ISSN 1045-6333

EFFICIENCY WAGES, TOURNAMENTS, AND DISCRIMINATION: A THEORY OF EMPLOYMENT DISCRIMINATION LAW FOR "HIGH LEVEL" JOBS

David Charny*
Mitu Gulati**

Discussion Paper No. 182

3/96

Harvard Law School
Cambridge, MA 02138

The Center for Law, Economics, and Business is supported by a grant from the John M. Olin Foundation.

*Professor of Law, Harvard Law School.
**Law Clerk for Judge Sandra L. Lynch, 1st Circuit Court of Appeals.

Efficiency Wages, Tournaments and Discrimination: A Theory of Employment Discrimination Law for "High-Level" Jobs[1]

David Charny[2] and G. Mitu Gulati[3]

Despite a generation of experience and intensive debate, the proper contours of employment discrimination law -- and related policies of affirmative action -- remain profoundly controversial. Just as the simpler forms of overt *de jure* discrimination have given way to concerns with more subtle forms of discrimination--"glass ceiling" exclusion, discrimination in layoffs, promotions and benefits, conditions of work, and sexual harrassment -- so the search for rationales has turned to more complex models of the workplace.

Although economists and sociologists continue to dispute the role of discrimination law in contributing to blacks' economic progress in the 1960s and 1970s,[4] there is an emerging consensus that the "success story" of that period --whatever its causes -- is unlikely to be replicated in the next decade.[5] In particular, the circumstances that might have enabled the

[1] We are grateful for the comments of Yochai Benkler, Einer Elhauge, Howell Jackson, Louis Kaplow, Sandra L. Lynch, Steve Shavell, and participants in the Seminar in Law and Economics at Harvard Law School. We owe a special debt to David Wilkins for his incisive suggestions and his encouragement throughout our work on this project. Research support was provided to David Charny by the Olin Foundation Law and Economics Reserch Program and to Mitu Gulati as an Olin Fellow at the Harvard Law School; we thank the Olin Foundation for its generosity.

[2] Professor of Law, Harvard Law School.

[3] Law Clerk for Judge Sandra L. Lynch, 1st Circuit Court of Appeals.

[4] Compare James J. Heckman and Brook S. Payner, Determining the Impact of Federal Antidiscrimination Policy on the Economic Status of Blacks: A Study of South Carolina, 79 **Am. Econ. Rev.** 139 (1989) and John J. Donohue III and James J. Heckman, Continuous versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks, 29 **J. Econ. Lit.** 1603 (1991) with William M. Landes, The Economics of Fair Employment Laws, 76 **J. Pol. Econ.** 507, 544-45 (1968), and **Richard A. Epstein, Forbidden Grounds, The Case Against Employment Discrimination Laws** (1992).

[5] See, e.g., John J. Donohue III and James J. Heckman, Re-Evaluating Federal Civil Right Policy, 79 **Geo. L. J.** 1713, 1729-30 (1991); Andrew Hacker, **Two Nations: Black and White, Separate, Hostile, Unequal** 5-25 (1991).

law to be effective during that period are unlikely to recur.

Several factors are of particular importance. First, the overt, acknowledged use of discriminatory criteria -- often based on "Jim Crow" laws -- has been dismantled; and the more overt "neutral" barriers to minorities' employment -- the battery of apparently discriminatory "objective" performance tests -- have been removed either as invalid or too costly to validate in court. Further progress in eliminating discrimination must aim at forms of discrimination that are more subtle, and therefore more difficult to eradicate.[6] Second, the manufacturing sector, in which the greatest progress was made through quotas and training programs, is shrinking, and so offers diminished opportunities to young minority workers entering the labor market for the first time. 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 -- the professional and technical jobs that form the desirable core of the new "information economy."[7] In particular, the legal system cannot effectively second-guess the validity of managers' personnel decisions in jobs where job performance is evaluated on "subjective" -- i.e., context-based -- performance criteria.[8]

For these reasons, it appears that the central challenge to the further "success" of the anti-discrimination norm is the difficulty of applying its legal conceptions to jobs where performance is difficult to monitor. Our argument analyzes this problems. The central claim is that that discrimination is likely to be self-perpetuating in these jobs -- even in the absence of overt bias against minority groups -- and that current law is likely to be ineffective in eradicating this self-perpetuating discrimination.

The core of the argument is an analysis of how discrimination

---

[6] See e.g. David Strauss, The Law and Economics of Racial Discrimination in Employment, 79 Geo. L. J. 1619 (1991) (arguing on this basis for the use of numerical solutions such as quotas).

[7] See **Good for Business: Making Full Use of the Nation's Capital** (Fact Finding Report of the Federal Glass Ceiling Comission, Washington D.C.) 16-17(1995)(hereinafter "Glass Ceiling Report")(studying the effect of the "glass ceiling" on disadvantaged groups in the context of the growth of the service sector and the shift in employment from manufacturing to service industries); **Martin Conroy, Faded Dreams: The Politics and Economics of Race in America** 198 (1994)(discussing economic gains for blacks in the 1960s and 1970s that have not been replicated in the 1980s and are unlikely to be in the 1990s because in part of the shift in the economy from manufacturing to higher education and skill requiring jobs).

[8] The problem was first identified and extensively analyzed, in the legal literature, in Elizabeth Bartholet, Application of Title VII to Jobs in High Places, 95 Harv. L. Rev. 945 (1982); see also Book Review, J. Hoult Verkerke, Free to Lose, 105 Harv. L. Rev. 2080 (1992); Barbara J. Flagg, Fashioning a Title VII Remedy for Transparently White Subjective Decisionmaking, 104 Yale L. J. 2009 (1995); Ramona L. Paetzold and Rafael Gelu, Through the Looking Glass: Can Title VII help Women and Minorities Shatter the Glass Ceiling?, 31 Hous. L. Rev. 1517 (1995).

works to secure job opportunities for minority job applicants. Previous economic analysis of discrimination law have depended upon two central analytic features: [1] a simple market clearing model for labor markets, in which wage equals marginal product, and wages fall until the market clear; and [2] a homogeneous view of the firm, in whch information about the qualities and efforts of workers are perfect and in which hiring and firing are costless.[9] These assumptions, although permitting legal analysts to reach incisive conclusions on the basis of relatively simple and rigorous economic models,[10] are inconsistent not only with ordinary observation of the workplace and of labor markets, but also with the generally received view among economists of how labor markets work. It is now a commonplace among economists that this simple model is seriously at odds with how many sectors of the labor market operate.[11]

For these sectors, labor economists have developed a variety of "efficiency-wage" and "tournament" models. 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 firms can draw. Second, once workers are hired, because it is costly to monitor them, or to fire them, firms must adopt complex hierarchical structures in order to provide incentives for enhanced worker effort.[12] As we shall argue, these twin 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 homogeneous labor conditions across firms.

As a methodological matter, the key claim of this article is for the importance of the organizational structure of firms (and the impact of this structure on the labor markets in which these firms hire) as key features in analyzing the causes of employment discrimination and the effects of anti-

---

[9] See, e.g., Epstein, supra note 4, 31-47; John J. Donohue III, Is Title VII Efficient?, 134 U. Pa. L. Rev. 1411 (1986); Richard A. Posner, The Efficiency and Efficacy of Title VII, 136 U. Pa. L. Rev. 517 (1987); Richard A. Posner, **The Economics of Justice** 351-363 (1981). An admirable exception is the use of efficiency wage theory in the recent, path-breaking article, Michael Selmi, Testing for Equality: Merit, Efficiency, and the Affirmative Action Debate, 42 U.C.L.A. L. Rev. 1251, 1301-09 (1995). However, Selmi applies efficiency-wage theory to analyze affirmative action, rathern than to provide a model for the causes of discrimination; on the affirmative action issue, our analysis departs substantially from Selmi's.

[10] 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 regulation. Compare, e.g., Epstein, supra note 4, and Posner, supra note 9, with Donohue, supra note 9.

[11] See e.g. Jeremiah Cotton, Labor Markets and Racial Inequality, in **Labor Economics: Problems in Analyzing Labor Markets** 183-209 (William Darity, Jr., ed. 1993) (discussing the inability of a number of standard economic models to explain persistent labor market inequality).

[12] See part III infra (discussing efficiency-wage and tournament models).

discrimination law.[13] Instead of resting on broad generalities or abstractions about the features of the employment markets and the policies of employers, analysis should extend to a more refined understanding of how work structures -- for example, 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 potential discrimination. We present a model of discriminatory conduct that for substantial sectors of the workplace - particularly, high-skill or knowledge-intensive jobs in which the discretionary, skill-based component of work creates substantial difficulty in monitoring individual workers' performance. The analysis that we develop applies, for example, to law firms, investment and commercial banks, brokerage houses, management and engineering consulting firms, accounting firms, universities and managerial and social service bureacracies in government and large corporations.

Our argument proceeds in four parts. Part I analyzes the current understanding of economic impacts of employment discrimination law. Our contention here is that the diverse extant approaches to discrimination law share a common feature: they all maintain that either market competition or a direct legal ban on discrimination are in theory sufficient to eliminate discriminatory employment outcomes. Part II analyzes the impact of employment discrimination law using an efficiency-wage--tournament model. Here we show that, with this more sophisticated understanding of employment decisions, we should expect to find discriminatory employment outcomes which are robust against market competition. Part III explores empirical predictions that our model might make about the past and future impact of discrimination law. In particular, we demonstrate that the discrimination that we describe is not reached by current law -- i.e., that it cannot be eliminated merely by a direct ban on discrimination. Parts II and III thus establish that the current understanding ignores an important cause of discriminatory employment decisionmaking. Part IV draws conclusions for policy and for further research.

