Market Affirmative Action

ROBERT COOTER*

This paper applies the economic theory of regulation to laws forbidding discrimination or requiring affirmative action. Perfect competition causes discriminators to pay for segregation. Market failures can shift the cost of discrimination to its victims. The most efficient remedies are the ones already developed by economists for other problems. Discriminatory cartels can be prohibited or undermined. Discriminatory signals can be overcome by supplementing market information. External effects of prejudice can be internalized by tax-subsidies. In general, any target for diversity can be achieved more efficiently by transferable rights than by quotas.

I was much cheered, on my arrival [at prison], by the warder at the gate, who had to take particulars about me. He asked my religion and I replied "agnostic." He asked how to spell it, and remarked with a sigh: "Well, there are many religions, but I suppose they all worship the same God." — Bertrand Russell

In response to a report on affirmative action, President David Gardner of the University of California quipped, "If you think you know the solution, you don’t understand the problem." People with a solution are like Bertrand Russell’s warder — they underestimate the depth of the disagreement. Intellectuals cannot resolve disputes that are deeper than ideas, but economists can improve the debate


2. Former President Gardner made this remark in response to a presentation of the affirmative action plan of the Berkeley campus to the Regents of the University of California. I thank Michael Heyman, who witnessed the event, for confirming the facts.
by identifying the most efficient means for pursuing any particular policy.

The laws that forbid discrimination and promote affirmative action regulate product and labor markets. Economics has developed an extensive critique of regulation. According to this critique, market-like instruments should replace bureaucratic rules wherever possible. Substituting the former for the latter promotes efficiency and liberty by lowering the cost and coercion of achieving policy goals. This Article applies the economic critique of regulation to antidiscrimination law and explores the possibility of substituting market affirmative action for bureaucratic rules. The Article concludes that substitution would increase efficiency and liberty, and also dilute the law’s symbolic condemnation of discrimination. Unfortunately, economic analysis has no theory of the symbolic and education function of law.

I. ECONOMIC CRITIQUE OF REGULATION

A policy science predicts the effect of alternative policies on shared values. For example, applied economics predicts how alternative policies affect efficiency and distribution. Most political disagreements are buttressed by false beliefs about how policies impinge upon values. By correcting such errors, science improves the quality of policy debate or even imposes a framework upon it.

Economists reached a consensus in the 1970s concerning the framework for analyzing government regulations. According to this framework, a prima facie case for regulation requires demonstrating a market failure and a policy to correct it. Tests for market failure are developed in general equilibrium theory; corrective policies are evaluated by cost-benefit techniques. A conclusive case for regulation requires a further demonstration that the proposed remedy would succeed politically, rather than being subverted by interest groups. Tests for political failure are developed in collective choice theory.

This consensus over the framework of debate obviously did not end disagreements among economists over regulation. Left-liberal economists stress market failures, and conservative economists stress political failures. However, the arguments between the two sides join because they proceed within the same general framework. The joining of the arguments focuses research on questions that both sides consider decisive.

Two examples, which are especially relevant to this article, illustrate the effect of economics on the debate about regulation. First, George Stigler published an article in 1964 purporting to show that federal securities laws do not increase the value of stocks. He concluded by advocating repeal of the federal securities laws and abolition of the Securities and Exchange Commission. Although his recommendations were not adopted, his paper helped to shift the debate about securities laws, which is now conducted in academic circles in terms of the "efficient market hypothesis."

As a second example, Charles Schultze proposed to replace "command and control" regulations with "market-like instruments." To illustrate the distinction, the state can reduce air pollution by imposing quantitative restrictions on polluters ("commands backed by sanctions") or by creating a limited number of transferrable emission rights ("markets in emissions"). The environmental lobby initially defended quantitative restrictions and opposed markets, using arguments that relied upon moral intuition. ("Why should you pay someone to stop dumping garbage in your back yard?") More recently, however, environmental organizations like the Sierra Club and the Environmental Defense Fund have come to appreciate that voters will "buy" cleaner air if it costs less. The search by environmentalists for cheaper solutions to pollution encompasses innovative proposals like markets in emission rights, which were given a statutory foundation in the 1990 amendments to the Clean Air Act.

As these examples suggest, most economists agreed in the 1970s and 1980s that government officials were retarding the economy with


7. SCHULTZE, supra note 4.

heavy-handed regulations. While agreeing about the necessity of reform, economists disagreed over its direction and scope. Thus, George Stigler wanted to repeal most regulations and Charles Schultze wanted to improve them.9 The diversity of views among economists provided intellectual sanction for reforming politicians who repealed some regulations and reformed others. However, few economists found anything good to say about “command and control” regulations.

Contrary to these trends, legislatures, courts, and administrators have imposed new “command and control” regulations since 1964 for the stated purpose of eliminating current discrimination or undoing its past effects upon various social groups, including racial minorities, women, the elderly, and handicapped people.10 These regulations apply to hiring and promoting (“employment discrimination”), and the sale of goods and services (“refusal to deal”). Although complex and uncertain in application, the effect of the laws are far-reaching. New rights have been created for employees, job applicants, and consumers. Many organizations have adopted targets for the social mix of employees and implemented procedures to handle disputes and complaints.

Various scholars have applied the economic theory of regulation to these laws and policies, but none has achieved the comprehensiveness of Richard Epstein’s Forbidden Grounds.11 This technical yet passionate book calls for legislatures to repeal all statutes suppressing discrimination in economic transactions and abandon the field to common law. Thus, Epstein’s prescription for antidiscrimination laws parallels Stigler’s prescription for securities laws. Against Epstein, John Donohue has argued that Title VII probably increased economic efficiency on the whole.12

As good as Epstein’s book is — and it is very good — its policy prescription is more prophetic than practical, because it cannot succeed in the U.S. without an unforeseeable change in the distribution of political sentiment and power. Furthermore, Epstein’s book poses the risk of identifying the economic theory of regulation exclusively with abolition, in which case the defenders of antidiscrimination laws

9. See SCHULTZE, supra note 4; STIGLER, supra note 4.
11. Id.
and affirmative action policies will feel compelled to oppose economics. This would be unfortunate in my view because economics has much to offer both sides in the policy debate over discrimination. To become more widely accepted in this debate, economics must be viewed as a flexible engine of analysis that can serve various political tendencies.

This Article contributes to this goal in two ways. First, I develop the simple, pure economic theory of regulation as applied to discrimination in economic transactions. Second, I develop the general lines of the case for reform, which involves replacing command and control regulations with market-like instruments to overcome discrimination and pursue affirmative action targets. (Hence the phrase "market affirmative action.")

II. DISCRIMINATION AS INEFFICIENCY

To discriminate among wines by taste, smell, and color displays refinement, but to discriminate against people by race, sex, religion, ethnicity, age, or disability may violate morality and law. In the United States and other countries, law restricts the criteria that can be used when employers fill jobs, universities award scholarships, or retailers sell commodities. Fairness in competition requires that the criteria for sorting winners from losers measure performance on dimensions relevant to the activity in question, such as speed, accuracy, comprehension, endurance, and originality. Performance is not measured by traits of persons such as race, gender, ethnicity, or age. Antidiscrimination laws prohibit the unfairness of sorting winners from losers by traits rather than performance. The absence of discrimination gives everyone an equal opportunity to compete for offices, jobs, commodities, privileges, and honors, regardless of their personal traits.

Performance in economic life is usually measured by productivity. Consequently, discrimination in economic life usually consists in sorting people according to traits rather than productivity.


14. See, for example, Rawl's discussion of "pure procedural justice" in JOHN RAWLS, A THEORY OF JUSTICE 83-90, 86 (1971).
people are sorted by traits rather than productivity, industrial efficiency diminishes. Consequently, discrimination as defined here necessarily reduces efficiency in production. Conversely, non-discrimination maximizes the productivity of organizations.

"Discrimination" is a term of opprobrium in ordinary speech. Most practices that satisfy the efficiency definition of discrimination are widely condemned. To illustrate, "ceiling quotas" on Jews, which lowered the productivity of universities, are widely condemned. However, the efficiency definition of discrimination admits the possibility that some practices falling within its scope may be widely condoned. To illustrate, the exclusion of non-Indians from jobs in tribal government on Indian reservations is widely condoned.

