
David Charny*
G. Mitu Gulati**

Although economists and sociologists continue to dispute the role of discrimination law in contributing to blacks’ economic progress in the 1960s and 1970s,¹ there is an emerging consensus that the “success story” of that period is unlikely to be replicated in the next decade.² The circumstances that enabled the law to be effective during that period are unlikely to recur. Several factors are of particular importance. First, the overt use of discriminatory criteria—often based on Jim Crow laws—has been dismantled and the less overt, “neutral” barriers to minorities’ employment—the battery of apparently discriminatory “objective” performance tests—have been removed either as invalid or as too costly to validate in court. Further progress in eliminating discrimination must aim at forms of dis-

*Professor of Law, Harvard Law School.
**Acting Professor of Law, UCLA School of Law.

Thanks for helpful suggestions and criticisms are due to Yochai Benkler, Devon Carbado, Catherine Fisk, Howell Jackson, Jerry Kang, Louis Kaplow, Gillian Lester, Bill Rubenstein, Richard Sander, Steve Shavell, and Seana Shiffrin, and to participants in the law and economics workshop at Harvard Law School. We owe a special debt to David Wilkins for his insights and encouragement at every stage, and for generously sharing with us his important ongoing research on minorities in the legal profession. The Harvard Law School and the Olin Program for Law and Economics at Columbia Law School provided funding to Professor Charny; the Olin Program for Law, Economics and Business at Harvard Law School provided funding to Professor Charny and Professor Gulati.


discrimination that are more subtle, and therefore more difficult to eradicate. Second, the manufacturing sector, which made great progress through quotas and training programs, is shrinking and so offers diminished opportunities to young minority workers first entering the labor market. Third, it appears unlikely that the success of discrimination law for relatively routine jobs in the manufacturing sector can be extended to high-level jobs—those professional and technical jobs that form the desirable core of the new information economy. In particular, the legal system cannot effectively second-guess the validity of managers' personnel decisions in jobs where job performance is evaluated on subjective performance criteria.

See Christopher Edley Jr., Not All Black and White: Affirmative Action, Race, and American Values 100–01 (1996); David A. Strauss, The Law and Economics of Racial Discrimination in Employment: The Case for Numerical Standards, 79 Geo. L.J. 1619 (1991) (arguing on this basis for the use of numerical solutions such as quotas). But see Thernstrom & Thernstrom, supra note 1, at 19, 520 (expressing skepticism about the problem of subtle discrimination).

Our analysis focuses on the effects of discrimination on racial minorities. We believe, however—and this will become apparent from later discussion—that much of what we say applies to women in high-wage jobs. Indeed, we have taken a number of our insights from studies done on the progress, or lack thereof, of women in high-wage jobs. It is worth noting, however, that we have not considered the applicability of our model to age discrimination. Age discrimination is fundamentally different from other types of discrimination since aging is something that occurs for all of us. Further, many of the decision-makers in high-wage jobs are older, and hence are less likely to hold stereotypes against the old that are of the type that are held against young minority or women workers. Finally, and most importantly, our analysis is one that attempts to study the effects of discrimination on an employee's entire career. The problem of age discrimination is more of a snapshot problem. Cf. Richard Posner, Aging and Old Age 320 (1995) ("[T]he kind of 'we-they' thinking that fosters racial, ethnic and sexual discrimination is unlikely to play a large role in the treatment of the elderly worker"); Christine Jolls, Hands-Tying and the Age Discrimination in Employment Act, 74 Tex. L. Rev. 1813 (1996) ("Old age has a temporal and, most critically, a universal element . . . that is lacking in the categories covered by Title VII.").


The problem was first identified and extensively analyzed, in the legal literature, by Elizabeth Bartholet's landmark article, Application of Title VII to Jobs in High Places, 95 Harv. L. Rev. 945 (1982). See also Barbara J. Flagg, Fashioning a Title VII Remedy for Transparently White Subjective Decisionmaking, 104 Yale L.J. 2009 (1995); Ramona L. Paetzold & Rafael Geltu, Through the Looking Glass: Can Title VII Help Women and Minorities Shatter the Glass Ceiling?, 31 Hous. L. Rev. 1517 (1995); Tracy Anbinder Baron, Comment, Keeping Women Out of the Executive Suite: The Courts' Failure to Apply
We contend here that the central challenge to further success of the anti-discrimination norm is the difficulty of applying its legal conceptions to jobs where performance is difficult to monitor. Our central claim is that discrimination is likely to be self-perpetuating (in part because of actions and choices by employees) in these jobs, and that current law, with its exclusive focus on intentional discrimination by the employer, is likely to be ineffective in eradicating this self-perpetuating discrimination.

The core of the argument is an analysis of how discrimination affects the job application process. In the legal literature, the most influential economic analysis of the anti-discrimination norm generally has depended upon three central analytic features: (1) a simple market-clearing model for labor markets, in which wage equals marginal product, and wages fall until the market clears; (2) a homogeneous view of the firm, in which information about the qualities and efforts of workers is perfect and in which hiring and firing is costless; and (3) a conception of discrimination as simply animus-based. These assumptions, although permitting legal analysts to reach incisive conclusions on the basis of relatively simple and rigorous economic models, are inconsistent not only with ordinary ob-


In terms of the difficulty of detecting discrimination these days, Richard Sander writes (in the context of housing discrimination):

The raw, undisguised discrimination so prevalent before Title VIII is now quite uncommon. Thirty years ago, banks, realtors, and housing developers frequently had explicit, undisguised policies barring blacks from white housing. Homeowners widely honored racial covenants even after 1953 when they were ruled unenforceable. Such overt discrimination is unusual today; when a black is the victim of discriminatory treatment, he or she is likely to not even know it.


Of course, even within the simpler models, there is ample room for controversy about the fundamental issues underlying discrimination law. In particular, the key issue is whether market forces are sufficiently powerful to erode discrimination without legal
servations of the workplace and of labor markets, but also with the generally received view among economists of how labor markets work. Conventional economists now see that this simple model is seriously at odds with the operation of many sectors of the labor market.9

For these sectors, labor economists have developed a variety of "efficiency-wage" and "tournament" models.10 These models share two essential features. First, firms find it profitable to set wages higher than the market-clearing wage, creating a pool of unemployed workers from which they can draw.11 Second, because it is costly to monitor or fire workers once they are hired, firms must adopt complex hierarchical structures that provide incentives for enhanced worker effort.12 These two features of a sophisticated microanalytic view of the firm suggest a fundamentally different view of discrimination law from those based on market-clearing models and homogenous labor conditions across firms.

As a methodological matter, the key claim of this Article is that the organizational structure of firms, and the impact of this structure on the labor markets in which these firms hire, is an important feature in analyzing the causes of employment discrimination and the effects of anti-discrimination law.13 Instead of resting on broad generalities or abstractions about features of the employment markets and the policies of employers, discrimination analysis should extend to a more refined understanding of work structures, such as definitions of job qualifications, patterns of promotion, degrees of inter-firm mobility by workers, and compensation policies. These structures powerfully affect the firms' propensity to discriminate and workers' responses to substantial sectors of the workplace. This Article focuses on high-skill or knowledge intensive jobs in which the discretionary, skill-based

---


10 See generally Milgrom & Roberts, supra note 5, at 248–443.

11 See id. at 252.

12 See infra Part II (discussing efficiency-wage and tournament models as tools for analyzing employment discrimination).

13 See, e.g., Elizabeth Chambliss, New Partners with Power?: Organizational Determinants of Law Firm Integration 203–10 (1992) (unpublished Ph.D. dissertation, University of Wisconsin (Madison)) (on file with author). Chambliss's empirical study of elite law firms finds that the organizational characteristics of a firm have a significant effect on the integration of the firm's attorney workforce. Id. See also GLASS CEILING REPORT, supra note 5, at 21 (articulating the need for research into the organizational structure of firms in order to properly understand the "glass ceiling" effect); Elizabeth Chambliss, Organizational Determinants of Law Firm Integration, 46 Am. U. L. Rev. 669, 672–74 (1997).
component of work creates substantial difficulty in monitoring individual workers’ performance. The analysis we develop applies, for example, to law firms, investment and commercial banks, brokerage houses, management and engineering consulting firms, accounting firms, universities, and governmental and large corporate bureaucracies.\footnote{We hope that the ideas articulated here will find application to a wide variety of high-wage job contexts. Nevertheless, in terms of providing examples that illustrate our claims, we primarily look to the world of elite law firms. The reasons for this are twofold. First, the context of elite law firms is one we have previously studied and we are more familiar with the relevant empirical data. Second, there has been little empirical work applying efficiency-wage and tournament concepts to the problems of discrimination in other high-wage sectors.}

Our analysis calls for a rethinking of current discrimination law and policy. Current anti-discrimination doctrine is fundamentally ill-suited to rooting out the types of unequal opportunity that our analysis identifies. Attempts by certain judges to remedy the problem by tweaking the pleading requirements and burdens of proof under the existing case law will not eradicate discrimination in the long run. We suggest that the best approach is to restructure the law to induce employers to take a proactive role in eliminating inequality. This would involve restructuring compensation, job assignments, and career pathways to provide equal opportunities for professional achievement to all employees.

I. The Current Understanding of the Role of Discrimination Law

The law inevitably proceeds on the basis of a theory, explicit or implicit, of what causes discrimination. Despite intense controversy about doctrinal detail, the doctrine’s basic conception is clear and largely unchallenged. The law bans overt discrimination, regardless of whether the discriminator is acting according to taste or a rational or irrational cognitive judgment.\footnote{One exception might be the business necessity defense to a disparate impact claim which in effect allows a business to engage in de facto discrimination in instances where a decrease in cost turns out to be accomplished by an effectively discriminatory hiring practice. However, cost cannot, except in the extreme circumstances of a bona fide occupational qualification, justify the use of explicitly discriminatory criteria.} Consequently, minorities now have the same incentives to acquire human capital as do “typical” workers, and so the social losses from the secondary effects of discrimination vanish as well. The task of the law is to solve the practical problem of how to identify discrimination. This should be easy in the standard analyses, because discrimination consists of refusing to hire or promote workers who are as qualified as the typical workers who are hired or promoted.\footnote{In some statistical discrimination models, minority workers are assumed to be, on average, less qualified than whites. See, e.g., Dennis J. Aigner & Glen G. Cain, Statistical Theories of Discrimination in Labor Markets, 30 INDUS. & LAB. REL. REV. 175, 175–87 (1977). Once the use of race is banned, however, minority applicants—in theory—face the} Disparate treatment and
disparate impact analyses only differ in detail as to the way presumptions, inferences, and burdens of persuasion are allocated. The gravamen of most discrimination litigation today is the question whether the employer is passing up (either intentionally or as a result of a neutral procedure) qualified minority applicants for less qualified nonminority applicants.\footnote{This often occurs because the alleged discrimination is so subtle that it is not easily visible to a factfinder. The factfinder must attempt to infer the discrimination from the differential in qualifications between the person who was allegedly hired on a preferential basis (e.g., the white male) and the person who was allegedly discriminated against (e.g., the female and/or black applicant). See Bray v. Marriott Hotels, 110 F.3d 986, 998 (3d Cir. 1997) (Alito, J., dissenting).}

The approach of the law follows naturally from the generally accepted accounts of the causes of discrimination. These accounts emphasize three causes: prejudice (a "taste" for discrimination); stereotypes about workers' abilities and training; and cultural norms.

\section{"Taste" for Discrimination}

The cornerstone of the modern discussion has been the model developed by Gary Becker.\footnote{GARY S. BECKER, THE ECONOMICS OF DISCRIMINATION (2d ed. 1971). For a summary of Becker's theory, see RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW, \$ 26.1 (4th ed. 1992).} In Becker's analysis, employers have an exogenously given taste for discrimination: that the worker belongs to a group that the employer dislikes adds a cost, beyond the money wage, to employing that worker. The employer is not willing to pay the same money wage for the work of the disliked group as for the work of the typical worker.\footnote{For example, if the employer has a distaste for women, then the employer will be willing to pay a female applicant less than an equally skilled male applicant for the same job, because the woman imposes the additional cost on the employer stemming from this distaste. In other words, if the cost of this distaste is $500, and the wage paid to the male worker is $10,000, the employer will see the woman worker as costing him $10,500, even though she gets a money wage of only $10,000. Because the woman worker costs more for the same wage, she won't be hired, and the woman will find herself unemployed while the man will be employed (unless the woman is willing to work for $9,500).}

Critics of this conception of "taste" have emphasized—correctly, we believe—that some economists have relied upon an oversimplified view of the psychological or social causes of discrimination. Instead, discriminatory conduct might be motivated, for example, by a desire of the predominant group to maintain its status. See Richard H. McAdams, Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination, 108 Harv. L. Rev. 1003, 1036–63 (1995). Or discriminatory conduct might be motivated by the tendency of dominant groups unconsciously to bias their evaluations of outsiders. See Deborah L. Rhode, Occupational Inequality, 1988 Duke L.J. 1207. For the purposes of our analysis, the essential lessons of these more sophisticated conceptions of the psychology of discrimination are well captured by Becker's economic model. Under the economic analysis, the crucial point is that some firms act, for whatever reason, as if
The key implication of this simple approach is that discriminating firms will not survive in the long run. An employer who does not discriminate will be able to hire members of disfavored groups at wages below those that must be paid to typical workers. This nondiscriminatory firm can then produce at lower cost, and so drive discriminating firms out of the market. If discriminators will eventually be driven out of the market, then anti-discrimination legislation is unnecessary. ²⁰ Indeed, critics argue on the basis of Becker's model that anti-discrimination legislation distorts the market away from a nondiscriminatory equilibrium. ²¹ The distortions arise because the chance of being sued under the anti-discrimination legislation increases the costs that an employer incurs when he hires a minority worker. For example, if the expected cost of a discrimination lawsuit is $500 for each woman worker hired, the cost of hiring a woman becomes $500 higher than the money wage that she is paid. ²² Within the premises of Becker's model, defenders of legal regulation are left to emphasize the interim costs of waiting for discrimination to be eradicated by market processes—in particular, the psychic harms suffered by victims of discrimination and the social loss that results when discrimination reduces the incentives for a group to work or invest in human capital. ²³