I. The current understanding of the role of discrimination law: a concise restatement

---

[13] See e.g. Elizabeth Chambliss, **New Partners with Power? Organizational Determinants of Law Firm Integration** 203-10 (1992) (doctoral dissertation at the University of Wisconsin on file with authors). Chambliss' empirical study of "elite" law firms that 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 7, at 21 (articulating the need for research into the organizational structure of firms in order to properly understand the "glass ceiling" effect).

What causes discrimination in employment decisions? To provide a basis for our own analysis, we identify three central strands of the current understanding.

1. **"Taste" for discrimination**. -- The cornerstone of the modern discussion has been the model developed by Gary Becker.[14] In Becker's analysis, employers have an exogenously given taste for discrimination: that the worker belongs to a group that the employer dislikes is stipulated to add a cost to employing that worker, beyond the money wage. In effect, the employer is not willing to pay the same money wage for the work of the disliked group (e.g., women or blacks), as for the work of the typical worker.[15]

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; in turn, the nondiscriminatory firm can produce at lower cost, and so drive discriminating firms out of the market. If the discriminators will be eventually driven out of the market, then anti-discrimination legislation is unnecessary. Indeed, critics argue on the basis of Becker's model that the anti-discrimination legislation distorts the market away from a non-discriminatory equilibrium.[16] The distortions arise because the fear of being sued under the anti-discrimination legislation increases the costs that

---

[14] Gary S. Becker, **The Economics of Discrimination** (2d ed. 1971). For a brief summary of Becker's theory, see Richard A. Posner, **Economic Analysis of Law**, Section 27.1 (3d ed. 1986).

[15] For example, if the employer has a distaste for women, and a woman applicant and a male applicant are equally skilled, then the employer will be willing to pay the woman a wage lower than the man, for the same job, since the woman imposes the additional cost on the employer of causing 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 hence in the short run, the woman will find herself unemployed, while the man will be employed (unless the woman is willing to work for $9,500).

Critics of the conception of "taste" have emphasized --correctly, we believe -- that some economists have relied upon an over-simplified 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 status, see Richard H. McAdams, Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination, 108 **Harv. L. Rev.** 1005, 1036-63 (1995), or the tendency of dominant groups unconsciously to bias its evaluations of outsiders, e.g., Deborah 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. For the economic analysis, the crucial point is that some (but not all) firms act --for whatever conscious or unconscious reason -- as if it were more costly to them to hire minority workers. As discussed in text, these firms will be at a competitive disadvantage and eventually be driven from the market.

[16] See Richard A. Posner, An Economic Analysis of Sex Discrimination Laws, 56 U. Chi. L. Rev. 1311 (1989) and Epstein, supra note 4.

an employer incurs when he hires a minority worker. For example, if the expected cost of a discrimination lawsuit is $500, per woman worker hired, the cost of hiring a woman becomes $500 higher that the money wage that she is paid.[17]) 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: particularly, 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.[18]

2. Statistical (rational) discrimination. – These models[19] begin with the assumption that firms (rationally) expect job applicants from certain groups to be less productive than typical 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.[20]

---

[17] 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 non-discriminatory employer will in the presence of anti-discrimination laws, face a non negligible cost in the form of a possible discrimination lawsuit.

[18] A definitive analysis following this approach is John J. Donohue, Is Title VII Efficient? 134 U. Pa. L. Rev. 1411 (1986); For a discussion of the relation between human capital and gender discussion, 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).

[19] The first important paper was Edmund Phelps, The Statistical Theory of Racism and Sexism, 62 Am. Econ. Rev. Papers and Proceedings 287, 292-294 (1973); see also Kenneth J. Arrow, Models of Job Discrimination, and Some Mathematical Models of Race Discrimination in the Labor Market, in Racial Discrimination in Economic Life, 83-102, 187-204 (Anthony H. Pascal ed. 1972)and Kenneth J. Arrow, The Theory of Discrimination, in Discrimination in Labor Markets, 3-33 (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).

[20] For 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 $1000 for being an engineer, and by another $1000 for her being from MIT. Ris-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 as a factor of production. Dennis J. Aigner and Glenn C. Cain, Statistical Theories of Discrimination in Labor Markets, 30 Indus. & Lab. Rel. Rev. 175, 175-187 (1977). However the

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 judgments about groups might correct themselves over time.[21] Further, even if these expectations are correct in the short run, an entrepreneur can profit in the long run by correcting these differences and producing an effective workforce at lower cost than other employers. For example, enterprising firms might find it profitble to make the additional investment needed to select the most productive members of less productive groups, or to train potentially productive members from these groups.[22]

In other circumstances, however, statistical discrimination is indefinitely self-perpetuating. First, statistical judgments of employers may be "robust" because they accurately reflect exogenously (socially) determined differences in the decisions of members of various groups to acquire job skills or to commit themselves to the workforce.[23] Second, statistical stereotypes may act as self-fulfilling prophecies: discrimination reduces incentives for its victims to acquire human capital, such as work skills or reputation; because the individual is judged on the basis of her membership in a group rather than her individual merit, the incentive for

---

assumption of risk averse firms is not very believable in a world of nearly perfect capital markets with investors who hold diversified portfolios.

[21] For example, during periods of labor shortage, firms will be forced to hire from the discriminated against group, and then will discover that its assumptions about this group have been false. Expectations will be revised.

[22] See generally **William Darity Jr., What's Left of the Economic Theory of Discrimination?**, in **The Question of Discrimination** 335, 335-45 (Steven Shulman and William Darity, Jr., eds. 1989); **Thomas Sowell, Ethnic America** (1981); Barry Chiswick, The Earnings and Human Capital of American Jews, 18 **J. Hum. Res.** 313, 334 (1983).

[23] Richard A. Posner, An Economic Analysis of Sex Discrimination Laws, 56 **U. Chi. L. Rev.** 1311 (1989), Gary S. Becker, Human Capital, Effort and the Sexual Division of Labor, 3 **J. Lab. Econ.** S33 (1985), and Solomon W. Polachek, Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure, 63 **Rev. Econ. & Stat.** 60 (1981), apply human capital analysis to the issue of the sexual division of labor and male female wage differentials. Becker's insight here was to treat the household as a unit of production (as opposed to the individual) where different members specialised 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, and so earn less at their market jobs. Women will self-select into low effort market jobs and because of the high opportunity costs they face (i.e., forfeiting the opportunity to raise children and do other househld activities that we exogenously assume they are better at than men) will quit jobs earlier than men. For discussions of the comparative quit rates of women and men, see Jeremy I. Bulow and Lawrence H. Summers, A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment, 4 **J. Lab. Econ.** 376 (1986); Francis D. Blau and Lawrence M. Kahn, Race and Sex Differences in Quits by Young Workers, 34 **Indus. & Lab. Rel. Rev.** 563 (1981); Samuel Issacharoff & Elyse Rosenblum, Women and the Workplace: Accomodating the Demands of Pregnancy, 94 **Colum. L. Rev.** 2154, 2159-79 (1994).

self-improvement is destroyed.[24]

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.[25] That is, only the firm currently employing the workers knows about their productivity. As outside firms do not know about the worker' productivity, they cannot make competing wage offers and so the market mechanism is short-circuited. This phenomenon depends on two conditions. First, the members of the minority group have a harder time signalling ability, and so their employers can conceal their abilities from other employers more easily (consider the situation where traditional publicly available signals, such as schooling and test scores, are regarded to be less reliable for disadvantaged workers). Secondly, certain high level jobs signal visibility to the general marketplace better than others (CEOs of a company and their performance are more visible to the external observer than janitors) The 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 these minority workers (since the other firms find it more difficult to observe the minority workers and their abilities)

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 is not self-correcting. The strategy used by the 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, then, the social costs of discrimination are two-fold: job assignments are inefficient, and the stock of human capital is too small. In particular, the model explains why members of the "invisible" groups might choose professions where visibility upon training is the highest (e.g. sports, certain high technology fields like engineering and nursing where skill and effort levels are easy to monitor).

3. **Workplace cultures, worker preferences, 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 (i.e., in accord with productivity as it would be

[24] Shelly Lundberg and Richard Startz, Private Discrimination and Social Intervention in Competitive Labor Markets, 73 **Amer. Econ. Rev.** 340, 340-347 (1983).

[25] Paul Milgrom and Sharon Oster, Job Discrimination, Market Forces, and the Invisibility Hypothesis, 102 **Q. J. Econ.** 453, 453-476 (1987).

perceived by a fully informed manager interested solely in maximizing returns to the firm). A third approach to persistent discrimination identifies reasons for which the firm might find it more costly to hire an individual minority worker -- i.e., reasons for which the minority worker might be less productive (net of labor costs) than a "typical" worker, at least if the firm's labor force is composed predominantly of typical workers. Costs to integrating the workplace might arise from the different exogenously given attributes of groups, such as language skills,[26] common culture or tastes,[27] or the reluctance of typical workers to help to train minority workers or to cooperate with them.[28] These costs of integrating might cause segregation to persist.[29] Alternatively, minority workers would be paid lower wages by firms where "typical" workers predominated, because hiring them imposes not only a money wage cost on an employer, but also a cost of integration.[30]

Of course, members of the minority group may strive to reduce these costs -- for example, by learning the language of culture of the predominant group.[31] Once a minority-group member has acquired the ability to look like the majority group, he will not be discriminated against. Hence the returns to certain 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 an initial endowment of resources and the educational opportunities required for the necessary training.[32]

4. **The role of anti-discrimination law**. -- In these conceptions, the role of anti-discrimination law is clear. The law bans overt

---

[26] See Kevin Lang, A Language Theory of Discrimination, 101 **Q. J. Econ.** 363, 365 (1986).

[27] See Epstein, supra note 4, at 61-69.

[28] E.g., Assar Lindbeck and Dennis J. Snower, Cooperation, Harassment, and Involuntary Unemployment: An Insider-Outsider Approach, 78 **Am. Econ. Rev.** 167, 187 (1988).