Organizations mix people by traits in various ratios, with some organizations being relatively homogeneous and others being relatively heterogeneous. As defined above, discrimination occurs when mixing in an organization does not go far enough to maximize production. People who are willing to sacrifice productivity and lose income in order to reduce mixing have "a taste for separation." Economists often take the satisfaction of preferences as the appropriate goal of public policy. However, there is much dispute about whether satisfying the taste of people for separation is an appropriate policy goal.

Equal opportunity to compete in economic transactions conflicts with freedom of contract. Complete freedom of contract implies the right to deal or not to deal with anyone for whatever reason, including their personal traits. In contrast, antidiscrimination laws attempt to prohibit certain motives from affecting economic transactions. Antidiscrimination laws necessarily interfere with freedom of contract.15 In general, liberty rights conflict with equality rights, because the former create a sphere of individual autonomy and the latter intrude into it.16

---

15. Their enactment, however, does not necessarily logically imply a much wider infringement demanded by some proponents. Equal opportunity to compete does not imply an equal probability of winning, since the probability of winning a fair competition increases with the contestant's skill. Equal opportunity does not imply equal outcomes, since some win and others lose in any competition. Equal opportunity does not imply distributive justice, since a fair competition may not distribute outcomes ideally. Equal opportunity does not imply that the most deserving person wins, since natural abilities and luck affect performance on any occasion.

III. SEPARATION AND MIXING IN PERFECT COMPETITION

In this section I review how competition in markets with free contracts undermine discrimination and determines the extent of separation, segregation, and diversity. I begin by reviewing how competition undermines discrimination. In general, people who have a particular preference bear the cost of satisfying it in a perfectly competitive economy. Consequently, competition causes people who prefer segregation to bear any additional cost that discrimination imposes on goods, services, or employment.

To see why, first consider labor markets. Discriminatory employers constrain themselves by refusing to hire or promote people with disfavored traits. The constraint imposes higher costs to obtain the same quality of labor. In perfect competition, lower cost producers eliminate higher cost producers. Thus, perfect competition eliminates discrimination by employers.

To illustrate, a professional football team, which will go unnamed, recruited the best available white players in the 1950s, but refused to recruit blacks. This discriminatory team competed against other teams that recruited the best available players regardless of race. Over time, the discriminatory team’s popularity and profits plummeted as it lost more and more of its games, so it eventually abandoned discriminatory recruitment.

Now consider how competition affects discriminatory employees, as opposed to discriminatory employers. Imagine a world whose people are blue or green, in which some blues refuse to work with greens, but otherwise people are nondiscriminatory. Workers of different color substitute perfectly for each other on the job, except that organizations employing discriminatory blues must pay the extra cost of segregating them from greens. Thus, the value of a discriminatory worker to an employer equals the value of any nondiscriminatory worker minus the incremental cost of segregation. Competition in the labor market aligns each worker’s wages with his value to employers. The perfectly competitive wage of discriminatory workers thus equals the wage of equivalent workers minus the incremental cost of segregation. In general, perfect labor market competition imposes the cost of segregation upon workers who demand it.

These facts are depicted in Figure 1. The horizontal axis indicates the quantity of labor and the vertical axis indicates the wage rate. Workers are distinguished into those who discriminate, indicated by

a subscript, d, and those who do not discriminate, indicated by a subscript, n. The curves $S_d$ and $S_n$ indicate the quantity of labor each group will supply as a function of the wage. The demand curves $D_d$ and $D_n$ indicate the value of the two kinds of labor to employers. Initially, assume that the cost of segregation is nil, so both kinds of labor are equally valuable to employers and they receive the same wage, $w_d = w_n$. Now assume that segregating the workplace becomes costly. As the cost of segregation increases, the demand curve for discriminatory labor shifts down from $D_d$ to $D'_d$ as shown, and the discriminatory wage falls from $w_d$ to $w'_d$. The reduction in use of discriminatory labor causes an increase in demand for nondiscriminatory labor, as indicated by the upward shift in demand from $D_n$ to $D'_n$. Consequently, a gap opens in the wage of the two groups, with discriminatory labor receiving the lower wage $w'_d$ and nondiscriminatory labor receiving the higher wage $w'_n$.

Fig 1: Discriminatory Employees

I have explained why perfect competition causes discriminatory workers to pay for segregation. In general, perfect labor markets impose an increase in the cost of production upon anyone whose demand for special working conditions causes it. Consequently, perfect labor markets would also impose the cost of additional integration upon workers who demand more of it than would maximize the firm’s productivity.

Now I turn from labor markets to markets for goods and services. A similar argument can be made about the refusal to deal as was made about employment discrimination. First consider discriminatory sellers and nondiscriminatory buyers. If sellers refuse to deal
with some buyers, the discriminatory sellers may experience additional costs. In perfect competition, all goods sell at cost, so discriminatory sellers will charge more than nondiscriminatory sellers for the same good. Nondiscriminatory buyers will purchase from the sellers with the lowest prices. Thus, perfect competition eliminates discriminatory sellers, just as it eliminates discriminatory employers. For example, a restauranteur who insisted on segregated dining facilities might have higher costs, which nondiscriminatory patrons would refuse to bear.

Now consider the case of discriminatory buyers. Once again, product markets strictly parallel labor markets. Specifically, consumers who prefer discriminatory sellers will pay a surcharge for the products they buy relative to nondiscriminatory consumers. The surcharge will equal the additional cost of segregating buyers. For example, diners who discriminate will pay the extra cost of segregating their facilities.\(^{18}\)

I have explained that people who demand segregation pay the cost of it in competitive markets. In turn, demand for segregation conveys differential advantage to various groups in the labor market. To illustrate, airlines apparently believe that a significant group of their customers prefer to be served by pretty, young stewardesses. An airline that caters to this taste may dismiss stewardesses as they reach middle age, thus incurring higher labor turnover and training costs. These higher costs will be paid by passengers in the form of higher ticket prices. However, discriminatory demand will also bid up the wage for pretty, young stewardesses, who will thus benefit from passenger discrimination.\(^{19}\)

IV. DIVERSITY AMONG ORGANIZATIONS AND IN ORGANIZATIONS

So far I have discussed discrimination in markets for labor, goods, and services. Exchange occurs in markets, and production occurs in

---

\(^{18}\) A related question is whether the satisfaction of discriminatory preferences should ever count as a social benefit. Some economists, who are true to the tradition of Bentham, count all preference satisfaction as equally valuable, regardless of whether the preferences are immoral; but others disagree. For discussion and citations, see Jeff L. Lewin & William N. Trumbull, *The Social Value of Crimes*, 10 Int'l Rev. L. Econ. 271 (1990).

\(^{19}\) United States courts have held that customer preferences cannot justify discrimination in favor of young women in the hiring of airplane stewardesses. E.g., Pan Am. World Airways, Inc. v. Diaz, 442 F.2d 385 (5th Cir. 1971), *cert. denied*, 404 U.S. 950 (1971). In *Diaz*, males alleged that Pan American World Airways discriminated in hiring airline staff. The court held that discriminatory preferences of customers could not justify discrimination in hiring airline staff.
organizations. Now I consider the extent to which people from diverse social groups will obtain advantages from mixing together in organizations.

Economic organizations join people together in a structure of roles and personal relationships to produce goods and services. Performing the necessary roles requires cooperation, and motivating the individuals who fill them requires competition. When combining cooperation and competition, disputes often arise over distributional issues like advancement and wages. When people in the organization share the same outlook, disputes are easier to avoid or resolve, and, as a result, the costs of decision making are lower. Furthermore, clustering people together who have the same tastes increases efficiency in the supply of public goods. To illustrate, achieving a consensus on interior decorating in the workplace is easier and cheaper to implement if all the workers like bright colors and clean edges. Thus, efficient governance and the supply of public goods to the workplace argue for clustering workers into homogeneous groups.

However, combining people with different outlooks in the same organization also has advantages. The division of labor causes people to view the production process from different perspectives. People with different perspectives tend to notice things that would be missed by people with the same perspective. Greater perceptiveness increases an organization’s flexibility and adaptability. Thus, heterogeneity, which increases the cost of decision making, may improve its quality.