B. Statistical (Rational) Discrimination

These models²⁴ begin with the assumption that firms rationally expect job applicants from certain groups to be less productive than typical

²⁰ See Milgrom & Roberts, supra note 5, at 343.
²² Ideally the $500 penalty would be imposed only on employers who displayed a taste for discrimination. The $500 penalty would be for this illegitimate taste that otherwise would have reduced the woman employee's wage by $500. However, this ideal situation occurs only in a world of certainty where the regulatory authority can detect discriminators exactly and set penalties at the right levels. The Posner-Epstein critique of discrimination law is based on the assumption that even a nondiscriminatory employer will, in the presence of anti-discrimination laws, face a nonnegligible cost in the form of a possible discrimination lawsuit.
²³ Donohue, supra note 7, provides a definitive analysis following this approach. For a discussion of the relation between human capital and gender, see Edward J. McCaffery, Slouching Towards Equality: Gender Discrimination, Market Efficiency, and Social Change, 103 YALE L.J. 595 (1993); see also Reuben Gronau, Sex-related Wage Differentials and Women's Interrupted Labor Careers—the Chicken or the Egg, 6 J. LAB. ECON. 277 (1988).
²⁴ The first important paper was Edmund S. Phelps, The Statistical Theory of Racism and Sexism, 62 AM. ECON. REV. 659, 659–61 (1972); see also Kenneth J. Arrow, Models of Job Discrimination, in RACIAL DISCRIMINATION IN ECONOMIC LIFE 83 (Anthony H. Pascal ed., 1972); Kenneth J. Arrow, Some Mathematical Models of Race Discrimination
applicants. For example, one might expect women as a group to have higher quit rates than men because they tend to have children or blacks to have lower educational levels than whites because of educational discrimination. These assumptions may not hold true in the case of an individual applicant for a job. But employers do not have the resources to examine the individual qualifications of applicants, and instead make judgments based on group affiliations.25

Unlike Becker's model, which analyzes a stable taste for discrimination, these models predict that discriminating firms will disappear only under very specific circumstances. Incorrect judgements about groups might correct themselves over time.26 Alternatively, even if expectations are correct in the short run, an entrepreneur might be able to profit in the long run by correcting these differences and producing an effective workforce at lower cost than other employers. These enterprising firms might find it profitable to make the additional investment needed to select the most productive members of less productive groups.27 In the long run, discriminating firms would face higher costs and go out of business.

In other circumstances, however, statistical discrimination is indefinitely self-perpetuating. First, statistical judgments of employers may be robust because they accurately reflect socially determined differences in the decisions of members of various groups to acquire job skills or to commit themselves to the workforce.28 Second, statistical stereotypes may act as

\[\text{in the Labor Market, in Racial Discrimination in Economic Life} \text{ 187 (Anthony H. Pascal ed., 1972); Kenneth J. Arrow, The Theory of Discrimination, in Discrimination in Labor Markets 3 (Orley Ashenfelter & Albert Rees eds., 1973) (discussing how statistical discrimination arises where it is expensive to collect information that would change an employer's perceptions about workers from a particular group).}\]

25For example, a black woman engineer from MIT is a member of 4 groups. A potential employer may devalue her expected productivity by $500 for being black, by another $500 for being a woman, raise it by $1,000 for being an engineer, and by another $1,000 for her being from MIT. Risk-averse firms are even more likely to discriminate. White male workers are the commodity about which they have the most information, and hence the commodity that produces the least uncertainty when hired. See Aigner & Cain, supra note 16. However, the assumption of risk-averse firms is not very believable in a world of nearly perfect capital markets with investors who hold diversified portfolios.

26For example, during periods of labor shortage, firms will be forced to hire from the discriminated against group, and then will discover that their assumptions about this group have been false.


28Gary S. Becker, Human Capital, Effort, and the Sexual Division of Labor, 3 J. LAB. ECON. S23 (1983), Solomon W. Polachek, Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure, 63 REV. ECON. & STAT. 60 (1981), and Richard A. Posner, supra note 21, apply human capital analysis to the issues of the sexual division of labor and male-female wage differentials. Becker's insight here was to treat the household, as opposed to the individual, as a unit of production where different
self-fulfilling prophecies. Because the individual is judged on the basis of her membership in a group rather than her individual merit, discrimination reduces incentives for its victims to acquire human capital.²⁹

Consider, for example, the situation in which the employer has accurate information about the productivity of his workers, but other firms (prospective employers) rely on externally observable status as a signal.³⁰ That is, only the firm currently employing the workers knows about their productivity. As outside firms do not know about the worker's productivity, they cannot make competing wage offers and so the market mechanism is short-circuited. This phenomenon depends on two conditions. First, because members of a minority group have a harder time signaling ability,³¹ their employers can conceal their abilities from other employers more easily. Secondly, certain high-level jobs signal visibility to the general marketplace better than others (e.g., the performance of a CEO of a company is more visible to the external observer than that of a janitor). These two assumptions produce the result that firms pay their minority workers lower wages than white male workers of the same skill levels. The rationale is that the firm has no incentive to pay the minority workers higher wages, since there is no fear that other firms will bid away those workers.

This analysis implies not only that minority "invisibles" are paid lower wages on average, but also that they are less likely to be promoted. Because wage differentials increase as ability increases, the most able "invisibles" invest too little in human capital. Further, this statistical discrimination does not correct itself. The strategy used by an individual firm is optimal for it, and hence the entrepreneur cannot enter the market with a better strategy to outcompete the discriminators. In this analysis,

members specialized in different tasks. The next step in the analysis was to take as exogenously given the assignment of women to effort intensive home activities such as child rearing. It follows then that women with children have less energy to spend on market activities, earn less in the market, and quit jobs earlier than men do.


³¹ For instance, traditional, publicly available signals, such as schooling and test scores, are regarded to be less reliable for disadvantaged workers than for "typical" workers.
the social costs of discrimination are twofold: job assignments are in-
efficient and the stock of human capital is too small. The model also
explains why members of the “invisible” groups might choose professions
where visibility upon training is the highest (e.g., sports and certain
high-technology fields like engineering and nursing where skill and effort
levels are easy to monitor).

C. Workplace Cultures, Worker Preference, and the Costs of Integration

In taste and statistical discrimination models, the firm does not make
its employment decisions in accord with the actual productivity of the
particular worker. Rather, the firm acts either on its exogenously given
tastes or on statistical generalizations that are accurate, if at all, only for
a group as a whole. A third theory of persistent discrimination identifies
reasons for which each minority worker might be less productive (net of
labor costs) than a “typical” worker. Costs of integrating the workplace
might arise from the different exogenously given attributes of groups,
such as language skills,32 common culture or tastes,33 or the reluctance of
typical workers to help to train minority workers or to cooperate with
them.34 These costs of integrating might cause segregation to persist.35
Alternatively, minority workers would receive lower wages from firms
where “typical” workers predominated, because hiring them imposes not
only a money-wage cost on an employer, but also a cost of integration.36

Of course, members of the minority group may strive to reduce these
costs. They may, for example, learn the language or culture of the pre-
dominant group.37 Once a minority-group member has acquired the ability
to “look” like the majority group, he will avoid discrimination.38 Hence

32 See Kevin Lang, A Language Theory of Discrimination, 101 Q.J. Econ. 363, 366
(1986).
33 See Epstein, supra note 1, at 61–69.
34 See Assar Lindbeck & Dennis J. Snower, Cooperation, Harassment, and Involuntary
35 See Sunstein, supra note 7, at 22. Epstein’s argument also works along these lines.
He expects segregated workplaces in equilibrium where different groups face high costs
of integrating with others. This would occur in a world where people have discriminatory
tastes, but there are no anti-discrimination laws. See Epstein, supra note 1, at 61.
36 Epstein explains this in part as the reduced costs of enforcing informal (nonlegal)
workplace norms if one has workers with common backgrounds (such as race or sex). See
Richard A. Epstein, The Status-Production Sideshow: Why the Antidiscrimination Laws
37 See Darity, supra note 27, at 340–56; Lang, supra note 32, at 367.
38 “Looking” like a majority group member does not necessarily require a physical
resemblance. Rather, the reference is primarily to adopting the attitudes, beliefs and
interests of the majority. For minority group members at a firm, this may mean pretending
not to be concerned with issues of discrimination in the workplace or civil rights issues
the returns from investment in types of education that enable the member of the minority group to fit into the dominant culture are higher for minority group members than majority group members. This effect is of economic consequence, however, only if minority groups have the initial endowment of resources and the educational opportunities required to invest in or purchase the necessary training. Without these resources, culture-based discrimination, like statistical discrimination, may be self-sustaining.

This survey of the current understanding of persistent employment discrimination makes clear why the law operates with a simple “act/harm” model of discrimination. On each of the extant theories, the law can remedy the evils at stake by simply banning discriminatory conduct. Such a ban should be readily enforced because discrimination, in current theories, entails preferring a less qualified worker in a favored group over a more qualified worker in the discriminated against group. Once the law forces employers to hire the “most qualified” workers, the secondary evils of discrimination also disappear. Faced with a market that equally awards opportunities to qualified workers, members of discriminated-against groups will have every incentive to acquire relevant education and training. As a doctrinal matter, therefore, controversy under the current model inevitably focuses on the detail of determining appropriate “qualifications” grounds for employment decisions, either under the Hicks^40 disparate treatment model or under the Wards Cove^41 standard of business justification of disparate impact.

We argue that these simple models are inadequate to understand current extant discrimination, and that, consequently, the legal structures regulating discrimination are inadequate to eliminate it.\footnote{\textsuperscript{42}}

\footnote{39}{\textsuperscript{39}}Darity discusses how the differential abilities of different immigrant groups to integrate themselves are a function of initial endowments of human or other capital, contrasting Jewish, Asian, and West Indian immigrants to African American descendants of slaves and Chinese American descendants of railroad workers. See Darity, supra note 27. The differential access to capital between blacks and other groups in America is startling. See Peter P. Swire, \textit{Equality of Opportunity and Investment in Creditworthiness}, 143 U. Pa. L. Rev. 1533, 1544–46 (1995) (revealing that only 44.6\% of black households have checking accounts, as compared to 80\% with the rest of the population; this result remains true for other kinds of financial assets).}

\footnote{40}{\textsuperscript{40}}See St. Mary's Honor Center v. Hicks, 509 U.S. 502 (1993).}

\footnote{41}{\textsuperscript{41}}See Wards Cove Packing Co. v. Antonio, 490 U.S. 642 (1989).}

\footnote{42}{\textsuperscript{42}}We acknowledge, and indeed seek to emphasize, that the earlier models of discrimination, such as Becker's taste model, were a function of the type of discrimination that existed at the time, i.e., taste-based animus. Understandably, therefore, the anti-discrimination laws that were designed at that time (e.g., the Civil Rights Act of 1964) were designed to combat the problem that was then predominant. Part of the exercise of this Article is to realize that the types of discrimination-based problems that we face now are
II. An Analysis of Employment Discrimination in Elite Jobs

In this Part, we explain how self-perpetuating discrimination may arise through mechanisms more complex and resilient than those identified by current theory. The basic theme of our analysis is that, in higher level jobs, difficulties in monitoring (or, equivalently, discretion by managers in evaluating workers) create powerful opportunities for discriminatory hiring and promotion by firms. Unlike the forms of discrimination ascribed to taste or some inaccurate statistical generalization, such policies are able to survive at equilibrium.

The purpose of this section is to build more sophisticated considerations of nonmarginal product analysis into the standard theories of discrimination described. We begin with a brief general statement of the intuitions behind nonmarginal product models of the labor market and the firm. We then develop in more detail the conception that we will use for our own analysis. Finally, we consider how discrimination arises in our setting.

A. Nonmarginal Product Models of Labor Markets

In the standard marginal product model of labor markets, the model encountered in "Econ 101," firms match workers with jobs. They first hire the most productive workers for a job, and then dip into the market to hire successively less productive workers. They stop hiring at the point where the additional product from the next worker is less than the cost of that worker—hence, the standard notion that wage equals marginal product. As we have already noted, the job-worker productivity matching story has dominated both the legal and economic analysis of discrimination in hiring.

Nonetheless, labor economists have long noted the anomalies that appear when one tests marginal product models. For example, firms do not reduce their wage offer even when they hire unemployed workers who are willing to work at salaries below those paid to the firm’s current employees. This creates a pool of unemployed workers who are willing
to work at wages equal to their marginal products, and who would at once be hired by firms who acted in accord with marginal product models. Another set of anomalies includes the phenomenon of up-or-out career paths. These are characterized by sharp jumps in salary upon promotion (e.g., from associate to partner), and, conversely, the firms' firing of relatively less productive workers, rather than simply reducing their compensation to the actual marginal product.

A variety of efficiency-wage or tournament models, or a combination of the two, can explain these anomalies. The key to these models is that either workers or the firm are imperfectly informed. The models can best be classified in terms of the type of information deficit faced by the firm and the mechanism that the firm uses to overcome it.

I. Adverse Selection

In adverse selection models, firms are ignorant ex ante about some relevant characteristics of the workers whom they hire. For example, they cannot discern whether a worker is honest or dishonest, or diligent or lazy. The firms attempt to solve this informational problem by paying wages above the market-clearing wage to attract a superior work force. A firm that pays higher wages than other firms will be able to select the better applicants from the pool (to the extent such selection is possible) and to hold on to more of its good workers than other firms. In effect, firms pay wages so much higher than the hypothetical market-clearing wage that jobs are rationed. A firm that cuts wages to employ currently unemployed workers will lose its "good" workers and will attract only the "bad" workers, who cannot get jobs at firms offering higher wages.  