[29] See Sunstein, Why Markets Don't Stop Discrimination, 8 **Soc. Phil. & Policy** 22 (1991). Epstein's argument also works along these lines. He expects segregated workplaces in equilibrium where different groups face high costs of integrating with others(in a world where people have discriminatory tastes, but there is no anti-discrimination law) see Epstein, supra note 4.

[30] Epstein explains this in part as the reduced costs of enforcing informal (non-legal) 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 are Still a Mistake, 108 **Harv. L. Rev.** 1085 (1994).

[31] See Darity, supra note 22, at 340-356; Lang, supra note 26.

[32] Darity discusses how the differential abilities of different immigrant groups to integrate themselves are a function of initial endowments of capital (whether human capital or other), contrasting Jewish, Asian and West Indian immigrants to African American descendants of slaves and Chinese-American descendants of railroad workers. See Darity, supra note 22.

discrimination, regardless of whether the discriminator is acting according to taste or a rational or irrational cognitive judgment.[33] In turn, the previous victims of discrimination now face the same incentives to acquire human capital as do "typical" workers, and so the social losses from that secondary effect of discrimination vanish as well. The task of the law is to solve the practical problem of how to identify discrimination. However, this should be easy in the standard analyses, because discrimination consisted of refusing to hire or promote workers who were as qualified as the typical workers who are hired or promoted.[34] Indeed, that is the approach of current law, though disparate treatment and disparate impact analysis differ in detail as to the way presumptions, inferences and burdens of persuasion are allocated, the gravamen of most discrimination litigation is the question whether the employer is passing up (either intendedly or as a result of a "neutral" procedure) qualified minority applicants for less qualified non-minority applicants. Thus the basic structure of the law largely tracks current understandings of the underlying structure of discrimination.

We shall argue that the received models are inadequate to understand current extant discrimination; correspondingly, the legal structures regulating discrimination are in adequate to eliminate it.

II. A New Analysis of Employment Discrimination in "Elite" Jobs

In this section, we explain how self-perpetuating discrimination may arise through mechanisms more complex and resilient than those identified by current theory. The basic themes of our analysis are that, in high level jobs, difficulties in monitoring (or, equivalently, discretion by maangers in in evaluating workers) creates powerful opportunities for discriminatory hiring and promotion by firms. Unlike the forms for discrimination ascribed to taste or some statistical generalization, these survive at equilibrium.

We begin with a brief general statements of the intuitions behind non-marginal product models of the labor market and the firm. We then develop in more detail the conception that we will use for our own

[33] A caveat: in effect, the business necessity defense to a disparate impact claim in effect allows a business to engage in de facto discrimination in instances where an decrease in cost turns out to be accomplished by a hiring practice that has the effect of discriminating against a minority group. However, cost cannot --except in the extreme circumstances of a bona fide occupational qualification -- justify use of explicitly discriminatory criteria.

[34] In some statistical discrimination models, of course, minority workers are on average less qualified than whites. See e.g. Aigner & Cain, supra note 20. Once explicit use of race is banned, however, minority applicants face the same incentives to invest in human capital as other applicants, and the statistical discrepancy in qualifications disappears.

analysis, and finally consider how discrimination arises in our setting. The project of this section is to build the more sophisticated considerations of non-marginal product analysis into the standard theories of discrimination described -- taste, signalling, workplace culture, and feedback effects on human capital investment.

A. Non-marginal product models of labor markets

In the standard marginal product model of labor markets, firms match workers with jobs: they hire first the most productive workers for a job, and then dip into the market to hire successively less productive workers; they will stop hiring at the point where the additional product from the next worker is less than the salary to be paid for that worker: hence the standard notion that "wage equals marginal product." The job-worker productivity matching story similarly has predominated, explicitly or implicitly, both the legal and economic analysis of discrimination in hiring.

Nonetheless, labor economists have long remarked numerous anomalies that appear when one tests marginal product models. For example, rational profit maximizing firms have not reduced wages when it appeared that the firms could save labor costs by hiring unemployed workers who would work at salaries below those paid to currently employed workers.[35] In particular, it is puzzling -- and important for our analysis -- that there persists a pool of unemployed workers who are willing to work at wages equal to their marginal products. Other anomalies included sharp jumps in salary upon promotion (e.g., from associate to partner), or, analogously, a firm's policy of firing relatively less productive workers rather than simply reducing their compensation to the actual marginal product. These anomalies can be explained by a variety of "efficiency-wage" or "tournament" models, all of which reflect that either workers or the firm are imperfectly informed.[36] The most important approaches, which provide the basis for our analysis, are these:

1. <u>Moral hazard.</u>[37] -- Here, the basic premise is that firms set

---

[35] For example, in the late 1980s and early 1990s, elite law firms hired fewer lawyers at the same high salaries they were paying earlier. See David B. Wilkins and G. Mitu Gulati, Why Are There So Few Black Lawyers in Corporate Law Firms? ((1995) (forthcoming, California Law Review).

[36] <u>See, e.g.</u>, **Robert M. Solow, The Labor Market as a Social Institution**, 28-33 (1990). In addition to Solow's book, useful general surveys of "efficiency wage" models are **Efficiency Wage Models of the Labor Market** (George A. Akerlof and Janet L. Yellen eds. 1986) and **Andrew Weiss, Efficiency Wages: Models of Unemployment, Layoffs, and Wage Dispersion** (1990).

[37] <u>See, e.g.</u>, George A. Akerlof and Lawrence F. Katz, <u>Workers' Trust Funds and the Logic of Wage Profiles</u>, 104 **Q.J. Econ.** 525 (1989); Jeremy I. Bulow and Lawrence H. Summers, <u>A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment</u>, 4 **J. Lab. Econ.** 376 (1986); Carl Shapiro and Joseph E. Stiglitz, <u>Equilibrium Unemployment as a Worker Discipline Device</u>, 74 **Am. Econ. Rev.** 433 (1984).

wages to solve monitoring problems in the firm. In the standard model, it is costly to monitor workers, and firms use discharge as the penalty for workers' slacking off. The worker's effort, in turn, is a function of the wage minus the opportunity cost of losing that wage (i.e. what other wage is available out there if the worker loses this job). The firm pays the worker a wage higher than the hypothetical market-clearing wage (i.e., where the demand and supply curves intersect); consequently, the worker suffers an economic loss at dismissal, and is thereby deterred from shirking.[38] 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.[39]

2. **Adverse selection**.[40] -- Firms pay higher-than market clearing wages to attract a superior work force: the firm that pays higher wages than other firms will be able to select the better applicants from the pool and hold on to more of its good workers than other firms. In effect, firms pay wages so high 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.[41]

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; salaries may be individually negotiated. But firms often tie their hands, as to large components of compensation, with lock-step or job-

---

[38] 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 Jeremy I. Bulow and Lawrence H. Summers, A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment, 4 **J. Lab. Econ.** 376 (1984).

[39] This trade-off 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)

[40] See, e.g., Andrew Weiss, Job Queues and Layoffs in Labor Markets with Flexible Wages, 88 **J. Pol Econ.** 526 (1980); Alan B. Krueger and Lawrence H. Summers, Efficiency Wages and the Inter-Industry Wage Structure, 56 **Econometrica** 259 (1988); Jane L. Yellen, Efficiency Wage Models of Unemployment, 74 **Am. Econ. Rev.** 200 (1984).

[41] 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'd attract only the high risk projects, since the lower risk projects would prefer to borrow from the lower interest rate lender (so would the high risk project, but he has a lower chance of receiving a loan than a low risk project). See Oliver Blanchard and Stanley Fisher, **Macroeconomics**, (1992)(discussion of Stiglitz's model of credit rationing)

determined compensation schemes.[42] First, they might do so because they cannot readily distinguish between "good" and "bad" workers. Second, workers apparently value equity: the firm might have to raise compensation for all workers to compensate them for the utility loss of having to observe other workers make more money.[43] Third, lock-step compensation is a way for the firm to commit itself to workers not to seek opportunistic revisions in compensation levels after workers become locked into the firm. Workers demand such a commitment because they often lack the information to assess the reasonableness of specific, case-by-case compensation adjustments made by management (monitoring each other is expensive and difficult for workers too) .

3. Other efficiency-wage analyses. -- Higher wages may elicit greater effort from workers, or reduce labor costs, for other reasons as well. Higher wages may decrease worker turnover, increase the size of the applicant pool, or secure worker loyalty by inspiring workers to make an "altruistic" commitment to the firm.[44]

4. Tournament models.[45] -- In these models, firs solve the moral hazard/shirking problem with a "carrot" rather than a "stick." Rather than threatening to fire workers who shirk (thereby causing them to sacrifice supra-market wages, the firm offers a reward for consummate efforts -- commonly a cash bonus or promotion to high-compensation position.)

---

[42] Cf. David H. Maister, Lockstep and the Power of Social Control, **Am. Law.**,Sept. 1994.

[43] See, e.g., George A. Akerlof, Labor Contracts as Partial Gift Exchange, 97 **Q. J. Econ.** 543 (1983); cf. John Rawls, **Theory of Justice** 530-41 (1971)(human tendency to assign importance to relative standing).

Professor Selmi argues that affirmative action serves the purpose of reducing discriminatory behavior in the workplace, and hence improving efficiency, since a non-discriminatory workplace is viewed as fairer, and hence effort inducing. See Selmi, supra note 8. While we agree with Professor 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 it is effort reducing as opposed to effort enhancing), and discrimination is seen as fair (i.e., statistical discrimination), and hence induces effort in the majority of workers.

[44] See, e.g., Alan B. Krueger and Lawrence H. Summers, Efficiency Wages and the Inter-Industry Wage Structure, 56 **Econometrica** 259 (1988)(discussing turnover as a function of the level of wages); George A. Akerlof, Labor Contracts as Partial Gift Exchange, 97 **Q. J. Econ.** 543 (1982) (mutual altruism between the firm and the workforce).

