs. One other important limitation of the correspondence test method relates to the types of jobs available for testing. The type of application procedure used in correspondence tests—sending resumes by mail—is typically reserved for studies of administrative, clerical, and other white-collar occupations. The vast majority of entry-level jobs, by contrast, more often require in-person application procedures. For jobs such as busboy, messenger, laborer, or cashier, for example, a mailed-in resume would appear highly out of place. Any study of the low-wage labor market then would require in-person application procedures. While in-person audit studies also face a restricted range of job openings, in-person application procedures allow for a substantially wider pool than can be achieved through paper applications alone.

In-person audits

The use of in-person audits, as opposed to mail-in resumes, represents a more elaborate simulation of the hiring process.¹⁸ In-person employment audits involve the use of matched pairs of individuals (called testers) who pose as job applicants in real job searches. Applicants are carefully matched on the basis of age, height, weight, physical attractiveness, interpersonal style, and any other employment-relevant

characteristics to which employers may respond in making hiring decisions. As with correspondence tests, resumes are constructed for each applicant that reflect equal levels of schooling and work experience. In addition, the in-person presentation of confederate job seekers must be carefully controlled. Testers must participate in extensive training to familiarize themselves with the details of their profile and to learn to present themselves to employers according to a highly structured protocol. Daily debriefings are necessary to ensure that the implementation of each test proceeds according to plan (see the appendix). Though in-person audits are time-consuming and require intensive supervision, the approach offers several advantages over correspondence studies. In-person audits provide a clear method for signaling race (through the physical presentation of job applicants); they allow for a wide sample of entry-level job types (which often require in-person applications); and they provide the opportunity to gather both quantitative and qualitative data, with information on whether the applicant receives the job as well as how he or she is treated during the interview process. For those readers interested in the nuts and bolts of audit design, I include a detailed methodological appendix at the end of this article.

*In-person employment audits involve the use of
matched pairs of individuals (called testers)
who pose as job applicants in real job searches.*

Table 1 presents the results of a number of recent audit studies conducted in cities across the country. Each study comes to the same basic conclusion—that race matters in hiring decisions. Estimates of the magnitude of discrimination do, however, vary across studies, with whites anywhere from 1.5 to 5 times more likely to receive a callback or job offer relative to equally qualified black applicants.¹⁹ Differences across cities may account for some degree of variation (with Washington, D.C., demonstrating the highest levels of discrimination), as well as differences in the specific design of each experiment (e.g., the level of education presented, the gender of testers, the outcomes measured, etc.). Relative to in-person audits, the correspondence test shows less evidence of discrimination, in part because call-back rates are much lower overall in response to mailed applications.²⁰ At the conclusion of this article, I consider some future directions for audit research that would improve the continuity across studies to allow for more straightforward interpretations of the nature of variation in discrimination across labor markets.

TABLE 1
SUMMARY OF RECENT AUDIT STUDIES OF RACIAL
DISCRIMINATION IN EMPLOYMENT^a

Audit Study	Application Method	Location	Sample Size	Interacting Characteristics	Gender of Testers	Education on Resumes	Outcome Measure	Ratio (White/Black)
Pager and Western (2005)	In-person	New York	522	Criminal record	Men	High school diploma	Callback or job offer	2.0
Bertrand and Mullainathan (2004) ^b	Correspondence	Chicago; Boston	1,323	Skill level; neighborhood	Men and women	High school diploma to college grad	Callback	1.5
Pager (2003) ^c	In-person	Milwaukee	350	Criminal record	Men	High school diploma	Callback	2.4
Bendick, Jackson, and Reinoso (1994) ^d	In-person	Washington, D.C.	149	None	Men and women	Two years' college	Job offer	5.0
Turner, Fix, and Struyk (1991) ^e	In-person	Chicago; Washington, D.C.	476	None	Men	High school diploma	Job offer	3.0

a. A similar (more inclusive) table was put together by Marc Bendick Jr. and is available at <http://www.bendickegan.com/pdf/Bendick%20SPSS1%202004%20Testing.pdf>, Table 1. For a similar overview of correspondence studies in comparative context, see Riach and Rich (2002).

b. The reported sample size for the Bertrand and Mullainathan (2004) study is derived from their Table 2. This study finds no difference in the rate of discrimination between Boston and Chicago. The study found no gender differences in overall callback rates (1.52 vs. 1.49 for males and females, respectively); when comparing males and females in sales jobs, however (male resumes were sent almost exclusively to sales positions), they found a much larger race effect for men (1.52 vs. 1.22 for males and females in sales jobs, respectively).

c. Estimates of race differences are based on between-pair comparisons.

d. The white/black ratio for this study is calculated by multiplying the percentage of testers who were interviewed (48.3 and 39.6 percent for whites and blacks, respectively) by the percentage of testers who, given an interview, were offered jobs (46.9 and 11.3 percent, respectively). Bendick, Jackson, and Reinoso (1994) found a somewhat higher incidence of discrimination against African American women than men.

e. Turner, Fix, and Struyk (1991) found rates of discrimination that are significantly higher in Washington, D.C., than in Chicago.

Variation notwithstanding, in all cases we come to the conclusion that race has large effects on employment opportunities, with a black job seeker anywhere between 50 and 500 percent less likely to be considered by employers as an equally qualified white job applicant. The matched designs allow us to separate speculation about applicant qualifications (supply-side influences) from the racial attributions or biases of employers (demand-side influences). While these studies remain silent on the many supply-side factors that may also contribute to the employment difficulties of young black men, they speak loud and clear about the significance of employer demand in shaping the opportunities available to young black and white job seekers. Before applicants have an opportunity to demonstrate their capabilities in person, a large proportion are weeded out on the basis of a single categorical distinction. Results from audit studies thus provide one clear source of evidence with which to address debates about the contemporary relevance of discrimination.

*We come to the conclusion that race
has large effects on employment opportunities,
with a black job seeker anywhere between
50 and 500 percent less likely to be
considered by employers as an equally
qualified white job applicant.*

Critiques of the Audit Method

While most researchers view the audit methodology as one of the most effective means of measuring discrimination, the approach is not without its critics. Before assuming that a field experimental design automatically confers high levels of both internal and external validity, the possible vulnerabilities of the audit methodology deserve careful consideration. Economist James Heckman is among the most vocal critics of the audit methodology, particularly when used to study the effects of race. Heckman's primary criticism focuses on the problems of effective matching.²¹ The validity of an audit study relies on its success in presenting two otherwise equally qualified job applicants who differ only by race. Given the vast number of characteristics that can influence an employer's evaluation, however, it is difficult to ensure that all such dimensions have been effectively

controlled. Because, race is not something that can be experimentally assigned, we must believe that audit researchers have been successful in identifying and matching on all relevant characteristics—something that, according to Heckman, leaves substantial room for bias. Heckman's primary critique focuses on the problem of unobservables—those characteristics “unobservable to the audit study [researchers], but . . . at least somewhat visible to the prospective employer and acted on in hiring . . . decisions” (Heckman 1998, 109). According to Heckman, blacks and whites (at the population level) may differ in the average and/or distribution of important characteristics. As an example, consider a hypothetical case in which whites on average have a faster response time in interview interactions than blacks. That is to say, the delay in seconds between a question posed by an interviewer and the initiation of response is shorter on average for whites than for blacks. (To be sure, response time is just one potential example, and I emphasize that it is a case that to my knowledge has no empirical basis. Heckman himself does not suggest any concrete examples of potentially relevant unobservables that could affect hiring outcomes; but it is instructive to consider a concrete hypothetical case for the purpose of clarity.) Because any difference in response time would be extremely subtle, it may not be immediately recognizable to researchers and may even register for employers only at a subliminal level. Nevertheless, if this trait produces an incremental advantage for the individual with a faster response time—because he or she is perceived as sharper or more engaged—we may mistake the employer's response for discrimination when in fact nonracial evaluations are driving the differential response.

