

POLITICAL ECONOMY RESEARCH INSTITUTE

PERI

University of Massachusetts Amherst

Screening for Honesty and Motivation in the Workplace: What Can Affirmative Action Do?

Elaine McCrate

February 2013

*This paper was presented as part of
a September 2011 Festschrift Conference in
honor of Thomas Weisskopf.*

WORKINGPAPER SERIES

Number 310

**POLITICAL ECONOMY
RESEARCH INSTITUTE**

Gordon Hall
418 North Pleasant Street
Amherst, MA 01002

Phone: 413.545.6355
Fax: 413.577.0261
peri@econs.umass.edu
www.peri.umass.edu



PREFACE

This working paper is one of a collection of papers, most of which were prepared for and presented at a festschrift conference to honor the life's work of Professor Thomas Weisskopf of the University of Michigan, Ann Arbor. The conference took place on September 30 - October 1, 2011 at the Political Economy Research Institute, University of Massachusetts, Amherst. The full collection of papers will be published by Elgar Edward Publishing in February 2013 as a festschrift volume titled, *Capitalism on Trial: Explorations in the Tradition of Thomas E. Weisskopf*. The volume's editors are Jeannette Wicks-Lim and Robert Pollin of PERI.

Since the early 1970s, Tom Weisskopf has been challenging the foundations of mainstream economics and, still more fundamentally, the nature and logic of capitalism. That is, Weisskopf began putting capitalism on trial over 40 years ago. He rapidly established himself as a major contributor within the newly emerging field of radical economics and has remained a giant in the field ever since. The hallmarks of his work are his powerful commitments to both egalitarianism as a moral imperative and rigorous research standards as a means.

We chose the themes and contributors for this working paper series, and the upcoming festschrift, to reflect the main areas of work on which Tom Weisskopf has focused, with the aim of extending research in these areas in productive new directions. The series is divided into eight sections, including closing reflections by our honoree himself, Professor Weisskopf. Each section except for the last includes comments by discussants as well as the papers themselves.

The eight sections are as follows:

1. Reflections on Thomas Weisskopf's Contributions to Political Economy
2. Issues in Developing Economies
3. Power Dynamics in Capitalism
4. Trends in U.S. Labor Markets
5. Discrimination and the Role of Affirmative Action Policies
6. Macroeconomic Issues in the United States
7. Applications of Marxist Economic Theory
8. Reflections by Thomas Weisskopf

This working paper is 2 of 6 included in Section 5.

- *Jeannette Wicks-Lim and Robert Pollin*

Screening for Honesty and Motivation in the Workplace: What Can Affirmative Action Do?

Elaine McCrate

The University of Michigan was the belly of the beast during the U.S. Supreme Court's 2003 deliberations on affirmative action in higher education admissions. The screening of applications to Michigan's undergraduate and law programs was the center of the legal controversy, clearly because screening constitutes one of the critical gateways to opportunity in education and employment.

Tom Weisskopf was in the Michigan belly in 2003, as a scholar-participant in the struggle to keep affirmative action alive. Among the many contributions of his work on affirmative action, Tom has advocated more careful and multidimensional screening of admissions and employment candidates in order to establish marginalized groups in the economic mainstream, as well as to promote better admissions and business practices overall.

Affirmative action has certainly resulted in more careful scrutiny of candidates; the empirical literature has borne this out repeatedly. Better screening may be one of the reasons that Deshpande and Weisskopf (2010), in their painstaking empirical study of the Indian railways, uncovered no evidence that affirmative action impairs productivity. They concluded that the particularly strong Indian form of affirmative action, reserving jobs for stigmatized groups, may in fact improve economic performance, augmenting the earlier results of others concerning the neutral-to-positive effects of the comparatively weak American style of affirmative action (Conrad, 1995; Holzer and Neumark, 2000).

Yet the perception remains that affirmative action does reduce productivity. There are strong productivity-related stereotypes of both African Americans and Dalits ("untouchables"). These stereotypes often concern cognitive skills. But in this paper I try to understand how employers screen for some quite different and particularly visceral and stigmatizing stereotypes of blacks – those of dishonesty and poor motivation, attitude, and work ethic; I also inquire how that affects the employment of African Americans.¹ While I will have much less to say about the stereotyping and screening of Dalits, I note in passing that a growing tendency to stereotype about motivation may be replacing the traditional pollution taboos (Jodhka and Newman, 2007).

Another few words about Tom before I continue: while we are here mainly to honor his distinctive blend of exacting research, unwavering political commitment, and just plain human decency, we should also pay tribute to his teaching which always incorporated the same values. Tom taught for many years in the Residential College of Literature, Science and the Arts at the University of Michigan, in the Social Theory and Practice Concentration; he was also the Residential College Director for ten years. The Social Theory and Practice Concentration describes its mission as "support[ing] students in developing the analytical and practical skills necessary for active engagement in the world and for building careers that promote equality and responsible citizenship" (University of Michigan, 2011). Tom and the Residential College chose each other because of

their mutual willingness to examine questions of power and inequality that don't get explored in the main-stream economics curriculum.

This paper is a tribute to Tom's teaching as well as his research – it's a compilation of some notes I've been accumulating while I've been teaching about gender and racial inequality. I like it as a teaching piece first because it leaves some unsettled theoretical questions about statistical discrimination, the theory that I invoke here, and some unsettled practical questions about affirmative action. Moreover, the material helps students understand the ways in which notions about honesty and work ethic, as well as other aspects of merit, are in great part socially determined, selectively bestowing deservingness and social kinship upon different groups.² The persistence of these particular stereotypes makes students think hard, which is what Tom always pushed them to do.