46 The economist Joseph Stiglitz uses a similar explanation to explain credit rationing. Lenders offer credit at below market-clearing interest rates, because if any one lender were to offer rates above the market-clearing level, he would attract only the high-risk projects. This is so because the low-risk projects would prefer to borrow from the low-interest rate lender (so would the high-risk projects, but they have a lower chance of receiving a loan than the low-risk projects). See Olivier Jean Blanchard & Stanley Fischer, Lectures on Macroeconomics, 478–88 (1989) (concisely explaining Stiglitz’s model of credit rationing); see also Joseph E. Stiglitz & Andrew Weiss, Credit Rationing in Markets with Imperfect Information, 71 Am. Econ. Rev. 393 (1981).

or who do not want the competitive environment of the high-wage job.\textsuperscript{47}

On the adverse selection logic, the question naturally arises why firms do not simply pay the “good” workers higher salaries than the “bad” workers. To some extent, of course, they do. For example, large bonus payments and commissions are common for high-level workers in service industries; in addition, salaries may be individually negotiated. But firms often tie their own hands with lockstep or job-determined compensation schemes.\textsuperscript{48} This is true for several reasons. First, they cannot readily distinguish between “good” and “bad” workers. Second, because workers apparently value equity, the firm might have to raise compensation for all workers, in order to compensate them for the utility loss of having to observe other workers make more money.\textsuperscript{49} Third, lockstep compensation enables the firm to document plainly the fact that it is not seeking opportunistic revisions in compensation levels after workers become locked into the firm. Workers demand such a showing because they often lack the

Hylton and Rogeau ultimately reject this construct (at least on taste grounds) because they continue to think that lenders who exercise their discriminatory taste will lose out to those who do not. \textit{See} Hylton & Rougseau, \textit{supra}, at 259. The rationale for their conclusion is that those exercising discriminatory taste are making worse decisions. \textit{See id.} We disagree. The point Hylton and Rougseau seem to miss is that they are talking about low-information markets, which give rise to a need to ration credit in the first place. In a low-information market, some set of applicants will look roughly similar. Among these applicants, it should not hurt a lender to use race in choosing credit recipients. \textit{Cf.} Peter Swire, \textit{The Persistent Problem of Lending Discrimination: A Law and Economics Analysis}, 73 Tex. L. Rev. 787, 818 (1995) (suggesting, in contrast to Hylton and Rougseau, that it is in the middle category of applicants for credit—i.e., the category where applicants look roughly the same—that loan officers exercise their tastes for discrimination, and that because of monitoring problems the exercise of these tastes is not driven out by market forces); A. Munnell, \textit{et al.}, \textit{Mortgage Lending in Boston: Interpreting HMDA Data}, 86 Am. Econ. Rev. 25 (1996) (finding evidence that while “good” candidates for mortgages would get loans no matter what their color, there is a significant amount of discrimination in choosing who among the “middle” range of candidates should get loans, i.e., among candidates who have some flaws in their credit history, but are not hopeless); \textit{but cf.} GARY S. BECKER & GUITY NASHAT BECKER, \textit{The Economics of Life} 119–20 (1997) (criticizing the Munnell/Boston Federal Reserve study).

\textsuperscript{47} \textit{See} Milgrom & Roberts, \textit{supra} note 5, at 341.


Selmi argues that affirmative action serves the purpose of reducing discriminatory behavior in the workplace, and hence improving efficiency, since a nondiscriminatory workplace is viewed as fairer, and hence effort inducing. \textit{See} Selmi, \textit{supra} note 7, at 1301–09. While we agree with Selmi’s proposition as a theoretical matter, as a practical matter affirmative action is probably seen by a large number of white workers as unfair, and, therefore, reduces effort as opposed to enhancing it, while discrimination (i.e., statistical discrimination) is seen as fair, and hence induces effort in the majority of workers.
information necessary to assess the reasonableness of specific, case-by-case compensation adjustments made by management.

2. Moral Hazard

In moral hazard models, the firms have imperfect information about how workers actually behave on the job. The firms set wages to solve this monitoring problem. In the standard model, where it is costly to monitor workers, firms use discharge as the penalty for a worker's slacking off. The firm pays the worker a wage higher than the hypothetical market-clearing wage; consequently, the worker suffers an economic loss at dismissal, and is thereby deterred from shirking. In this framework, employers set wages by reference to how much effort wage levels will induce from workers as compared to how much effort the firm could induce by direct monitoring.

Moral hazard models sometimes postulate other reasons for which a supramarginal product wage would elicit greater effort from workers. For instance, higher wages may induce workers to stay at the firm and may create an atmosphere of altruism in which workers make extra efforts as a voluntary contribution to the group effort at the firm.

3. Tournament Models

In these models, firms solve the moral hazard/shirking problem with a carrot rather than a stick. Instead of threatening to fire workers who

---


51 Along the high-wage incentive lines, a recent IMF working paper reports empirical data that shows how raising the wage of government employees from 100% of the manufacturing wage to 200% produces a significant lowering in the amount of corruption. See Economic Focus, ECONOMIST, Aug. 16, 1997, at 60 (describing Caroline Van Rijckegehem & Beatrice Webber, Corruption and the Rate of Temptation: Do Low Wages in the Civil Service Cause Corruption, IMF Working Paper, May 1997).

52 Bulow and Summers model workers who have lower opportunity costs of being fired (such as women who have the option of household activities and child rearing). According to Bulow and Summers there are two sectors: the low-monitoring, high-wage sector, and the high-monitoring, low-wage sector. Equally skilled workers could be employed in either sector, depending on how the employer perceives their incentive to shirk. See Bulow & Summers, supra note 28.

53 This tradeoff between sanction size and probability of detection is familiar in the literature on optimal sanctions. See A. Mitchell Polinsky, An Introduction to Law and Economics 75–86 (1989).

54 Tournament models are developed in Edward P. Lazear & Sherwin Rosen, Rank-Order Tournaments as Optimum Labor Contracts, 89 J. POL. ECON. 841 (1981); see also Lorne Carmichel, Firm-Specific Human Capital and Promotion Ladders, 14 BELL J. ECON.
shirk, thereby causing them to sacrifice supramarket wages, the firm offers a reward for consummate efforts—commonly a cash bonus or promotion to a high-compensation position. Workers in effect participate in a tournament to win, competing against each other in the efforts they make for the firm. Tournament models have particular appeal in our context because they help to explain the structural features of promotion to partnership, vice-presidency, or tenure that are common in law firms, consulting firms, brokerage houses, and universities: careful evaluation over a period of years and sharp changes in status and compensation for successful firm members, even when the transition seems to mean little change in actual productivity.

B. A Model of the "High-Level" Employer: Difficulties with Monitoring and its Consequences

Our model of the firm draws on the economic mechanisms developed by analysis of efficiency—wages and tournaments. The key feature of firms on which we focus—firms that are increasingly predominant in a knowledge-based or postindustrial economy—is the delegation of wide discretion to individual workers and work teams over decisions about how a job is to be done. High discretion, in turn, means low monitoring, which in turn requires the use of the various indirect incentive devices that we have already described in general terms. We now turn to a precise description of how these devices appear in the firms with which we are concerned.


55 See MILGROM & ROBERTS, supra note 5, at 367.

56 Similarly, the cost of not winning the tournament can be quite high. It often involves job loss, as in the case of up-or-out partnership decisions at law firms. See id. at 376–82 (describing job ladders in hierarchies as often being elimination tournaments, where the prize at each stage is not only the bonus or higher salary, but also the option to continue competing).

57 The emergence and development of this type of firm has received wide attention among economists and sociologists. Important discussions of these developments include Charles F. Sabel, Learning by Monitoring: The Institutions of Economic Development, in THE HANDBOOK OF ECONOMIC SOCIOLOGY 137 (Neil J. Smelser & Richard Swedberg eds., 1994) and EMPLOYMENT RELATIONS IN A CHANGING WORLDLY ECONOMY (Richard Locke et al. eds., 1995).
We deal with two stages of the employment process. The firm first hires a few workers from a large pool of job applicants. Then, after a period of employment, the firm decides whether to promote the worker (e.g., to partnership).

The crucial characteristic of the firms we analyze is that they find it expensive to observe workers' proficiency for the job or their effort level at work. Because workers' traits are costly to observe, large numbers of applicants or workers may be roughly indistinguishable to the firm, despite differences in workers' talent, training, and effort levels.

The decision of how closely to scrutinize applicants reflects a trade-off. The firm wants the best workers it can get, given the jobs in question, but it also wants to minimize the cost of interviewing and evaluating potential workers. More extensive evaluation processes incur greater costs, but also increase the chances of hiring a better worker. The firm positions itself at the point at which the tradeoff is optimal.

For purposes of exposition, our discussion will assume a normal statistical distribution of work abilities. Further, we will assume that signals of extremes in ability—very high degree of competence (applicants whom we shall call "superstars") and very low degree of competence ("low-achievers")—are fairly reliable. Signals from workers grow less reliable (more noisy) as those workers' level of quality approaches the mean from either direction. That is, the farther one is from the mean, the more reliable the signal of one's ability is. It is relatively cheap to spot "superstars" and "low-achievers," and relatively costly to rank precisely any two candidates chosen from among the average applicants. The closer to average are candidates, the more costly it is to try to measure the real differences in their abilities.

On balance, the firms that we analyze make hiring decisions based on cursory examinations of the workers—examinations that do not cost as much as a full-scale evaluation. At the hiring stage, workers present evidence of characteristics they have acquired, which signal to the employer their individual ability levels and potential effort levels (e.g., grades). Promotion decisions involve a greater degree of inquiry, but still involve the assessment of noisy signals. Correspondingly, job applicants have limited resources to devote to looking for work, and so look for jobs at only a limited number of firms. Consequently, applicants sort themselves among firms. Workers who signal low ability know that they are unlikely to get a job with a firm that gets an abundance of high-signaling or average-signaling applicants, despite the fact that the signals are only roughly correlated with ability. Even a limited opportunity to observe ability should be enough for firms to separate out the group of best

---

58 Even probationary periods provide only a very limited opportunity for evaluation and, for many firms, serve primarily as recruitment devices.
qualified applicants, thereby inducing applicants to sort themselves among firms. 59

Once a worker comes to a job, the firm wishes to induce the worker to exert effort while minimizing the firm’s expenditures on direct supervision of workers. While the firm obtains some direct information about effort and competence, this information is quite noisy. 60 For example, law firms may observe the outcomes of trials or negotiations, but they would find it prohibitively expensive to observe in detail an associate’s research, client interviews, and so forth. Such intensive monitoring in effect would require the firm to do the associate’s work all over again. 61 In “learning by monitoring” firms, to take another example, each work team is given substantial independence in setting performance standards as well as choosing production methods; the details of its decisions could only be checked by another team which, in effect, did the work from scratch. 62 As in the models discussed above, 63 the firms that we analyze here induce effort by paying a higher-than-market wage, and threatening to discharge or hold back shirking workers. 64 The wage is high enough so that even a low probability that the firm will discover shirking is a sufficient deterrent. 65 Alternatively, the firm may offer a large (supra-marginal-product) reward to successful workers. Commonly, this takes the form of promotion garnering a large increase in compensation. 66

59 High wages attract high-ability workers and serve a sorting function, as do low wages. See, e.g., Weiss, supra note 45, at 531–32; Yellen, supra note 45, at 200.
60 See, e.g., Milgrom & Roberts, supra note 5, at 404–06.
61 See Wilkins & Gulati, supra note 44, at 525–30.
62 See supra note 57.
63 See supra Part II.A.
64 For example, Walter Y. Oi, Employment Relations in Dual Labor Markets (“It’s Nice Work If You Can Get It”), 8 J. LAB. ECON. S123, S134 (1990), explains that supernormal wages serve to deter workers from shirking because the penalty for being caught shirking is the loss of the supernormal wage in an economy with high unemployment, rather than the loss of a market wage in a full employment economy. Firms also have more indirect sanctions, such as giving bad references for quitting workers to future employers. Cf. William A. Klein & John C. Coffee Jr., Business Organizations and Finance 35 (6th ed. 1996) (noting that a violation of a duty of loyalty might justify the imposition of punitive damages or other special remedies). In a market where high-wage jobs are a scarce commodity and information is costly to acquire, a signal that indicates low ability (such as a bad reference or the fact that one was laid off) can sharply reduce a worker’s chances of finding a similar job.
65 With high-level, high-paying jobs, even a very low probability of discharge may suffice to induce effort. Consider the example provided by the elite law firm summer associate experience. Summer associates may work very hard, and be very anxious about the prospect of an offer, even though the probability of not getting an offer is low. The reason is that the consequences of not getting an offer are drastic. It becomes extremely difficult to get a job at another elite firm, because a second elite law firm faced with an excess supply of labor would prefer not to hire someone who has already signaled himself as low-quality, even if the signal is very noisy.
66 See supra Part II.A.3.
Thus, our firms, like those described by the standard efficiency-wage models, set wages and employment policy to maximize the amount of effort they can extract from a worker. Wages are not determined exclusively by the external labor demand and supply functions. Firms set wages above the market-clearing wage. Consequently, there are a large number of unemployed workers who would like to work at the current wages and are qualified to do so. When employers decide to hire someone new, they have to choose out of this group of unemployed workers. The larger the group of unemployed workers, the easier for an employer to find a worker to fit the job description in question.

Thus, firms that we describe either pay higher-than-market wages to ward off shirking, or offer tournament bonuses to induce extra effort. So long as the extra benefit in terms of inducing effort or attracting high ability workers is greater than the extra wages, this is a rational profit-maximizing strategy, and firms will continue to use it. Wages far above the hypothetical market-clearing wage lead to considerable involuntary unemployment. For example, at the wages paid by elite law firms, many unemployed lawyers, as well as lawyers consigned to non-elite firms, are willing and able to do elite firm work. Yet the elite firms choose not to lower their wages. The factors we have enumerated—difficulty in detecting quality, inflexible efficiency/tournament wages, and a resulting pool of unemployed qualified workers—set the stage, as the next section explains, for a substantial degree of self-perpetuating discrimination in hiring and promotion.