A related problem emerges if blacks and whites differ on key characteristics, not on the average, but in the level of dispersion. To continue with the same example, imagine a case in which blacks and whites each have a mean response time of 0.5 seconds, but blacks demonstrate greater heterogeneity along this dimension than whites. Differential results may then be observed depending on the overall qualifications of the testers relative to the requirements of the job. If testers are highly qualified relative to the positions they apply for (which tends to be the case in audit studies), differential dispersion on any key variable will favor the group with lower dispersion (because a smaller proportion of applicants in this group will be at the low end of the tail relative to a high-dispersion group).

Heckman's critique raises some important considerations and surely encourages a more rigorous scrutiny of the audit methodology. In each case, it is worth considering when and how these concerns can be effectively addressed. Heckman's concern is that if, on average, blacks and whites differ in the mean or variance on any unobserved productivity-related variable, estimates from matched-pair studies will be biased by design. If auditors were randomly drawn from the population and matched on a rote basis according to readily measurable characteristics, this critique would surely be valid. It is a mistake, however, to assume that the researcher is at a necessary disadvantage relative to the employer in identifying productivity-related characteristics. In fact, the researcher is herself or himself an employer in the planning and implementation of an audit study.

The job of a tester is not an easy one, and finding a suitable team to complete this type of project requires extensive screening and careful selection. The job requires solid writing skills (for the written narratives that follow each audit), good communication skills (to communicate the necessary information in an interview, to make a good impression on the employer); high levels of motivation (to keep up day after day), reliability (to accurately conduct and report each test), navigation skills (to find locations throughout the city); and an endless number of other qualifications. Thus, apart from the more explicit traits of height, weight, race, and age, researchers must search for testers who can perform well in an intensely demanding position.²² As an employer, the researcher must identify subtle cues about applicants that indicate their ability to perform. Whether or not these cues are explicit, conscious, or measurable, they are present in a researcher's evaluation of tester candidates as they are for employers' evaluations of entry-level job applicants. Like employers, researchers are affected by both objective and subjective/subconscious indicators of applicant quality in their selection and matching of testers, in ways that should ultimately improve the nuanced calibration of test partners.

A related concern of Heckman has to do with the possibility that matching (even when done successfully) may itself produce distortions in the hiring process. Because audit partners are matched on all characteristics that are most directly relevant to the hiring process (education, work experience, physical appearance, interpersonal skills, etc.), employers may be forced to privilege relatively minor characteristics simply out of necessity of breaking the tie. "By taking out the common components that are most easily measured, differences in hiring rates as monitored by audits arise from the idiosyncratic factors, and not the main factors, that drive actual labor markets" (Heckman 1998, 111). If employers care only marginally about race, but are confronted with two applicants equal on all other dimensions, race may take on greater significance in that particular hiring decision than is true under more normal circumstances, when evaluating real applicants who differ according to multiple dimensions.

Again, this critique is an important one, though in this case one that can be addressed more easily. If the only outcome of interest in an audit study is whether an applicant gets the job, Heckman's concern is certainly relevant. If forced to choose a single hire, employers will use whatever basis for differentiation exists, whether that particular attribute is valued highly or not. Audit studies that measure callbacks as an outcome variable, by contrast, avoid situations in which employers can choose only one applicant. In fact, employers typically interview an average of eight applicants for each entry-level job they fill. If race is only a minor concern for employers, we would expect both members of an audit pair to make it through the first cut. To the extent that race figures prominently even in these early rounds of review, we can infer that race is invoked as more than a mere tie-breaker. In these cases, the evidence of race-based decision making is quite strong.²³

A third important critique of the audit methodology raises the problem of experimenter effects, or the possibility that the expectations or behaviors of testers can influence the audit results in nonrandom ways. For example, if a

tester expects to be treated poorly by employers, he or she may appear more withdrawn, nervous, or defensive in interactions. The nature of the interaction may then create a self-fulfilling prophecy in which the tester experiences poor outcomes, but for reasons unrelated to the experimental condition (e.g., his or her race). Indeed, the possibility of experimenter effects represents one of the most serious threats to the validity of the audit experiment. While there is no way to conclusively rule out the possibility of experimenter effects, several precautions can be taken to minimize the problem. First, effective training and supervision are critical to the successful implementation of an audit study. Testers must be exceedingly familiar with their assumed profiles and the audit protocol, such that appropriate responses to employer queries become almost automatic. Extensive role-plays, videotaped interviews, and practice audits help testers to become comfortable with their role and to gain important feedback on their performance. Likewise, during the course of the fieldwork, daily debriefings and regular troubleshooting sessions are critical to identify any potential problems or to refine the protocol in ways that best suit the specifics of the study. Finally, after the fieldwork is completed, it is possible to conduct an indirect check on the problem of experimenter effects. Typically a significant proportion of tests are conducted with little or no in-person contact, either because the employer is not present or does not have time to meet with the applicant. By comparing audit outcomes for testers who did and did not interact with employers, we can assess the degree to which in-person interaction leads to a different distribution of results. If testers are acting in ways that fulfill their expectations of discrimination, we would expect outcomes for those tests conducted with interaction to show greater evidence of differential treatment than those without. If the results show no difference, or show weaker evidence of differential treatment, we can be more confident that experimenter effects are not driving the results.²⁴ As a final note, it is worth reiterating that a key advantage of correspondence tests (relative to in-person audits) is the ability to present matched pairs of resumes to employers without the use of real testers. That these studies typically also demonstrate consistent evidence of discrimination provides one further reassurance that the outcomes from in-person audit studies are not merely the product of mismatched testers or participants' enacted expectations (Bertrand and Mullainathan 2004).

Reconciling Competing Measures of Discrimination

If employment discrimination is indeed as great a problem as the results of field experiments suggest, how can we reconcile this conclusion with competing evidence demonstrating a small or nonexistent wage gap between equally qualified blacks and whites? As mentioned earlier, recent analyses of large-scale survey data indicate that after statistically controlling for a wide range of individual characteristics (cognitive ability, in particular), most or all of the wage gap

between young black and white men can be eliminated. The implication of this line of research is that discrimination plays little role in determining the economic attainment of young men. How then can we account for the substantial evidence of discrimination indicated by the audit results?

I consider this discrepancy from several perspectives. In this discussion, I do not interrogate the validity of specifications used in existing studies of wage inequality, taking at face value the reliability of the analyses. It should be acknowledged, however, that a wide range of estimates of the black-white wage gap exist within the survey literature, with numerous studies reporting a large and persist racial gap (see, e.g., Cancio et al. 1996; Neumark 1999). Likewise, analysis following the Neal and Johnson (1996) paper has found some evidence that the original results overstate the extent to which cognitive ability can account for racial disparities in wages. Carneiro, Heckman, and Masterov (2003), for example, found that adjustments for years of schooling at the time the respondent's cognitive ability was measured lead to the reemergence of a substantial wage differential. Tomaskovic-Devey, Thomas, and Johnson (2005, 76) found that, while wages measured in early adulthood show little evidence of racial inequality (because there is little wage dispersion to begin with), the racial wage gap then grows across the life course, reaching 14 percent by the time these men reach forty (controlling for cognitive ability and other person-specific characteristics). It thus remains an open question exactly how much of the wage gap can be explained by individual attributes like cognitive ability. Nevertheless, this discussion takes as its starting point the argument that, after extensive controls, analyses of black-white wage disparities can be largely explained by observed individual characteristics. In contrast to evidence from audit studies, these survey results suggest little reason to be concerned with the problems of discrimination.