I examine stereotypes about honesty and work ethic through the lens of statistical discrimination theory (SDT), for two reasons. First, information about honesty and attitude is exceptionally poor, and as such SDT, with its explicit attention to the screening process, should be informative about the consequences and persistence of these stereotypes. Second, however, there seem to be some empirical anomalies here. The first puzzle is that some common screening devices, such as commercially available integrity tests, give no reason to believe that black workers will steal more from their employers or be less conscientious. The pass rates and mean scores of black and white test takers are most often statistically indistinguishable (Sackett and Harris, 1985; Sackett and Wanek, 1996; Ones et. al., 1993; Berry, Sackett and Wiemann, 2007), yet many employers believe blacks are less honest and motivated than whites. Furthermore employers act on these stereotypes. In an experimental study, testers who mentioned a (feigned) criminal background on their resumes got significantly fewer callbacks than those who did not, and *the effect was stronger for blacks* (Pager, 2007).

Yet (here is the second puzzle) when employers use criminal background checks, they are more likely to hire black applicants, especially black men, even controlling for the racial composition of the applicant pool (Holzer, Raphael and Stoll, 2006). The criminal background check is based on a cumulatively biased process of patrolling, arresting, and punishing blacks more than whites for comparable behavior, especially drug offenses. Although black youth are no more likely than white to use drugs (Bureau of Justice Statistics, 2009), the incentives for the police, as well as the relative lack of privacy for poor blacks, result in more drug arrests and more criminal records among blacks (Tonry, 1995). Some day someone should compare actual hiring rates for blacks who have been screened with integrity tests (seemingly racially neutral) rather than criminal background checks (clearly biased), but I have not found a dataset that would make that possible. Nonetheless, it is still curious that criminal background checks seem to promote black employment. Why?

The paper has five subsequent sections. Because honesty and motivation have remained in the background in economic theory until fairly recently, the first section summarizes their importance for employers, and employers' stereotypes about them. Because SDT is centrally concerned with the screening process (an aspect of the theory which does not appear in the standard textbook treatments), I summarize and evaluate the most important variants of the theory in the second section. In the third section, I survey the most common mechanisms employers actually use to screen for honesty and motivation, including integrity tests and criminal background checks, and their consequences for black employment. In the fourth section, I inquire why the

market does not punish statistical discriminators. Finally, I return to the question of affirmative action and screening for honesty and motivation.

EMPLOYERS' PERCEPTIONS OF MOTIVATION AND HONESTY

Managers are unequivocal about the significance of work ethic and attitude (Bowles, Gintis, and Osborne, 2001; Huang and Cappelli, 2006). Similarly, businesspeople worry a lot about employee theft, although the threat is often exaggerated, with greatest concern expressed by the retail sales, food service, warehousing, banking, and medical services industries (Dickens, *et. al.*, 1989; Sackett and Harris, 1984; Murphy, 1993). The National Retail Security Survey placed the value of employee theft in the U.S. retail industry alone at \$15.1 billion in 2001 (about 0.8% of sales, and nearly three times the cost of larceny reported to the FBI) (Hollinger and Davis, 2002; Federal Bureau of Investigation, 2002). Anonymous surveys have identified a large number of employees who admit to occasional theft, although it tends to be infrequent and to involve relatively small amounts.³ Honesty also matters to employers because its measure is strongly correlated on personnel tests with conscientiousness (which may be part of what employers have in mind when they are discussing motivation, work ethic, or attitude) (Ones, Viswesvaran, and Schmidt, 1993; Sackett and Wanek, 1996; Hogan and Brinkmeyer, 1997; Berry, Sackett and Wiemann, 2007)⁴

I note two things about employers' views about the relationship between these characteristics and race. First, their views are quite heterogeneous, but, second, their perception of blacks is decidedly negative relative to other ethnic groups. Interviews in the early 1990s of Chicago employers seeking to fill unskilled, entry-level positions, found that the “employers view[ed] inner-city workers, especially black men, as unstable, uncooperative, dishonest, and uneducated” (Kirschenman and Neckerman, 1991: p. 204). They characterized black workers as having a “bad work ethic,” being “lazy and unreliable,” and having “a bad attitude” (*ibid.*, p. 213). “When asked directly whether they thought there were any differences in the work ethic of whites, blacks and Hispanics, 37.7 percent of the employers ranked blacks last, 1.4% ranked Hispanics last, and no one ranked whites there. Another 7.6 percent placed blacks and Hispanics together on the lowest level; 51.4% either saw no difference or refused to categorize in a straightforward way” (*ibid.*, p. 210).

Similarly, in their interviews of employers of entry-level workers in four major U.S. cities, Moss and Tilly (2001: p. 97) found that by far the greatest complaint about black workers was that “blacks have lagging motivation” (33.4% of the employers agreed with that). Furthermore, employers viewed black motivation much more negatively than they did that of other ethnic groups. Doubtless many employers portrayed themselves as less disparaging of blacks than they actually were, but the data illustrate the main points: a greater tendency to negatively stereotype blacks, and most likely some heterogeneity in employer views.⁵

THE ROLE OF TESTS IN SDT

All SDT starts with the problem of imperfect information. While most discussions of SDT focus on cognitive ability, it is even more likely that employers cannot directly observe “habits of action and thought that favor good performance in skilled jobs, steadiness, punctuality, responsiveness, and initiative” (Arrow, 1974b, p. 97) – and honesty. Even more so in unskilled jobs, these habits are paramount, precisely because there are so few

other skill requirements. (Is the night janitor, with keys to all the offices, going to leave valuable property and records alone?) Work ethic and honesty are critical when monitoring is difficult.

An extended digression on SDT is in order, since screening is a central part of the theory, which is not discussed in the standard textbook treatment. There are two main traditions of SDT, one that focuses on testing error (Phelps, 1972; Lundberg and Startz, 1983), and one that emphasizes prior stereotypes (Arrow, 1973; Coate and Loury, 1993). In each version, costly tests and screening processes are central to the story. Therefore we need to understand the theoretical role of tests – formal and informal – and to explore their actual use and effectiveness with respect to honesty and motivation.