[45] Tournament models are developed in Edward P. Lazear and Sherwin Rosen, Rank Order Tournaments as Optimum Labor Contracts, 89 **J. Pol. Econ.** 841 (1981); Barry J. Nalebuff and Joseph E. Stiglitz, Prizes and Incentives: Towards a General Theory of Compensation and Competition, 14 **Bell J. Econ.** 21 (1983); James M. Malcolmson, Work Incentives, Hierarchy, and Internal Labor Markets 92 **J. Pol Econ.** 486 (1984); Lorne Carmichael, Firm-Specific Human Capital and Promotion Ladders, 14 **Bell J. Econ.** 251 (1983). The model is applied to law firms in the important study, Marc Galanter and Thomas Palay, **Tournament of Lawyers: The Transformation of the Big Law Firm** (1991). See also Kevin A. Kordana, Note, Law Firms and Associate Careers: Tournament Theory versus the Production-Imperative Model, 104 **Yale L. J.** 1907 (1995)(critique of the application of tournament models to law firms).

Workers in effect participate in a "tournament" to win, competing against each other in terms of the efforts that they make for the firm. The tournament models have particular appeal in our context because they help to explain the structural features of promotion to partnership or vice-presidency or tenure that are common in law firms, consulting firms, brokerage houess, and universities: careful evaluation over a period of years, sharp changes in status and compensation for successful firm members (even when the transition seems to mean little change in actual productivity).[46]

B. Further Specification of an Employment Model

With this background, we develop a model for analyzing the effects of anti-discrimination law. We deal with two stages of the employment process. The firm first hires a few workers from a large pool of job applicants. After a period of employment, the firm decides whether to promote the worker (e.g., to "partnership").

1. Difficulties with monitoring. -- The crucial characteristic of the firms that 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 the firm 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 decides 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 ("duds") are fairly reliable; signals from workers grow less reliable (more "noisy") as those workers' quality approaches the mean (from either direction--"average" workers). That is, the farther one is from the mean the more reliable is one's quality signal. Intuitively, this means that it is relatively cheap to spot "superstars" and "duds", and relatively costly to rank precisely any two candidates chosen among the "average" applicants or workers.

On balance, the firms that we analyze decide to make hiring

[46] Similarly the cost of not winning the tournament can be quite high -- often involving job loss, as in the cases of "up-or-out" partnership decisions at law firms.

On balance, the firms that we analyze decide to make hiring decisions based on cursory examinations of the workers -- examinations that do not cost as much as a full-scale evaluation.[47] At the hiring stage, workers present the evidence of characteristics they have acquired, which signal to the employer their individual ability levels and potential effort levels (for example, high grades). Promotion decisions involve a much 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 signalling or average signalling applicants, despite the fact that the signals we are talking about are only very roughly correlated with ability. Thus, even a limited ability to observe ability should be enough for firms to separate out the group of best-qualified applicants, and thereby induce applicants to sort themselves among firms.[48]

2. **Inducing effort.** -- Once a worker comes to a job, the problem for the firm is how 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." For example, law firms may observe the outcomes of trials or negotiations, but they would find it unduly expensive to observe in detail the 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.[49] As in the models discussed above,[50] the firms that we analyze here induce effort by paying a higher-than-market wage, and threatening to discharge or hold back shirking workers.[51] The wage is sufficiently high so that even a low

---

[47] Even probationary periods provide only a very limited opportunity for evaluation, and, for many firms, serve primary as recruitment devices.

[48] High wages attract high ability workers and serve a sorting function, as do low wages. See, e.g., Andrew Weiss, Job Queues and Layoffs in Labor Markets with Flexible Wages, 88 **J. Pol. Econ.** 526, 531-32 (1980), Janet L. Yellen, Efficiency Wages of Unemployment, 74 **Am. Econ. Rev.** 200 (1984).

[49] See Wilkins and Gulati, supra note 35.

[50] Infra, part II [discussing non-marginal product models]

[51] 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. 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 the worker's chances of finding a similar job.

probability that the firm will discover shirking is sufficient for deterrence.[52] Alternatively, the firm may offer successful workers a large (supra-marginal-product) reward: for example, promotion garners a large increase in compensation.[53]

3. Determination of wages and employment levels. -- Effort is an increasing function of the wage a worker is paid. Firms try to maximize the amount of effort they can extract from a worker, and minimize the wage they have to pay; wages are not determined exclusively by the external labor demand and supply functions. Because firms set wages above the market-clearing wage, 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 it is for an employer to find a worker to fit the job description in question.

4. Summary.-- Thus, firms that we describe pay higher-than-market wages to ward off shirking, or offer "tournament" bonuses to induce extra effort. This is a rational profit maximizing strategy so long as the extra benefit in terms of inducing effort, or attracting high ability workers, is greater than the extra cost in terms of wages firms will keep utilizing this strategy. Wages far above the hypothetical market-clearing wage mean considerable involuntary unemployment. For example, at the wages paid by "elite" law firms, there are lots of unemployed and "non-elite" law firm lawyers willing and able to do "elite" firm work.[54] Yet the "elite" firms choose not to lower their wages.[55] These factors -- difficulty in detecting quality, inflexible efficiency/tournament wages, and a resulting pool of available unemployed qualified workers -- set the stage, as the next section explains, for a substantial degree of self-perpetuating discrimination in hiring and promotion.

---

[52] With high-level, high-paying jobs, even a very low probability of discharge may suffice is to induce effort. Consider the example is provided by the 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 signalled themselves (even if the signal is very noisy) as low quality (i.e., the signal is that another law firm did not consider this person worthy of an offer, even at its low levels of monitoring).

[53] See supra part I (discussing tournament models).

[54] See Wilkins and Gulati, supra note 35.

[55] Id. at _.

C. Discrimination at hiring

Will firms discriminate[56] when choosing among job applicants? Consider a firm that faces an applicant pool that includes "typical" and "minority" applicants. Minority applicants are members of a group that has historically suffered from discrimination in education and job opportunities. We analyze the firms' decisions assuming taking the existence of this past societal discrimination as given. This means that there will be a sectoral baseline level of statistical discrimination -- a level which reflects the influence of societal discrimination on the human capital of applicants. (We emphasize, however, that paradoxically the discrimnation that we describe will prove to be stable, one 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 would also persist; and it would cause an insufficient investment in human capital by minority workers.

In particular, on the basis of our description of hiring and wage determination, we make five claims. 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 do in Becker's standard model.[57] Second, even absent a taste for discrimination, the rational firm will use use minority group status as a signal if the mean achievement level of minority group members is less than of the "typical" group. Third, minorities end up with reduced total compensation because as a group they are valued less than the "typical" workers. Fourth, and perhaps most counter to intuition, individuals from minority groups are less likely to be hired even if their mean achievement level is the same as the mean achievement level of the "typical" group. (This exacerbates the effects of the statistical discrimination which occurs when the mean achievement level of the minority group is less than than of the "typical" group.) This occurs because of the self-reinforcing interactions between the strategies that workers adopt in the face of discrimination, the reliability of the signals that they generate as a result, and the mechanisms that the firms use to respond to the relative unreliability of minority applicants' signals. Fifth, firms that engage in discrimination are robust against market competition.

---

[56] A note on the definition of discrimination is in order. Our approach is straightforward: when we say that firms discriminate, we mean that the firms make hiring and promotion decisions for minority individuals different from the decisions that the firm would make for individuals who were identical except that they were "typical" (non-minority). Intent is not part of the definition.

[57] See Richard A. Posner, The Economics of Justice 352 (1985)(describing Becker's model).

Consider each effect in turn.

1. <u>No economic loss from discrimination based on taste</u>. -- If the members of the firm dislike members of some minority groups or otherwise get pleasure from discriminating, they will engage in discrimination: 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 or duds, so that distinguishing among them is more costly than it is worth), and the interviewer has discretion, he can follow his taste.[58] Indeed, the firm would want its hiring decisionmakers to follow their prejudices, as this would increase the decisionmakers' enjoyment of the job and so reduce the money wages that they would demand from the firm. The firm would use providing its decisionmakers with the opportunity to exercise prejudice 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 their decisionmakers' discriminatory tendencies.

2. <u>Endogenously determined rationality of minority group status as a signal</u>. -- Even putting aside background social deprivation inflicted on minority groups, statistical discrimination will sustain itself vis-a-vis job applicants in this sector of labor market. Members of minority groups who plan to enter the sector -- <u>e.g.</u>, those who have begun professional 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.

The simple causal chain of self-perpetuating statistical discrimination clearly extends to the job sectors that we describe. Minority group status is costless to observe; and the false-negative judgments that may result are virtually costless 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 vis-a-vis comparable non-minority applicants, and this affects their incentives to acquire job qualifications.[59]

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

---

[58] The firm suffers a loss from passing over a minority superstar, but by definition this is statistically a very small part of the total applicant pool and of the total number of job applicants who are hired.

[59] As we explain below, an additional source of statistical discrimination is the degree of noisiness of minority applicants' signals relative to signals of non-minority applicants.

3. Depressed wages 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: they obtain jobs at wages reduced to reflect the firm's estimate of their (lower) 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; their human capital investment is largely wasted.[60] And, because the minority workers are not employed, they never have a chance to prove that the discriminatory stereotype (under-achievement) is wrong as applied to them, this gradual ameliorative mechanisms for changing firms' attitudes does not work; the stereotypes cannot be expected benignly to disappear.

4. Altered distribution of minority group members' investments 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 the total amount, of investment in human capital.