Before we dismiss discrimination as "the problem of an earlier era," however, it is worth considering under what circumstances these discrepant findings could be meaningfully reconciled. First, it is important to keep in mind that the employment relationship is characterized by a number of discrete decisions, including hiring, wage setting, promotion, and termination. Discrimination may affect all, none, or some of these decisions. Varying incentives or constraints characteristic of different employment decisions can mediate the emergence of discrimination in important ways. For example, there is reason to believe that decisions about whom to interview and whom to hire may be more susceptible to discriminatory bias relative to those decisions made at later stages of the employment relationship. Both economic theories of statistical discrimination and social-psychological theories of unconscious bias predict that discrimination will be most pronounced when objective information about the target is limited or unreliable.²⁵ Indeed, the amount of information employers have about applicants at this point of introduction is at a minimum. We would expect, then, that, whether exerted consciously or not, underlying assumptions about race and productivity will be most likely to color evaluations of blacks at earlier stages in the employment process (i.e., hiring) than at later stages (i.e., wage setting/termination decisions), when more objective performance indicators have become available.²⁶ Likewise, we

would expect to see the effects of discrimination reflected in differential employment rates rather than wage rates.

In addition to information asymmetries that affect employers' perceptions about workers, workers' perceptions of employer decision making is likewise most limited at the point of hire. Uncertainty about the competing applicant pool, about the employer's preferences, and about the job itself makes acts of discrimination particularly difficult to detect at the initial point of contact. At later stages, by contrast, workers have more information with which to compare their treatment to others in comparable positions. Employers concerned about detection, or even accusations of discrimination, will thus be safer eliminating black applicants early on.

Finally, aside from the distinct conditions that characterize different stages of the employment process, it is important also to consider their interdependence. If individuals who have been refused employment opportunities are excluded from estimates of wage disparities, hiring discrimination against blacks will result in a more select sample of black wage earners than whites. In fact, barriers to labor market entry—including but not limited to hiring discrimination—will lead young black men to remain unemployed longer and may cause them to drop out of the labor force altogether. In this case, our estimates of wage disparities may be distorted by the large numbers of black men missing from the sample of wage earners. Trends in labor force participation indeed show high levels of labor force nonparticipation among young black men and a growing black-white disparity in rates of joblessness.²⁷ Because individuals who are not working and not looking for work are excluded from standard economic analyses, increases in labor force nonparticipation among blacks can substantially distort conventional measures of racial wage disparities.²⁸ According to Western and Pettit (2005, 573), "By 1999, the high rate of black joblessness inflated black relative earnings by between 7 and 20% among working age men, and by as much as 58% among young men." According to this and other analyses, black-white wage equality is in large part an artifact of decreasing labor force participation among the most disadvantaged young black men. Without effectively accounting for the processes that precede labor force participation—such as discrimination, discouragement, incarceration, or other sources of selection—wage estimates can account for only one incomplete picture of the larger employment process.

Discrepancies between wage estimates and measures of discrimination at the point of hire may then reflect one (or both) of two processes: First, incentives to discriminate at the point of hire are greater than those at later stages, due to information asymmetries that affect both employer and worker perceptions. In this case, wage estimates and hiring discrimination estimates may both represent accurate reflections of discrimination at different stages of the employment relationship. Second, discrimination at the point of hire may distort wage estimates by contributing to the large numbers of young black men who are unemployed or who drop out of the labor force altogether. In this case, wage estimates reflect only the more "select" members of the black population, artificially reducing the contrast with less select white workers. In either scenario, discrimination at the point

of hire remains an active barrier to employment for young black men. Indeed, the magnitude of the results shown here, across a wide range of studies, suggests that barriers to labor market entry are likely to represent a serious constraint on the achievement of economic self-sufficiency among young black men today.²⁹

Directions for the Future

This article has sought to familiarize readers with the audit methodology and to consider how we might reconcile the findings from audit studies with those from recent analyses of survey data. But there is much more to be done. In this final section, I consider several directions for future research that could substantially improve our estimates and interpretation of audit results measuring employment discrimination.

First, it would be useful for future research to develop a standardized audit framework that could be replicated across testing sites and over time, similar to the model pursued in recent housing audits (Turner et al. 2002). Though several researchers have conducted multicity studies, no researcher has attempted to include more than two sites, thus limiting our comparative perspective on discrimination across labor markets and over time. Second, the introduction of additional experimental variables (e.g., skill, education, written references, etc.) would allow researchers to calibrate the effects of race against other key labor market determinants. The effects of race matters *relative to what*? As shown in Table 1, a few prior studies have included one or more variables in addition to race. Pursuing such designs would help to translate the effects of race into another meaningful metric. We might ask, for example, How do the advantages of whiteness compare to the advantages of having a high school degree? How many years of additional work experience would a black applicant need to be competitive with an otherwise comparable white applicant? Though experimental designs are constrained in the number of conditions that can be included in any single study, the inclusion of key comparison variables would greatly contribute to the interpretation and translation of audit results.

Finally, additional research should make efforts to empirically map the findings from audit studies onto population surveys of job search and employment patterns. This next step is important in helping us to assess how the prevalence of discrimination encountered by testers corresponds to discrimination experienced by real job seekers with similar characteristics. Indeed, Heckman (1998) warns us not to interpret the findings from audit studies as accurate measures of the prevalence of discrimination in everyday life. "The impact of market discrimination is not determined by the most discriminatory practices in the market, or even by the average level of discrimination among firms, but rather by the level of discrimination at the firms where ethnic minorities or women actually end up buying, working and borrowing. It is at the margin that economic values are set. . . . Purposive sorting within markets eliminates the worst forms of discrimination" (pp. 102-3). Heckman's argument suggests that minority workers

are likely to avoid applying for jobs at discriminatory firms, thus reducing actual experiences of discrimination. Provided there are a sufficient number of employers who are willing to hire minorities, blacks and whites can sort into different labor markets without experiencing any direct discrimination.

How can we assess the validity of these claims? First, we need better information about how and where black and white job seekers search for work and how similar or different the distribution of employers encountered in real searches is to that achieved in the simulated search process of audit studies. This information can help us understand how the audit results map onto the direct experiences of actual job seekers.³⁰ Second, we need more information about how black job seekers make decisions about search behavior. Do blacks consciously avoid employers or industries in which discrimination is known to be prevalent? This information can help us understand the extent to which perceived discrimination may shape or constrain the search process of minority job seekers in ways that indirectly affect employment outcomes. And finally, what distortions in labor market sorting might discrimination produce, apart from its direct effects on hiring decisions? Economic models assume that, in the presence of discrimination, blacks and whites can sort into different labor markets with little consequence. This assumption is true only if those employers willing to hire blacks are no different (in compensation, security, number of vacancies, opportunities for promotion, etc.) from those employers who prefer to hire whites. By contrast, previous research suggests that occupational segregation (or crowding) within labor markets (whether due to discrimination or self-selection) is often associated with lower wages, less job security, longer search times, and/or reduced labor force participation (e.g., Parcel and Mueller 1983, chap. 5). Assessing the possible indirect effects of discrimination, in addition to any direct effects on hiring decisions, would provide a more complete understanding of the role of discrimination in contemporary labor markets.