C. Discrimination in Hiring

Will firms discriminate when choosing among job applicants? Consider a firm that faces an applicant pool including “typical” and minority applicants. Minority applicants are members of a group that has historically suffered from discrimination in education and job opportunities. In analyzing the firms’ decisions, we take the existence of past societal discrimination as given. Past discrimination establishes a baseline level

---

67 See Wilkins & Gulati, supra note 44, at 530–35.
68 See id. at 531.
69 A note on the definition of discrimination is in order. When we say that firms discriminate, we mean that the firms make hiring and promotion decisions for minority individuals differently from the corresponding decisions for individuals who are identical except that they are “typical” (nonminority). Intent is not part of the definition.
70 It has been empirically documented that some minority groups have a strong perception that there is a serious problem of discrimination in society today. See, e.g., Stephen E. Haydon, Comment, A Measure of Our Progress: Testing for Race Discrimination in Public Accommodations, 44 UCLA L. Rev. 1207, 1214 (1997) (citing two studies in which racial minorities ranked discrimination as the greatest problem facing American society). Minorities are more likely to hold this perception than majority group members.
of statistical discrimination—a level which reflects the influence of societal discrimination on the human capital of applicants. Paradoxically, the discrimination we describe will prove stable, once in place, even if the background societal discrimination is eliminated. Alternatively, we might assume that firms discriminate on the basis of taste—i.e., dislike of some groups—or high costs of integration of minority groups into the workplace. In our model, this discrimination, like statistical discrimination, would persist and cause an insufficient investment in human capital by minority workers.

Our analysis here proceeds in two stages. At the first stage, we identify the fundamental features of discrimination in an efficiency-wage or tournament setting. These differ in three important respects from discrimination in a market where “wage equals marginal product.” First, it does not hurt the firm economically to indulge a taste for discrimination, so long as this taste is exercised in choosing between identical candidates. Firms that engage in taste-based discrimination do not disappear, as they would in Becker’s standard model. Second, even absent a taste for discrimination, the rational firm will use minority group status as a signal if the mean achievement level of minority group members is less than that of the “typical” group. Third, minorities end up with reduced total compensation because as a group they are valued less than the “typical” workers. Having explained these basic claims, we then turn to a more detailed and inevitably more speculative analysis of minority workers’ responses to this discriminatory dynamic. In particular, we shall stipulate that workers choose among various career “strategies,” and argue that minority workers radically alter their choice of strategies in response to the discrimination that we describe. Our model of strategic choices points to further consequences of discrimination in the efficiency-wage and tournament settings. Most importantly, and perhaps most counterintuitively, individuals from minority groups are less likely to get hired, even if their mean achievement level is the same as the mean achievement level of the “typical” group. This result occurs because of the self-reinforcing interactions between the strategies workers adopt in the face of discrimination, the reliability of the signals they generate as a result, and the mechanisms that the firms use to respond to the relative unreliability of minority applicants’ signals. This in turn exacerbates the effects of the statistical discrimination, which occurs when the mean achievement level of the minority group is less than that of the “typical” group. Once minorities’

because discrimination today is the product of subtle choices by majority group members (who may not even realize that they, or others, are discriminating) and self-defeating strategies chosen by minority group members themselves.

71 See Posner, supra note 7, at 352 (describing Becker’s model for explaining how the market will push out firms that practice employment discrimination).
strategic choices are taken into account, it appears even more strongly, then, that firms engaging in discrimination are robust against market competition.

We consider each effect individually.

1. Firms Suffer No Economic Loss from Discrimination among Average Workers Based on Taste

If members of a firm dislike members of some minority groups or otherwise attain pleasure from discriminating, they will engage in discrimination. In this case, discriminatory hiring does not reduce the profits of the firm. Making a choice on the basis of discriminatory animus that is not work-related hurts profits only if the firm hires an inferior worker when a superior one is available. But when employees look alike (that is, are neither “superstars” nor “low-achievers,” so that distinguishing among them is more costly than it is worth), and the interviewer has discretionary hiring power, he can follow his taste.\textsuperscript{72} Indeed, the firm would want its hiring decisionmakers to follow their prejudices, as this would increase the decisionmakers’ enjoyment of the job and thus reduce the money wages they would demand from the firm. The firm would use the opportunity for decisionmakers to exercise their prejudices as a way of compensating those employees. Similarly, if discrimination results from unconscious bias or social bonding, the firm has no incentive to attempt to correct its decisionmakers’ discriminatory tendencies.\textsuperscript{73}

2. Endogenously Determined Rationality of Minority Group Status as a Signal

Even putting aside background social deprivation inflicted on minority groups, statistical discrimination will sustain itself vis-a-vis job applicants in high-wage sectors of the labor market. Members of minority groups who plan to enter this sector—e.g., those who have begun training to enter a profession—will be induced to make smaller investments in human capital than “typical” applicants because they expect to make lower returns from their investment in training.

This simple causal chain of self-perpetuating statistical discrimination clearly extends to the job sectors we describe. Minority group status is easy and essentially costless to observe, and the false negative judg-

\textsuperscript{72} The firm suffers a loss from passing over a minority “superstar,” but by definition, this is statistically a very small part of both the total applicant pool and the total number of hired applicants.

\textsuperscript{73} See Epstein, supra note 1, at 60–69 (explaining that discrimination increases racist employees’ morale).
ments that may result are of negligible cost to the firm because there is a surplus of equally competent people. The use of race as a distinguishing factor at the interview stage is a value-increasing strategy for a firm. A firm that does not use minority status as a signal incurs higher labor costs. In turn, minority workers know that they are discriminated against in regard to comparable nonminority applicants, and this affects their incentives to acquire job qualifications.\textsuperscript{74}

Our analysis adds several distinctive features to this pattern. We consider these in the next two subsections.

3. \textit{Depressed Employment for Minority Workers}

In the standard market-clearing model of the labor market, the effect of statistical discrimination is to reduce wages for minority workers, who obtain jobs at wages reduced to reflect the firm’s lower estimate of their productivity. In our analysis, the use of the signal is more onerous than in the standard models. Minority group members cannot overcome the statistical judgment by agreeing to work at lower wages, because wages are inflexible. Instead, minority workers simply join the pool of the unemployed, and their human capital investment is largely wasted.\textsuperscript{75} Furthermore, since minority workers are not employed, and therefore never have a chance to prove that the discriminatory stereotype of underachievement is wrongly applied to them, this gradual ameliorative mechanism for changing firms’ attitudes does not work. Thus, the stereotypes will not easily disappear.

4. \textit{Altered Distribution of Minority Group Members’ Investments in Human Capital}

The prospect of discrimination alters incentives towards human capital investment and work effort in peculiarly complicated and perverse ways—by altering the distribution, not just the total amount, of investment in human capital.

\textsuperscript{74} As we explain below, an additional source of statistical discrimination is the degree of noisiness of minority applicants’ signals relative to the signals of nonminority applicants. See infra Part II.C.4.

\textsuperscript{75} Two observations soften our conclusion. First, though wages are inflexible, it may be possible for minority workers to obtain jobs by accepting reductions in nonwage compensation, such as inferior perks, less staff support, or uninteresting assignments that are unlikely either to improve their skills or to increase their status in the firm and the profession. Second, in keeping with the usage in efficiency-wage literature, we use the term “unemployment” here to mean that highly trained lawyers are not employed by elite law firms (i.e., literally “underemployment”). These workers may well gain employment outside the law or by performing work in the nonelite sector.
To understand how this happens, consider further the choices of minority applicants faced with discrimination.\textsuperscript{76} The average minority job applicant can then pick among four potential strategies:

\textbf{a. High-Risk Strategy:} The applicant picks projects that would signal her to be a "superstar" if she does well, but which carry a high possibility of failure, since she is actually of average ability (failure then signals low-achiever status). These high-risk projects and strategies\textsuperscript{77} are also more demanding than most others. Using this strategy, the applicant of average ability makes high effort, but enjoys a low probability of success (though success may carry high rewards) and a high probability of failure.\textsuperscript{78}

\textbf{b. High-Effort Strategy:} Applicants who adopt this strategy pick projects that demand exceptionally high effort, but somewhat increase the possibility of being labeled "superstar" or "average," but decreases the probability of being labeled "below average." The important differences from a high-risk strategy, then, are that a high-effort strategy reduces rather than increases the probability of failure (this is what makes it less risky); but it may increase the probability of "superstar" success less than would a high-effort but low-risk strategy.\textsuperscript{79}

\textbf{c. Average Strategy:} The applicant picks activities requiring average ability in which she is likely to signal that she is of average ability and, conversely, faces a small risk of appearing a "low-achiever."

\textbf{d. Low-Effort Strategy:} Here the applicant makes a minimal investment in signaling job qualifications, knowing that this strategy is likely to signal employers that she is of low ability, but also providing her with the opportunity to obtain gains from leisure.\textsuperscript{80}

\textsuperscript{76} See generally Edley, supra note 3, at 229–33; Thernstrom & Thernstrom, supra note 1, at 387; Haydon, supra note 70, at 1214.

\textsuperscript{77} One high-risk strategy might be for a summer associate to take on a large number of projects with uncertain due dates. Assuming that she can only produce one project result at a time, she will hope that no more than one project becomes due on the same day (since only one project would be completed and the others would necessarily fail). If she is able to complete an exceptionally large number of projects at the same time, her supervisors may notice her large number of billed hours and label her a "superstar." Yet if more than one project comes due at the same time and her failure invokes a client's displeasure with the firm, the cost to the employee will be very high, especially if these projects could have all been successfully completed had the employee not overloaded herself.

\textsuperscript{78} It is not a high-risk strategy for a real "superstar" to choose projects that will signal her as a "superstar," since she has the abilities to complete the projects successfully.

\textsuperscript{79} It is important to realize that the amount of effort employees exert at these elite high-wage firms is extremely high. Therefore, when we talk about "high-effort" we mean extremely high effort that enables one to stand out among all those who are exerting high effort anyway. For example, the hypothetical summer associate who attempts to become friends with all the partners might be seen as exerting a high-effort strategy. This high-effort strategy, however, could become a high-risk one when the summer associate is seen as a glad-hander, and thereby, offends a partner.

\textsuperscript{80} Note that because all signals are noisy, some applicants who follow the low-effort
How will job applicants choose among these strategies? Minority applicants know in advance that they will be at a disadvantage at the interview stage. To get a job, a minority worker must signal that she is a much better choice than the nearest “typical” applicant because the interviewer will (a) choose the minority applicant only if she is clearly better (and most probably, only if she is a “superstar”), and (b) be more suspicious of the minority applicant’s qualifications, because the minority applicant had skewed incentives to acquire qualifications. The minority applicant knows that she has to signal “superstar” if she wants to secure the same job for which the “typical” worker needs only to signal average.

Faced with these hiring patterns, “typical” applicants generally have no reason to pursue a high-risk strategy. They are likely to be hired on the average strategy, which carries a much lower risk of failure. Due to discrimination, however, average minority applicants are unlikely to be hired if they pursue the average strategy, and hence, choose either the high-risk or the low-effort strategy. This choice depends on whether they are willing to make the very substantial extra effort that the high-risk strategy entails—a choice that may be irrational considering the low probability of success under that strategy.

strategy do secure jobs. In other words, there is still a small probability that a low-effort applicant will be lucky, acquire an average or even “superstar” signal, and get a job.

The concept of people pursuing high-risk strategies when faced with failure if they behave normally (i.e., make an average strategy choice) has been applied elsewhere. See Susan Rose-Ackerman, Risk Taking and Ruin: Bankruptcy and Investment Choice, 20 J. LEGAL STUD. 277, 286–92 (1991) (analyzing the strategy choices made by financial institution managers in the face of bankruptcy); see also Belcfine v. Aloe, 112 F.3d 633, 639 (3d Cir. 1997) (describing a Pennsylvania statute aimed at ameliorating some of the skewed incentives identified by Rose-Ackerman’s paper). Two further remarks on our “strategic” framework are in order. First, workers make professional decisions on many grounds other than signaling strategy. For example, professionals may choose to work on projects—regardless of consequences for “signals”—to gain valuable experience or to satisfy an internalized standard of excellence. Nonetheless, an economic model can provide insight by isolating the effects of one substantial causal force in a complex social setting. Cf., e.g., TOMAS J. PHILIPSON & RICHARD A. POSNER, PRIVATE CHOICES AND PUBLIC HEALTH: THE AIDS EPIDEMIC IN AN ECONOMIC PERSPECTIVE 7–8 (1993) (claiming that economic models predict “central tendencies” rather than “totality of individual behavior”). Coauthor Gulati’s work with David Wilkins has detected some of these tendencies in law firm professional patterns, see Gulati, supra note 44, and the discussion infra, Part II.E; we hope that our argument will lead to further empirical research.

Second, our analysis might be said to “essentialize” the traits of the workers whom we analyze, by looking at the consequences of a single, once-and-for-all choice of strategy. Workers’ ambitions and professional plans will change with time and context. Again, however, to discern the central patterns that result from the dynamics that we describe, it is useful as a first step to assume that each worker has made a single choice.

Minority applicants on average are likely to prefer high-risk to high-effort strategies. High-risk strategies carry a greater probability of being labeled a “superstar,” despite the greater probability of being labeled a “low-achiever.” For example, coauthor Gulati, who graduated from law school recently, has suggested in the context of extracurricular activities at Harvard Law School that entering the voluntary moot court competition is a high-risk strategy, trying to get very high grades is a high-effort strategy, participating in
These choices change the distribution of achievement among groups of job applicants, independent of their effects on the mean level of talent. Minority applicants have lower incentives to invest in signals that would place them on par with "typical" applicants because minority applicants know that if a choice between workers is made, they will lose out. In aggregate, minority applicants choose strategies that are likely to produce either "superstar" or "low-achiever" signals.\(^3\) This change in distribution of signals in turn affects the evaluation of minority applicants. Even if the average level of achievement for minority applicants is the same, the difference in distribution, and correspondingly, in noisiness of signals, indicates the perpetuation of statistical discrimination.

Once this distribution of talent is established, discriminatory hiring outcomes result from the differential distribution of group members among "superstar"- and average-achievement cohorts. The employment prospects for minority group members worsen when one considers the likely distribution of favorable and unfavorable signals among job applicants.