To understand how this happens, consider further the choices of minority applicants faced with discrimination. The "average" minority job applicant can then pick among four potential strategies:

(i) high-risk strategy -- The applicant picks projects that would signal her to be a "superstar" if she does well, but for which, since she is actually of average ability, carry a high possibility of failure (failure then signals low ability or "dud" status). These projects/strategies[61] are also more demanding than most other. On this strategy, the applicant makes high effort, but enjoys a low probability of success (though success may

[60] Two observations soften our conclusion. First, though wages may not be flexible, it may be possible for minority workers to obtain jobs by accepting reductions in non-wage compensation. For example, minority lawyers might be hired, but then receive less desirable assignments: assignments that are uninteresting, or are unlikely to increase their skills or gain them status in the firm or the profession. Similarly the minority lawyers might receive inferior perqs or staff support. Second, note that "unemployment" here means that highly trained lawyers are not employed in the sector of "elite" law firms. They may gain employment outside the law, or doing legal work in the "non-elite" sector.

[61] One high-risk strategy might be to take on a large number of projects with uncertain due dates. One might hope that it never occurs that more than one project becomes due on the same day (assuming that one can only produce project result at a time, and if two or more come due, the others fail). If one is able to do an exceptionally large number of projects at the same time the higher-ups may notice the large number of hours billed, but if more than one project comes due at the same time and it brings upon the firm a client's displeasure, the cost to the employee will be very high--especially if these are simple projects that could have all been successfully completed if the employee hadn't overloaded herself.

carry high rewards), and a high probability of failure.[62]

(ii) high-effort strategy -- The applicant picks projects that demand exceptionally high effort, but increase somewhat the prospect of being labeled a "superstar."

(iii) average strategy -- The applicant of average ability picks activities that require average ability, and in which she faces a high probability of being able to signal that she is of average ability, and a low risk of appearing to be a "dud."

(iv) shirking strategy -- Here the applicant of average ability makes a minimal investment in signalling job qualifications, knowing that this strategy is likely to result in a signal to employers that she is of low ability, but also obtaining gains from leisure. (Note that, because all signals are noisy, some who follow the shirking strategy do get jobs. In other words there is still a small probability that a "shirker" will be lucky and acquire an "average" or even "superstar" signal and get a job)

How will job applicants have chosen among these strategies? Minority applicants, before they ever get to the interview stage, know that they are going to be at a disadvantage there. To get a job they have to signal that they are much better than the nearest "typical" applicant because the interviewer will (a) choose the minority worker only if she is clearly better -- most probably, only if she is a "superstar," and (b) be more suspicious of the minority worker's qualifications, because the minority worker had skewed incentives to acquire qualifications. The minority applicant knows that she has to signal "superstar" if she wants the same job that the "typical" worker needs only to signal "average."

Faced with these hiring patterns, "typical" applicants will tend to choose the "average" strategy, while the minority applicant most likely chooses either the high-risk or the shirking strategy for acquiring credentials.[63] Typical applicants generally have no reason to pursue a "high-risk" strategy; they will be hired on the average strategy, which carries a much lower risk of failure. Because of discrimination, however, "average" minority applicants are likely not to be hired if they pursue the average strategy. They choose, then either the high-risk or the shirking strategy. The choice depends on whether they are willing to make the very substantial extra effort that the high-risk strategy entails -- a choice that may not be rational when one considers the low probability of success

---

[62] It is not a high-risk strategy for a real "superstar" to choose projects that will signal her to be a "superstar"--since she in fact has the abilities to successfully complete the projects that would show one to be a "superstar."

[63] The concept people pursuing high risk strategies when faced with failure if they behave normally ("average" strategy choice) has been applied elsewhere. See Susan Rose-Ackerman, Risk Taking and Ruin: Bankruptcy and Investment Choice, 20 J. Legal Stud. 277 (1991)(analyzing the choices of strategies made by financial institution managers in the face of bankruptcy).

under that strategy.[64]

These choices effect a change in the distrbution of achievement among groups of job applicants, independently of the effect it has 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 they know that if they are on par, and a choice between workers is made, minority applicants will lose out. In aggregate, the minority applicants choose strategies that are likely to produce either "superstar" signals or "dud" signals.[65] This change in distribution of signals to invest in, in turn affects the evaluation of minority applicants; consequently, even if the minority average level of achievement is the same, difference in distribution [and correspondingly in noisiness of signals] means that statistical discrimination is perpetuated.

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 prospects for minority group members becomes even worse when one considers the likely distribution of favorable and unfavorable signals among job applicants.

Consider a group of 100 job applicants, of whom 10 signal that they have outstanding abilities, 80 signal that they are average, and 10 signal that they are low-ability/duds. Assume that the firm needs to hire 30 people for the job. It will hire the outstanding 10, and 20 of the average group. The 10 low ability people will not be hired. Among the 80 average, the firm is <u>indifferent</u> as to which 20 it hires. The signals these 80 average applicants present at their interviews look basically the same to

---

[64] Minority applicants on average are likely to prefer high-risk to high-effort strategies because high-risk ones carry greater probability of being labeled a superstar, although also greater probability of being labeled a dud, than high-effort ones. For example, co-author Gulati who graduated from law school recently has suggested in the context of activities at Harvard Law School, the voluntary moot court competition is a high-risk strategy; trying to get very high grades is a high-effort strategy, and participating in voluntary legal aid work or a non-competitive law journal is an average strategy, and doing nothing is a shirking strategy.

[65] The following (semi-autobiographical) example expands on the one suggested in the footnote 64. In the context of Harvard Law School, activities that come to mind (since they are ones that co-author Gulati participated in) are the upper round moot court competition, legal aid work at the legal services center or legal aid bureau, doing editorial work on a non-competitive law journal. Assume with us that one important <u>reason</u> people participate in these activities is to obtain a signal that they can later take to the labor market.

With the moot court competition, lots of people participate, the effort levels required are high, and the probability of success is low (the signal to the job market of being the "best oralist" at the finals is probably a "superstar" signal). In terms of penalties, you lose out on time to study for your classes, and quite possibly get lower grades than your peers. Hence moot court fits the category of a high risk activity. With the legal aid bureau, the legal services center, a non-competitive law journal, so long as you fulfill their time commitment, you are assured of a signal to the labor market. There is no failure. Yet this signal is not going to signal that you are a "superstar" as perhaps being the best oralist in the moot court finals does--but it does signal that you are not a "shirker" or a "dud".

the interviewers (given the low ability of the interviewers to gauge ability in these industries–on account of high monitoring costs).[66] As a result, the interviewers can discriminate.

The result is that there will be statistical discrimination even if minority and non-minority mean achievement levels are the same. Two types of causal chains are at work here. First, 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 of prior existing biases, the exercise of which among the averages does not hurt profits) -- i.e., a world in which the distribution of achievement signals is the same as the distribution of ability/talent. As a result of the discrimination in a world of scarce jobs, the "average" minorities 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 to be 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 than it is to do so for "typicals." Achievement signals are now distributed so that members of minority groups more likely to fall at the extremes than as are "typicals" (most of whom are correctlty signalling themselves to be "average", whereas the minorities are producing a large number of "noisy" and "false" signals) -- a distribution which is likely to obtain when individuals' investment in credentials is affected by discrimination.[67]

Furthermore, as a result of average members of a discriminated against group picking strategies that will show them as "superstars" when they are not, employers will be more skeptical of a member of a discriminated against group who signals that he is a "superstar." The

[66] For example, a law firm interviewing on campus at Harvard Law School typically sees a grade report and a one page resume. Maybe they can identify the people who graduate with straight As as "superstars," and the ones with straight Cs as duds, but the vast majority of applicants will be in the B+ to B range, within which distinctions don't mean anything. We hypothesize that to most "elite" law firms students in the B+ range appear more or less interchangeable in terms of their potential success as lawyers. Hence it does not matter which particular law students are hired in the B+ range. Of course, law firms may use other criteria for hiring, such as social class, apparent fit with corporate culture and so forth; even so, this will probably leave a sense of choosing fairly arbitrarily among a number of equally qualified applicants for associate positions. 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 and Gulati, supra note 35.

[67] For minority workers/applicants there will be no point in making an effort to accumulate higher credentials unless they are "superstar" credentials. The typical "average" applicant has an incentive to signal himself as a superstar as well, but the typical applicants make less effort than "average" minority applicants, because the "average typical" applicants perceive fewer gains from falsely signalling themselves as "superstars" than do "average" minority candidates (relative to if the "average typicals" signal themselves as "average").

reason for this is that the employer recognizes 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:[68] the apparent high achievement is the outcome of a successful gamble.[69] Hence the firm will discount signals given by minorities who signal themselves to be "superstars." In other words the employer will find these signals less believable than if they were coming from a "typical."

5. Survivial of discrimination revisited: firms that do not engage in statistical discrimination may fail. -- In a competitive market if minority status were a valid criteria, then a firm is foolish not to use it. The firm that does not use minority status will make lower profits than the others, and hence will 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 either the behavior of competitors or that of minority workers (as a group). On the whole, minority workers will still see the market as discriminating against them.

6. Implications. -- Over all, our analysis paints a grim picture. Its main features are these:

(a) In our model individual firms can indulge their "taste" for discrimination -- or, more generally, act on discriminatory cognitive biases or social identifications --without hurting their monetary profits. In turn, this discrimination alters minority applicants' incentives to obtain training and thereby changes the distribution of "signals" for minority applicants. As a result of this distribution, firms would be rational to apply their hiring criteria to the disadvantage of minority workers by interpreting minority workers' qualifications with greater suspicion. As these hiring policies become the rational decision for a firm faced with identical or close-to-identical workers, firms that follow comparatively non-discriminatory policies are driven out of the market over time.

(b) Minority workers who are hired will be better qualified for their jobs than their non-minority colleagues, because they are hired only when they are able to signal that they are unambiguously better than the nearest "typical" applicant.[70]

(c) The economy as a whole faces an underutilization of talent.

---

[68] Professor Selmi cites numerous studies that detail how "negative acts of a member of an outgroup [minority] . . ., are frequently attributed to personal disposition, while positive acts are discounted as the product of luck or special circumstances." Selmi, supra note 9, at 1285 n. 111.

[69] Recall that non-minority groups are less likely to go to for these high-risk strategies (a) because they have not reason to take the extra risk of failure-- 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.