The competition between homogeneous and diverse firms is exemplified by the competition between American and Japanese corporations. Large American corporations often have an ethnically diverse board of directors, which gives Americans an advantage in international business networks. In contrast, large Japanese corporations have homogeneous boards of directors, which lowers the transaction costs of governance. International competition is now testing which

---

20. This insight is fundamental to the “new industrial organization,” which conceives of the firm as a principal-agent problem. See, e.g., Jean Tirole, The Theory of Industrial Organization (1988).

21. Henry Hansmann found that cooperatives tend to succeed when their members are homogeneous so that decision costs remain low. See Henry Hansmann, When Does Worker Ownership Work? ESOPs, Law Firms, Codetermination, and Economic Democracy, 99 Yale L.J. 1749 (1990).

22. In a famous article, Charles Tiebout argued that mobility would naturally cause people to separate into local jurisdictions in which the residents have similar tastes with respect to public goods, thus increasing efficiency in the supply of public goods. In effect, I am arguing that this conclusion also applies to the workplace. See Charles M. Tiebout, A Pure Theory of Local Expenditures, 64 J. Pol. Econ. 416 (1956); see also Wallace E. Oates, Fiscal Federalism (1972).

23. This insight is formalized in propositions like this: If individuals observe a variable with error, and the error is normally distributed with zero mean, then the best predictor of the true value of the variable is the mean of the individual observations.
kind of corporation is more efficient.

I have described a few of the advantages and disadvantages of homogeneity and heterogeneity in organizations. The reader can probably think of many more. Organizations that compete in the marketplace try to combine persons and practices in the most profitable way, and the resulting mix defines the firm's "culture." In competition, firms with successful corporate cultures multiply and unsuccessful ones disappear.

It is conceivable that one type of corporate culture would be so successful as to eliminate all the others in a fair competition, thus leaving all firms with the same mix of persons and practices. However, this outcome seems as unlikely in economic competition as in biological evolution. At any point in time, we expect competing firms to vary significantly from each other in corporate culture, including the way they mix social groups. In other words, we expect diversity among organizations with respect to diversity within organizations.

Most citizens of a large country like the United States feel the need for institutions and identities that occupy an intermediate position between the state and the individual. Important parochial attachments include community, region, race, ethnicity, and religion. Intermediate organizations such as residential communities, churches, and private clubs are seldom representative of the nation as a whole with respect to measures of diversity applied in discrimination cases. In the case of Indian reservations, restricting residence and employment to tribal members increases ethnic homogeneity.24 Few people would advocate the enforcement of centrally determined targets for diversity upon these intermediate institutions.

Diversity among organizations can arise from separation or segregation. Indeed, distinguishing one from the other is difficult philosophically and practically.25 When perfect competition imposes the

24. To illustrate, the Vice Chairman of the Pasqua Yaqui tribe told me in the summer of 1991 that his greatest concern was to get Yaqui children out of integrated classrooms.

25. For an intelligent discussion of this problem by an economic historian, see Jennifer A. Roback, Plural But Equal: Group Identity and Voluntary Integration, 8 Soc. Phil. & Pol'y, Spring 1991, at 60. Roback writes: How integrated should society be? To answer this question, two prior questions must be addressed. First, what function do ethnic groups perform for their members? This question focuses on the benefits of ethnic identity: distinctiveness and separateness. Second, why do we want ethnic integration? This question focuses on the benefits of ethnic interaction: mixing and transformation. Id. at 60. Roback goes on to ask how to determine the appropriate rate at which to trade off ethnic identity and group solidarity against ethnic integration and social cohesion. Her answer is that government, which cannot get the price of tangible commodities right,
cost of separation or segregation upon the parties who demand it, the case for interference by the state is weak. The strong case for state regulation arises when market failures enable the people who demand segregation or separation to transfer their cost to others. Before turning to market failures, however, I consider the possibility that the state might want to alter the level of mixing under perfect competition. If the state intervenes, it should pursue its policy goal efficiently, which leads to the next topic.

V. PROMOTING MIXING BY TARGETS, TAXES AND TRANSFERABLE RIGHTS

Many markets are “workably competitive,” which means that they function much like the ideal type of a perfectly competitive market. Suppose that politicians and policy makers wish to achieve more mixing than workable competition yields, or to achieve it very quickly.26 One way to pursue such a goal is by imposing “targets” or “quotas” on organizations. “Quotas” are apparently unconstitutional in U.S. law and “targets” are apparently constitutional,27 although the difference between them is elusive.28 Fortunately, this difficult distinction is unimportant for the purposes of this paper. Instead of contrasting targets and quotas, I will contrast non-transferable targets and transferable targets.29

What happens under workable competition if government imposes targets upon all organizations in order to achieve uniform mixing? For example, what happens if government requires fifty percent of the workers in each organization to be male and fifty percent to be female? Like any quota system, employment targets are insensitive to underlying differences in costs, which causes inefficiencies.

To depict this fact, Figure 2 shows average production costs on the vertical axis and the proportion of female workers on the horizontal axis for two hypothetical firms in different industries. According to the figure, production costs are minimized in Firm A when thirty percent of its workers are female, and production costs are minimized in Firm B when forty percent of its workers are female. Thus,

cannot possibly get this trade-off right. So, government should remain neutral towards race. Id. at 61, 78-80.

26. John Donohue argues that perfect competition might take too long to eliminate discriminatory practices. See Further Thoughts, supra note 12; Is Title VII Efficient?, supra note 12.

27. In the Bakke case, Justice Powell held that racial or ethnic targets to achieve diversity are permitted in an educational institution, but quotas are unconstitutional. See Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265 (1978).


in perfect competition without regulation, the proportion of female workers in Firms A and B will equal thirty and forty percent, respectively. To be concrete, the proportion of women graduating from American law schools is higher than Ph.D. programs in economics, so Firm B might be a law firm, which can attract female lawyers easily, and Firm A might be an economic forecasting firm, which has difficulty attracting female economists.

![Fig 2: Production Costs & Female Workers](image)

Now suppose that government wants to increase female employment in the two firms to fifty percent. One way to do so is by imposing a fifty percent target on both firms. However, replacing male with female workers may cost one firm more than the other. Alternatively, the policy goal could be fifty percent female employment on average over the two firms. Implementing this goal at least cost might require one firm to have more than fifty percent female workers and the other to have less.

The two approaches are contrasted by extending the previous example. To keep the numbers simple, assume that each firm employs one hundred people. Thus, Firm A employs thirty women and Firm B employs forty women in the initial situation without regulation, for total employment of seventy women. The government’s target requires female employment of one hundred, or an increase in female employment of thirty. The horizontal axis in Figure 3 measures the
increase in female employment in Firm A from left to right, and the 
increase in Firm B is measured from right to left. At the extreme 
left, Firm A replaces zero males with females and Firm B replaces 
three males with females. At the extreme right, the opposite applies — 
Firm A adds thirty females and Firm B adds zero females. Con- 
sequently, each point on the horizontal axis represents a different 
way of distributing an increase of thirty in total female employment 
between the two firms.

Figure 3: Distributing Additional 
Female Workers Across Firms

Figure 3 permits a direct comparison between uniform and cost-
minimizing targets. According to the figure, the cost-minimizing way 
to distribute the increase of thirty female workers between the two 
firms is indicated by the intersection of the two cost curves, which is 
the point where the marginal cost of adding an additional female 
worker is the same for both firms. At the cost minimizing point, 
Firm A replaces five male workers with females, and Firm B re-
places twenty-five male workers with females. Thus, the minimal 
cost of adding thirty additional female workers is indicated by the 
areas I+II+III in Figure 3.

However, if a uniform target is imposed, Firm A must increase its 
female employment by twenty and Firm B must increase its female 
employment by ten, even though Firm B can replace male with fe-
male employees at a lower marginal cost than Firm A. The total cost 
of adding thirty additional female workers by uniform targets is 
I+II+III+IV. Compared to the minimum cost approach, targets
achieve the goal of adding thirty female workers at extra cost corresponding to area IV.

Ideally, government administrators would respond to differences in labor markets by adjusting the targets to take account of the pool of qualified workers and other labor market conditions. For example, the female target for Firm B would be higher than for Firm A to reflect the greater number of women in law as opposed to economics. However, the economic critique of regulation builds upon the insight that the best regulations for society are seldom best for the regulators, and even officials who have incentives to impose the best regulations seldom have enough information. To illustrate, officials might impose fifty percent targets on each firm because doing so maximizes the need for regulators, or because the officials lack the necessary information to adjust the targets in response to conditions in local labor markets.