Consider a group of one hundred job applicants, of whom ten signal that they have outstanding abilities, eighty signal that they are average, and ten signal that they have low abilities. Assume that the firm needs to hire thirty people for the job. It will hire ten outstanding applicants and twenty of the average group. The ten low-ability applicants will not be hired. Among the eighty average applicants, the firm is indifferent as to which twenty it hires. The signals these eighty average applicants present during their interviews look basically the same to the interviewers, given the interviewers’ difficulty in gauging ability in these industries because of high monitoring costs.\(^4\) Consequently, the interviewers can discriminate in choosing among the average applicants.

\(^3\) The following example expands on the one suggested in note 82, supra. Assume that one important reason that law students participate in these activities is to obtain a signal that they can later take to the labor market. With the moot court competition, many people participate, the required effort levels are high, and the probability of success is low (the signal to the job market of being the "best oralist" at the final is probably a "superstar" signal). In terms of penalties, moot court competitors lose out on time to study for classes and quite possibly get lower grades than their peers. Hence, moot court fits the category of a high-risk activity. With the legal aid work or a noncompetitive law journal, so long as participants fulfill the necessary time commitment, they are assured a signal to the labor market. Failure is not possible. Although this signal will not identify legal aid and journal participants as "superstars," it will signal that they are not "low-achievers."

\(^4\) For example, a law firm interviewing on campus at Harvard Law School typically sees a grade report and a one-page resume. Maybe firms can identify the people who graduate with straight As as "superstars" and the ones with straight Cs as "low-achievers," but the vast majority of applicants will be in the B-to-B+ range, within which such distinctions mean little. We hypothesize that most elite law firms view students in the B-to-B+ range as more or less interchangeable in terms of their potential success as lawyers. Hence, it does not matter which particular law students are hired from that range.
The result is that there will be statistical discrimination even if minority and nonminority mean achievement levels are the same. Even if mean achievement of minorities is equal to that of "typicals," and most of them correctly signal themselves to be average, they will be much less likely to be hired because of taste discrimination or prior existing biases, the exercise of which among the average applicants does not hurt profits. In a world of scarce jobs, discrimination encourages average minorities to adopt skewed strategies. The skewed strategy is in the self-interest of the individual minority who is unlikely to get a job if she correctly signals herself as average, but has a better chance if she falsely signals herself as a "superstar." Across the economy, however, firms now are even more likely to discriminate because correlating achievement signals of minorities to their true abilities has become harder to do than it is for "typicals." Achievement signals are now distributed so that members of minority groups, who are producing a large number of noisy and false signals, are more likely to fall at the extremes than are "typicals," most of whom correctly signal themselves as average. This distribution is likely to occur when individuals' investment in credentials is affected by discrimination.\footnote{Selmi cites numerous studies that describe how nonminority group members see negative minority actions as stemming from their personalities, while positive actions are seen as stemming from chance or circumstances. See Selmi, supra note 7, at 1285 n.111.}

Furthermore, since average members of a minority group pick strategies that falsely signal them as "superstars," employers are more skeptical of a member of a discriminated against group who signals that she is a "superstar." Employers recognize that average members of these groups have incentives to undertake high-risk strategies that will show them to be "superstars" when, in fact, they were merely lucky.\footnote{Recall that nonminority group members are less likely to attempt these high-risk strategies because (a) they have no reason to take the extra risk of failure, since they have a relatively high probability of obtaining good jobs even if they turn out to be average; and (b) the high-risk strategies require greater effort.} Because employers recognize that apparent high achievement is often the outcome of a successful \textit{gamble},\footnote{For minority workers/applicants, there will be no point in making an effort to accumulate high credentials unless they are "superstar" credentials. Thus, while the "typical" applicant has an incentive to signal herself as a "superstar," she makes less effort than the average minority applicant to signal this to employers because she perceives fewer gains from falsely signaling herself as a "superstar" than does the average minority candidate.} the firm will discount signals given by minorities who signal themselves to be "superstars." In other words, the employer

Of course, law firms may use other criteria for hiring, such as social class, apparent fit with corporate culture, and so forth; even so, choosing fairly among a number of equally qualified applicants for associate positions will be arbitrary. For a detailed discussion of how the vast majority of applicants to elite law firms are indistinguishable, especially in the context of the largely routine work they are expected to perform, see Wilkins & Gulati, supra note 44, at 549–52.
will find these signals less believable than if they were coming from a "typical."

5. *Survival of Discrimination Revisited: Firms that Do Not Engage in Statistical Discrimination May Fail*

Once these strategic choices are taken into account, it becomes particularly clear that an economically rational firm will engage in discrimination. The firm that did not discriminate on the basis of minority status would make lower profits than its competitors, and would be driven out of business. Further, one maverick firm cannot change the incentives that minority workers face. One firm has too small of an impact on the market to change the behavior of its competitors. On the whole, minority workers will still see the market as discriminating against them.

6. *Implications*

Overall, our analysis paints a grim picture. Its main features are these:

(a) In our model, individual firms that indulge their taste for discrimination—or, more generally, act on discriminatory cognitive biases or social identifications—will not suffer a loss of profits. In turn, this discrimination alters minority applicants' incentives to obtain training and thereby changes the distribution of their signals. As a result of this skewed distribution, even firms without a taste for discrimination would be rational to interpret minority workers' qualifications with greater suspicion, putting minority workers at a competitive disadvantage. Firms that follow comparatively nondiscriminatory policies will be driven out of the market over time.

(b) Minority workers who are hired will be better qualified for their jobs than their nonminority colleagues because they are hired only when they are able to signal that they are unambiguously better than the nearest "typical" applicant.88

(c) The economy as a whole faces an underutilization of talent. Jobs are not occupied by the workers best suited to fill them. Nonminority workers with low levels of talent accumulate too much human capital and minority workers with high talent accumulate too little human capital relative to what they would in a system with no discrimination.

88 Recall, however, that this result is softened if minority workers can be given on average different tasks with different perquisites than "typicals" (albeit with the same job titles and wages).
D. Discrimination at the Promotion Stage

Discrimination at the promotion stage follows the same basic pattern as discrimination at the hiring stage. In promotion, as in hiring, the firm faces a tradeoff between accuracy of assessment and costs of information gathering. Making distinctions among the workers who signal themselves as average is not cost-effective.\(^89\) Thus, as at hiring, if the firm derives some utility from discriminating, it does not hurt the firm for it to do so. The profits of a firm that discriminates among its applicants or workers in the middle range will be no lower than those of a firm that does not discriminate. It will survive in the long run, contrary to the predictions of simple discrimination theory.

The greatest difference between the initial hiring and the promotion stages lies in the incentives that influence the workers’ choices among high-risk, high-effort, average, and shirking strategies. As discussed earlier, workers face a choice of projects. They can pick the high-risk or high-effort projects, where success will show them to be “superstars,” or the “shirking” projects that give them plenty of leisure time.\(^90\) We speculate that the main differences between the training (pre-application) setting and the work (pre-promotion) setting is that work, compared with education and job training, probably (a) provides many more high-risk projects and many fewer high-effort (but “superstar” signaling) projects; and (b) incorporates greater penalties for shirking. Both factors would drive the minority worker towards high-risk projects, even more than in the hiring setting. Picking an average project is of no use, since a worker who signals himself as average will be subject to discrimination when it comes time for promotion. But again, as more minority workers pick high-risk projects, their signals become noisier since success simply may be the result of a lucky gamble.

On the other hand, the presence of anti-discrimination law can skew a particular minority worker further towards shirking. The reason is that a low-monitoring firm does not know if a minority worker has a viable discrimination lawsuit. The small probability that a minority candidate has a potential lawsuit makes the firm (which has not collected very much information on the worker at that point) reluctant to fire the minority candidate (before the promotion stage, at which point the firm does collect a paper record that should give it protection from suit). This effect pushes

\(^89\) See, e.g., Galanter & Palay, supra note 54, at 96–97 (describing increased difficulties of monitoring performance in large, highly specialized, elite law firms); Milgrom & Roberts, supra note 5, at 407 (noting that measuring employee performance is difficult and expensive in professions like accounting and law).

\(^90\) As an example, consider associates at a law firm who can exercise some choice about effort levels and about which projects to work on. Some cases require high ability or high effort, while others require average or low ability.
the minority candidates further towards skewed strategies; they are less likely to be fired if they shirk or fail at high-risk projects, but they also remain less likely to be promoted (since the firm will have collected adequate information to protect itself at the promotion stage).\footnote{Cf. Turner, supra note 2, at 468–69 ("Title VII is now principally used as a form of job protection by incumbent employees rather than as a weapon to open employment opportunities and jobs for African-Americans and other protected groups.") (citation omitted).} Moreover, the same worker who knows she is going to face discrimination at the promotion stage has an incentive to shirk and take low-ability-requiring projects, since she suspects that she will not make partner anyway.

As a result of these considerations, minority workers' strategic choices are skewed towards high-risk and shirking strategies. As we have seen, this skewing of choices in turn makes it harder for employers to interpret the signals that they get from quality of work.\footnote{See Ayres & Siegelman, supra note 38, at 1490. In discussing the firing effects of disparate impact liability on hiring quotas, the authors assume that "black workers are more variable in their productivity on the job than are whites." The paper, however, does not offer an account of this greater variability. See id. (emphasis added).}

\textit{E. Applications}

Does the world fit the predictions of this model? Consider, first, the example of the elite law firm and, second, more general evidence about the incidence of discrimination.

\textit{1. Application to Elite Law Firms}

Large, elite corporate law firms provide archetypal contexts for the model that we have developed.\footnote{Our discussion here draws in part on work discussed in Wilkins & Gulati, supra note 44. In particular, Black Lawyers drew on David Charny & Mitu Gulati, Efficiency Wages, Tournaments, and Discrimination: A Theory of Employment Discrimination Law for "High Level" Jobs, John M. Olin Center for Law, Economics, and Business, Harvard Law School, Discussion paper no. 182 (March 1996), an earlier presentation of the basic model presented herein. See Wilkins & Gulati, supra note 44 (providing some suggestive evidence for the model's approach in its analysis of law firm data).} At the entry level, hiring decisions are based on minimal information: scrutiny of one page resumes, brief on-campus interviews, and only slightly more extended office visits. Grades and other standardized criteria place applicants in a standard distribution, with a few outstanding prospects at each end of the spectrum. Most applicants look fairly similar to one another.\footnote{See Wilkins & Gulati, supra note 44, at 546–52.} Although what happens after applicants are hired is somewhat more varied and controversial, observers of the large law firm have tended to conclude that firms do not
closely monitor associates over the period extending until partnership decisions are made.\textsuperscript{95} It is difficult to monitor highly specialized lawyers, who exercise a great deal of discretion in the tasks that they perform.\textsuperscript{96} Moreover, monitoring involves assigning high-level lawyers to supervise employees whose productivity may only be a small fraction of that of the monitoring senior associate or partner. In short, intensive monitoring will not justify its cost.\textsuperscript{97}

Instead, firms must provide other incentives to obtain consummate effort from associates despite relatively low degrees of monitoring. One strategy is to pay associates wages high enough to create a strong economic penalty if a job is lost. Large rewards for those who survive the partnership tournament present an alternative way of creating strong incentives.

Low-level monitoring combined with tournament rewards and penalizing job losses define the experience of associates, particularly as they are evaluated over the course of their employment.\textsuperscript{98} In particular, the tendency to demand "superstar" performance from minorities at elite law firms has received substantial attention.\textsuperscript{99} Higher attrition rates and greater dissatisfaction among women and minority associates could also reflect the "superstar" demands placed on them.\textsuperscript{100}

\textsuperscript{95} In her study of integration in elite law firms, Elizabeth Chambliss explains that:

[Lawyers' work tends to be non-routine, even within the same area of practice, and different kinds of cases require different kinds of work products. Thus, there rarely is an objective basis for comparing the quality of different lawyers' work. Furthermore, lawyers in elite law firms work in teams, which complicates the assessment of individual contributions.]

Chambliss, supra note 13, at 691–92. \textit{See generally} Marc Galanter & Thomas Palay, \textit{The Transformation of the Big Law Firm, in Lawyers' Ideals/Lawyers' Practices} 48 (Robert Nelson et al. eds., 1992). Although we agree with and do ourselves make the claim that elite firms are low-monitoring at any given point in time, we believe that firms tend to have collected a considerable amount of information to draw upon in evaluating employees by the time firms make the promotion-to-partnership decision.

\textsuperscript{96} \textit{See generally, e.g.,} Peter D. Sherer, \textit{Leveraging Human Assets in Law Firms: Human Capital Structures and Organizational Capabilities, 48 Indus. & Lab. Rel. Rev.} 671 (1995) (discussing the increased degree of specialization at large, highly leveraged law firms).

\textsuperscript{97} \textit{See} Wilkins & Gulati, supra note 44, at 527–28; \textit{see also} Milgrom & Roberts, supra note 5, at 407.

\textsuperscript{98} \textit{See} Wilkins & Gulati, supra note 44, at 530–39, 565–68.

\textsuperscript{99} \textit{See}, e.g., Interim Report of the Committee on Minority Employment of the San Francisco Bar Association (1993). Elizabeth Chambliss, in her empirical study, \textit{supra} note 13, at 133, 145, 189, 203–10, found that two strong determinants of elite law firm integration for black lawyers were (i) the availability of black clients, and (ii) the proportion of black decisionmakers at the firm (i.e., partners). In the context of elite law firms, client contacts are a sure signal of what they value for "superstars," and the significant predictive effect of the proportion of black decisionmakers indicates the existence of taste discrimination.

\textsuperscript{100} \textit{Cf.}, e.g., Frederick H. Bates & Gregory C. Whitebread, \textit{Do Something Different:}
2. Predictions About the Incidence of Discrimination

In industries where differences in effort and ability are hard to observe, enforcement of anti-discrimination laws is likely to be harder as well.\textsuperscript{101} In the past few decades, the United States has been moving towards an economy dominated by service-oriented businesses, i.e., ones where effort and ability are harder to evaluate. Unfortunately, there is no more than a minimal amount of data on the particular incidence of discrimination in these types of jobs.