[70] Recall, however, that this result is softened if minority workers can be given on average different tasks with different perqs than "typicals" (albeit with the same job titles and money wages).

The best workers for particular jobs are not utilized. Non-minority 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.

D. Affirmative action in hiring

Affirmative action[71] in the sectors that we describe will diminish firms' productivity only to a very small extent: the firm can implement affirmative action by selecting minority candidates from among the large pool of workers who signal themselves as average, and produce a workforce which is more or less as qualified as the one that it had under the non-quota hiring scheme. In theory, the extra "weight" given to minority candidates could be set at the level necessarily to counter the amount of discrimination that we describe; this in turn would resolve as well the problem of minority workers' facing skewed incentives.

Under our analysis, affirmative action has mixed effects on minority workers. On the one hand, affirmative action 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.") To that extent, minority candidates will therefore shift towards the strategies the place them among the average; they have less reason to pursue the high-risk strategies, that labeled them either "superstars" or "duds."[72]

On the other hand, affirmative action may harm minority workers when employers use low levels of monitoring in choosing between workers. Affirmative action would tend to make hiring decisionmakers more suspicious that the signals given by a black worker are "noise" as opposed to being legitimate: 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

---

[71] We define affirmative action as choosing a minority candidate over a "typical" when the two are indistinguishable in terms of other qualifications.

[72] Stephen Coate and Glenn Loury, Antidiscrimination Enforcement and the Problem of Patronization, 83 **Am. Econ. Rev.** 92 (1993), develop a model in which an affirmative-action policy discourages workers in the favored minority group from efforts. In effect, affirmative action supplements the effects of extant statistical discrimination. 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. See also, Glenn C. Loury, Why Should We Care About Group Inequality? in **The Question of Discrimination**, 268, 285 (Steven Shulman & William Darity, Jr., eds. 1989).

acquiring training.

On balance, the widespread use of affirmative action provides confirmation for the basic features of the model. In an important empirical sense, the use of race or gender as an affirmative hiring criterion is entirely symmetric with the use of these factors as a negative criterion: both uses of minority status reflect that, in instances of low monitoring, fuzziness of qualification signals, and efficiency wage or tournament incentive structures, use of "minority" signals is a low-cost personnel strategy for the firm.[73] Indeed, that affirmative action was so widespread that its most overt, apparently abusive feature -- race-norming --had to be banned by legislative action reflects the strong economic pressure towards use of "minority" status.[74]

Finally, it is important to emphasize that the widespread use of affirmative action standards changes fairly drastically one's intuitions about the strategies that will be pursued. With "elite" law firms hiring at "elite" schools, for example, it seems likely that the widespread use of affirmative action has eliminated distortions towards high risk strategies for minority law students. There is less reason to believe that this is the case, however, for promotion to partnership, as discussed in the next section.

E. Discrimination at the Promotion Stage

Discrimination at the promotion stage displays the same basic structure as discrimination at the hiring stage. At promotion, as at hiring, the firm faces the tradeoff between accuracy of employment assessment and costs of information gathering. Firms will economize on information costs by not spending the resources that would be needed to make perfectly accurate distinctions among the workers who signal themselves as average; such distinctions are not cost-effective.[75] Thus, as at hiring, the firm may find itself without reliable grounds to differentiate among those whom it

---

[73] Selmi, in his discussion of affirmative action, supra note 9, at 1301, suggests that affirmative action may serve as a type of "gift," in the sense discussed by Akerlof, supra note 43, and thereby efficiently induces minority workers to make greater efforts than they would otherwise make. The suggestion is provocative; it seems unlikely, however, that the second-order effects on productivity that affirmative action would accomplish, under this theory, would suffice to overcome otherwise large differences in worker qualification; hence, Selmi's theory is perhaps best understood as supplementing our discussion.

[74] In part, this pressure came from the usefulness of affirmative action hiring as a way of avoiding lawsuits, as Justice Scalia has argued in his employment discrimination opinions. See, e.g., Johnson v. Transportation Agency, Santa Clara County, 480 U.S. 616, 646-49 (1987). That firms would adopt this strategy to avoid the relatively low present value cost of a lawsuit also suggests the relative fuzziness of perceived signals.

[75] See, e.g., Marc Galanter and Thomas Palay, Tournament of Lawyers: The Transformation of the Big City Law Firm 50 (1992)(describing increased difficulties of monitoring performance in large, highly specialized elite law firms).

might promote. Hence, as in the hiring case, if the persons making the promotion decision derive some utility from discriminating, it does not hurt the firm for them to do so. The profits of a firm that discriminates 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 firms that we expect to be driven out of the market are those that pick a low ability worker or average ability worker over one of high ability.)

Similarly, the analysis of statistical discrimination extends from the hiring to the promotion case. The greatest difference between the initial hiring and the promotion stages lies in the choice among high-risk, high-effort, average, and shirking strategies. As discussed earlier, we see workers as faced with a choice of projects. They can pick the "high-risk" projects where success will show them to be "superstars", or the "shirking"[76] projects that give them lots of leisure time. Picking an "average" project is 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 pick high risk projects, the signal becomes noisier, because success may be the result simply of a lucky gamble.[77]

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-signalling") projects; and (b) incorporates greater penalties for shirking. Both factors drive the

[76] Once a minority worker is hired (and decides to shirk as opposed to pursuing a high-risk strategy), she is still likely to exert the amount of effort necessarry to not be fired, because loss of a high paying job is costly, even if one is not going to be promoted. On the other hand, the presence of anti-discrimination law can skew this particular minority towards further shirking (or high risk behavior). The reason is that in a low monitoring world the firm does not know whether of not this minority has a viable discrimination lawsuit or not. The small probability every minority candidate has a potential lawsuit makes the firm reluctant to fire the minority candidate, and hence can push the minority candidates further towards skewed strategies (because they less likely to be fired if they are caught shirking or failing at high-risk projects, but also are less likely to be promoted).

[77] As an example, consider associates at a law firm who can exercise some choice about effort levels and about projects/cases to work on. Some cases require high ability (or high effort), others average or low ability. For a worker the choice of a high-ability case is high-risk because she is going to find it very difficult. However, if she does succeed this will signal her to be of high ability as opposed to average ability, and hence give a chance for promotion that she would not have had if she had signalled herself as average. On the flip side this same worker who knows she is going to face discrimination at the promotion stage if she signals herself as average has an incentive to shirk and take low ability requiring projects, since she knows she will not make partner anyway. The result of members of discriminated against groups picking these skewed strategies is that it becomes harder for employers to interpret the signals that they get from quality of work: When a member of a discriminated-against group wins a particularly hard case, the employer will be more suspicious that this is actually an average ability worker who had good luck in winning her case, than the employer would be if the person winning the case had been from a group not discriminated against.

minority worker is driven towards high-risk projects, even more than in the hiring case.

Indeed, some high-risk "projects" are available only to minority applicants. Consider the woman who must decide how "aggressive" to be on the job. "Aggressiveness" – in the form of overt ambitiousness, energy, search for opportunities to make an impression, and so forth – may lead to a "superstar" label; but, in women, it runs a high risk of offending decisionmakers in predominant male firm.[78] In short, this is a high-risk strategy, but may appeal to many women who see that high-effort or average strategies are likely to leave them invisible among the pack of mediocre male candidates, some of whom will be promoted.

F. Application to elite law firms

Large elite corporate law firms provide archetypal contexts for the model that we have developed. At the entry level, hiring decisions are based on minimal scrutiny: scrutiny of one page resumes, brief on-campus reviews 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 each other.[79]

Although what happens after applicants are hired is somewhat more various and controversial, observers of the large law firm have tended to conclude that firms do not closely monitor associates over the period to partnership.[80] With large numbers of lawyers, dispersed at branch offices, and highly specialized with individual expertise and with a great deal of discretion in the tasks that they perform:[81] lawyers doing complicated and specialized work make monitoring difficult. Moreover, monitoring involves taking high level lawyers, and using them to supervise employees whose productivity may only be a small fraction of the monitoring partner, and little affected by intensity of monitoring. In short,

---

[78] An important example is Price Waterhouse v. Hopkins, 490 U.S. 228 (1989) in which an accountant is passed over for partnership because she was too "aggressive" and not "feminine" enough; there was no dispute about the high quality of her work.

[79] See Wilkins and Gulati, supra note 35.

[80] This is the conclusion of the leading study, see generally Marc Galanter and Thomas Palay, The Transformation of the Big Law Firm, in **Lawyers' Ideals / Lawyers' Practices** 48 (Robert Nelson ed. 1992). In her study of integration an "elite" law firms, Elizabeth Chambliss explains that, "individual performance with elite law firms cannot be evaluated according to objective technical criteria. 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. Furthermore, lawyers in elite law firms work in teams, which complicates the assessment of individual contributions. See Chambliss, supra note 13

[81] See, e.g., Leverage Stuctures at the Elite Law Firms, **Indus. & Lab. Rel. J.** (1995)(discussing the increased degree of specialization at the large highly leveraged law firms).

intensive monitoring will not be cost-justified.[82]

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 surive the partnership "tournament" present an alternative way of creating strong incentives for consummate effort.

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.[83] In particular, the tendency to demand "superstar" performance from minorities has received substantial attention.[84] Higher attrition rates and greater dissatisfaction among women and minority associates could also reflect this greater demands for performance placed on them.[85]

G. Summary

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

---

[82] See Wilkins and Gulati, supra note 35.

[83] Id.