Research on discrimination poses numerous complications, with issues of measurement of central concern. This discussion sought to provide an overview of some of the recent debates about the relevance of discrimination in contemporary labor markets, the measurement tools used to study discrimination, and the varying results in the empirical literature. While little consensus remains among researchers about the appropriate techniques for studying discrimination, active comparisons across studies can help to shed light on the relative strengths and weaknesses of existing approaches. While no research method is without flaws, careful consideration of the range of methods available helps to match one's research question with the appropriate empirical strategy. This article focused primarily on the strengths and limitations of the audit methodology for studying discrimination. Although the audit design cannot address all relevant aspects of labor market disadvantage, it can provide strong and direct measures of discrimination at the point of hire, a powerful mechanism regulating the broader array of labor market opportunities. Future research should extend this focus to include a broader perspective on the employment process, from search

decisions to hiring behavior to wages, tenure, and promotion, comparing findings across studies for a more complete picture of discrimination in labor markets today.

Appendix

The Implementation of an In-Person Audit Study

This appendix is intended to provide a general orientation to the nuts and bolts of designing and implementing an in-person audit study. This discussion addresses the selection and matching criteria necessary for ensuring high-quality and well-aligned applicant pairs, training and supervision requirements, outcome measures, and the ethics of audit research.

Matching. The selection of testers who will play the role of job applicants is one of the most critical components in the design of an employment audit and arguably one of the most time-intensive. Testers must be chosen based on personal attributes that make them individually well qualified to perform what can be a highly demanding job requiring a substantial degree of autonomy, but they must also be chosen based on personal attributes that make them a good match for another well-qualified individual (their test partner). Taking into account the wide range of characteristics employers may pay attention to in evaluating applicants, testers should be matched on concrete factors: such as age, height, weight, and level of education; in addition to more subjective criteria: articulateness, ease of personal interaction, physical attractiveness, and nonverbal communication style. Though the relevance of these characteristics may vary by job type or employer, they are all nevertheless potentially influential in hiring decisions and thus must be considered in deciding on potential matches. Taking all these considerations into account, it is not unusual to interview between eighty to one hundred applicants for each tester hired.

The matching process itself is an art as much as it is a science (an issue that has provoked criticism by some) (Heckman and Siegelman 1993). While a number of psychometric scales exist to measure personality attributes, verbal ability, and so on, certain intangible qualities are arguably more important in making a first impression. Including a wide range of external evaluators (individuals not directly involved in the research project) can provide important feedback about the holistic impressions formed by each potential tester and the degree of similarity between proposed pairs.

Note that audits of contexts other than employment require less attention to physical appearance and personality characteristics. Housing audits and audits of consumer markets, for example, are typically based on a far narrower (and easier to control) set of tester characteristics. Likewise, requirements are less stringent when treatment conditions can be randomly assigned. In testing the effects of a

criminal record, for example, testers can alternate which individual presents himself as the ex-offender over the course of the study, thus evening out any unobserved differences within the tester pair. If one tester is slightly more attractive, for example, in certain cases he or she will be a slightly more attractive offender and in other cases a slightly more attractive nonoffender. Any individual differences will even out if each tester serves in the treatment and control condition in an equal number of cases.^a In testing the effects of race, by contrast, the treatment condition cannot be randomly assigned. The quality of the matches thus becomes extremely consequential, as race can be fully confounded with any other individual characteristic. To the extent that differences will persist, researchers should err in the direction of choosing black testers with slightly more desirable attributes. Results will then represent a conservative test of discrimination.

Constructing resumes. Once tester pairs have been matched, they are assigned resumes reflecting equal levels of education and experience. Substantial thought must go into choosing high schools and neighborhoods that have similar reputations and student/resident compositions; likewise, work histories must be developed to reflect not only equal amounts of prior work experience but also similar types of work experience.^b In addition to pretesting resumes to assess their comparability, ideally resume types can be assigned independent of treatment condition (e.g., any given resume will be used by both black and white testers, to control for any unmeasured differences). In some cases, the resume will be the only point of contact between the tester and the employer (e.g., in cases where the person in charge of hiring is not present at the time of the test, and the tester leaves a resume); it is thus important that all relevant information can be effectively conveyed on this single-page document.

Training. No matter how carefully matched two testers may be, they can only act as successful as audit partners if they learn to interact with employers in similar ways. A wide range of questions can come up in the course of a conversation or interview with an employer, and testers must be prepared to share similar information and communicate similar types of profiles in their descriptions of past (fictitious) experiences. Before starting actual fieldwork for an audit study, testers typically participate in an extensive training period during which they rehearse the content of their profile, the appropriate way to phrase answers to interview questions, and work on aligning their responses with those of their test partner. Training can consist of videotaped mock interviews, practice interviews with employer confederates, and practice audits with real employers. In addition to the initial training period, daily debriefings with testers can help to identify problems that may arise or additional content that needs rehearsing.

Problems of implementation. With any field experiment, the unpredictabilities of the real world often interfere with carefully planned research designs.

Traffic can back up (or public transportation can break down), making it impossible for one tester to make it to an employer at the specified time; a job can get filled in between the time the two testers come to apply; a tester may run into someone he knows during an audit; an employer may know the manager of a fictitious job listed on the tester's resume. The key to maintaining the integrity of the experimental design lies in the ability to respond quickly to unexpected happenings and to constantly tweak existing protocols to take account of new situations. In cases where the protocol appears not to have been fully (or effectively) implemented, the test should be cancelled. While it is impossible to catalogue the countless number of potential disruptions that may arise, researchers must be vigilant throughout the course of the study. Effective and continual supervision of the testing process is one of the most important elements of a successful audit study.

Supervision. The quality of the data from an audit study depends on the degree to which testers effectively follow the established protocol. And yet evaluating testers' performance is difficult, since the majority of the testers' work is completed with no direct supervision. To monitor the quality of the testing process, a number of formal procedures can be put into place. First, immediately following each visit to an employer, testers are typically required to fill out an extensive summary form, including a large number of closed-ended questions (e.g., job title, race/gender/age of employer, screening tests required?⁹ asked about criminal background?⁹ etc.). In addition, testers write a lengthy open-ended narrative, describing their contact with the employer and the content of interactions they had during the test. These summary forms allow researchers to monitor the relative experiences of tester pairs and to identify any anomalies in the testing experiences that may confound measurement of the treatment variable. Second, the researcher (or project manager) should be available for daily debriefings with each of the testers, following the completion of each day's work. On occasions where something unexpected occurs, the project manager should be contacted immediately. Third, weekly meetings can be useful to allow testers the opportunity to brainstorm together about how to make the logistics of testing as efficient and controlled as possible. And finally, spot checks of tester performance can provide helpful tools for surveillance and continued training. For example, researchers can arrange for testers to unknowingly apply for jobs with confederate employers (i.e., employers who are collaborators of the researcher), to allow for an external assessment of their performance. Arranging for hidden cameras to record these spot checks can provide an additional training tool, as the audit team can watch and discuss the videotapes to identify differences in presentation style between tester pairs. The vast majority of problems that arise in the course of fieldwork for an audit study are relatively minor and can be resolved quickly, provided effective monitoring. It is only when problems continue unchecked that they can pose a significant threat to the validity of the research.