In contrast, “market-like instruments” of regulation economize on the amount of information that officials require to enforce them. In terms of Figure 3, market-like regulation will induce Firm A to add five female workers and Firm B to add twenty-five. The mechanisms that can achieve this goal include targets supported by tax-subsidies or transferable employment rights. I will explain both of them, without discussing the formidable practical obstacles to their adoption.

The tax-subsidy solution requires each firm to pay a tax on “excess” workers of the disfavored type and receive a subsidy for “surplus” workers of the favored type. To illustrate by our previous example, assume that policy makers adopt the target of fifty percent female workers in the combined professions of law and economics. Each firm would receive a target of fifty percent females in its workforce in these professions. A firm that fell short of the target would be charged a tax on its excess male workers, and a firm that exceeded the target would receive a subsidy for its surplus female workers. For example, if Firm A is an economic forecasting company that employs one hundred economists and thirty of them are female, then it must pay a tax on its twenty “excess” male economists. If the tax is high enough, Firm A will respond by replacing male economists with females. In general, the higher the tax-subsidy, the more male employees will be replaced with females. It is not

30. For example, the University of California at Berkeley sets employment targets for faculty hiring in each department based upon the pool of qualified graduates in the field. Thus, the law school has higher targets for women than the economics department.
hard to see that a tax-subsidy rate exists, and can be identified easily, which achieves any target for increased female employment at the least cost to society.

To illustrate, return to the example in which the goal is adding thirty female workers in Firms A and B, and each firm is assigned a target of fifty percent female employees in professional jobs. The tax must be paid on shortfalls relative to the target, and the subsidy will be received for surpluses relative to the target. Initially the tax-subsidy is set at a low level and gradually raised until the goal of thirty additional female workers is achieved. Figure 3 indicates that this goal will be achieved by setting the tax-subsidy at $p^*$. To see why, consider the response of Firm A and Firm B when the tax-subsidy equals $p^*$. If the tax is set at $p^*$, Firm A will prefer to add up to five female workers to avoid paying the tax. However, after it adds five female workers, it is cheaper to pay the tax than to add more females. Consequently, after adding five females, Firm A will pay a tax on the remaining fifteen “excess” male employees, for total tax payments of $15p^*$. Similarly, Firm B will add ten additional females to achieve its target and avoid paying the tax. However, Firm B will then take advantage of the subsidy by hiring additional females out to the point where the additional cost equals the subsidy, which occurs when Firm B adds a total of twenty-five females. Thus, Firm B will receive a subsidy of $p^*$ on its fifteen “surplus” female employees, for a total subsidy of $15p^*$. Notice that in equilibrium the total taxes paid by Firm A equal the total subsidies enjoyed by Firm B.

An alternative policy that achieves the same results in principle is to create a transferable right to employ disfavored workers. Firms would be required by law to own as many legal rights as they employ workers of the disfavored type. To illustrate by our previous example, assume that policy makers adopt the target of fifty percent female workers in the combined professions of law and economics. Thus, the regulators must create a total number of rights to employ male lawyers or economists equal to fifty percent of the current workers in those jobs. These rights must be allocated initially, say by gift or auction. It matters little from an efficiency perspective how they are initially allocated. Once allocated, firms would buy and sell the rights.

31. In a bitter satire, Derrick Bell imagines the effects of a government bill authorizing whites to purchase permits allowing them to discriminate against blacks in business. The purpose served by the imagined permits is to allow for the satisfaction of discriminatory preferences at a price. See Derrick Bell, Faces at the Bottom of the Well: The Permanence of Racism 47-64 (1992). Jerry Mashaw contemplates a market to discriminate, analogous to a market to pollute. See Jerry L. Mashaw, Against First Principles, 31 San Diego L. Rev. 211, 231-37 (1994); Mashaw, supra note 29.
To illustrate concretely, suppose that regulators initially give Firm A and Firm B male employment rights equal to fifty percent of their workers. Thus, each firm receives the right to employ fifty males, for a total of one hundred males. However, the two firms initially employ 130 males combined. To comply with the law, the two firms must now replace thirty male workers with females. The firms will bargain with each other to try to accomplish this goal in the cheapest way. If bargaining replicates a competitive market, Firm A will offer to pay $p^*$ to Firm B for each male employment right. Firm B will respond by selling the right to employ fifteen males to Firm A. So in equilibrium, Firm A will increase its female workers by five and purchase fifteen male employment rights from Firm B, whereas Firm B will increase its female workers by twenty-five and sell fifteen male employment rights to Firm A.

Notice that tax-subsidies and transferable employment rights achieve the identical outcome. Specifically, Firm A increases female employment by five, Firm B increases female employment by twenty-five, and Firm A pays 15$p^*$ to Firm B.  

Apparently, a tax-subsidy scheme is in place in Germany with respect to employment of the handicapped. Employers are assigned a target for hiring disabled workers based upon the proportion of handicapped in the working population. Employers who fail to hire their target of the handicapped make payments to the government which passes the funds along to employers who hire more than their share. New Jersey has experimented with a similar approach to low income housing located in middle class communities.

VI. DISCRIMINATORY POWER

I have discussed discrimination in markets with perfect competition and the achievement of affirmative action targets in perfectly competitive labor markets. Although the perfectly competitive model
describes powerful forces at work in the economy, part of the historical differences in wages between blacks and whites, or between women and men, are usually attributed to discrimination. Rather than confirming the prediction that discriminators paid for it, empirical studies suggest that the targets of discrimination in the United States historically received lower wages than others with equivalent skills, and that civil rights laws helped raise the income of blacks.\textsuperscript{35} Given the evidence, discriminatory practices in this country cannot be explained fully by the model of perfect competition.\textsuperscript{36}

In subsequent sections of this Article, I will consider several market failures that might explain how discriminators shift the burden of segregation to its victims. First I develop a model of discrimination based upon power, not competition. Just as producers collude to fix prices and obtain monopoly profits, so social groups sometimes collude to obtain the advantages of monopoly control over markets. To enjoy the advantages of monopoly, a social group must reduce competition from others by excluding them from markets. In this way, the more powerful social group can shift the cost of segregation to its victims, and more costs besides, so that the victims of discrimination are worse off and the discriminators are better off.

To illustrate, recall the hypothetical example in which some blues discriminate against greens, and perfect competition causes the discriminatory blue workers to bear the cost of segregation. Now suppose that discriminatory blue workers organize themselves and acquire enough power to disrupt the workplace. The blues could use this power to threaten employers who failed to discriminate against greens. Faced with the power of the blues, employers might find that they could maximize their profits by avoiding disruption, even at the cost of segregating workers and confining greens to lower level jobs. This example describes circumstances in which segregation reduces productivity and its victims bear the cost.

The consequences of discriminatory power in the market for skilled and unskilled labor are depicted in Figure 4. The demand for skilled labor is indicated by the curve labelled $D_s$, and the supply of skilled labor by greens, blues, and the sum of greens and blues, is

\textsuperscript{35} The empirical evidence is reviewed in \textit{Forbidden Grounds}, \textit{supra} note 10, at 242-66. In the usual approach, "human capital" variables explain part of the wage-gap between sexes or races, and the unexplained residual is attributed to discrimination. Gillian Hadfield has suggested that the "human capital" variables should be viewed as results of discrimination in non-market institutions such as families and schools. See Gillian K. Hadfield, \textit{Households at Work: Beyond Labor Market Policies to Remedy the Gender Gap}, 82 Geo. L.J. 89 (1993).

\textsuperscript{36} Martin Katz argues that past discrimination could produce wage-disparities between the races that are relatively permanent, in spite of high levels of competition. These conclusions depend upon special assumptions such as barriers to entry. See Martin J. Katz, \textit{The Economics of Discrimination: The Three Fallacies of Croson}, 100 Yale L.J. 1033 (1991).
indicated by the curves $S_g$, $S_b$, and $S_g + S_b$, respectively. In the absence of discrimination, the wage for all skilled workers equals $w$. The demand for unskilled labor is indicated by the curve labelled $D_u$, and the supply of unskilled labor (blue and green) is indicated by the curve $S_u$. In the absence of discrimination, the wage for unskilled workers is $w_u$.