Two of Phelps' three accounts of SDT have received much attention. In his first account, there is no racial difference in expected quality, but the variance of white testing error is lower than the variance of black testing error. In other words, there is a massive communication failure between white employers and black jobseekers. There are several reasons, and some experimental evidence, suggesting that error variance could be higher for blacks on some kinds of screening devices. Personal interviews are particularly susceptible to cultural miscommunication. Blacks are less likely than whites to have personal contacts within firms that could provide information about applicant quality with relatively low error. Finally, social psychologists have found that one of the more automatic components of human cognition is the tendency to see same-group members as more heterogeneous and out-group members as more homogeneous (Fiske, 1998).

Since high-quality blacks cannot communicate their integrity or work ethic to potential employers as effectively as whites, employers weight the individual component of productivity less, and the racial component more, than they would with better signals. Another way to see this is with an extreme example. Suppose that blacks and whites apply for the same job taking the same test (say, an interview). White employers assume that black and white workers have the same distribution of motivation, but “blacks look alike” to them in the interview. Since white employers can't read black signals at all, black interview scores are all equal to the average (the test is useless in screening blacks), and black productivity is measured with greater error than white. Below-average blacks are rated too high and above-average blacks are rated too low. Since the white employer cannot see the variance in actual black motivation, but does see the variance in white motivation, there will typically be some white workers rated above average, and essentially not forced to compete with above-average black workers who are rated merely as average.

In the second Phelps variant of SDT, employers believe that average black qualifications are lower than white qualifications for a given test score, because of black disadvantages in upbringing. In this case, blacks get fewer offers because, given black social disadvantage, employers believe their test scores are likely to overestimate actual black productivity. Phelps suggested that “skin color or sex is taken as a proxy for relevant data not sampled” (p. 659). In other words, employers believe test scores are biased upward because of omitted variables which are correlated with race.

This version is close to the standard textbook treatment of SDT. As such, it warrants some special scrutiny. No one has made a careful empirical case that low social background (or race) is associated with greater theft or shirking at work. Nor is there ever likely to be a convincing one. Explanations based on social background would have to be distinguished from reasons that are specific to the workplace. Any observed group differ-

ences in the probability of stealing or shirking at work will depend on aspects of the employment relationship that are difficult to measure: the fallback position of workers in the event of dismissal, the degree of monitoring and the opportunity for theft, the size and structure of personnel incentives, and workers' perceptions of the firm's commitment to its employees. Each of these is likely to be correlated with race and socioeconomic background.

The correlation between social background and honesty or work ethic is plausibly either negative or positive: social disadvantage can be both a cause of and a deterrent to theft or low effort at work. First consider the case for the claim that social or racial disadvantage is a deterrent to workplace theft. (1) Black alternative employment prospects are worse than those of whites, so the value of the job is likely to be greater for blacks. (2) Blacks are underrepresented in jobs that present lucrative opportunities for fraud (including what fraud examiners ever so delicately refer to as the "misappropriation of assets" for personal use) – bookkeepers, internal auditors, fiscal officers, and executive officers.⁶ (3) There is also a racial difference in the probability of detection: blacks work less independently and are monitored more closely than whites, mostly because they have jobs in low-trust occupations. Misconduct is riskier for blacks, who will be caught stealing or shirking more often than whites of comparable rectitude and diligence. Black workers also have less control over their schedules, so they are more likely to be noticed and penalized at work for lapses in punctuality, relative to whites with similar records. Greater monitoring of blacks is an incentive to comply with employers' expectations of honesty and effort (McCrate, 2006). Thus there are several reasons to expect a negative correlation between social or racial disadvantage, and workplace malfeasance. Researchers have found such a negative correlation in laboratory studies of ethical behavior (Piff, *et al*, 2012).

However, disadvantage might also be positively correlated with shirking or theft. Because blacks are monitored more closely than whites, they will be caught and punished more often for the occasional petty theft or slacking that a fairly large number of employees of all descriptions engage in. While punishment sometimes has a disincentive effect, it can backfire when it is frequent or severe: morale falls, and shirking and dishonesty have increased in some reciprocity experiments (Murphy, 1993; Fehr and Gächter, 2000; Bewley, 2000; Nagin, *et. al.*, 2002).

So, returning to Phelps' argument about social disadvantage and testing error, the *prima facie* case for discounting black test scores is not good; there is ample (perhaps more) reason to think that black social disadvantage leads to underestimates of honesty and work ethic, as well as overestimates. It is clear, then, that *in the interpretation of black and white test scores, observers selectively view the omitted data that is correlated with race.*

Similarly, if intercultural miscommunication makes black signals noisier (Phelps' error variance story), this doesn't mean that anyone's imperfect signals are interpreted through a passive, neutral receiver. The problem is not just high error variance for blacks, but the way that agents handle noise when race is involved. Thus even though observers may not be willing to categorize interviewees as dishonest or lazy solely on the basis of race, they seem to form hypotheses about group-specific dispositions that they then regard as confirmed by ambiguous signals. A listless white job candidate, for example, may be seen as tired or having a bad day, while a similar black candidate may be viewed as unmotivated (Darley and Gross, 1983).

On the subject of *ex ante* stereotypes, Phelps' models are silent. This lacuna is apparent in the error variance model's symmetric treatment of black scores far from the expected value, be they high or low. Because of the high black error variance, the model posits that employers simply weight the individual component of productivity less, and the group component more: high black signals are downgraded, and low blacks signals are *upgraded* toward the mean. However, social psychologists have also found that perceptions of outgroup homogeneity do not generally revert to the general population mean; rather they are associated with stereotyping. People who believe there is little outgroup variance also make stereotypic judgments about outgroup members more readily than people who perceive more outgroup variance (Fiske, 1998). And these stereotypes strongly influence the interpretation of ambiguous signals.