Quantitative and qualitative outcomes. One of the attractive features of the in-person audit design is its ability to measure a wide range of outcome variables, reflecting a range of applicant experiences. The primary outcome variable is typically a quantitative indicator of positive response by the employer: a callback or job offer. In addition, however, the audit process can detect a number of more subtle indicators of differential treatment. In some cases, for example, testers are channeled into jobs other than the ones originally advertised (e.g., the job ad was for a retail sales clerk, but the employer offers the tester a job as a stock boy). In other cases, employers may express revealing comments in the course of a conversation or interview. Tracking the level of attention, encouragement, or hostility testers elicit can provide important information about the experiential aspects of the job seeking process. Indeed, by observing the kinds of treatment testers receive in their ongoing job searches, one can identify the experiences that may lead certain workers to become discouraged from seeking work altogether.

Testing for litigation versus research. One of the common questions about the audit methodology concerns how it can be used to reduce the problems of discrimination. The audit method was initially designed for the enforcement of antidiscrimination law. Testers were used to detect racially discriminatory practices among real estate agents, landlords, and lenders, providing the evidence necessary to pursue litigation.^c Audit studies for research purposes, by contrast, are oriented not toward a specific intervention but rather toward obtaining accurate measures of the prevalence of discrimination across a broad sector or metropolitan area. The difference between these two types of studies is further reflected in the design of the study. Testing for litigation requires multiple audits of the same employer (or real estate agent, etc.) to detect consistent patterns of discrimination by that particular individual and/or company. Testing for research, by contrast, typically includes no more than a single audit per employer, with discrimination detected through systematic patterns across employers, rather than repeated acts of discrimination by a single employer. The distinction here is important in what we can tell from audit studies intended for research purposes. In these studies, *it is not possible to draw conclusions about the discriminatory tendencies of any given employer.* Indeed, even in the complete absence of discrimination, an employer confronted with two equivalent candidates will choose the white applicant in 50 percent of cases. Using a single test of each employer, therefore, does not allow for individual-level assessments of discrimination; only by looking at systematic patterns across a large number of employers can we determine whether hiring appears influenced by race or other stigmatizing characteristics.^d The point of research-based audit studies, therefore, is to assess the prevalence of discrimination across the labor market, rather than to intervene in particular sites of discrimination. While the objective is different, research audit studies provide important information about discriminatory practices that can support calls for strengthening antidiscrimination policy or other policy initiatives designed to protect vulnerable workers.

Ethics of audit research. Discussions of audit studies inevitably lead to questions about the ethics of research of this kind. Audit studies require that employers are unwittingly recruited for participation and then led to believe that the testers are viable job candidates. Contrary to the ethical standards for research established by the federal government, this design does not allow for the use of informed consent by research subjects for participation and often avoids debriefing subjects after the study's completion. How then are audit studies permitted to take place? As it turns out, there are specific criteria that regulate the use of research of this kind, and a well-designed audit study can arguably meet each of them. Below I provide a discussion of the relevant concerns and potential solutions to the ethical problems posed by research of this kind.

The use of deception in social science has long been met with suspicion. While individual researchers may feel they can clearly distinguish between appropriate and improper research practices, examples from the past indicate that researchers' individual judgments may not always conform to the standards of the discipline (e.g., Milgram 1974). Because of past transgressions, legislation concerning the use of human subjects now governs all social science research and includes, as one of its fundamental criteria, the use of informed consent from all research participants.^e In the case of audit studies, however, the nature of the research requires that subjects remain unaware of their participation, and the condition of informed consent therefore cannot be met.

While current federal policy governing the protection of human subjects strongly supports the use of informed consent, there is recognition that certain types of research that fail to obtain formal consent can be deemed permissible. According to the regulations, a human subjects institutional review board (IRB) "may . . . waive . . . informed consent provided (1) the research involves no more than minimal risk to human subjects; (2) the waiver or alteration will not adversely affect the rights and welfare of the subjects; (3) the research could not practicably be carried out without the waiver or alteration; and (4) whenever appropriate, the subjects will be provided with additional information after participation."^f Each of these conditions can arguably be satisfied in the context of audit studies of discrimination. While there are potential risks to subjects, reasonable efforts can be made to reduce the costs to subjects and thereby impose only minimal risk.

Most audit research poses two primary potential risks to subjects: (1) loss of time and (2) legal liability. In the first case, subjects are asked to evaluate a pair of applications submitted by phony applicants. Time spent reviewing applications and/or interviewing applicants will therefore impose a cost on the subject. Most employment audit studies limit their samples to employers for entry-level positions—those requiring the least intensive review—in part to minimize the time employers spend evaluating phony applicants. Entry-level positions are typically filled on the basis of cursory overviews of applications and limited personal contact (Fix and Struyk 1993). Contact with subjects is thus minimal, consisting of requesting an application and/or answering a few brief questions. Audits of higher-skill jobs, by contrast, impose a greater burden on employers,

as the hiring process for such positions typically requires a greater investment of time and effort.^g

A second potential risk posed by audit research is the potential for employers and/or firms to be held liable for discrimination if evidence were to be publicly released as to their performance in the audit. In fact, as mentioned above, the evidence provided by audit studies intended for research cannot support claims of discrimination against any individual employer. Nevertheless, efforts must be taken to protect employer identities so that even association with a study on discrimination cannot be made. To this end, identifying information should be kept in a secure location, and any publicly released publications or presentations should omit all identifying characteristics of individuals and firms.

The issue of debriefing subjects following the completion of the audit study is a complicated one. Though typically IRB protocol supports the debriefing of subjects whenever possible, in certain cases acknowledging the occurrence or nature of a research study is deemed undesirable. It could be argued, for example, that subjects could be placed at greater risk should their behavior, as a result of the audit study, fall under greater scrutiny by superiors. For human resource personnel or managers who are thought to be discriminating, the consequences may be more serious than if no attention were brought to the audit whatsoever. While the chances that negative consequences would result from this research in any case are very small, some IRB committees take the view that eliminating the debriefing stage is the most prudent strategy. The purpose of audit research is *not* to harm individual employers. Rather, the research seeks to improve our understanding of the barriers to employment facing stigmatized groups in their search for employment.

As a final matter, it should be emphasized that the ethics of audit research is not only of concern in a university context. The legal standing of testers has likewise received close scrutiny by the courts. In fact, the issue of testing has reached the highest judicial body, with the United States Supreme Court upholding the standing of testers in its 1982 decision.^h A more recent ruling by the 7th Circuit Court again upheld the standing of testers in cases of employment discrimination, broadening their endorsement of this methodology. In each of these rulings, the courts have been primarily concerned with the use of testing for pursuing litigation against employers (rather than for pure research, as is the case here). Implicit in these holdings, however, is the belief that the misrepresentation involved in testing is worth the unique benefit this practice can provide in uncovering discrimination and enforcing civil rights laws. According to former Equal Employment Opportunity Commission (EEOC) Chairman Gilbert Castellas, "Using employment testers in a carefully controlled manner is an important tool for measuring the presence or absence of discrimination. If we can shed light on barriers to fair hiring in entry-level jobs, which are the gateway to self-sufficiency and economic independence, we will have made an important step in assuring equal opportunity for everyone."ⁱ Indeed, despite certain burdens imposed by audit studies, the ultimate benefit from research of this kind extends far beyond

the contribution of a single study. Rigorous and realistic measurement of discrimination is fundamental to understanding and addressing persistent barriers to employment facing members of stigmatized groups.