Now consider how discrimination changes wages. If skilled blues exclude skilled greens from the market for skilled labor, the supply falls from $S_g + S_b$ to $S_b$, and the skilled wage rises to $w'$. Discrimination forces greens to work as unskilled labor. The additional greens entering the unskilled labor market swell the supply from $S_u$ to $S'_u$, which causes wages to fall from $w_u$ to $w'_u$. Thus, discrimination increases wages for skilled blue workers, and lowers wages for unskilled blues and all greens.

**Fig 4: Discriminatory Power**

Discrimination as depicted in Figure 4 divides blues against each other by increasing the wage of skilled blues and decreasing the wage of unskilled blues. However, the unskilled blues could also use discrimination to their advantage if they obtained power in the market for unskilled labor. For example, unskilled blues might distinguish the tasks of unskilled labor into two types, “blue work” and “green work.” If more greens seek unskilled work than blues, and if
demand is higher for “blue work” than for “green work”, then segre-
gating tasks will cause the wage of unskilled blue workers to rise
above the wage of unskilled green workers.

This market analysis can be applied to Title VII of the Civil
Rights Act of 196437 and the Age Discrimination in Employment
Act (ADEA).38 These laws prohibit employment discrimination
based on race, color, religion, sex, national origin, or age. In prac-
tice, most workers fall within its protection except for young white
males. Complaints of discrimination must be filed with the Equal
Employment Opportunity Commission which investigates them.
Thus, for example, a company that wrongfully denied a job to some-
one two years ago might be ordered to hire the person and pay compen-
sation equal to the difference between his current wage and the
higher wage in the better job for two years. In terms of Figure 4, a
skilled green worker who was forced to accept unskilled employment
could sue for the difference between \( w_s \) and \( w_u \). The fact that the law
limited damages to back-pay discouraged lawyers from taking small
cases on a contingency fee, but their reluctance may be overcome by
revisions in Title VII in 199139 that brought employment discrimina-
tion closer to tort law by broadening damages.40

40. My thanks to Drew Days and Pam Karlan for helping me get the facts right
about current law. Donohue's analysis of data on Title VII suits discloses a paradox that
developed with time. Employment discrimination against women and minorities almost
certainly declined between 1970 and 1989, yet filings of employment discrimination cases
increased over 20 times. See John J. Donohue III & Peter Siegelman, The Changing
happened? Many of the original suits were brought against hiring practices that discrim-
inated against classes of people. For example, blacks were effectively prevented from
becoming firemen in Birmingham, Alabama. See Samuel Issacharoff, When Substance
Mandates Procedure: Martin v. Wilks and the Rights of Vested Incumbents in Civil
Rights Consent Decrees, 77 CORNELL L. REV. 189, 195 (1992). The success of these suits
and the abatement of discrimination for other reasons causes more minorities and women
to move into better jobs and more integrated work environments. This progress towards
opportunities greatly increased the possibility for a new wrong: discriminatory firings
or lay-offs. So a new tort was born, “wrongful discharge,” which can be asserted under
Title VII or the ADEA by a person who is fired. As time passed, the character of Title
VII complaints changed from discriminatory hiring of classes to discriminatory firing of
individuals. Econometrics shows that a good predictor of the number of employment dis-
crimination suits against firms at any point in time is the number of their employees
protected under law and the state of the economy. The more protected employees and the
worse the state of the economy, the more challenges against layoffs as discriminatory.
Donohue III & Siegelman, supra.
VII. ANTI-DISCRIMINATION AS ANTITRUST

In general, a group with the power to reduce competition from others can benefit itself, whether the group is defined by race, religion, gender, or industry. Discriminatory social groups are much like cartels, and a discriminatory norm is analogous to a price-fixing agreement. Thus, the analysis and attack upon discriminatory market power can borrow much from monopoly theory and antitrust law. I will develop this parallel.41

Cartels are unstable because each member can increase its profits by defecting from the group. For example, the Organization of Petroleum Exporting Countries (OPEC) tried to fix prices, but countries like Algeria secretly discounted oil in order to sell more of it.42 As a cartel becomes large, detecting and preventing such "cheating" by members becomes harder. Without legal backing and formal enforcement of their agreements, large cartels like OPEC collapse.43

Similarly, social groups can exert power to increase their wages by restricting competition in the labor market, but individuals can profit from violating the restrictions. To illustrate by the example in Figure 4, individual blue workers who cease to support segregation actively can enjoy the discriminatory wage $w'_s$ and avoid the inconvenience, expense, or danger of participating in industrial disruptions. In economic jargon, individual blues have an incentive to "free ride" with respect to discriminatory norms by withholding enforcement effort. Furthermore, an employer can reduce his wage bill for skilled workers by hiring greens, who would otherwise have to work at unskilled jobs, and paying a wage more than $w'_u$ and less than $w_s$. So the self-interest of employers and blues as individuals does not prompt them to sustain discriminatory norms. In general, sustaining discriminatory norms requires the collusion of many people, which presupposes sanctions to enforce the discriminatory norms. Informal sanctions such as gossip, ostracism, and boycotts can operate spontaneously,
especially when a culture stresses group solidarity. In the past, many Americans used informal sanctions to punish individuals who failed to keep the races separate or women "in their place." However, the informal sanctions were probably not enough to sustain segregation without being buttressed by formal laws.

Although cartels are inherently unstable, the U.S. antitrust framework does not merely withhold enforcement from contracts to create cartels. In addition, the original U.S. legislation, which was enacted at the end of the nineteenth century, outlaws cartels and other "conspiracies against trade." The courts have interpreted the law to prohibit certain collusive practices ("per se prohibitions"), such as retail price maintenance, regardless of whether collusion occurred in fact. These prohibitions greatly increase the difficulty of sustaining a cartel. Similarly, U.S. civil rights laws prohibit business practices involving "disparate treatment" of those persons belonging to any one or more protected classes. The illegality of conducting certain business transactions with the intent to discriminate greatly increases the difficulties involved in explicit discrimination, especially in large organizations.

Over the years, the effective scope of antitrust law expanded from banning cartels to suppressing monopolies. A monopoly can arise even without collusion or engaging in practices that are illegal per se. For example, monopoly power can be "thrust upon" a producer due to economies of scale in production. Such monopolies are evaluated for their legality in the United States according to a balancing test. The balancing test is intended to determine whether the savings in cost from scale economies outweigh the risk to the public of having only one or two producers. Balancing tests have their own history that I cannot discuss in detail, but a relevant episode is the.

48. There is a certain irony in the fact that large cartels are easiest for the government to detect and prosecute, but large cartels also are the ones that tend to fall apart even without government prosecution.
51. A compact and lucid discussion of the balancing test is in National Bancard Corp., 596 F. Supp. at 1256-68.
rise and fall of the "structural approach." In the 1970s, the anti-
trust authorities prosecuted some very large manufacturers with the
intention of restructuring whole industries in order to increase the
number of producers. However, this approach was deemed a failure
and abandoned in the 1980s for a variety of reasons.

Antidiscrimination law has a history with some similarity to anti-
trust law. At first the government focused its prosecutorial efforts
upon explicitly discriminatory practices. In these cases, the plaintiff
had to prove the existence of disparate treatment by the employer.
In 1971, however, the law evolved further and the U.S. Supreme
Court developed the concept of "disparate impact." A practice can
have an illegal disparate impact in the absence of discriminatory in-
tent. The illegality of the outcome is identified by a pattern sug-
gesting that a protected group has been unreasonably disadvantaged
by a business practice. Thus, the concept of disparate impact in an-
tidiscrimination law bears a certain resemblance to "monopoly struc-
ture" in antitrust law. Whether the courts and the political process
will deem the past twenty-three years of disparate impact analysis a
success or a failure remains to be seen.

When antitrust laws block cartels, the industry may try to circum-
vent the law through the help of regulators. For example, airlines are

52. For accounts of the "structural approach," see Joe S. Bain, International
Differences in Industrial Structure: Eight Nations in the 1950s 134-54 (1966);
Richard E. Caves, American Industry: Structure, Conduct, Performance 17-37
(4th ed. 1977); F.M. Scherer, Industrial Market Structure and Economic Per-
formance 81-150 (2d ed. 1980).