[84] For example, the **Interim Report of the Committee on Minority Employment of the San Francisco Bar Association** (1993) stated: "One partner in management reflected, 'Every time a minority fails the bar, it's like a bomb goes off. That's why it is so important to look for superstars.' Another managing partner said, 'We're all trying to pull in too few people. ... A few unfortunately incidents ... have soured people.'" Elsewhere in the Report, indicating the higher degrees of pressure minority associates face, a minority partner said, "I have to hit it out of the ballpark every time I come to bat." Another said, "Any mistake by a non-Asian minority is quickly viewed as evidence that they are not as talented. White associates can make mistakes without that presumption geing applied.' One minority associated said, "You can't have any negatives," "Being a minority means that you have to prove yourself at every step. ..." In part, as some of the comments recognize, this may be in part an artefact of affirmative discussion (as discussed supra, section E): "They assume I was hired because of my race and that I am incompetent," said one minority associate.

Elizabeth Chambliss, in her empirical study, supra note 13, at 133, 145, 189, 203-210, finds that two strong determinants of "elite" law firm integration for black lawyers are (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 points in the direction of the existence of "taste" discrimination is some percentage of the decisions made at the law firm.

[85] See, e.g., Robert L. Nelson, Of Tournaments and Transformations: Explaining the Growth of the Large Law Firms, 1992 **Wisc. L. Rev.** 733, 736 (1992); Note, The Underrepresentation of Hispanic Attorneys in Corporate Law Firms, 39 **Stan. L. Rev.** 1403, 1412 (1987); Ann J. Gellis, Great Expectations: Women in the Legal Profession, A Commentary on State Studies, 66 **Ind. L. J.** 941, 950 (1991); Frederick Bates and Gregory C. Whitebread, Making a Commitment to Minority Lawyers, 1990 **A.B.A. J.**, 78, 82.

[1] Our analysis emphasizes the importance of discretionary evaluation and difficulties in monitoring, in sustaining discriminatory equilibria.

[2] The analysis shows how the discrimination is self-perpetuating in part by influencing the career strategies that workers adopt in response to discrimination.

[3] In particular, our argument explains the interactions between modes of evaluation and the strategies that workers adopt; responses to the strategies are reflected in types of evaluation [in, e.g., responding to "fuzziness" of signals].

[4] The analysis "endogenizes" the noisiness or fuzziness of signals in response to changes in strategy that might be caused by discrimination.

[5] The analysis takes into account the effects of under-employment in creating a pool of workers which permits firms to discriminate without suffering economic loss.

III. Implications for Discrimination Law and Policy

We identify three types of empirical implications that derive from our analysis. First, the analysis suggests characteristics of sectors where discrimination would be prevalent. Second, the argument has implications for the effects of discrimination law: if anti-discrimination law is to be effective -- i.e., if the law should be used to regulate the type of discrimination that we have described --how should the law be structured? Third, the analysis suggests a set of empirical predictions about litigation under the current law: it predicts what types of lawsuits would be brought now, what kinds of arguments would be made in them, and what the range of likely outcomes would be.

A. 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.[86] In the past few decades the United States has been moving towards an economy dominated by service-oriented firms. In many service sectors, effort and ability are hard to observe; firms are likely to use efficiency-wage or tournament structures to induce effort and to attract high-ability workers. Given that differences in effort and ability are difficult for the firm's insiders to observe, workers' performance will be even harder for

[86] See, e.g., Barbara Flagg, supra note 8; Note, Civil Rights and Professional Wrongs: A Female Lawyers Dilemma, 73 Tex. L. Rev. 1419, 1436 (date)(discussing the case of female lawyers bringing Title VII lawsuits on their own behalf).

an outside government agency to evaluate in determining the validity of a discrimination claim. Anti-discrimination law and competitive market forces work together to deter discrimination in sectors where effort and ability are readily observed; conversely, difficulties in monitoring mean that both the anti-discrimination effects of both legal enforcement and market forces will be attenuated.[87]

Evidence collected in the last few years by Professor James Heckman of the University of Chicago and a number of his colleagues, and studies conducted by the Urban Institute using auditing techniques, are consistent with predictions of our analysis. Heckman and his colleagues, in a series of papers, have studied the impact of anti-discrimination laws on black economic progress in the South. Their findings have been that federal anti-disrimination laws were significantly responsible for economic progress made by blacks in the South during the period 1960-1980.[88] However, the progress was observed predominantly in low-skill manufacturing jobs, such as those in the textile industry. Interestingly, Heckman and his colleagues find that improvements in educational levels have little or nothing to do with the gains for blacks that they observed. In fact, they go so far as to say that their empirical results show a negative correlation between education and economic progress for the segment of society that they studied.

In contrast, data from the National Longitudinal Survey of Youth for the years 1979-1988 indicate that young black men faced substantially lower incentives to obtain an education than young white men of equal levels of ability.[89] Maxwell's study supports the hypothesis that in a majority of sectors requiring education-generated expertise, black workers expect to face discrimination.

Perhaps the most interesting studies of discrimination for the purposes of our analysis are those conducted by the Urban Institute.[90] In

---

[87] See Glass Ceiling Report, supra note 7, at 29.

[88] See Richard J. Butler, James J. Heckman, and Brook S. Payner, The Impact of the Economy and the State on the Economic Status of Blacks: a Study of South Carolina, in **Markets in History: Economic Studies of the Past**, 262-64 (David Galenson ed. 1989); John J. Donohue III and James J. Heckman, Continuous versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks, 29 **J. Econ. Lit.** 1603, 1643 (1991); James J. Heckman and Brook S. Payner, Determining the Impact of Federal Antidiscrimination Policy on the Economic Status of Blacks: A Study of South Carolina, 79 **Am. Econ. Rev.** 138, 177 (1989).

[89] Nan Maxwell, 47 **Indus. & Lab. Rel. Rev.** 249, 249-64 (1994). See also **Richard B. Freeman, Labor Markets in Action**, 128 (1989)(the perception of discrimination held by black youths affects their incentives to work and increases their on-the-job absence rates).

[90] See Ronald B. Mincy, The Urban Institute Audit Studies: Their Research and Policy Context, in **Clear and Convincing Evidence: Measurement of Discrimination in America** (Michael Fix and Raymond J. Struyk eds. 1993).

these studies, pairs of equally qualified black and white applicants, and hispanic and white applicants, apply for jobs. The jobs on which the studies focus are primarily in retail and service sectors, and required at least a high school degree; the simulated applications all indicated a college degree.[91] 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 -- in terms of our analysis -- the non-white auditors received a greater degree of unfavorable treatment when they applied for service and management higher-wage jobs.[92]

This empirical evidence -- taken as a whole --substantially favors the hypotheses that [1] market forces and anti-discrimination laws combined have worked to remove "blatant" discrimination -- i.e., discrimination which is readily observed by outsiders; but that [2] a substantial amount of discrimination remains in higher-skill jobs where there remains substantial discretion in applying quality standards -- i.e., where discrimination is more difficult for outsiders to observe.[93]

B. Current doctrine is ineffective to regulate the type of discrimination described here

In disparate treatment litigation,[94] after the plaintiff meets the minimal 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 how that the actual reason for the employment decision was discriminatory animus. This approach is cleary 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,[95] evaluation of job prospects or job performance.

[91] In the studies, the auditors tended to be qualified or overqualified for the jobs to which they applied, and yet were not hired: these are sectors in which jobs are rationed, i.e., the labor market is not clearing.

[92] Mincy, supra note 94, at 175.

[93] In broad terms, similar interpretations of the empirical evidence are presented in John J. Donohue and James J. Heckman, supra note 5, and in David Strauss, supra note 6; however, these commentators do not present a more specific analysis of what might account for the inability of law or market forces to touch discrimination in the types of jobs that we have analyzed.

[94] See McDonnell Douglas v. Green, 411 U.S. 792 (1973); Texas Dep't of Community Affairs v. Burdine, 450 U.S. 248, 253-56 (1981).

[95] 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 (non-discriminatory) ground.

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 the plaintiffs prove that they were better qualified than comparable white applicants or promoted individuals.[96] In our analysis, the minority candidates are no better qualified: the collectively discriminatory impact comes from individual decisions among equally qualified applicants.

One turns, naturally, to statistical analysis of hiring patterns under either a disparate treatment or disparate impact theory. At first cut, one might have higher hopes for the effectiveness of this approach. 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 aggregate, for firms in a number of sectors of the economy.[97]

However, even if the same statistical showing can be made for individual firms,[98] this showing does not suffice to establish liability. The additional requirements of proof would be difficult to meet for the discriminatory effects predicted by our analysis.

For a pattern-and-practice disparate treatment claim, defendants would offer a neutral explanation of the pattern of hiring or promotion (job qualification difference, at least plausibly established along the lines already discussed); plaintiffs then would be required to supplement statistical proof with further evidence of discriminatory intent.[99] 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.

Disparate impact doctrine proceeds differently. Requiring no demonstration of discriminatory intent, disparate impact doctrine instead imposes a broad requirement of reasonableness on all employment practices that have statistically demonstrated discriminatory impacts.[100] At least two doctrinal barriers, however, would prevent disparate impact analysis from condemning the hiring and promotion structures that we have described.

---

[96] See, e.g., Mesnick v. General Elec. Co. 950 F.2d. 816, 825 (1st Cir. 1991).

[97] See Hacker, supra note 5, at 60-85.

[98] It may be harder to make this showing for firms than for 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, David H. Kae, Is Proof of Statistical significance Relevant? 61 Wash. L. Rev. 1333 (1986); Ramona Paetzold, Problems With Statistical Significance in Employment Discrimination Litigation, 26 New Eng. L. Rev. 395 (1991).

[99] The leading case is Hazelwood School District v. United States, 433 U.S. 299 (1977); see also Teamsters v. United States, 431 U.S. 324 (1977).

[100] Of course, the precise standard to be met remains sharply controversial after the 1991 Civil Rights Act.