-
- a. Note that even in cases where the experimental condition can be randomly assigned, it is nevertheless desirable to match testers as closely as possible, so as to minimize extraneous “noise” in the comparisons of tester outcomes.
 - b. Typically resumes are constructed to reflect a range of entry level work experience, including, for example, jobs in sales, restaurants, and manual labor.
 - c. In these discrimination cases, testers serve as the plaintiffs. Despite the fact that the testers themselves were not in fact seeking housing (or employment) at the time their application was submitted, their treatment nevertheless represents an actionable claim. This issue has received close scrutiny by the courts, including rulings by the highest federal courts. The U.S. Supreme Court upheld the standing of testers in its 1982 decision (*Havens Realty Corp. v. Coleman*, 455 U.S. 363, 373 [1982]). A more recent ruling by the 7th Circuit Court again upheld the standing of testers in cases of employment discrimination, broadening their endorsement of this methodology.
 - d. This feature has certain desirable properties from the perspective of gaining approval from an institutional review board (i.e., university ethics committees). Concerns about confidentiality and risks to employers are reduced when no single participant can be identified as a discriminator.
 - e. DHHS CFR45.46.116.
 - f. 56 *Federal Register* 117, p. 28017, June 18, 1991.
 - g. In the present research, I further limit imposition on employers by restricting audits to the first stage of the employment process. In most cases, then, I look only at whether or not an employer invites the tester for an interview, rather than including the interview and job offer stages as well. Limiting the research design to the initial process can thus further reduce the burden to subjects.
 - h. *Havens Realty Corp. v. Coleman*, 455 U.S. 363, 373 (1982).
 - i. This statement was drawn from a press release issued on December 5, 1997, and can be found at <http://www.eeoc.gov/press/12-5-97.html>.

Notes

1. *EEOC v. Target Corporation*, case no. 02-C-146. As of this writing, the case was pending appeal at the 7th Circuit Court, case no. 04-3559.

2. Connerly (2000, 20-21).

3. Dovidio and Gaertner (2000), for example, measured racial attitudes and discrimination at two points in time (late 1980s and late 1990s). They found substantial declines in self-reported racial prejudice. Evidence of discrimination, by contrast, remained stable. To test for discrimination, the researchers performed a simulated hiring experiment in which subjects were asked to evaluate the application materials for black and white job applicants of varying qualification levels. When applicants were either highly qualified or poorly qualified for the position, there was no evidence of discrimination. When applicants had acceptable but ambiguous qualifications, however, subjects were nearly 70 percent more likely to recommend the white applicant than the black applicant. This finding was consistent across the two time periods.

4. In 1942, only 32 percent of Americans believed that “white students and black students should go to the same schools”; by 1995, this proportion increased to 96 percent. In 1944, 45 percent believed that blacks “should have as good a chance as white people to get any kind of job”; by 1972, this proportion had increased to 97 percent (Schuman et al. 1997, 104-5).

5. As one simple empirical measure of this trend, I calculated the number of articles included in *Sociological Abstracts* that have the words “race” or “racial” and “discrimination” in their title relative to the proportion that merely reference “race” or “racial.” Nearly 20 percent fewer articles in the period 1986 to 2002 reference the word “discrimination” in their title relative to those articles about race written between 1963 and 1985; among those written in 2003 to 2004, nearly 40 percent fewer articles about race directly indicate an emphasis on discrimination in their titles. In recent years, therefore, an explicit emphasis on discrimination seems increasingly uncommon in sociological research. Political resources devoted to the problems of racial discrimination have likewise declined. For example, the proportion of Equal Employment Opportunity Commission (EEOC) cases addressing racial discrimination declined steadily in the 1990s, relative to increases in claims focusing on discrimination by gender or disability (Donohue and Siegelman 2005).

6. Note, however, that the black-white test score gap remains large and statistically significant.

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

8. The most common explanation for black disadvantage is “a lack of motivation or willpower,” with over half of white respondents endorsing this view. By contrast, fully two-thirds of black respondents believe that discrimination is an important explanation, with “poor quality education” representing the second most common choice. Author’s calculations from the 2000 General Social Survey.

9. Fiske (1998); Bodenhausen (1988); Trope and Thomson (1997); Banaji, Hardin, and Rothman (1993). Despite the progressive changes in racial attitudes generally, research indicates that the content of racial stereotypes has changed little over time (Devine and Elliot 1995); what has changed is the conscious effort on the part of nonprejudiced individuals to inhibit the activation of these stereotypes (Devine 1989). While these conscious strategies have successfully resulted in a substantial reduction in the expression of racial bias, actions taken under pressure or in cognitively demanding situations remain vulnerable to the influence of implicit racial attitudes (Gilbert and Hixon 1991).

10. Random assignment helps to remove the influence of any respondent characteristics that may affect their outcomes by breaking the link between respondent characteristics and selection into treatment conditions.

11. For a review of experimental field experiments in international contexts, see Riach and Rich (2002).

12. The present discussion focuses on the case of racial discrimination, but these methods can be readily applied to studies of discrimination on the basis of gender, age, neighborhood, and numerous other social categories.

13. In fact, very few correspondence studies have been conducted in the United States. This approach has been more widely used in European and Australian contexts. See Riach and Rich (2002) for a review.

14. White male names triggered a callback rate of 9.19 percent, compared to 6.16 percent among black male names.

15. Fryer and Levitt (2004, 786) reported that “Blacker names are associated with lower income zip codes, lower levels of parental education, not having private insurance, and having a mother who herself has a Blacker name.”

16. See Bendick, Jackson, and Reinoso (1994). It would be undesirable, however, to use only extracurricular activities to signal race. This subtle cue would likely be missed by many employers in the course of their cursory review.

17. To the extent that applicants presenting “race-neutral” extracurricular activities are not assumed to be white in 100 percent of cases, more conservative results will be obtained. For an example of this approach, see Dovidio and Gaertner (2000).

18. For an in-between approach using telephone contact (with voice and style of speech signaling race, class and gender), see Massey and Lundy (2001).

19. When comparing the results across studies, I find it useful to calculate relative differences (ratio tests) rather than percentage point differences. Because baseline response rates differ across studies, ratio measures allow for more straightforward comparisons.

20. In the Bertrand and Mullainathan (2004) study, even among high-skilled white applicants the callback rate was less than 11 percent. Lower callback rates can depress evidence of differential treatment. If, for example, 5 percent of employers call back *all* applicants as a matter of policy, the resulting contrast would be based on a very small number of employers who conduct any type of screening at the resume submission stage.

21. Heckman (1998, 107-11). Elsewhere, Heckman and Siegelman (1993) identified five potential threats to the validity of results from audit studies: (1) problems in effective matching, (2) the use of "overqualified" testers, (3) limited sampling frame for the selection of firms and jobs to be audited, (4) experimenter effects, and (5) the ethics of audit research. Each of these issues is addressed in detail in Pager (forthcoming, Appendix 4A). See also the series of essays published in Fix and Struyk (1993). In addition to the criticisms expressed by Heckman, audit studies are often costly and difficult to implement and can only be used for selective decision-points (e.g., hiring decisions but not promotions).

22. Given these extensive demands, it is common for researchers to screen between fifty and one hundred applicants (already selected on age, race, and gender) before finding a single matched pair.