53. The abandonment of the structural approach to antitrust adjudication in the
1980s is sometimes called the "victory of the Chicago school." See, e.g., Andrew I. Gavil,
Teaching Antitrust Law in Its Second Century: In Search of the Ultimate Antitrust
changes in the FTC's merger guidelines in 1982. FTC Statement Concerning Horiz-
ontal Mergers, reprinted in 4 Trade Reg. Rep. (CCH) ¶ 13,200, at 20,901-06 (June
15, 1993). Another indication of change is successive editions of Scherer's book Indus-
trial Market Structure and Economic Performance. Scherer, supra note 52. Also, the
ABA's Section of Antitrust Law has discussions on current developments in antitrust in
its volume on the annual meeting. See, e.g., Symposium, The State of Antitrust, 61

54. I have stressed the role of discrimination by a group in reducing competition
from others. Another possibility is that a firm with market power will use a trait like race
or sex to discriminate in pricing among consumers. The latter type of discrimination
would fall under the ban of the Federal Trade Act. More generally, Ian Ayres has sug-
gested that antitrust law could be used to attack discrimination as an "unfair or decep-
tive trade practice." Ian Ayres, Fair Driving: Gender and Race Discrimination in Retail

development of the law is in chapter 10 of Epstein's book. Forbidden Grounds,
supra note 10, at 182-204.
forbidden to collude in setting prices, but they had much influence with the Civil Aeronautics Board and apparently used it to impose the cartel price upon many routes. Similarly, southern whites actively used the power of state and local government to reduce competition from blacks through the “Jim Crow” legislation that was enacted in the closing decades of the nineteenth century.

Antidiscrimination laws can also be used to benefit a social group by reducing competition from others. To illustrate by Figure 4, suppose the greens, who were the historic victims of discrimination in our hypothetical example, acquire legislative power and enact laws mandating preferential hiring of greens. For example, the law might mandate filling job openings in various categories with greens until sixty percent of the workers are green. (Perhaps sixty percent of the population is green.) In job categories where the target binds, blues cannot compete with greens for jobs, which causes the green wage to rise above the blue wage for equivalent workers. These arguments underlie the claim that affirmative action is reverse discrimination.

The phrase “rent seeking” refers to the efforts of people to secure laws that convey monopoly power and profits upon themselves. A standard prescription for preventing rent-seeking is to remove the issue from ordinary politics by constitutionalizing it. For example, constitutional guarantees of private property inhibit politically influential people from using the state to appropriate the property of others for themselves. Similarly, constitutional guarantees against discrimination can reduce rent-seeking by social groups. On the other hand, the creation of vague and uncertain constitutional rights by courts can unleash extensive rent-seeking through litigation.

I have shown that social groups, including racial and ethnic groups, are paradigmatic interest groups in many respects. Like other interest groups, they seek to collude and redistribute wealth to themselves by inefficient restrictions on competition. However, self-interest and morality often prompt individuals to evade these restrictions. So discriminatory social groups suffer the same problems of instability as any other cartel. To sustain discriminatory norms, evaders must be punished by a combination of informal sanctions and formal laws. By repealing these laws and undermining these sanctions, law can cause the discriminatory norms to disintegrate. Constitutional protection against discrimination, like constitutional

57. FORBIDDEN GROUNDS, supra note 10, at 91-97.
59. Regulatory takings as rent-seeking is a theme in Richard Epstein’s book. Id.
guarantees of property, can facilitate competition and preclude wasteful efforts to redistribute income among social groups by political means.

VIII. DISCRIMINATORY SIGNALS AND ASYMMETRICAL INFORMATION

I first considered discrimination in the context of competition, and then I considered market power. Now I consider a different kind of market imperfection, specifically, imperfect information on the part of buyers and sellers. To understand the problem of imperfect information, I begin with a familiar example concerning insurance against automobile accidents. Insurance companies classify drivers into broad groups and set premiums according to the probability that the average driver will have an accident. For example, young drivers cause more accidents on average than old drivers, and young males cause more accidents on average than young females. The gender and age of policy holders, which are cheap for insurance companies to discover, predict the riskiness of drivers with sufficient accuracy to be useful for setting insurance rates. So insurance companies charge higher premiums for being young and male.

"Good signal" is the name economists give to a characteristic that predicts accurately on average and is cheap to observe. In transactions with imperfect information, the parties search for good signals to reduce their uncertainty. Examples of good signals include the smell of a peach, the weight of a football player, the megahertz of a computer chip, the class rank of a law student, the rating of a bond, and the brand name of an automobile.

Now I turn to signaling in labor markets. Just as insurance companies know little about individual policy holders, so employers know little about job applicants. In choosing among them, employers rely upon signals to predict performance. For example, a job applicant with a college degree can easily provide the employer with a copy of his transcript. The college degree may signal traits like intelligence that the employer values. Education effectively signals intelligence because more intelligent people can acquire education more easily and cheaply than less intelligent people.

The original models of job-market signaling concerned the "rat-
race” that could arise when the signal had no intrinsic value. For example, suppose that certain employers value native intelligence but not education. Students might over-invest in education, not to learn anything useful, but merely to signal their intelligence (“credentials race”). In the rat-race models, people over-invest in an observable variable to signal a fixed trait.

Discriminatory signaling inverts the “rat-race” models. In discriminatory signaling, a fixed trait like gender or race signals an unobserved variable. To illustrate, men are physically stronger than women on average, so some employers reject all female applicants for jobs requiring strength. By adopting such policies, an employer will often make mistakes like rejecting a strong woman and accepting a weak man, just as an automobile insurance company sometimes over-charges safe males and under-charges dangerous females. If these mistakes cost less than gathering more individualized information, the use of the signal maximizes profits, and competition will reinforce the discriminatory practice. Conversely, if the cost of these mistakes exceeds the cost of gathering more individualized information, then the use of the signal is inefficient and competition will eliminate it.

People cannot acquire a fixed trait. To illustrate, few people will alter their race or sex in order to improve their job prospects, even with modern surgery. Consequently, discriminatory signals cannot produce a wasteful rat-race. Instead, sorting by traits can produce the opposite effect: under-investment in human capital. If employers attribute to each individual the average productivity of members of the group having his traits, then the benefit of investments that increase an individual's productivity will accrue in part to the group. Consequently, each individual will have a tendency to under-invest in acquiring productivity-increasing skills. The tendency to under-invest may be strongest in groups that are the victims of discrimination.

Although discriminatory signals can cause inefficiency, the usual objection to them is that they are unfair. The unfairness consists in judging individuals by averages. Suppose that government prohibits employers from using certain signals. For example, a statute might give strong women the right to sue employers who hire men exclusively for jobs that require strength. If the prohibited signals are inefficient, the law bans what competition will eliminate anyway. If the prohibited signals are efficient, the law augments the cost of production, which someone must bear. Competition drives the market

63. See Stewart Schwab, Is Statistical Discrimination Efficient?, 76 AM. ECON. REV. 228 (1986); see also Is Title VII Efficient?, supra note 12.
price of a good down to the cost of producing it. Thus, a reduction in an industry's efficiency typically causes the consumers of its products to pay higher prices.

To illustrate, assume that gender efficiently signals the physical strength of job applicants. If the law bans the use of this signal and the prohibition is effective, rational employers will adopt the best substitute for the banned signal. The best substitute may be a direct measure of physical strength, or the best substitute may be another signal, such as the applicant's height, weight, and age. In any case, competition will pass the higher cost of sorting job applicants on to consumers in the form of higher product prices.\(^4\)

Banning discriminatory signals is not usually the cheapest way to undo their effects, because a ban forces employers to rely upon a more costly substitute for every job applicant. To illustrate, assume that gender efficiently signals strength, and a ban on its use compels employers to test strength directly. The direct test will have to be administered to every job applicant, including all the males, which will increase costs significantly.