Arrow's distinctive contribution was the recognition both of the role of costly screening processes and *a priori* stereotypes. Unlike Phelps, Arrow recognized that agents may handle noise differently when race is involved; in his model, stereotypes are independent elements of contemporary racial inequality. The employer does not know whether a worker has invested in necessary work habits or skills; the employer believes that blacks are less likely to have done so; and the employer's investment in a worker (minimally, a screening cost) is wasted if the worker has not made the necessary investment in herself. Under these circumstances, there will be a racial wage or employment differential, and a lower rate of return to investment for blacks. Coate and Loury's model (1993), which also starts with stereotypes as a given component of the economic landscape, concludes with potentially multiple stable equilibria; among these, some subset is discriminatory, with outcomes depending on prior beliefs or stereotypes.⁷

SDT, in all its forms, certainly has its critics. Skeptics maintain that firms still have the incentives, and in the long run, the means, to identify individual high performers regardless of race. Uncertainty should stimulate a market for information, at least some firms will identify productive differences *within* groups, the informational value added by race will approach zero, and competitive markets will punish firms that act on false stereotypes (Darity, 1989 and 1998; Aigner and Cain, 1977; Cain, 1986). There is some empirical support for this. For example, Stoll *et al.* (2004) and Raphael *et al.* (2000) found that black hiring officers hire somewhat more black workers (after controlling for the black application rate), despite, or because of, the fact that they are more likely than white hiring officers to work in firms with stricter hiring requirements and screening methods. Black hiring officers may do a better job reading own-group signals, and employers may reduce cultural miscommunication or the tendency to act on stereotypes by using black hiring agents. (There are of course other obvious explanations for this, such as networks). Since there are some methods for distinguishing between good and bad workers, however imperfectly, and since there is some apparent heterogeneity among employers in their susceptibility to stereotypes, firms who inaccurately stereotype blacks will be at a competitive disadvantage.

The two key questions that arise from this criticism of SDT are (1) whether any of the tests can override the putative value of race as a signal, that is, whether they can fill the informational vacuum that tends to be filled by race, and (2) whether markets actually punish less discerning firms. I address each question in the following two sections.

NOISY SIGNALS AND RACE

I will consider several of the most common tests that employers use to assess attitude and honesty: interviews, integrity tests, and criminal background checks. I am assuming that the employer is trying to evaluate a job applicant rather than a worker who has been observed on the job for some time.

First, the most common screening device, the ubiquitous personal interview, is also the most problematic. Holzer (1987) found that interviews (as well as reference checks and probationary periods) had little predictive value for a manager's rating of a new hire's productivity relative to other employees; if anything, the relationship was generally negative. In the personnel psychology literature, their validity is also very low, typically below 0.2 (Herriott, 1989; Cook, 1988; Schmitt and Chan, 1998; Arvey and Campion, 1982)⁸ However, despite a fair amount of research on bias in interviewing, personnel psychologists have found only weak or contradictory evidence on the possibility of racial bias in interviews (see Cesare, 1996; Lin, Dobbins and Farh, 1992; and Campion and Arvey, 1989).

Next, consider the many commercially available paper-and-pencil integrity tests (also called honesty tests), which proliferated after the national ban on polygraph testing for employment in 1988. The mission of integrity tests expanded rapidly to include the prediction of a multitude of counterproductive workplace behaviors, not just dishonesty. In the early 1990s, about 6,000 organizations administered about five million integrity tests annually (Camara and Schneider, 1994). Twenty-eight percent of retail employers responding to the National Retail Security Survey (2005) used them. Such tests inquire about past theft, attitudes toward theft, attitudes toward risk, and so forth. Scores on these tests are also associated with conscientiousness (Ones, Viswesvaran, and Schmidt, 1993; Sackett and Wanek, 1996; Hogan and Brinkmeyer, 1997; Berry, Sackett and Wiemann, 2007).

The voluminous literature on integrity tests has concluded that there are no significant differences in pass rates or mean scores between blacks and whites (Ones et. al., 1993; Sackett and Harris, 1985; Sackett and Wanek, 1996; Berry, Sackett and Wiemann, 2007). While this says nothing about bias, it has been a major selling point for the producers of the tests, since it protects employers from scrutiny by the EEOC. The usual EEOC trigger is a black-white pass-rate ratio lower than 80%; sometimes statistical measures of significant difference are used as well (Joy, 1991, p. 82).

But there are lots of questions about the validity of integrity tests. The first limitation of integrity tests is that they only approximate the test characteristic under clinical, not real-life, conditions. There is inevitably an underlying decision rule about how behavior under test conditions corresponds to behavior in specific work environments, with specific technologies, systems of monitoring, penalties for poor performance, work norms, supervisors, and labor market conditions.

Second, the validity of the tests depends critically on the assumption that prospective employees cannot game the test. However, much research shows that many occupational honesty tests are fakeable. Tests often (but not always) include lie scales to catch the fakers (Cooper and Robertson, 1995; Murphy, 1993), but these are vulnerable to manipulation by a skilled test taker.⁹ Sackett and Wanek (1996) noted that in several recent studies, when research subjects were instructed to fake an integrity test, they could raise their scores.

Guastello and Rieke (1991) found that integrity test scores are positively correlated with scores on lie scales, and that predictive validity fell substantially when correcting for faking.

Good integrity tests are significantly better than alternatives such as personal interviews. But their contributions are modest. In a comprehensive meta-analysis of 180 studies of 25 honesty tests, Ones, Viswesvaran and Schmidt (1993) found the mean operational validity of integrity tests for predicting supervisory ratings of job performance to be .41. However, among the 180 studies of the 25 integrity tests that Ones *et al* used (yielding 665 validity coefficients), Sackett and Wanek (1996) considered the most compelling to be those that tried to predict workplace performance, used a job applicant sample, and used non-self-report criteria. Only 79 of the 665 validity coefficients satisfied these criteria, and the average validities of this subset ranged from .09 (for the studies of tests focused most directly on predicting theft) to .27 (for tests predicting counterproductive workplace behavior more broadly defined). In the domain of employment testing, these validities are mediocre to bad.

Finally, there has been very little *independent* evaluation of integrity tests for validity. Most validation studies of honesty tests have been conducted by the test publishers themselves (Sackett and Wanek, 1996), and the raw materials of integrity test validation research are in most cases proprietary information (Camara and Schneider, 1994).