23. Indeed, we see evidence of more discrimination in audit studies testing actual job offers. This could be due to the kinds of tiebreaker effects discussed by Heckman (1998), though it may also result from the fact that job offers are more consequential, and thus employers may exert their preferences more forcefully at this final stage.

24. See Pager (2003, Appendix A) for an example of such a test.

25. Aside from active assumptions about general productivity by race, mere uncertainty can likewise lead to bias. Specifically, the problem of erroneous statistical discrimination is aggravated by racial disproportionality among employers. Generally, individuals have access to more and more reliable information about members of their own group (whether due to familiarity with their neighborhoods, schools, social networks, or simply due to a greater ability to recognize individuating information) (Strauss 1991; Anderson 1990). If the information white employers have about black applicants is seen as less reliable (simply as a result of lesser familiarity), risk-averse employers will be less inclined to consider these workers (Aigner and Cain 1977). Pervasive occupational and residential segregation by race may contribute to the preservation of inaccurate assumptions and/or the simple enhancement of uncertainty (Arrow 1998; Tomaskovic-Devey and Skaggs 1999).

26. See Altonji and Pierret (1997). While there is evidence that on-the-job evaluations may continue to be affected by racially biased perceptions, these effects can be mediated to some degree by objective performance indicators (Castilla 2005).

27. Current Population Survey data show that in 2000, joblessness among black men, aged twenty to thirty-five, with a high school diploma, was around 22 percent compared to less than 10 percent for white men of the same age and education.

28. For example, Neal and Johnson (1996) included a correction for labor force participation by assigning all nonparticipants a wage of zero. This correction produces a race coefficient nearly double in size to the original (.134 versus .072) and reduces the amount of the race gap explained by cognitive ability from roughly 70 to 60 percent (pp. 881-85). See also Butler and Heckman (1977); Mare and Winship (1984); Western and Pettit (2005); Chandra (2000); Fairlie and Sundstrom (1997).

29. This discussion focuses on discrimination against African Americans without college education, as the majority of audit studies focus on the experiences of job candidates with no more than a high school degree. Patterns of discrimination would likely differ at higher levels of the occupational hierarchy, with college-educated blacks less likely to experience barriers to access and more likely to experience channeling or barriers to mobility within the organizational setting (see Feagin and Sikes 1994; Collins 1989, 1993; Grodsky and Pager 2001).

30. Previous research, for example, indicates that roughly 20 to 25 percent of search time is spent on contacts generated by newspaper advertising, with friends and relatives and direct contact of firms by applicants representing much more common sources of new employment (Holzer 1988). At the same time, minorities appear more successful in job searches generated by general newspaper ads than through other means (Holzer 1987). Some have argued that the samples generated for audit studies (primarily from classified ads) yield a conservative test of discrimination: firms who wish to discriminate, it is argued, are more likely to advertise job openings through more restrictive channels than the metropolitan newspaper, such as through referrals, employment agencies, or more selective publications (Fix and Struyk 1993, 32).

These claims could be usefully tested by drawing more direct comparisons between samples of firms contacted by real black and white job seekers and those included in audit studies.

References

- Aigner, Dennis J., and G. Cain Glen. 1977. Statistical theories of discrimination in labor market. *Industrial and Labor Relations Review* 30:749-76.
- Altonji, Joseph G., and Charles R. Pierret. 1997. Employer learning and statistical discrimination. NBER Working Paper 6279, National Bureau of Economic Research, Cambridge, MA.
- Anderson, Elijah. 1990. *Streetwise: Race, class, and change in an urban community*. Chicago: University of Chicago Press.
- Arrow, Kenneth J. 1998. What has economics to say about racial discrimination? *Journal of Economic Perspectives* 12 (2): 91-100.
- Ayres, Ian, and Peter Siegelman. 1995. Race and gender discrimination in bargaining for a new car. *American Economic Review* 85 (3): 304-21.
- Banaji, M. R., C. Hardin, and A. J. Rothman. 1993. Implicit stereotyping in person judgment. *Journal of Personality & Social Psychology* 65 (2): 272-81.
- Bendick, Marc Jr., Lauren Brown, and Kennington Wall. 1999. No foot in the door: An experimental study of employment discrimination. *Journal of Aging and Social Policy* 10 (4): 5-23.
- Bendick, Marc, Jr., Charles Jackson, and Victor Reinoso. 1994. Measuring employment discrimination through controlled experiments. *Review of Black Political Economy* 23:25-48.
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review* 94 (4): 991-1013.
- Bodenhausen, Galen. 1988. Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology* 55 (5): 726-37.
- Butler, Richard, and James Heckman. 1977. The government's impact on the labor market status of black Americans: A critical review. In *Equal rights and industrial relations*, ed. Farrell Bloch, 235-81. Madison, WI: Industrial Relations Research Association.
- Cancio, A. Silvia, T. David Evans, and David J. Maume. 1996. Reconsidering the declining significance of race: Racial differences in early career wages. *American Sociological Review* 61:541-56.
- Carneiro, Pedro, James J. Heckman, and Dimitriy V. Masterov. 2003. Labor market discrimination and racial differences in premarket factors. NBER Working Paper #10068, National Bureau of Economic Research, Cambridge, MA.
- Castilla, Emilio. 2005. Gender, race, and meritocracy in organizational careers. In *Culture and inequality workshop*. Princeton, NJ: Princeton University Press.
- Chandra, Amitabh. 2000. Labor-market dropouts and the racial wage gap: 1940-1990. *American Economic Review* 90 (2): 333-38.
- Collins, Sharon. 1989. The marginalization of black executives. *Social Problems* 36:317-31.
- . 1993. Blacks on the bubble: The vulnerability of black executives in white corporations. *Sociological Quarterly* 34:429-48.
- Connerly, Ward. 2000. *Creating equal: My fight against racial preferences*. San Francisco: Encounter Books.
- Cross, Harry, Genevieve Kenney, Jane Mell, and Wendy Zimmerman. 1990. *Employer hiring practices: Differential treatment of Hispanic and Anglo job seekers*. Washington, DC: Urban Institute Press.
- Devine, Patricia. 1989. Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology* 56:5-18.
- Devine, P. G., and A. J. Elliot. 1995. Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin* 21 (11): 1139-50.
- Donohue, John J., III, and Peter Siegelman. 2005. The evolution of employment discrimination law in the 1990s: A preliminary empirical investigation. In *Handbook of employment discrimination research*, ed. L. B. Nielsen and R. L. Nelson, 261-85. Dordrecht, the Netherlands: Springer.
- Dovidio, John F., and Samuel L. Gaertner. 2000. Aversive racism and selection decisions. *Psychological Science* 11 (4): 315-19.