In the absence of regulations, the victims of discriminatory signals may have private remedies. To illustrate by the preceding example, if employers have no irrational prejudice against hiring women for jobs requiring strength, but gender efficiently signals strength, then strong women would probably find it in their interests to undergo direct tests and provide employers with the results. Consequently, selective use of direct testing would occur without government intervention in the labor market.\(^5\)

The objection to non-intervention in this example is that female applicants would have to bear the cost of a test that men need not

\[^{64}\text{Behind these remarks lies a complicated incidence theory developed in public finance. For a survey, see Richard A. Musgrave & Peggy B. Musgrave, Public Finance in Theory and Practice 171-98 (2d ed. 1976).}\]

\[^{65}\text{Another solution to discriminatory signaling is inhibited by current antidiscrimination laws. To see why, consider the case of the smart, lazy law student with mediocre grades who becomes inflamed with ambition after graduation. To overcome the signal given by his mediocre record, he offers to work for half pay for a year if a major law firm will give him a chance. At the end of the year, the firm can either fire him or retain him at the regular wage. Thus, the law student assumes the risk that his record signals to prospective employers. Similarly, the victims of discriminatory signaling might want to assume the employer's risk in order to be given a chance. For example, a female lawyer might offer to work for lower wages for a year in order to demonstrate that she can be just as effective as males. Title VII and the ADEA have created a tort of wrongful discharge that members of the protected classes can assert against employers. The existence of this tort greatly increases the risk to an employer who wants to give people the chance to prove their worth.}\]
take. Regulations would be required to overcome this objection. The employer might be required to test directly the strength of any applicant requesting it. Or the state might provide direct tests of strength without charge. In general, the economic strategy for correcting discriminatory signals is to increase the flow of information to the market so that relying upon them is unnecessary. This approach usually suggests a cheaper solution than banning the use of the discriminatory signal. The savings in cost is obtained by augmenting the information on potential victims of discrimination, without forcing the gathering of more information on everyone.

Having discussed signaling in markets, I offer some remarks about signaling in organizations. Most organizations contain a hierarchy of roles, and the members compete to rise in the hierarchy. Competition for promotion is called an “internal labor market” because it is so similar to competition for employment. An employee who works long enough in an organization gets to know its people and practices, which may enable him to produce more than he could in alternative employment. Consequently, an employee’s wage often exceeds what he could earn in the next best employment. Organization-specific knowledge ties the worker to the firm and the firm to the worker. Instead of quitting, a disgruntled worker often complains. Assessing the merit of such complaints is difficult because productivity is difficult to measure in organizations. Imperfect information about productivity causes the same reliance upon signals in promoting as in hiring. Consequently, the analysis of market signaling developed above for job applicants also applies to promotions.

Many social critics believe that decision-makers frequently rely upon false signals that reflect social stereotypes, not accurate averages. Competition can teach a sharp lesson to businesses that rely upon false signals. Decision-makers whose prosperity depends upon the accuracy of their perceptions are better situated than social critics or legislators to penetrate myths. However, competitive pressures are blunted in many organizations, especially in the public sector or the not-for-profit sector. This is where economic theory predicts that false signals are most likely to persist.

IX. EXTERNALITIES AND TIPPING POINTS

When each person’s action depends upon what others do, the interdependency of behavior can create instabilities. To illustrate, if each buffalo follows the one in front, the whole herd may run over the cliff. Similarly, “white flight” can destabilize an integrated school or neighborhood. In this section I analyze the instabilities created by interdependent preferences towards mixed social groups.
I begin with a simple model of white flight. Assume that an all white neighborhood consists of one hundred families who can be ranked according to their attitudes towards residential integration. At one end of the ranking, the one hundredth white family would move out of the neighborhood if one black family moved in. Similarly, the ninety-ninth family would move out if two black families moved in. Proceeding down the ranking, the first family would move out when ninety-nine black families had moved in.

I also assume that blacks have a continuous distribution of attitudes towards living in neighborhoods with whites. Some black families would be willing to move into an all-white neighborhood, many black families would be willing to live in an integrated neighborhood, and some black families would be unwilling to live in a neighborhood with any whites.

Now assume that demand and supply in the housing market is such that whenever a house becomes vacant in a particular white neighborhood, more black families want to buy it than white families. It is not hard to see that if one black family moves into the white neighborhood, a process will be set in motion that may not end until all whites have moved out. Specifically, if one black family moves in, the one hundredth white family will move out. Now the house of the one hundredth white family must be sold. The buyer of the vacant house is more likely to be black than white, so it is likely that two black families will now reside in the neighborhood. As a result, the ninety-ninth white family will move out. Now it is likely that there will be three black families in the neighborhood and the ninety-eighth white family will move out. The process continues until the neighborhood is all black.

The tragedy of this situation is that many of the whites in the neighborhood may positively value residential integration, and many of the blacks who move in may also value it. In spite of much positive sentiment favoring integration, it is unachievable by unrestricted sales in a free market. Instead, the integrated neighborhood inexorably unwinds and becomes segregated.

Notice that this model’s dynamics make no particular assumptions about the source of attitudes towards integration. For example, the attitudes of whites or blacks in the model may reflect skin prejudice, cultural pride, class consciousness, or fears about housing values. The dynamics of the model are the same regardless of the attitudes

66. This model is based upon Thomas C. Schelling, Sorting and Mixing: Race and Sex, in Micromotives and Macrobehavior 135 (1978).
underlying expressed preferences towards residential integration.

The most familiar economic models compare equilibria ("comparative statics"), whereas this model describes a dynamic path. Indeed, this model probably describes the actual dynamic in American cities after World War II when many neighborhoods went from all white to all black. It also probably describes forces at work in many public schools today. Since the model is unfamiliar and important, I will develop it better with the help of a graph.

Figure 5 graphs the attitudes towards integration of residents in a neighborhood that is all white. The horizontal axis in Figure 5 shows the proportion of white residents in a neighborhood who plan to move out. The vertical axis shows the proportion of black residents. Thus, the graph shows the proportion of white residents who would plan to move out as a function of the proportion of black residents who move in.

To illustrate its interpretation, suppose that the graph consisted of a single point at the northwest corner, corresponding to the value (0%, 100%). This would indicate that no white family in the neighborhood would plan to move out even if all the other residents were black. Conversely, suppose that the graph consisted of a single point at the southeast corner of the graph, corresponding to the point (100%, 0%). This would indicate that all whites in the neighborhood would plan to move out if only one percent of the neighborhood became black. The curved line in Figure 5 represents a more realistic case in which there is a continuous distribution of sentiment towards integration. I will analyze the dynamics of white flight created by the distribution of preferences represented by the curved line in Figure 5.

As the curve in Figure 5 is constructed, one hundred percent of the white residents would be willing to remain if less than twelve percent of the residents were black. Thus, up to twelve percent of the white families can move out and be replaced by blacks without provoking white flight. However, once the proportion of black residents reaches twelve percent, white flight begins, as can be seen by considering some points on the graph. Point A indicates that more than twelve percent of the white residents would plan to move out if the neighborhood were twelve percent black. This is an unstable situation in which integration starts to unwind. At point B, approximately fifty percent of the white residents would be planning to move out if twenty-five percent of the residents were black. At point C, seventy-five percent of the whites would be planning to move out if fifty percent of the residents were black. And so the flight goes on until no whites remain in the neighborhood.
To complete the model, another graph could be drawn showing the distribution of black preferences towards integration. In this graph, the vertical axis would show the proportion of whites in the neighborhood, and the horizontal axis would show the proportion of blacks who would be willing to move into the neighborhood. The model is left incomplete for the sake of simplicity.

A neighborhood characterized by Figure 5 is one in which eighty-eight percent or more whites is an unstable equilibrium, and zero percent whites is the only stable equilibrium. The stability conditions can be stated precisely in terms of the diagonal line in Figure 5. The model assumes that whenever a house becomes vacant, it is more likely to be purchased by a black than a white family. So long as the neighborhood is at a point where the curved line representing white attitudes towards integration lies below the diagonal line, more whites plan to move out in response to the existing proportion of black residents. This is a disequilibrium. There is an equilibrium at any point where the curved line representing white attitudes towards integration touches the diagonal line. To have a stable equilibrium, the curved line must intersect the diagonal line from below.
Tragic Segregation

I will describe a neighborhood as “spontaneously integrated” if the preferences of its residents can be graphed as a curve in Figure 5 that intersects the diagonal line from below. Otherwise, the neighborhood spontaneously segregates, as depicted by the curved line in Figure 5. Spontaneously segregated neighborhoods are like a tragic drama in which social laws lead inexorably to an end that no one wants. The outcome frustrates the desire of many people of both races to live in an integrated community.