Stephan and Abigail Thernstrom present data that shows that although the black-white gap in high school education in math and writing skills was closing somewhat in the pre-1988 period, it has widened considerably since.\textsuperscript{102} The widening gap in the types of skills integral to most service sector jobs is consistent with the hypothesis that blacks have lower incentives to invest in the skills necessary for success in the newly im-
important service sectors than do whites. Nan Maxwell, in turn, finds that data from the National Longitudinal Survey of Youth for the years 1979–1988 indicates that young, black men face substantially lower incentives to obtain an education than young, white men of equal levels of ability. Maxwell’s study is consistent with the hypothesis that in a majority of sectors requiring education-generated expertise, black workers expect to face discrimination.

The most interesting studies of discrimination for the purposes of our analysis are those conducted by the Urban Institute. In these studies, pairs of equally qualified black and white applicants, and Hispanic and white applicants, applied for jobs. The jobs on which the studies focused were primarily in retail and service sectors and required at least a high school degree. The simulated applications all indicated a college degree. The audit studies in Chicago, Washington, D.C., and San Diego found that Hispanic and black auditors received unfavorable treatment in a substantially larger number of instances than did their white counterparts. Most importantly for our analysis, the non-white auditors received a greater degree of unfavorable treatment when they applied for higher-wage jobs in service and management.

This evidence, scant though it is, favors the hypothesis that: (1) market forces and anti-discrimination laws have combined to remove blatant discrimination, i.e., discrimination that is readily observable by outsiders; but that (2) a substantial amount of discrimination survives in higher-skill jobs

---


105 See id. at 170. In the studies, the auditors tended to be qualified or overqualified for the jobs to which they applied, and yet were still not hired. Jobs are rationed in these sectors, i.e., the labor market is not clearing.

106 See id. at 173.

107 See id. at 175.


Yelnosky's primary rationale is that discrimination in these jobs is easier to detect than in high-level jobs. See Yelnosky, supra, at 414–15. But the fact that this type of discrimination is easy to detect also means that market forces can work to penalize and eventually eradicate such behavior. See William Julius Wilson, The Declining Significance of Race: Blacks and Changing American Institutions 106 (2d ed. 1978) (stating that "[S]tudies of unemployment in the urban core reveal that blacks do not
that are characterized by discretion in hiring and promotion and where discrimination is thus difficult to detect.\textsuperscript{109}

III. Affirmative Action

A. Affirmative Action in Hiring

So far, to preserve the simplicity of our model, we have omitted any discussion of affirmative action. However, to complete the picture, we now address how affirmative action, particularly among elite firms, necessarily affects the causal mechanisms that we describe.

As a preliminary matter, it is worth noting that the widespread use of affirmative action\textsuperscript{110} provides some confirmation for the basic features of our model. Interestingly, the use of race or gender as both an affirmative hiring criterion and a negative criterion reflect that, in instances of low monitoring, fuzzy signals, and efficiency-wage or tournament incentive structures, the use of minority signals is an efficient personnel strategy for a firm.\textsuperscript{111} Indeed, that affirmative action was so widespread that its most blatant form—race-norming—had to be banned by legislative action\textsuperscript{112} reflects the strong economic pressure towards use of minority signals once firms face potential legal liability for hiring disparities.

Affirmative action\textsuperscript{113} in the sectors that we describe will diminish firms' productivity only to a very small extent. A firm that selects minority candidates from among the large pool of workers who signal themselves as average will produce a workforce that is more or less as qualified as the one that it would have had under a nonquota hiring scheme. In theory,
the extra weight given to minority candidates could be set at the level necessary to counter the amount of discrimination we describe. This would resolve the problem of minority workers’ facing skewed incentives as well.

Under our analysis, affirmative action policies have mixed effects on minority workers. For instance, affirmative action policies may cause minority workers to shift their choice of job qualification strategies. The main effect of affirmative action will be that more average minority workers are hired. (Presumably, the “superstars” are hired already, although affirmative action might raise the quality of the firm willing to hire the “superstars.”) Minority candidates will therefore shift towards the strategies that signal them as average; they will have less reason to pursue the high-risk strategies that labeled them either “superstars” or “low-achievers.”

We emphasize that the widespread use of affirmative action standards will change, fairly drastically, the strategies that will be pursued by minority workers and applicants. With elite law firms hiring at elite schools, for example, it seems likely that the widespread use of affirmative action has largely eliminated distortions towards high-risk strategies for minority law students. There is less reason to believe that this is the case, however, in the promotion setting, for reasons to which we now turn.

At the same time, affirmative action will tend to make hiring decisionmakers more suspicious that the signals given by a black worker are noise as opposed to legitimate signals. By hypothesis, as a result of affirmative action, workers of average or low ability now make up a larger portion of the minority workforce. This would be the case especially if affirmative action programs at the initial hiring stage were so generous as to include substantial numbers of applicants who had pursued shirking strategies in acquiring training.

**B. Affirmative Action in Promotion**

To understand how affirmative action works at the promotion stage of our analysis, we further specify the strategic problem that employees face between hiring and promotion. The building blocks of the analysis

---

114 Stephen Coate and Glenn Loury develop a model in which an affirmative action policy discourages workers in the favored minority group from expending efforts. See Stephen Coate & Glenn Loury, Antidiscrimination Enforcement and the Problem of Patronization, 83 Am. Econ. Rev. 92 (1993). In effect, affirmative action supplements the effects of extant statistical discrimination. *Cf.* Glenn C. Loury, *Why Should We Care about Group Inequality?*, in *The Question of Discrimination*, 268, 285 (Steven Shulman & William Darity Jr., eds., 1989). Our analysis develops this approach by identifying more complex effects that affirmative action might have on the career paths of minority groups, and on the evaluation of minority applicants by firms.
are low monitoring, discretion for employees, high wages, and a tournament for promotion. The period between hiring and promotion is typically lengthy. At elite firms it can be many years.

As described above, the process up to the promotion decision is a low monitoring process, similar to the hiring process. At the time of promotion, however, the firm has had an opportunity to observe an employee’s project choices and performance over a number of years.\(^{115}\)

The preceding section described the effects of a perceived discriminatory environment on the project choices of minority workers. In many cases a discriminatory environment might produce self-defeating strategies: minority workers would choose primarily high-risk or shirking strategies.\(^{116}\) These strategic choices, in the aggregate, will reduce the likelihood of a minority worker being promoted, evaluated against a baseline of workers not subject to discrimination. The strategies are not likely to have yielded much success by the end of the semi-probationary period between hiring and promotion when the firm analyzes an employee’s project choices. The firm tends to rely on readily observable information in drawing inferences about workers, and, therefore, concludes that minority workers will fail to make the promotion. The perception of discrimination on the part of the minority workers becomes a self-fulfilling prophecy as it skews their strategies towards high-risk or shirking strategies that produce a self-defeating result.

If one adds anti-discrimination law into the mix, this can skew strategies of minority workers even further. The existence of anti-discrimination law may make many employers more reluctant to fire their minority employees in the pre-promotion period. Further, minority workers value the scarce high wages they are receiving, and think they have a low probability of promotion anyway. These circumstances push minority workers towards the safe, job-preserving (but nonpromotion) low-effort strategy.

If, as some scholars argue, affirmative action causes minority workers to doubt their abilities further,\(^{117}\) then this will push them even further

---

\(^{115}\) Cf. Wilkins & Gulati, supra note 44, at 542 (contrasting the information available at promotion to that available at hiring).

\(^{116}\) We stylize our analysis to talk primarily about employees choosing projects. In reality, however, even in high-discretion workplaces, there is both employee choice and project manager choice. This becomes more true when employees have been at the firm for long periods of time, as they develop reputations based on their initial project choices and performances. As far as we can tell, this twist does not weaken our story in any way. Indeed, factoring in project manager choice between employees strengthens our hypothesis.

\(^{117}\) See Antonin Scalia, The Disease as Cure, 1979 Wash. U. L.Q. 147, 154–55; Shelby Steele, A Negative Vote on Affirmative Action, N.Y. Times, May 13, 1990, § 6 (Magazine), at 46. We are, however, skeptical that the skewing in minority strategies is caused primarily by self-doubt. We think it more likely that the skewing is caused by minorities perceiving
towards the safer (in terms of not being fired) shirking strategies that will undermine their likelihood of gaining promotion.

The story thus far would produce the conclusions: (a) abolish affirmative action in hiring, and (b) abolish anti-discrimination laws. That might work if the perception of discrimination by minority workers was not based on reality. But, of course, this is unsatisfactory when discrimination is a reality. Nor can the minority worker’s dilemma be resolved by affirmative action at hiring and promotion alone. At first cut, one might surmise that affirmative action in promotion would have a reverse skewing effect on minority workers’ strategy choices: that is, affirmative action, ideally, would lead members of the discriminated-against group to make the choices they would make were they not subject to discrimination.

However, firms are unlikely to implement affirmative action policies at the promotion stage, for two reasons. First, the firm by that point has concrete information about employees’ project choices and final performances from a number of years. Based on this information, the firm can concretely rank candidates for promotion. So, in a sense, even though this is a system of low monitoring, the firm has a high level of information at the point of making the promotion decision. To use affirmative action here would be contrary to the self-interest of the firm, given that it can make an educated guess about the better employee. Indeed, this conclusion appears to be consistent with our observation of elite firms. More than a few of them use some affirmative action in hiring (when they are

---

their abilities to be undervalued by others. One is more likely to have an accurate idea of one’s own skills than are outsiders who have limited information.

Along these lines, there is beginning to emerge concrete evidence that black students perform differently when they are told that they are being evaluated versus when they are told they are not being evaluated. Social psychologist Claude Steele has hypothesized that this skewing of performance is a result of the black students altering their behavior or tests in order to adjust to the outside perception that they are inferior. See Claude M. Steele & Joshua Aronson, Stereotype Threat and the Intellectual Test Performance of African Americans, 69 J. PERSONALITY & Soc. Psychol. 797, 808 (1995); Connie Leslie, You Can’t High-Jump if the Bar is Set Low: A New Prescription to Help Black Kids Succeed, NEWSWEEK, Nov. 6, 1995, at 82; Ethan Watters, Claude Steele Has Scores to Settle, N.Y. TIMES, Sept. 17, 1995, § 6 (Magazine), at 45.

118 The fact that at the promotion stage firms have a considerable amount of information about employees, information on which they can rank these employees, is not inconsistent with the claim that these firms employ low monitoring strategies. The point is that even with strategies that are low monitoring at any point in time, the firms will inevitably collect a considerable amount of information on the employee if the employee remains at the firm for a sufficiently long period of time (e.g., at the partnership stage that comes eight or ten years down the road). A pertinent set of firm strategies are analyzed in Susan Sturm & Lani Guinier, The Future of Affirmative Action: Reclaiming the Innovative Ideal, 84 CAL. L. REV. 953 (1996). This article arrives at a similar recommendation to ours on affirmative action, and is based in part on a similar conception of firm structure. Their approach, however, rests on a general critique of current meritocratic standards, while we have attempted to analyze a specific economic structure under which current approaches to incentives and monitoring produce discrimination.
choosing between equals), but do not do so at promotion, when they have concrete information about years of project choice. In sum, an individual firm will not use affirmative action at the promotion stage because it will suffer an economic loss from doing so. Second, with this concrete information as to employees’ project choices, a firm using affirmative action in promotion would be susceptible to a reverse discrimination lawsuit.

There is, however, a third option: affirmative action in the period between hiring and promotion. This would mean structuring the workplace in such a manner so as to ensure that minority workers do not self-select into a disproportionate number of self-defeating shirking or high-risk strategies.

Given the current state of the law, it is unlikely that a firm could put in place a system that focused on minority workers alone.\textsuperscript{119} Instead, however, the firm could institute a system with periodic checks to ensure that no set of employees (minority or not) adopts self-defeating strategies.\textsuperscript{120}

Putting in place a program to ensure that no group of employees is systematically skewed strategies should work. If managers or preassigned mentors are held accountable for ensuring that their employees do not fall into such patterns of project choice, we might see a correction of the problem of skewed project choice. Legally, such a program should be acceptable to both judges and politicians, because it is not geared to providing any group an advantage but only to ensuring the proverbial “level playing field” for all.

Will firms do this? Not on their own. Even firms that are aware of the self-defeating strategies that minorities often pursue may lack the incentive to do anything about it. Firms see workers’ strategy choices as (a) not caused by the firm, and (b) a product of affirmative action for these employees usually at earlier stages. Put differently, firms are likely to see no legal or moral obligation to do anything because they do not consider themselves responsible for causing this situation in the first


\textsuperscript{120} Cf. Edley, supra note 3, at 185–87 (describing the promotion processes in the Army where promotion boards are instructed to give extra scrutiny to the files of minority and women candidates so as to screen them for the possibility of past discrimination).
place.\textsuperscript{121} And firms certainly see no economic advantage from restructuring institutions that are highly profitable as is.

The question then is how to induce firms to ensure that their racially neutral processes between hiring and promotion do not produce disproportionately negative results for minority workers. One way to do this would be to give firms incentives to examine their workplace structures and construct solutions to ensure that the process between hiring and promotion (when the tournament winners are, so to speak, decided) is fair to all. These incentives could be cash incentives that firms would receive from the government if they could demonstrate that they had made progress in ensuring a fair path between hiring and promotion. Firms that availed themselves of this option would get the governmental reward and receive the positive publicity associated with a fair working environment. In effect, this would create a tournament among firms to ensure that they are fair, while still remaining competitive. Presumably, employees will migrate towards the firms that signal themselves as fairer, as evidenced by the government’s seal of approval. Further, the incentive system will give firms a positive reason to adjust the existing efficiency-wage/tournament systems to see if they can modify them to be equally efficient, but fairer.\textsuperscript{122}

An alternative would be a modified “bottom-line” defense, reviving for our special circumstances a version of the argument rejected by the Supreme Court in \textit{Connecticut v. Teal}.\textsuperscript{123} Under this modified defense, firms would defend against a disparate impact suit, not simply by pointing to their “bottom-line” employment figures, but rather by showing as well that the improved “bottom-line” figures are a product of a program designed to correct the distortion in minority employees’ professional choices. Under current law, this might not operate as a very powerful incentive to adopt such programs: the threat of disparate impact appears to have been sharply diminished in recent years, in part in response to restrictive Supreme Court decisions such as \textit{Wards Cove Packing Co. v. Atonio}.\textsuperscript{124} But if the standards for disparate impact suits were adjusted, as we suggest

\textsuperscript{121} We are obliged to Jerry Kang for pointing out that some firms may be motivated to alter their structures in response to moral arguments. However, even assuming the viability of moral arguments, we think these moral arguments are unlikely to work if firms see themselves bearing no moral responsibility for the creation of the problem in the first place.