First, it is arguable that the employment practice (imperfect monitoring) does not cause the discriminatory impact, at least in the way that the term "cause" is understood under current doctrine.[101] Our argument describes 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 – as in 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."[102] 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.[103] To be sure, the employer remains liable if its actions directly cause the disproportion in applicants --most obviously, if the employer harasses applicants from a minority group, or if the employer is known to discriminate against minority applicants in hiring. However, in our model, the employer does not take such actions: it does not harass 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 readily could defend their actions by invoking a routinely accepted business justification: the firm is making its hiring decisions based on workers' qualifications.[104] The practices that we describe are clearly distinguishable from the classic Griggs situation of job qualifications that are no "job-related." By hypothesis, in our description, the qualifications imposed are job-related; but they are differentially satisfied, in equilibrium, by the typical applicant and by minority applicants, because of the equilibrium established in efficiency-wage/tournament markets by past discriminatory practices.

---

[101] The requirement of causation is emphasized in Wards Cove Packing Co. v. Antonio, 490 U.S. 642 (1989) and codified with slight modifications by the Civil Rights Act of 1991 at title VII, section 703(k)(1)(B). The codification exempts plaintiffs from the requirement of identifying a particular employment practice that causes the disparate impact when elements of an employer's decisionmaking process are not capable of "separation for analysis." This proviso probably would not help plaintiffs out of the difficulty described in text: their problem is not that elements of the process are not capable of separation for analysis, but rather than none of the elements, taken separately or together, cause a disparate impact in the standard legal sense of causation.

[102] 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.

[103] E.g., EEOC v. Chicago Lamp Miniature Lamp Works, 947 F.2d 292 (7th Cir. 1991).

[104] See, e.g., Woody v. St. Clair County Comm., 885 F.2d 1557 (11th Cir. 1989); Lucas v. Burnley, 879 F.2d 1240, 1244 (4th Cir. 1989).

In short, the type of discrimination that our model described is actionable under neither disparate treatment nor disparate impact theories.

C. Current Doctrinal Controversies in Discrimination Law

In light of this general doctrinal analysis, we comment on how our analysis illumines central current debates in discrimination law.

1. Standards of proof for discrimination in "high level" jobs. -- 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.[105] The articulated rationale is that courts should not second-guess difficult and expertise-laden personnel judgments, such as law firm and accounting firm partnership decisions and university tenure decisions.[106]

Our analysis confirms that this approach is correct, though not for the reasons 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, say, the hiring criteria used by employers of haberdashers or stevedores. The higher standard of proof is justified, nonetheless, because it discourages suits by "average" (as contrasted to "superstar") plaintiffs. Under a preponderance-of-evidence standard, "average" workers had a shot at winning; because there were a large number of discriminatees in this cohort, there would be a large number of lawsuits, most of which would be lost at trial.[107] Raising the standard of proof to require "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 latter standard.[108] Though perhaps perverse from a broader policy perspective, this restriction of the number of suits is at least a sensible way of economizing on limited legal resources: it screens out a large number of

---

[105] E.g., Ezold v. Wolff Block Schorr and Solis-Cohen, 983 F.2d 509 (3rd Cir. 1992) (law firm partnership decision); Jackson v. Harvard University, 900 F.2d 464 (1st Cir. 1990) (tenure decision). Similar deference to partnership decisions is shown in Lucido v. Cravath Swaine and Moore, 425 F. Supp. 123 (S.D.N.Y. 1977); Wheeler v. Hurdman, 825 F.2d 257 (10th Cir. 1987).

[106] 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).

[107] Note that there are not a large nmber of suits at the hiring stage because the stakes are low relative to the costs of suit, cf. John J. Donohue & Peter Siegelman, The Changing Nature of Employment Discrimination Litigation, 43 Stan. L. Rev. 983, 984 (1991); however, once one gets to the stage of promotion to partnership or tenure, stakes increase substantially.

[108] 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.

claims that would lose under current doctrine.[109]

2. The "lack-of-interest" defense. -- As Vicki Schultz's important empirical work has shown,[110] the lack-of-interest defense has grown increasingly important as a defense in employment discrimination cases, particularly in service industries. Our analysis helps to explain why the argument should be powerful in explaining hiring discrimination. Under the lack of interest defense, the employer shows that the reason for the low percentage of minorities in the workforce is the low percentage in the hiring pool. Our model predicts that this should occur: minorities learn that their prospects for getting the job are dim, and so do not make the investment needed to qualify and apply. 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; but, as we have explained above, the type of discrimination that we describe is very difficult to show to be illegal, and so, correspondingly, cannot be cited by plaintiffs as preventing defendant employers from offering a lack-of-interest defense.

3. Affirmative action and quotas. -- Our analysis predicts that one would expect to see affirmative action (including quotas) implemented in many efficiency-wage/tournament settings. In particular, as we have explained, affirmative action (e.g., by weighting of minority scores, a practice now illegal[111]) can accomplish a substantial increase in minority hiring over the pre-quota practice; second, there is no evidence of substantial impairment of total productivity.[112] While data on the extent of voluntary affirmative action programs is difficult to come by -- in part because firms are reluctant to be explicit about affirmative action hiring -- the evidence suggests that voluntary affirmative action is relatively common in service and complex manufacturing jobs, where monitoring is relatively

---

[109] As explained in subpart B, 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 bring many of them to court -- unless the special "clear and convincing" standard for evidentiary proof is imposed.

[110] Vicki Schultz and Stephen Petterson, Race, Gender, Work and Choice: An Empirical Study of the Lack of Interest Defense in Title VII Cases Challenging Job Segregation, 59 **U. Chi. L. Rev.** 1073 (1992); see also Vicki Schultz, Telling Stories About Women and Work: Judicial Interpretation of Sex Segregation in the Workplace in Title VII Cases Raising the Lack of Interest Argument, 103 **Harv. L. Rev.** 1749 (1990).

[111] See section 106, 1991 Civil Rights Act, codified at 42 U.S.C. sectcion 20000e-2(l).

[112] See Jonathan S. Leonard, The Impact of Affirmative Action Regulation and Equal Employment Law on Black Employment, 4 **J. Econ. Persp.** 47, 61 (1991); Jonathan S. Leonard, Antidiscrimination or Reverse Discrimination: The Impact of Changing Demographics, Title VII, and Affirmative Action on Productivity, 19 **J. Hum. Resources** 145, 171-172 (1984)(both articles finding no effects of affirmative action on productivity).

difficult.[113]

At the same time, our analysis suggests that firms may choose to adopt affirmative action programs even when these programs are not socially beneficial, at least from an economic point of view.[114] A firm's decision to give weight to minority status no more enjoys a presumption of economic efficiency than the firms' (implicit) use of minority status as negative factor in hiring and promotion.[115]

IV. Conclusion

In part, our argument aims simply to change how legal analysts think about the sources of discrimination and the impacts of discrimination policy. In particular, we urge that analysis must focus more carefully, in a micro-analytic 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. Specific conclusions about the scope of our model's applicability must await further empirical research. Nonetheless, to put the argument in perspective, we conclude by commenting on the potential impliciations of the model for employment discrimination policy broadly conceived.

Our analysis puts the advocate of a "moderate" anti-discrimination policy in an uncomfortable position. The "moderate" position -- the predominant one in the legal establishment since the passage of Title VII -- relies on the rooting out of readily identifiable discriminatory motives and practices as an crucial means of insuring economic equity and economic progress for traditionally disadvantaged minority groups. In our view, however, the moderate approach leaves untouched firms employment policies whose effect is to perpetuate the employment disadvantage of minority groups, perhaps indefinitely into the future.

Our analysis, in turn, points towards two alternative, extreme solutions. The "deregulatory" alternative would dismantle discrimination

---

[113] See Hacker, supra note 5, at 35-71 (data and analysis). While affirmative action programs for more routine jobs are adopted simply to avoid litigation costs from discrimination suits, this explanation carries relatively less weight for higher-level jobs, where such suits (as to hiring) would be quite difficult to win under current law. See, e.g., Flagg, supra note 8.

[114] In a low monitoring environment where a firm knows that some of its internal decision makers may have acted towards minority employees in a discriminatory way, but cannot identify such behavior until too late, affirmative action may help counter the possibility of a lawsuit.

[115] In this respect, the analysis lends support to the powerful argument about firms' propensities toward reliance on affirmative action, developed by Justice Scalia's dissent in Johnson v. Transportation Agency, Santa Clara County, 480 U.S. 616, 646-49 (1987).

law and return to reliance on market mechanisms to provide for equal job opportunity. Our analysis suggests that such dismantling would, ironically, be necessary for the elimination of discrimination through market mechanisms. The market could stop the discrimination we describe if firms could employ minority workers at lower wages than typical workers doing the same jobs.[116] If wages paid by each firm are inflexible, however, then, firms that employed minority applicants probably would hire only minority workers.[117] Even if legal, such practices may be socially objectionable. Particularly, 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.

At the alternative "extreme" is widespread imposition of hiring quotas or targets. Firms might simply be required to bring the 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. Gradually, in either case, the legal policymaker might hope that the availability of job opportunities gradually would eliminate the different choices of education and effort levels by groups that are the victims of discrimination. However (even putting aside the various moral and political objections to "group-conscious" employment decisions) our analysis raises substantial questions about whether affirmative action would work to reduce discrimination. In particular, affirmative action programs 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 the extreme alternatives is how quickly they would work to reduce discrimination. Politically, it might seem that neither alternative --market-created apartheid or massive imposition of status-conscious hiring --is palatable. While both approaches, de facto, play important if illicit

[116] This is probably illegal under the Equal Pay Act. Even if not illegal, it would be widely disapproved. Firms might create separate tracks that would be filled mostly by minority workers -- such as "permanent associates" positions. Such a 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; such disparate treatment would probably be illegal. 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.

[117] 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); Note, The Jewish Law Student and New York Jobs: Discriminatory Effects in Law Firm Hiring Practices, 73 **Yale L. J.** 625 (1964); Alan Dershowitz, **Chutzpah** (1992)(describing Professor Dershowitz's personal experience with discrimination against Jews at the nation's "elite" law firms).

roles in our current employment law system, they are nonetheless viewed, as a rhetorical and cultural matter, with extreme disfavor. Discrimination 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.