One way to avoid this tragedy is the use of racially restrictive covenants (“ceiling quotas”) to stabilize integration. To illustrate, if fifty percent of the houses in the hypothetical neighborhood modelled in Figure 5 had enforceable deeds restricting ownership to whites, the neighborhood would have stabilized with fifty percent white families and fifty percent black families. Furthermore, the fifty percent white families and the fifty percent black families in the neighborhood would tend to be those with the most positive attitudes towards integration. However, racially restrictive covenants, which were used historically to keep blacks out rather than keeping whites in, have been struck down by the federal courts.\(^6\)

The continuing legal obstacles faced by ceiling quotas are illustrated by the example of Starrett City, which is a private housing project in Brooklyn whose construction was partly financed by the federal government.\(^6\) In 1987 the project had approximately 20,000 tenants, with whites occupying nearly sixty-five percent of the 5881 apartments, and the other thirty-five percent being occupied by blacks and Hispanics. The managers had set a racial target of sixty-five percent white and thirty-five percent nonwhite, which they defended on the grounds that it was necessary to maintain a stable, integrated community. There were, however, fewer white applicants for vacant units than nonwhite applicants. The National Association for the Advancement of Colored People (NAACP) challenged these targets in a suit brought in 1979. The settlement reached in 1984 stipulated that allocation by race should continue, but 174 additional units should be made available to nonwhite applicants. This settlement was subsequently challenged by the Civil Rights Division of the Justice Department, which contended that any such system of racial quotas, even one whose purpose is integration, violates the

---

67. In Shelly v. Kraemer, 334 U.S. 1 (1948), the Supreme Court found unconstitutional a covenant in a deed prohibiting the sale of the property to Negroes.
Federal Fair Housing Act. The view of the Justice Department prevailed in federal court which ordered an end to the racial quotas.\(^6\)

The federal government position on Starrett City was to ban quotas, regardless of the consequences for white flight. The opposite approach was recently taken by the school board for Des Moines, Iowa.\(^7\) White flight allegedly began with implementation of a "choice" program that permitted students to transfer from one school to another. The school board responded by prohibiting transfers by white students and continuing to allow transfers by all other students. In effect, the school board sought to stabilize the existing racial mix in each school by setting a quota of zero on white transfers.

As these examples suggest, the law deals with racial mixing differently in neighborhoods and schools. Most neighborhoods consist of privately owned residences, and most American children attend public schools. Legal officials have tried to prevent discrimination in sales or leasing of residences, without aiming at any particular outcome with respect to residential integration. For schools, the courts and administrators have often defined ideal patterns of racial mixing and assigned students to schools in an effort to achieve these ideals.

Can economists devise a way to stem white flight other than relying upon targets? From the viewpoint of economic theory, white flight is caused by an externality consisting in individual preferences over the racial mix in neighborhoods or schools. Economists typically propose to remedy externalities with tax-subsidies or transferable rights. In principle, these remedies could be applied to racial mixing. Thus, a tax might be assessed on housing sales that worsen a neighborhood's racial mix, or a subsidy might be paid on housing sales that improve the neighborhood's racial mix. Similar devices could be developed for education, such as school vouchers that increase in value when the enrollment of a pupil in a particular school improves its racial balance.

Alternatively, a solution using transferable property rights might

---

69. A follow-up story reported that Starrett City was avoiding the effect of the court order by not filling any vacant apartments from its waiting list of applicants who fit the poverty criteria, black or white, but instead keeping them vacant until someone applied who was above the legal definition of poverty. The middle class applicants who were above the poverty line were disproportionately black. Alan Finder, Starrett City Keeps Apartments Vacant Despite Waiting List, N.Y. TIMES, July 14, 1990, § 1, at 1. A more complete history is in Howard Husock, Subsidizing Discrimination at Starrett City, CTRY J., Winter 1992, at 48.

be developed. It might work like this: racially restrictive rights of residence would be issued by the government to property owners in participating neighborhoods. The distribution of rights would correspond to the government’s ideal racial mix for the neighborhood. To sell a property in the neighborhood to a prospective buyer, the owner would have to possess a right of residence corresponding to the prospective buyer’s race. For example, if the best prospective buyer of a property were black, but the owner of it possessed a “white occupancy right,” the sale could not be consummated without buying a “black occupancy right” from someone else.

A theoretical possibility, which I have not discussed, is that unregulated markets with negative externalities from mixing result in excessive integration relative to the efficient level. I simply note in passing that taxes could also eliminate this inefficiency.

CONCLUSION

The economic theory of regulation suggests that some laws and policies concerned with discrimination and affirmative action stifle enterprise unnecessarily, so they should be repealed or reformed. This paper develops the general lines of the argument for reform and reaches the following conclusions: (i) competition causes discriminators to pay the cost of segregation, not the people discriminated against; (ii) in competitive equilibrium there will be diversity among organizations concerning diversity within organizations; (iii) social groups that act like cartels can reduce competition, which benefits themselves and harms others; (iv) the discriminatory power of social groups can be attacked by forbidding practices motivated by discriminatory intent, increasing competition with practices that have disparate impact, and repealing laws that reduce competition among social groups; (v) supplementing market information can often undo the effects of discriminatory signals more cheaply than prohibiting their use; and (vi) targets for the mixing of social groups, including stemming white flight and integrating neighborhoods or schools, can be achieved at least cost to the public by tax-subsidies or transferable rights.

Oscar Wilde once remarked that a cynic is someone who knows the price of everything and the value of nothing.\(^7\) This aphorism condemns economics to cynicism. However, knowing the price of public policies is an essential step in deciding their value, even when disagreements run as deep as religion. Current antidiscrimination

\(^7\) OSCAR WILDE, LADY WINDERMERE'S FAN Act 3 at 52-3 (Methuen Student Editions 1985) (1892). The character Cecil Graham asks “What is a cynic?” to which Lord Darlington responds, “A man who knows the price of everything and the value of nothing.”
laws impose various prohibitions upon economic transactions. Prohibiting transactions rather than pricing them prevents the development of a legal market, so the costs of the policy motivating the prohibition remain unknown. Tax-subsidies, transferable rights, or other forms of "market affirmative action" could partly solve this problem and others.\textsuperscript{72}

If the disagreement over policy remained within the bounds of economic science, disputants would respect the general framework of the economic theory of regulation developed in this article. There is, however, a decisive reason why this will not happen. Law condemns practices by prohibiting them. Many people regard the public condemnation of discrimination as vitally important to social justice and public education. Efficient remedies for discrimination do not prohibit and condemn enough of it to satisfy its victims. To address these concerns, economics needs a theory of the expressive and educative power of law.\textsuperscript{73}

The educative role of law is to change values by inducing people to internalize norms. From an economic perspective, a norm becomes "internalized" when disobeying it has a psychological cost to the individual. To illustrate, many people now refrain from smoking in public places, even though the laws are seldom enforced formally. Apparently, smoking ordinances have been internalized by many smokers. Economics needs a predictive theory of internalization in order to analyze the educative role of law in crucial areas such as discrimination.\textsuperscript{74}

Market environmentalism originally had no place in federal legislation. However, heavy moralism in environmental law eventually yielded to a more pragmatic approach, one that acknowledges a place for market-like instruments. Perhaps environmentalists could take a more pragmatic approach because they could leave the condemnation of pollution to parents, schools, and intellectuals, rather than the law. Like market environmentalism in 1970, market affirmative action has no present, but it may have a future. A massive

\textsuperscript{72} Space has constrained me from saying much about the important problem of protecting market affirmative action from political corruption. For an example of collective choice theory applied to discrimination, see Lynn A. Baker, Direct Democracy and Discrimination: A Public Choice Perspective, 67 CHI.-KENT L. REV. 707 (1991).


\textsuperscript{74} For the start to developing such a theory, see Kenneth G. Dau-Schmidt, An Economic Analysis of the Criminal Law as a Preference-Shaping Policy, 1990 DUKE L.J. 1, 23; see also Cass R. Sunstein, Endogenous Preferences, Environmental Law, 22 J. LEG. STUD. 217 (1993).
change in attitudes and values has occurred in America concerning discrimination and separation. These changes are proceeding independently of law under the direction of parents, schools, churches, and intellectuals. I suspect that these people and organizations are more effective than law in inducing people to internalize new values. If the opponents of discrimination and proponents of affirmative action become more confident that their values are shared by most Americans, perhaps they will become more pragmatic and critical about the means of achieving their ends.