\textsuperscript{122} At this stage we have not thought through the question of how one would evaluate firms for having achieved the restructuring of their work environments. At first glance, however, one option would be to have firms choose their method of providing this information. They could produce evidence as to numbers or they could explain what checks and balances they had in place to ensure that their workplace did not disadvantage minority workers.

\textsuperscript{123} 457 U.S. 440 (1982).

\textsuperscript{124} 490 U.S. 642 (1989).
below, to accommodate scrutiny of the types of discrimination that we describe here, the prospect of a defense to such suits might be a powerful inducement to employers.

C. Summary

We have developed a description of economically robust discrimination in elite firms with low monitoring in the services and managerial sectors. The main differences from extant theories of discrimination are these:

(1) The analysis emphasizes the role played by discretionary evaluation and monitoring difficulties in sustaining discriminatory equilibria;

(2) Our analysis shows how discrimination perpetuates itself, in part, by influencing the career strategies that workers adopt in response to discrimination;

(3) In particular, our argument explains the interactions between strategies that workers adopt and corresponding modes of evaluation (e.g., in responding to noisy signals);

(4) The analysis takes into account the effects of underemployment in creating a pool of workers that permits firms to discriminate without suffering economic loss.

IV. "High-Level" Discrimination in Current Legal Doctrine

Our analysis of the causes of discrimination in high-level jobs calls for fundamental changes in the law if it is to be effective in rooting out discriminatory inequalities in economic opportunity. Here, we begin to undertake the project of reconceptualization of the fundamental elements of discrimination law—the requirements for pleading and proof in disparate treatment and disparate impact cases. We argue that the current standards of discrimination law, in both disparate treatment and disparate impact suits, are ill-suited to scrutiny of the types of discrimination that we have identified.

A. Pleading of the Disparate Treatment Case

The structure of pleading in the prima facie case is of immense practical importance: as discrimination cases are largely decided on summary judgment, these rules effectively decide the fate of the vast majority of discrimination lawsuits. 

Fisher v. Vassar College\textsuperscript{125} addresses the question what inferences should be drawn from the plaintiff's satisfaction of the requirements for

\textsuperscript{125} 114 F.3d 1332 (2d Cir. 1997) (en banc).
the prima facie\textsuperscript{126} case as established by \textit{McDonnell Douglas Corp. v. Green}.\textsuperscript{127} In particular, the question is whether the evidence is sufficient to produce a "more likely than not" inference of discrimination. Either the four \textit{McDonnell Douglas} factors were meant to identify evidence that, in the Court's view, would \textit{always} justify a "more likely than not" inference that intentional discrimination was the reason for the adverse job action, \textit{or} the Court intended to create an artificial but judicially mandated presumption for the plaintiff, even though there would be cases in which evidence satisfying the four-factor test would not actually justify an inference of discrimination.\textsuperscript{128}

The solution to this question should logically lie in the Court's language in \textit{McDonnell Douglas} and \textit{Burdine}. Unfortunately, the Court's opinions in these cases use the terms "inference" and "presumption" interchangeably\textsuperscript{129}—resulting in the conflict between the dissents, majority, and concurrence in \textit{Fisher}. Since the language in the opinions provides no answer, the next step is to examine the contexts underlying \textit{McDonnell Douglas} and \textit{Burdine} and see if the four-factor test was a mere formality or really produced an accurate inference of discrimination. The answer should give us insight into the Court's thought process in those early cases.

Our analysis sheds light on this question because it shows the kinds of inferences that could reasonably have been drawn from the facts in responding to the four factors \textit{at the time} the Court set out this test, i.e., in \textit{McDonnell Douglas} in 1976. Our analysis of the economic structure of discrimination suggests that the four factors were actually meant to make out a prima facie case based on a logical inference, and not to mandate a presumption based on a bare-bones satisfaction of four factors.

\textit{McDonnell Douglas} and \textit{Burdine}, decided nearly two decades ago, were two of the first cases to set out a framework within which Title VII was to be applied.\textsuperscript{130} Apart from the requirement that the plaintiff belong to a racial minority, he or she had to be qualified, had to have applied for

\textsuperscript{126}See \textit{id}. at 1336.

\textsuperscript{127}411 U.S. 792 (1973).

\textsuperscript{128}According to the majority in \textit{Fisher}, a prima facie case of discrimination can be made out even where the four \textit{McDonnell Douglas} factors do not add up to evidence from which a reasonable factfinder could infer that the employment decision was more likely than not based on the consideration of impermissible factors. See \textit{Fisher v. Vassar College}, 114 F.3d 1332, 1336–37 (2d Cir. 1997) (en banc).

\textsuperscript{129}A "presumption" is a legally mandated conclusion drawn from certain facts—here, the establishment of a prima facie case. See Candace S. Kovacic-Fleischer, \textit{Proving Discrimination after Price Waterhouse and Wards Cove: Semantics as Substance}, 39 Am. U. L. Rev. 615, 625 (1990). An "inference," in contrast, is a conclusion that can be logically drawn from the evidence. See \textit{id}. at 627. An inference, unlike a rebuttable presumption, does not "burst" or disappear because the presumption is an artificial construct put into place to induce the other side to speak. See \textit{id}. at 627–28.

\textsuperscript{130}The events at issue in the cases took place in 1964 and 1973, respectively. See
the job and been rejected, and the job had to have been kept open despite
the plaintiff's availability.\textsuperscript{131} As a threshold matter, there was more overt
discriminatory animus motivating employment decisions then than there
is today.\textsuperscript{132} However, it is the McDonnell Douglas factors themselves that
shed more light on the issue. First, there is the question of being "qualified."
In a lower-level job (e.g., a factory position involving working on a
production line), being "qualified" is likely to mean being the best person
for the job, something that can be determined with some certainty.\textsuperscript{133}
Second, there is the factor of the job being open despite both (a) the need
for someone qualified and (b) the availability of such a person (albeit of
the wrong race or gender as far as the employer is concerned). These were
the crucial factors in McDonnell Douglas and Burdine,\textsuperscript{134} and it appears,
in the context of overt animus against certain groups at that time, that the
satisfaction of these factors could and did produce a "more likely than
not" inference that the reason for the failure to hire was intentional
discrimination.

In the context of discrimination in high-level jobs today, however, the
McDonnell Douglas factors do not necessarily warrant an inference that
discrimination has occurred, at least not the animus-based discrimination
at which the factors were aimed. Contemporary cases, as we have argued,
will often differ. In the context of a scarce, high-wage, high-level job with
largely subjective qualifications there will, by definition almost, be an
abundance of "qualified" applicants.\textsuperscript{135} The question then is whether the
fact that one of these qualified applicants is hired produces a prima facie
case for all the others who are members of protected classes.

The Court in McDonnell Douglas and Burdine used the four-factor
test because of the substantial inference of intentional discrimination \textit{at the time}—thereby justifying the imposition of a rebuttable presumption
on the defendant. Satisfaction of the four-factor test was therefore a useful
rule of thumb for lower courts in deciding whether a prima facie case had
been made out. It no longer plausibly serves that role.

\textsuperscript{131} See McDonnell Douglas, 411 U.S. at 802.
\textsuperscript{132} See generally, Thernstrom & Thernstrom, supra note 1.
\textsuperscript{133} To use an extreme example, if the job is carrying bricks, it is fairly easy to set the
qualification as being the person who can carry the most bricks.
\textsuperscript{134} In McDonnell Douglas, the plaintiff, Green, was a black mechanic and lab
Technician who, even though qualified, was fired at the time that the company was actively
seeking employees for his position. See McDonnell Douglas, 411 U.S. at 794–96. In
Burdine, the plaintiff was a female accounting clerk who the company refused to consider
for a promotion even though the job was vacant and there were no other available qualified
candidates. See Burdine, 450 U.S. at 250. Eventually, six months later, one of the
plaintiff's former subordinates was hired for the position. See id. at 254 n.6.
\textsuperscript{135} Cf. Edley, supra note 3, at 180.
Title VII was put in place to eradicate easy-to-observe, overt racism in lower-level jobs. Now we are trying to apply it to hard-to-observe, subtle discrimination in high-level jobs. In \textit{Fisher}, the majority judges appear to have attempted to make it easier for plaintiffs to cross the initial hurdles of summary judgment in a discrimination case. The ultimate question in these cases is whether the adverse employment action was taken because of intentional discrimination. As we have explained, this is near impossible to prove where the decision involves a large number of similarly qualified individuals and subjective qualifications (unless, of course, explicitly discriminatory comments have been made).

However, even though a large percentage of these cases are doomed to failure, plaintiffs still have an effective weapon so long as they can get past summary judgement and threaten a trial. The reason for this is that not only do most employers not want the expense of a trial, but they often are hesitant to reveal their internal decision-making processes. Employers fear that embarrassing information will be revealed about their decision-making processes that will undermine the incentive system they have in place. Therefore, if the plaintiff-employee gets past summary judgment, the employer is likely to settle.

Given that summary judgment is where the real battle is often fought, the quantity and type of evidence one needs to get past it becomes crucial in determining whether certain types of suits can be brought. If the prima facie case requirement is de minimis and creates a substantive inference of discrimination (under \textit{Fisher}), then one can get past summary judgment even though one has a case that is ultimately a loser.

This "liberal" strategy, whatever its superficial appeal as tough on discrimination, thus perversely imposes de facto liability in many cases where no inference of discriminatory conduct is warranted and may make it more expensive to hire minority employees.\footnote{Cf. Ayers & Siegelman, \textit{supra} note 38, at 1488–89 (claiming that employers prefer employees who can be easily fired).} Further, the model shows that, if applied improperly (i.e., effectively resulting in protection for minority workers such that they may be fired only "for cause"), anti-discrimination laws can skew the strategies of minority workers towards shirking and help sustain a discriminatory equilibrium.

\textbf{B. Standards of Proof in Disparate Treatment Cases}

As we have shown, the current pleading standard for disparate treatment cases may lead to an inference of discrimination imposed on employers who have engaged in no discriminatory conduct—at least as discrimination is defined by current doctrine. When one turns to the standards of proof applied after trial, the opposite problem emerges:
current doctrine is unlikely to impose liability after trial for the types of discrimination that we have described.\textsuperscript{137} In effect, the doctrine too loosely lets the plaintiffs surmount the prima facie hurdle, only to have some of these plaintiffs lose on the merits even when their claims should be sustained.\textsuperscript{138}

In disparate treatment litigation,\textsuperscript{139} after the plaintiff meets the de minimis prima facie case, the defendant articulates a nondiscriminatory reason for its employment decision. The plaintiff wins the lawsuit only by introducing a preponderance of evidence to show that the actual reason for the employment decision was discriminatory animus. This approach is clearly unsuited to rooting out the type of discrimination that we have identified. In many instances, the employment decisions described by our model, though they have a discriminatory impact in aggregate, are justified under the disparate treatment standard. They are not based on discriminatory animus, but rather on a rational, albeit imperfect,\textsuperscript{140} evaluation of job prospects or job performance.

Even in those instances in which the hiring or firing decision was motivated by discriminatory animus, a victory in court will be almost impossible. Courts generally require, in the absence of a "smoking gun," that plaintiffs prove that they were better qualified than comparable white applicants or promoted individuals.\textsuperscript{141} In our analysis, the minority candidates are no better qualified; the collectively discriminatory impact comes from individual decisions among equally qualified applicants.

Presumably, a statistical analysis of hiring patterns under either a disparate treatment or disparate impact theory should detect the discrimination. After all, our analysis predicts a clearly identifiable statistical effect: disproportionately large hiring of whites from the pools of applicants, and disproportionate promotion of whites from the pool of employees eligible for promotion. In fact, these patterns are observable, in ag-

\textsuperscript{137} Of course, this anomalous contrast would not arise if the summary judgment standards, which we have already analyzed, accurately predicted the outcomes of trials. Because of the way McDonnell Douglas presumptions are now applied, however, they do not do so. If the presumptions at the summary judgment stage were revised to reflect the actual course of subsequent pleading and proof, the threat of excessive pretrial liability would disappear. However, the problem of underinclusiveness (no liability for discriminatory conduct) would remain.

\textsuperscript{138} Note that it will often be rational for plaintiffs to bring these suits, even if they correctly anticipate that they might lose on the merits if the case were not settled. See, e.g., Lucian Arye Bebchuk, A New Theory Concerning the Credibility and Success of Threats to Sue, 25 J. LEGAL STUD. 1 (1996).


\textsuperscript{140} Title VII disparate treatment analysis does not require that employment decisions be correct by some absolute standard, or even that they be rational—only that they be based on a permissible (nondiscriminatory) ground.

aggregate, for firms in a number of sectors of the economy. However, even if the same statistical showing could be made for individual firms, this showing would not suffice to establish liability. The additional proof requirements would be difficult to meet for the discriminatory effects predicted by our analysis. For a pattern-and-practice disparate treatment claim, defendants would excuse their hiring and promotion practices with a neutral explanation. Plaintiffs would then be required to supplement statistical proof with further evidence of discriminatory intent. For the job decisions that we have analyzed, plaintiffs cannot do so because employers have no intent to discriminate in the standard sense of the term.

An emergent approach to high-level discrimination cases raises the standard of proof for plaintiffs. In effect, plaintiffs must show "clear and convincing" proof of discrimination. The articulated rationale is that courts should not second-guess difficult and expertise-laden personnel judgments, such as law or accounting firm partnership decisions and university tenure decisions.