To summarize, the common properties of personal interviews and integrity tests are that they are not apparently racially biased (no disparate impact), but they are far from perfect signals. In the face of strong stereotypes about black criminality, are employers really likely to trust these signals?

Maybe employers trust criminal records more. Five-sixths of retail employers conduct criminal background checks, swamping the 28% of retail employers who use integrity tests (Hollinger and Langton, 2005). Legal barriers to the use of criminal records for employment screening have been falling and employers have been checking them more. With the pervasiveness of stereotypes about black criminality, the effect of this device on the employment of African Americans is ambiguous: more blacks than whites have been convicted of a crime, but the record check may reduce employers' reliance on race as a signal, which may or may not outweigh the effect of identifying more convicted blacks. According to one study, the net result is that employers who examine these records are 8.4% more likely to have hired a black applicant in the most recently filled position than employers who do not check the records, and the effect is stronger for employers who are unwilling to hire convicts (Holzer, Raphael, and Stoll, 2006). This result is possible in a statistical discrimination model only if firms exaggerate the extent of black criminality, that is, only if they hold an erroneous stereotype. This is only one study; its results need to be replicated with other datasets. However, for the sake of argument, I am going to proceed assuming that this result is robust. The fact that employers who use criminal background checks are more likely to hire blacks needs to be compared with the employment practices of employers who use screening devices which seem to be unbiased. However, we do not know whether employers who use criminal background checks hire more or less blacks than employers who use interviews or integrity tests.

To summarize, the hiring decision involves a substantial degree of fundamental uncertainty. When screening is especially costly or uninformative (as it is for the young), when workers can't reliably communicate infor-

mation about their productivity to employers, when the perceptual muck is especially thick, the inclination to rely on cheap and historically salient indices, such as race and racially biased criminal background checks, can be very strong. While interviews and integrity tests are likely to be preferred when the criterion is disparate impact, their predictive validity is low, and employers may be more likely to hire blacks when the signal is biased against blacks. (This is an unresolved question; no one has compared the hiring record of employers using background checks and the hiring record of employers using interviews or integrity tests.)

WHY DOESN'T THE MARKET PUNISH STATISTICAL DISCRIMINATORS?

Since employers vary in their ability to read, and their propensity to discount black signals of honesty and work ethic, those who are more competent observers should reap the reward of lower screening costs and more productive employees. However, this may not actually come to pass, if stereotypes are self-fulfilling. This kind of perverse feedback loop can happen through a couple different mechanisms. First, because employers expect blacks to perform worse, firms are more likely to hire blacks in low-trust positions, and to monitor them more closely. As a result employers see more malfeasance by blacks. They see what they expect to see (McCrate, 2006).

Second, Coate and Loury (1993) extended Arrow's insights to develop an account of endogenously self-confirming racial stereotypes in a job assignment model. The employer's judgment that a worker is qualified for a more desirable job with more discretion or autonomy is a function both of the applicant's noisy test score and the employer's stereotype about the probability of satisfactory qualifications among blacks. For workers, the net benefit of investing in the necessary traits depends on the employer's requisite cutoff score for the more autonomous position, which is set higher for blacks when the employer holds a stereotype that they are less likely to be qualified. Coate and Loury assume that there is little point to investing when the cut score is very high or very low, and thus the percentage of workers who are qualified for the more responsible job initially increases, then decreases with the cut score. Equilibrium occurs when employers' expectations about a group's qualifications are consistent with workers' investment decisions given the cut score, which is itself contingent on the employer's prior beliefs about the group's qualifications. There are potentially multiple stable equilibria in the model, and some of these are discriminatory, with different outcomes for different groups depending only on prior beliefs or stereotypes. If firms initially believe that blacks are less motivated than whites, blacks invest so as to confirm the expectations. Blacks come to have the traits that are associated with them, simply because others expect them to be that way. The source of the problem is the stereotype, which causes employers to structure interactions (such as screening) and incentives so as to produce exactly what they expect to see.¹⁰

For both reasons – greater monitoring of blacks, and different black investment as a result of expectations – we can expect self-reproducing stereotypes. In the words of Judith Butler (2004, p. 206), agents do not engage in “simple seeing, an act of direct perception, but [in] the racial production of the visible, the workings of racial constraints on what it means to ‘see’...a repeated and ritualistic production of blackness.”

If academics can figure out how employers' stereotypes end up reproducing themselves, one would suppose that an employer, with more at stake, would figure it out too. As the critics of SDT have pointed out, this lack

of vision is hardly profit-maximizing behavior. And as Loury (2002) points out, these employers should at least be willing to experiment.

So why the dearth of experiments? This is especially perplexing in the case of stereotypes about work ethic and honesty, since the similarity of black and white integrity test scores should provide additional comfort for those willing to experiment. Part of the problem is that people readily accept stereotype-confirming information as inherent in the nature or basic disposition of the target individual (“that’s just the way they are”); stereotype-discordant anomalies tend to be regarded as situational (Pettigrew, 1979; Fiske, 1998). Hence Loury’s notion of “biased social cognition”: if employers think that blacks are truly and exogenously less honest or motivated, employers do not perceive their own role in contributing to self-fulfilling racial stereotypes. They further see little reason to experiment.

Second, although stopping a self-confirming stereotype in its tracks is Pareto superior to leaving it unchallenged, Loury points out that most firms, as “competitive observers,” cannot affect worker investment incentives by examining their own stereotypes or job assignment practices, because they cannot affect overall social stereotypes by themselves. Loury (2002) further argues that “monopolistic observers,” such as very large firms, who do have the “power to create facts” (p. 40), are typically disinterested in experimenting with alternative explanations of racial inequality. The problem lies not so much at the level of *inference*, “the *quantitative* calculation of parameters from the available data,” as it does at the level of *specification*, “the *qualitative* framework guiding an agent’s data processing” (pp. 45-46), which is logically prior to inference.