- Fairlie, Robert W., and William A. Sundstrom. 1997. The racial unemployment gap in long-run perspective. *American Economic Review* 87 (2): 306-10.
- Farkas, George. 2003. Cognitive skills and noncognitive traits and behaviors in stratification processes. *Annual Review of Sociology* 29:541-62.
- Farkas, George, and K. Vicknair. 1996. Appropriate tests of racial wage discrimination require controls for cognitive skill: Comment on Cancio, Evans, and Maume. *American Sociological Review* 61:557-60.
- Farley, Reynolds. 1997. Racial trends and differences in the United States 30 years after the civil rights decade. *Social Science Research* 26:235-62.
- Feagin, Joe R., and Melvin P. Sikes. 1994. *Living with racism: The black middle-class experience*. Boston: Beacon.
- Fiske, Susan. 1998. Stereotyping, prejudice, and discrimination. In *The handbook of social psychology*, 4th ed., ed. Daniel Gilbert, Susan Fiske, and Gardner Lindzey, 357-411. Boston: McGraw-Hill.
- Fix, Michael, and Raymond J. Struyk, eds. 1993. *Clear and convincing evidence: Measurement of discrimination in America*. Washington, DC: Urban Institute Press.
- Freeman, Richard B., and Harry J. Holzer, eds. 1986. *The black youth employment crisis*. Chicago: University of Chicago Press, for National Bureau of Economic Research.
- Fryer, Roland G., Jr., and Steven D. Levitt. 2004. The causes and consequences of distinctively black names. *Quarterly Journal of Economics* 119 (3): 767-805.
- Gilbert, Daniel T., and J. Gregory Hixon. 1991. The trouble with thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology* 60 (4): 509-17.
- Grodsky, Eric, and Devah Pager. 2001. The structure of disadvantage: Individual and occupational determinants of the black-white wage gap. *American Sociological Review* 66 (August): 542-67.
- Hakken, Jon. 1979. *Discrimination against Chicanos in the Dallas rental housing market: An experimental extension of the housing market practices survey*. Washington, DC: U.S. Department of Housing and Urban Development.
- Harrison, Roderick J., and Claudette E. Bennett. 1995. Racial and ethnic diversity. In *State of the union: America in the 1990s*, vol. 2, *Social trends*, ed. R. Farley, 141-210. New York: Russell Sage Foundation.
- Heckman, James J. 1998. Detecting discrimination. *Journal of Economic Perspectives* 12 (2): 101-16.
- Heckman, James, and Peter Siegelman. 1993. The Urban Institute audit studies: Their methods and findings. In *Clear and convincing evidence: Measurement of discrimination in America*, ed. Michael Fix and Raymond J. Struyk, 187-258. Washington, DC: Urban Institute Press.
- Holzer, Harry. 1987. Informal job search and black youth unemployment. *American Economic Review* 77 (3): 446-52.
- . 1988. Search methods used by unemployed youth. *Journal of Labor Economics* 6 (1): 1-20.
- Holzer, Harry, Paul Offner, and Elain Sorensen. 2005. What explains the continuing decline in labor force activity among young black men? *Labor History* 46 (1): 37-55.
- Jencks, Christopher, and Meredith Phillips, eds. 1998. *The black-white test score gap*. Washington, DC: Brookings Institution.
- Lahey, Joanna. 2005. Age, women, and hiring: An experimental study. NBER Working Paper #11435, National Bureau of Economic Research, Cambridge, MA. <http://www.nber.org/papers/w11435>.
- Mare, Robert D. 1995. Changes in educational attainment and school enrollment. In *State of the union: America in the 1990s*, vol. 1, *Economic Trends*, ed. R. Farley, 155-213. New York: Russell Sage Foundation.
- Mare, Rob, and Christopher Winship. 1984. The paradox of lessening racial inequality and joblessness among black youth: Enrollment, enlistment and employment, 1964-1981. *American Sociological Review* 49:39-55.
- Massey, Douglas, and Garvey Lundy. 2001. Use of black English and racial discrimination in urban housing markets: New methods and findings. *Urban Affairs Review* 36:452-69.
- Milgram, Stanley. 1974. *Obedience to authority: An experimental view*. New York: Harper & Row.
- Murnane, Richard, John Willett, and Frank Levy. 1995. The growing importance of cognitive skills in wage determination. *Review of Economics and Statistics* 77:251-66.
- Neal, Derek, and William Johnson. 1996. The role of premarket factors in black-white wage differences. *Journal of Political Economy* 104 (5): 869-95.

- Neumark, David. 1996. Sex discrimination in restaurant hiring: An audit study. *Quarterly Journal of Economics* 111 (3): 915-41.
- . 1999. Wage differentials by race and sex: The roles of taste discrimination and labor market information. *Industrial Relations* 38 (3): 414-45.
- O'Neill, June. 1990. The role of human capital in earnings differences between white and black men. *Journal of Economic Perspectives* 4:25-45.
- Pager, Devah. 2003. The mark of a criminal record. *American Journal of Sociology* 108:937-75.
- . Forthcoming. *Marked: Race, crime, and finding work in an era of mass incarceration*. Chicago: University of Chicago Press.
- Pager, Devah, and Bruce Western. 2005. Discrimination in low wage labor markets. Paper presented at the annual meeting of the American Sociological Association, August, Philadelphia.
- Parcel, Toby L., and Charles W. Mueller. 1983. *Ascription and labor markets: Race and sex differences in earnings*. New York: Academic Press.
- Riach, P. A., and J. Rich. 2002. Field experiments of discrimination in the market place. *Economic Journal* 112 (November): F480-F518.
- Ridley, Stanley, James A. Bayton, and Janice Hamilton Outtz. 1989. Taxi service in the District of Columbia: Is it influenced by patrons' race and destination? Mimeograph, Washington Lawyers' Committee for Civil Rights under the Law, Washington, DC.
- Schuman, Howard, Charlottee Steeh, Lawrence Bobo, and Maria Krysan. 1997. *Racial attitudes in America: Trends and interpretations*. Cambridge, MA: Harvard University Press.
- Strauss, David A. 1991. The law and economics of racial discrimination in employment: The case for numerical standards. *Georgetown Law Review* 79:1619.
- Tomaskovic-Devey, Donald, and Sheryl Skaggs. 1999. An establishment-level test of the statistical discrimination hypothesis. *Work and Occupations* 26 (4): 420-43.
- Tomaskovic-Devey, Donald, Melvin Thomas, and Kecia Johnson. 2005. Race and the accumulation of human capital across the career: A theoretical model and fixed-effects application. *American Journal of Sociology* 111 (1): 58-89.
- Trope, Yaacov, and Erik P. Thomson. 1997. Looking for truth in all the wrong places? Asymmetric search of individuating information about stereotyped group members. *Journal of Personality and Social Psychology* 73 (2): 229-41.
- Turner, Margery, Michael Fix, and Raymond Struyk. 1991. *Opportunities denied, opportunities diminished: Racial discrimination in hiring*. Washington, DC: Urban Institute Press.
- Turner, Margery Austin, Stephen L. Ross, George C. Galster, and John Yinger. 2002. Discrimination in metropolitan housing markets: National results from phase 1 HDS 2000. U.S. Department of Housing and Urban Development, final report. http://www.huduser.org/Publications/pdf/Phase1_Report.pdf.
- Turner, Margery Austin, and Felicity Skidmore, eds. 1999. *Mortgage lending discrimination: A review of existing evidence*. Washington, DC: Urban Institute.
- Waldinger, Roger. 1999. *Still the promised city? African-Americans and new immigrants in postindustrial New York*. Cambridge, MA: Harvard University Press.
- Western, Bruce, and Becky Pettit. 2005. Black-white wage inequality, employment rates, and incarceration. *American Journal of Sociology* 111:553-78.
- Wienk, Ronald E., Clifford E. Reid, John C. Simonson, and Frederick J. Eggers. 1979. *Measuring discrimination in American housing markets: The housing market practices survey*. Washington, DC: U.S. Department of Housing and Urban Development.
- Wilson, William Julius. 1978. *The declining significance of race: Blacks and changing American institutions*. Chicago: University of Chicago Press.
- . 1987. *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press.
- Yinger, John. 1995. *Closed doors, opportunities lost*. New York: Russell Sage Foundation.