Our analysis suggests a rationale for this approach, distinct from the rationales articulated by courts. It may be implausible to think that judges are less competent to review law firm (or law school) hiring decisions than they are to review the hiring criteria used by employers of haberdashers or stevedores. The higher standard of proof might be justified, nonetheless, because it discourages suits by average (as contrasted to "superstar") plaintiffs. Under a preponderance-of-evidence standard, average workers had a better chance of winning, even though most of these lawsuits would be lost at trial. Raising the standard of proof to require

\[142\] See Hacker, supra note 2, at 60–85.

\[143\] It may be harder to make this showing for firms than for the whole sector because plaintiffs may not find data about enough decision points to produce statistically significant results of discrimination for an individual firm. See generally, D. H. Kaye, Is Proof of Statistical Significance Relevant?, 61 Wash. L. Rev. 1333 (1986) (examining and rejecting hypothesis testing in favor of other proposed methods); Ramona L. Paetzold, Problems With Statistical Significance in Employment Discrimination Litigation, 26 New Eng. L. Rev. 395 (1991) (criticizing hypothesis testing as a statistical tool in employment discrimination litigation).


\[145\] See, e.g., Wallace v. SMC Pneumatics, Inc., 103 F.3d 1394, 1400 (7th Cir. 1997) (case of a high-level manager, "unsuccessful managers, like unsuccessful generals, are quite likely to be held strictly liable for their failures"); Ezold v. Wolf, Block, Schorr and Solis-Cohen, 983 F.2d 509 (3d Cir. 1992) (law firm partnership decision); Jackson v. Harvard Univ., 900 F.2d 464 (1st Cir. 1990) (tenure decision). Similar deference to partnership decisions is shown in Wheeler v. Hurdman, 825 F.2d 257 (10th Cir. 1987), and Lucido v. Cravath, Swaine and Moore, 425 F. Supp. 123 (S.D.N.Y. 1977).

\[146\] See, e.g., Mesnick v. General Elec. Co., 950 F.2d 816, 825 (1st Cir. 1991) (describing the court's reluctance to second-guess a firm's decision where discrimination was not clearly evidenced).

\[147\] Note that there are not a large number of suits at the hiring stage because the stakes are low relative to the costs of suit. Cf. John J. Donohue III & Peter Siegelman, The
“clear and convincing evidence” discourages suits by average workers, who will almost certainly lose under a “clear and convincing” standard. Only “superstars” will bring suit under the latter standard.\footnote{Though perhaps perverse from a broader policy perspective, this disincentive is at least a sensible way of economizing limited legal resources; it filters out a large number of claims that would fail under current doctrine.} The real problem, as we have seen, is that the underlying doctrine is flawed.

C. Disparate Impact Doctrine

Because disparate impact doctrine imposes a broad requirement of reasonableness on all employment practices that have statistically demonstrated discriminatory impacts, it might appear better suited than disparate treatment doctrine to legal scrutiny of the hiring and promotion structures we have described.\footnote{However, at least three doctrinal barriers stand in the way. First, it is arguable that the employment practice (imperfect monitoring) does not cause the discriminatory impact, at least in the way “cause” is understood under current doctrine.} Our argument describes

\textit{Changing Nature of Employment Discrimination Litigation}, 43 STAN. L. REV. 983, 984 (1991) (explaining that the “vast majority” of current suits challenge firing decisions). However, once one gets to the stage of promotion to partnership or tenure, the stakes increase substantially.

Congress, through the Civil Rights Act of 1991, has provided for added incentives, in terms of attorneys fees and damages, that make it more attractive for plaintiffs' counsel to take on Title VII cases. See Turner, supra note 2, at 464–65. As far as we can tell, however, there has been no evidence that these added incentives have reduced the imbalance towards a disproportionate number of promotion discrimination cases over hiring discrimination cases.

\footnote{To state this point more formally, raising the standard of proof screens cases by reducing the number of suits that appear to be positive present value ex ante but that will actually end in judgments for defendants.}

\footnote{This is not to say that the substantive standard of liability should be formulated so that these claims should lose. The point is that most of them would not meet the substantive requirements of current law. The chance of winning is apparently great enough, however, to entice many of them to court—unless the special “clear and convincing” standard of evidentiary proof is imposed.}

\footnote{Our critique of pleading and proof requirements in disparate treatment is congruent with the important analysis of the prima facie case in Deborah C. Malamud, The Last Minuet: Disparate Treatment After Hicks, 93 MICH. L. REV. 2229 (1995). While Malamud powerfully argues that the prima facie case is generally too rigid to accommodate the inferential complexities of discriminatory intent, our argument has focused on describing a particular pattern of discriminatory conduct which disparate treatment doctrine is ill-suited to correct.}

\footnote{The requirement of causation is emphasized in \textit{Wards Cove Packing Co. v. Atonio}, 490 U.S. 642 (1989) and codified with slight modifications by the Civil Rights Act of 1991, § 703, 42 U.S.C. § 2000e-2(k)(1)(B) (1996). The codification exempts plaintiffs from the requirement of identifying a particular employment practice that causes the disparate impact when elements of an employer's decision-making process are not capable of "separation for analysis." See id. This proviso probably would not help plaintiffs out of the difficulty described in the text. Their problem is not that elements of the process
a causal chain in which an apparently neutral system of worker evaluation induces discriminatory outcomes through the system's effects on rationally chosen effort levels. Courts have suggested that in discrimination suits, much like the general tort context where joint conduct by two actors causes harm, a conscious, economically rational decision by the alleged victim may break the chain of proximate causation.\(^{152}\) For example, employers are not liable for discrimination when unequal hiring results from the decision of members of the group in question not to apply for the job.\(^{153}\) To be sure, the employer remains liable if its actions directly cause the disproportion in applicants, such as by harassing applicants from a minority group. However, in our model, the employer does not take such actions; it does not harass applicants or adopt a directly discriminatory approach to hiring. Once a discriminatory equilibrium is established among firms as a group, each individual firm continues to get discriminatory outcomes even though its own hiring criteria are entirely neutral.

Second, even if the use of such apparently neutral hiring criteria were condemned as discriminatory, firms could readily defend their actions by invoking a routinely accepted business justification: the firm is making its hiring decisions based on workers' qualifications.\(^{154}\) The practices that we describe are clearly distinguishable from the classic Griggs v. Duke Power Co.\(^{155}\) situation of job qualifications that are not "job-related." Here, the qualifications imposed are job-related but are differently satisfied by typical and minority applicants due to the equilibrium established in efficiency-wage/tournament markets by past discriminatory practices.

A third doctrinal barrier is the lack of interest defense. As Vicki Schultz's important empirical work has shown, the lack of interest defense has grown increasingly important in employment discrimination cases, particularly in service industries.\(^{156}\) Our analysis helps explain the power of this argument in explaining hiring discrimination. Under the lack of interest defense, the employer shows that the reason for the low percentage of minorities in the hiring pool is the low percentage in the workforce.

---

\(^{152}\) In tort, the same analytic outcome is sometimes achieved by concepts such as waiver or assumption of risk that have no direct application in anti-discrimination law.

\(^{153}\) See EEOC v. Chicago Lamp Miniature Lamp Works, 947 F.2d 292 (7th Cir. 1991).

\(^{154}\) See Woody v. St. Clair County Comm'n, 885 F.2d 1557 (11th Cir. 1989); Lucas v. Burnley, 879 F.2d 1240, 1244 (4th Cir. 1989).

\(^{155}\) 401 U.S. 424, 431 (1971) ("[N]either the high school completion requirement nor the general intelligence test is shown to bear a demonstrable relationship to successful performance of the jobs for which it was used.").

Our model predicts that once minorities learn that their prospects for getting the job are dim, they will not make the investment needed to qualify for employment. Of course, when the dearth of minority applicants is due to hiring discrimination, the lack of interest defense does not defeat liability under Title VII. However, as we explained above, the type of discrimination that we describe is very difficult to show to be illegal and so cannot be used by plaintiffs to prevent the defendant-employers from offering a lack of interest defense.

D. Doctrinal Alternatives

Although we leave the details of possible doctrinal reform to future work, our argument has permitted us to sketch out the main consequences of our model for doctrinal analysis. Even if discrimination law proves effective for overt forms of discrimination, the law will fail to root out the type of discrimination that our model has described. As we suggested in Part III.B, supra, the focus must be instead on developing mechanisms that specifically address discrimination in high level jobs by adjusting incentives and professional pathways. The solution must come in the form of an affirmative action incentive program for employers, which induces them to identify and rectify distorted incentives for minority workers.

The focus on mechanisms implemented with firms does not mean, however, that the law has no role to play. The threat of legal liability might provide a powerful inducement for firms to alter their personnel policies. Our doctrinal analysis suggests that the simplest way of exerting such pressure would be to adjust disparate impact doctrine to increase the probability that firms would be held liable for the structures that we have described. (This might involve, for example, eliminating barriers to these claims, such as the lack of interest defense.) Then, a modified "bottom-line" defense, as described in section III.B, would provide a legal incentive to develop corrective programs.

E. Conclusion

The mainstream understanding of anti-discrimination law relies on rooting out readily identifiable discriminatory motives and practices as a crucial means of insuring economic equity and economic progress for traditionally disadvantaged minority groups. In our view, however, this moderate approach leaves untouched, perhaps indefinitely, firms' employment policies which act to perpetuate the disadvantage of minority groups.

Our analysis, in turn, points to three alternative solutions. The deregulatory alternative would dismantle discrimination law and return to reliance on market mechanisms to provide for equal job opportunity. Our analysis suggests that such a dismantling would, ironically, accelerate the
elimination of discrimination since current anti-discrimination law impedes the market forces that would be needed to eliminate discrimination without government intervention. Market forces would remedy the discrimination we describe if firms could employ minority workers at lower wages than “typical” workers doing the same jobs. If wages paid by each firm were inflexible, however, then firms that employed minority applicants probably would hire only minority applicants. This result would be illegal, and, even if legal, socially objectionable. In particular, a society in which these practices were commonplace would be quite objectionable to many “moderates”; they prefer to tolerate the more covert, though more destructive and ineradicable, discrimination now accomplished in efficiency-wage markets.

A second solution to the problems we identify, one of a very different tenor than the first, is widespread imposition of hiring quotas or targets. Firms might simply be required to bring their proportion of minority workers up to the proportion in the relevant population. They would then either invest in additional training and monitoring for these workers, or, if that were not cost-justified, simply tolerate inferior performance from these workers. In either case, the legal policymaker would hope that the availability of job opportunities would gradually eliminate the skewed education and effort strategies currently pursued by discriminated against groups. However, even putting aside the various moral and political objections to group-conscious employment decisions, our analysis raises substantial doubts about the usefulness of affirmative action programs in reducing discrimination. This alternative might make it even more difficult for minority workers to signal the high quality of their work.

From the viewpoint of our economic analysis, then, the hard question about these two polar alternatives is how quickly they would work to reduce discrimination. While both approaches play important if illicit roles in our current employment law system, they are nonetheless viewed, as a rhetorical and cultural matter, with extreme disfavor. Discrimination

---

157 Firms might create separate tracks that would be filled mostly by minority workers, such as “permanent associates” positions. This statistical pattern likely would subject the firm to liability for discrimination. The general point is that any solution to the current discrimination problem that is economically rational for the firm will involve disparate (and relatively unfavorable) treatment for minorities during the transition period. In observing that discrimination law may impede progress towards racial equality, our views approach those of Epstein and Posner, although the logic of our argument is very different from theirs.

158 For example, Jews first overcame the barriers to hiring by the elite law firms by starting their own firms, with lawyers who were almost all Jewish. For a description of the Jewish experience, see Jerold S. Auerbach, Unequal Justice: Lawyers and Social Change in Modern America 50, 62 (1976); Alan M. Dershowitz, Chutzpah (1991) (describing Dershowitz’s personal experience with discrimination against Jews at the nation’s elite law firms); Note, The Jewish Law Student and New York Jobs: Discriminatory Effects in Law Firm Hiring Practices, 73 Yale L.J. 625 (1964).
law prefers to attempt to fit its regulatory tools within a "liberal" model, in which the act of discrimination is a wrong that inflicts a harm on an individual victim. The irony is apparent. From an economic perspective, the "liberal" approach, while condemning individual acts of discrimination, may serve to perpetuate deprivation of employment opportunities to members of minority groups.

We urge a third alternative—a new route out of the current impasse. In our conception, the problems of persistent discrimination in high-level jobs emerge from two related sets of factors: (i) discrimination is subtle, exercised in choices made between equals, and hard to detect; and (ii) legitimate perceptions of discrimination by minority group members result in large numbers of minorities choosing self-defeating strategies. Because the discrimination is both hard to detect and largely a function of strategy choices by minority group members themselves, a penalty mechanism that depends on the ability of a regulatory body (e.g., the EEOC or a court) to accurately detect some percentage of violators cannot work.

Instead we have suggested an employer-initiated restructuring of the mechanisms that foster this discrimination. Firms would be given incentives or rewards to identify themselves as having studied and restructured their workplaces to eliminate the problems of subtle biases in choices between equals, and the effects of certain groups choosing self-defeating strategies. Such a program should be legally viable (race- or gender-neutral) as long as it is framed in terms of ensuring an equal opportunity workplace; incentives for adopting such programs could come from both economic advantages and from the pressures exerted by a reconstructed employment discrimination doctrine.

Our proposal is tentative, but we hope to provide a starting point for a newer and more informed debate about discrimination and possible solutions to it. More importantly, however, our goal in this paper is not to suggest that we have the magic solution—we are skeptical that such a solution exists for the problems of discrimination in high-level jobs. Instead, we aim to alter how legal analysts think about the sources of discrimination and the impacts of discrimination policy. We urge that analysis must focus more carefully, in a microanalytic way, on the particular organizational structures of firms. Tournament and/or efficiency-wage compensation patterns, promotional pathways, interfirm patterns such as up-or-out promotion, and special statuses such as partnership or tenure are factors deserving of particular attention. Without such a focus, these complex practices, otherwise economically beneficent, will continue to stand in the way of full equality of opportunity for American workers.

159 See supra Part III.B.