AFFIRMATIVE ACTION

In the case of honesty and work ethic, there is close to a perfect storm of exceptionally strong stereotypes and exceptionally ambiguous information. The coordination problem stifles experimentation. What can affirmative action do? The literature on affirmative action demonstrates that one of the channels of higher productivity is better screening methods (Conrad, 1995; Weisskopf, 2004; Holzer and Neumark, 2000). But when the issue is honesty and work ethic, what exactly constitutes better screening?

The criminal background check is one screening device that seems to work as both the proponents and critics of SDT expect – in the sense that it reduces the value of race as a signal, gives employers enough confidence to actually hire more blacks among those who do not have criminal records, and seems to be diffusing as a fairly standard personnel practice, at least in the retail industry. According to Holzer, Raphael and Stoll (2006) there is also a positive association between the use of background checks and the practice of affirmative action. However, the fact remains that a criminal background check is racially biased. It may be that a biased signal, if required for all job applicants, is better than an apparently neutral signal, such as an integrity test score, for much the same reason that social scientists often prefer datasets with properties that bias them against the researcher’s stated conclusion.

On the other hand, the criminal background check also stigmatizes African Americans. This is most obvious for those with criminal records. Less obviously, as long as employers think of the black applicant with a clean record as the “exceptional” African American, the practice of checking records attests to, and reinforces, the stigmatization of all African Americans.

There may be a couple ways to improve the tradeoff between employer confidence in hiring blacks and the hardening of racial stigma. The empirical work in cognitive psychology would suggest more effort by whites to differentiate among whites, and to apply exacting tests to them as well as to the members of other groups. Schauer (2003) advocates screening everyone the same in order to avoid stigmatizing some; employers should check white applicants with as much care, and suspicion, as black. Employers could, for example, be required to use criminal background checks for all job applicants if they use them for any.

Weisskopf (2004) argues for a more nuanced and multi-dimensional evaluation system, other than just the mechanical use of a single test. However, this entails costly experiments, which most employers are unlikely to undertake voluntarily in the context of the coordination and specification problems that Loury described. Affirmative action, which has been shown to push employers toward better screening than they would otherwise choose, is one partial way out of the dilemma.

But my sense is that affirmative action will be hobbled in the absence of changes in the criminal justice system. We urgently need to change the entire biased system that generates criminal records. If this were changed, one hopes that employer confidence in criminal background checks does not depend on their racial bias. But if it does, we need to consider other measures. Full employment policies are critical. As labor markets tighten, employers become more willing to hire those with minor criminal records. Tight labor markets *make* employers experiment and adopt more nuanced methods of screening that help to overcome stereotypes. In addition, labor market intermediaries such as unions and churches, that work both with employers and clients with criminal records, can promote black employment under conditions of imperfect information and stereotyping. Finally, a strengthened affirmative action, perhaps closer to the Indian system of hard quotas (Weisskopf, 2004), may be necessary to address the dilemma of biased screening methods that promote black employment only because they are biased.

REFERENCES

- Aigner, Dennis J. and Glen G. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review* 30(2): 175-187.
- Arrow, Kenneth J. 1973. "The Theory of Discrimination," pp. 3-33 in Orley Ashenfelter and Albert Rees, eds., *Discrimination in Labor Markets*. Princeton: Princeton University Press.
- Arrow, Kenneth J. 1974a. "Some Mathematical Models of Race Discrimination in the Labor Market," pp. 187-203 in Anthony H. Pascal, ed., *Racial Discrimination in Economic Life*. Lexington, MA: Lexington Books.
- Arrow, Kenneth J. 1974b. "Models of Job Discrimination," pp. 83-102 in Anthony H. Pascal, ed., *Racial Discrimination in Economic Life*. Lexington, MA: Lexington Books.
- Arvey, R.D. and J.E. Campion. 1982. "The Employment Interview: A Summary and Review of Present Research." *Personnel Psychology* 35: 281-322.
- Association of Certified Fraud Examiners. 2004. *2004 Report to the Nation on Occupational Fraud and Abuse*.
[://www.acfe.com/fraud/report](http://www.acfe.com/fraud/report).
- Berry, Christopher M., Paul R. Sackett, and Shelly Wiemann. 2007. "A Review of Recent Developments in Integrity Test Research." *Personnel Psychology* 60: 271-301.
- Bowles, Samuel, Herbert Gintis and Melissa Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature* 39(4): 1137-1176.
- Bureau of Justice Statistics, *Sourcebook of Criminal Justice Statistics*,
[://www.albany.edu/sourcebook/pdf/t3622009](http://www.albany.edu/sourcebook/pdf/t3622009). (accessed January 2012).
- Butler, Judith. 2004. "Endangered/Endangering: Schematic Racism and White Paranoia." pp. 204-211 in Sara Salih, ed., *The Judith Butler Reader*. Malden, Mass.: Blackwell.
- Cain, Glen G. 1986. "The Economic Analysis of Labor Market Discrimination," in Orley Ashenfelter and Richard Layard, eds. *Handbook of Labor Economics*. Amsterdam: North-Holland.
- Camara, Wayne J. and Dianne L. Schneider. 1994. "Integrity Tests: Facts and Unresolved Issues." *American Psychologist* 49(2): 112-119.
- Camara, Wayne J. and Dianne L. Schneider. 1995. "Questions of Construct Breadth and Openness of Research in Integrity Testing." *American Psychologist* 50(6): 459-460.
- Campion, James E. and Richard D. Arvey. 1989. "Unfair Discrimination in the Employment Interview." pp. 61-73 in Robert W. Eder and Gerald R. Ferris, eds., *The Employment Interview: Theory, Research and Practice*. Newbury Park: Sage Publications.

- Cesare, Steven J. 1996. "Subjective Judgment and the Selection Interview: A Methodological Review." *Public Personnel Management*, 25(3): 291-306.
- Coate, Stephen, and Glenn C. Loury. 1993. "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review* 83(5): 1220-1240.
- Commission of Sponsoring Organizations of the Treadway Commission. 1999. *Fraudulent Financial Reporting: 1987-1007: An Analysis of U.S. Public Companies*. Mark S. Beasley, Joseph V. Carcello, and Dana R. Hermanson.
- Conrad, Cecilia. 1995. "The Economic Cost of Affirmative Action," in M.V. Lee Badgett and Margaret Simms, eds., *Economic Perspectives on Affirmative Action*. Washington, DC: Joint Center for Political and Economic Studies.
- Cook, Mark. 1988. *Personnel Selection and Productivity*. Chichester: John Wiley and Sons.
- Cooper, Dominic, and Ivan T. Robertson. 1995. *The Psychology of Personnel Selection: A Quality Approach*. London: Routledge.
- Darity, Jr., William. 1989. "What's Left of the Economic Theory of Discrimination?" in Steven Shulman and William Darity, Jr., eds. *The Question of Discrimination: Racial Inequality in the U.S. Labor Market*. Middletown, CT: Wesleyan University Press.
- Darity, Jr., William. 1998. "Intergroup Disparity: Economic Theory and Social Science Evidence." *Southern Economic Journal* 64(4): 805-826.
- Darley, John M., and Paget H. Gross. 1983. "A Hypothesis-Confirming Bias in Labeling Effects." *Journal of Personality and Social Psychology*. 44:20-33.
- Deshpande, Ashwini and Thomas E. Weisskopf, 2010. "Does Affirmative Action Affect Productivity in the Indian Railways?" Working paper.
- Dickens, William T., Lawrence F. Katz, Kevin Lang, and Lawrence H. Summers. 1989. "Employee Crime and the Monitoring Puzzle." *Journal of Labor Economics* 7(3): 331-347.
- Federal Bureau of Investigation. 2002. *Crime in the United States: 2001*. Washington, DC: U.S. Government Printing Office.
- Fehr, Ernst, and Simon Gächter. 2000. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14(3): 159-181.
- Fiske, Susan T. 1998. "Stereotyping, Prejudice, and Discrimination." Pp. 357-411 in Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, eds., *The Handbook of Social Psychology*. Boston: McGraw-Hill.
- Guastello, S.J., and M.L. Rieke. 1991. "A Review and Critique of Honesty Test Research." *Behavioral Sciences and the Law* 9:501-523.

- Herriott, Peter. 1989. "The Selection Interview." pp. 433-438 in Herriott, Peter, ed., *Assessment and Selection in Organizations*. Chichester: John Wiley and Sons.
- Hogan, J. and K. Brinkmeyer. 1997. "Bridging the Gap Between Overt and Personality-Based Integrity Tests." *Personnel Psychology* 50: 587-599.
- Hollinger, Richard C., and John P. Clark. 1983. *Theft by Employees*. Lexington, MA: Lexington Books.
- Hollinger, Richard C., with Jason L. Davis. 2002. *National Retail Security Survey: Final Report*. Gainesville, Florida: University of Florida.
- Hollinger, Richard C., and Lynn Langton. 2005. *2004 National Retail Security Survey: Final Report*. Gainesville: University of Florida.
- Holzer, Harry J. 1987. "Hiring Procedures in the Firm: Their Economic Determinants and Outcomes." Cambridge: National Bureau of Economic Research Working Paper No. 2185.
- Holzer, Harry and David Neumark. 2000. "Assessing Affirmative Action." *Journal of Economic Literature* 38:483-568.
- Holzer, Harry J., Steven Raphael, and Michael A. Stoll. 2006. "Perceived Criminality, Criminal Background Checks, and the Racial Hiring Practices of Employers." *Journal of Law and Economics* 49 (2): 451-480.
- Huang, Fali, and Peter Cappelli. 2006. "Employee Screening: Theory and Evidence." Cambridge, MA: National Bureau of Economic Research Working Paper 12071.
- Jodhka, Surinder S., and Katherine Newman. 2007. "In the Name of Globalisation: Meritocracy, Productivity and the Hidden Language of Caste." *Economic and Political Weekly* Oct. 13, 2007:4125-4132.
- Joy, Dennis S. 1991. "Basic Psychometric Properties of a Preemployment Honesty Test: Reliability, Validity, and Fairness." pp. 65-88 in Jones, John W., ed., *Preemployment Honesty Testing: Current Research and Future Directions*. Westport, CT: Quorum Books.
- Kirschenman, Joleen, and Kathryn M. Neckerman. 1991. "We'd Love to Hire Them, But...": The Meaning of Race for Employers." in Christopher Jencks and Paul E. Peterson, eds., *The Urban Underclass*. Washington, DC: Brookings Institution.
- Lin, Thung-Rung, Gregory H. Dobbins, and Jiing-Lih Farh. 1992. "A Field Study of Race and Age Similarity Effects on Interview Ratings in Conventional and Situational Interviews." *Journal of Applied Psychology* 77(3): 361-371.
- Loury, Glenn C. 2002. *The Anatomy of Racial Inequality*. Cambridge: Harvard University Press.
- Lundberg, Shelly, and Richard Startz. 1983. "Private Discrimination and Social Intervention in Competitive Labor Markets," *American Economic Review* 73(3): 340-347.

- McCrate, Elaine. 2006. "The Racial Gap in Workplace Autonomy." In Marlene Kim, ed. *Race and Economic Opportunity in the Twenty-First Century*. London: Routledge.
- Moss, Philip, and Chris Tilly. 2001. *Stories Employers Tell: Race, Skill, and Hiring in America*. New York: Russell Sage Foundation.
- Murphy, Kevin R. 1993. *Honesty in the Workplace*. Pacific Grove, CA: Brooks/Cole Publishing Co.
- Nagin, Rebitzer, Sanders, and Taylor. 2002. "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment." *American Economic Review* 92(4): 850-873.
- Ones, Deniz S., Chockalingam Viswesvaran, and Frank L. Schmidt. 1996. "Group Differences on Overt Integrity Tests and Related Personality Variables: Implications for Adverse Impact and Test Construction." Paper presented at the annual conference of the Society for Industrial and Organizational Psychology Conference.
- Ones, Deniz S., Chockalingam Viswesvaran, and Frank L. Schmidt. 1993. "Comprehensive Meta-Analysis of Integrity Test Validities: Findings and Implications for Personnel Selection and Theories of Job Performance." *Journal of Applied Psychology* 78(4): 679-703.
- Pager, Devah. 2007. *Marked: Race, Crime, and Finding Work in an Era of Mass Incarceration*. Chicago: University of Chicago Press.
- Pettigrew, T.F. 1979. "The Ultimate Attribution Error: Extending Allport's Cognitive Analysis of Prejudice." *Personality and Social Psychology Bulletin* 5:461-476.
- Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism," *American Economic Review* 62(4): 659-661.
- Piff, Paul K., Daniel M. Stancato, Stéphanie Côté, Rodolfo Mendoza-Denton, and Dacher Keltner. 2012. "Higher Social Class Predicts Increased Unethical Behavior." *Proceedings of the National Academy of Sciences of the United States of America* 109(9).
- Raphael, Steven, Michael A. Stoll, and Harry J. Holzer. 2000. "Are Suburban Firms More Likely to Discriminate Against African-Americans?" *Journal of Urban Economics* 48: 485-508.
- Rieke, Mark L. and Stephen J. Guastello. 1995. "Unresolved Issues in Honesty and Integrity Testing." *American Psychologist*: 458-459.
- Sackett, Paul R., and Harris, Michael M. 1984. "Honesty Testing for Personnel Selection: A Review and Critique." *Personnel Psychology* 37(2): 221-245.
- Sackett, Paul R., and Michael M. Harris. 1985. "Honesty Testing for Personnel Selection: A Review and Critique." pp. 236-276 in *Personality Assessment in Organizations*, ed. H. John Bernardin and David A. Bownas. New York: Praeger.

- Sackett, P.R. and J.E. Wanek. 1996. "New developments in the use of measures of honesty, integrity, conscientiousness, dependability, trustworthiness, and reliability for personnel selection." *Personnel Psychology* 49: 787-830.
- Sackett, Paul R., and Steffanie L. Wilk. 1994. "Within-Group Norming and Other Forms of Score Adjustment in Preemployment Testing." *American Psychologist* 49(11): 929-954.
- Schafer, Frederick. 2003. *Profiles, Probabilities, and Stereotypes*. Cambridge: Harvard University Press.
- Schmitt, Neal, and David Chan. 1998. *Personnel Selection: A Theoretical Approach*. Thousand Oaks: Sage Publications.
- Slora, Karen B. 1991. "An Empirical Approach to Determining Employee Deviance Base Rates." pp. 21-38 in Jones, John W., ed., *Preemployment Honesty Testing: Current Research and Future Directions*. Westport, CT: Quorum Books.
- Stoll, Michael A., Steven Raphael, and Harry J. Holzer. 2004. "Black Job Applicants and the Hiring Officer's Race." *Industrial and Labor Relations Review* 57(2): 267-287.
- Tonry, Michael. 1995. *Malign Neglect: Race, Crime, and Punishment in America*. New York: Oxford University Press.
- Weisskopf, Thomas E. 2004. *Affirmative Action in the United States and India: A Comparative Perspective*. London and New York: Routledge.

¹ I will use the terms "motivation," "work ethic," "attitude," "conscientiousness" and "honesty" almost interchangeably, partly because employers do (for example, Wal-Mart's notion of "time theft"), and partly because there is a strong association between "integrity" and "conscientiousness" in formal tests. Integrity test scores, however, are largely orthogonal to cognitive ability scores (Ones, Viswesvaran and Schmidt 1993).

² For example, the putative dishonesty of some working-class people pales in comparison to the mendacity and the fraud associated with most American financial crises since 1980 – dishonesty perpetrated by financial officers in the most elite circles of their industry.

³ Hollinger and Clark (1983) found that 41.8% of retail workers admitted to stealing; 32.2% of hospital employees, and 26.2% of manufacturing workers did so as well. See also Slora, 1991, and Nagin, *et. al.*, 2002.

⁴ Measures of dishonesty and other counterproductive work behaviors are also negatively correlated on these tests with agreeableness and emotional stability. Conscientiousness, agreeableness, and emotional stability are three of the "Big Five" dimensions of personality (Berry, Sackett, and Wiemann, 2007).

⁵ Jodhka and Newman (2007) observed similar regional stereotypes in India. For example, the owner of a small manufacturing firm remarked (p. 4128): "There is a great deal [of stereotyping] about Uttar Pradesh people. There is a constant mimicking of Bihari labourers. Lazy guys, come in drop in without work... The work I expect to be done in three minutes would probably take an hour and a half." Also, Jodhka and Newman reported some heterogeneity among employers, with some denying that caste and religion influenced their hiring decisions: "I haven't seen any kind of correlation between the religion of the person and his work. It is basically his caliber, attitude and his commitment that is seen" (p. 4126).

⁶ The Association of Certified Fraud Examiners (2004) estimated (very loosely) that employee fraud in the U.S. cost \$660 billion in 2003, about 6% of an organization's total revenues. This dwarfs the magnitude of employee theft in the retail sector (0.8% of sales (National Retail Security Survey)). Employee fraud also raises the spectre of bankruptcy, financial penalties, and delisting by national exchanges (Committee of Sponsoring Organizations of the Treadway Commission, 1999).

⁷ The Coate and Loury model involves endogenously self-confirming racial stereotypes. See Section IV.

⁸ In this context, validity means correlation between test scores and performance measures, usually supervisor ratings or specific measures of detected counterproductive activity.

⁹ Lie scales add up the number of socially approved but implausible responses. For example, a test might ask whether the respondent has ever said anything about someone behind his back that she would not be willing to repeat to his face.

¹⁰ A perverse feedback loop also follows from Phelps' error variance model. Because the relationship between test scores and unobservable productivity is stronger for whites than blacks, the return to investment in such productivity is lower for blacks. Highly productive blacks are discouraged by their relative position in the hiring queue; these incentives generate lower mean unobservable black productivity endogenously (Lundberg and Startz, 1983).