WHY MARKETS DON’T STOP DISCRIMINATION

BY CASS R. SUNSTEIN

I. INTRODUCTION

Markets, it is sometimes said, are hard on discrimination. An employer who finds himself refusing to hire qualified blacks and women will, in the long run, lose out to those who are willing to draw from a broader labor pool. Employer discrimination amounts to a self-destructive “taste” – self-destructive because employers who indulge that taste add to the costs of doing business. Added costs can only hurt. To put it simply, bigots are weak competitors. The market will drive them out.

On this account, the persistence of employment discrimination on the basis of race and sex presents something of a puzzle. And if markets are an ally of equality and a foe of employment discrimination, perhaps discrimination persists because of something other than markets. Perhaps labor unions are to blame; perhaps the real culprit is the extensive federal regulation of the employment market, including minimum-wage and maximum-hour laws and unemployment compensation. If competitive markets drive out discrimination, the problem for current federal policy lies not in the absence of aggressive anti-discrimination law, but instead in the absence of truly competitive markets.

If this account is correct, the prescription for the future of anti-discrimination law is to seek ways to free up employers from the wide range of governmental disabilities – including, in fact, anti-discrimination law itself. The argument seems to be bolstered by the fact that some groups subject to past and present prejudice – most notably, Jews and Asian-Americans – have made substantial progress in employment at least in part because of the operation of competitive markets.

In this essay, however, I argue that under plausible assumptions and in many settings, markets will not stop discrimination and that reliance on competitive pressures would be a grave mistake for a government intending to eliminate discriminatory practices. Indeed, markets are often the problem rather than the solution. They guarantee that discrimination will persist. To say this is not at all to deny what is undoubtedly true, that a market system will under certain conditions drive out discrimination and that a cartelized industry or a government

1 See generally Clint Bolick, Changing Course: Civil Rights at the Crossroads (New Brunswick: Transaction, 1988). At the moment I will not define the protean term “discrimination” precisely; instead, I will rely on common-sense intuitions to the effect that the term prohibits differential treatment of one person because of his or her group membership. I deal here with discrimination, thus understood, on the basis of race, sex, and disability. Other bases for discrimination – age, sexual orientation, religious belief – raise overlapping questions, but they also call up considerations that would complicate the analysis here if they were to be discussed in detail.
WHY MARKETS DON’T STOP DISCRIMINATION

intent on perpetuating discrimination will have extremely significant advantages over decentralized markets in accomplishing that end. But it is to say that for identifiable reasons and purely as a matter of fact, the enthusiasm for markets as an anti-discrimination policy is at best wishful thinking. 2

This claim is not a normative one. It is solely explanatory, and I do not make policy recommendations here. But the account I offer of the relationship between markets and discrimination does have implications for the appropriate understanding of an anti-discrimination guarantee. I conclude with a few tentative remarks on that subject.

II. WHY MARKETS PRODUCE DISCRIMINATION: STANDARD ACCOUNTS

As a provisional matter, let us define discrimination as a decision by a market actor to treat one person differently from another because that person is black or female. It is irrelevant for present purposes whether the differential treatment is based on hostility, fear, taste, unconscious devaluation, selective empathy and indifference, employer self-interest, unwarranted generalizations, or an accurate perception of facts about group members. This definition has a degree of vagueness, and it is certainly contestable whether government should ban all of the conduct that it identifies. But the definition has the advantage of tracking both current law and at least some ordinary intuitions about the problem.

On one view, the notion that competitive markets act against discrimination, so defined, might seem odd or even idiosyncratic. In the history of the United States as well as other nations, markets and discrimination have accompanied one another at numerous points and in numerous places. Surely governmentally-imposed discrimination, because of its centralized character, has especially egregious effects; 3 but to point to the comparative strength of collective controls is not to say that decentralized markets work in the opposite direction. Indeed, there is good empirical evidence that government programs have succeeded in reducing discrimination produced by the market. The enactment of civil rights legislation appears to have led to increases in black and female employment in various sectors. At least in general, the disputed question has to do not with the direction of change but with the degree of the improvement. 4

Moreover, study after study has shown that the market often devalues the

2 Even if this is so, the argument for anti-discrimination measures is not complete. It would be possible to acknowledge that markets will perpetuate discrimination and also to argue that governmental controls do not solve the problem, or impose large ancillary causes, or are inefficient, or are unjust. I do not accept these arguments, but will not deal with them here. Instead, I focus on the more narrow claim that markets are an effective anti-discrimination policy.

3 Consider also the fact that nonmarket societies quite frequently reflect pervasive discrimination. Discriminatory attitudes can be comfortably accommodated in both market and nonmarket economies.

products and enterprises of both blacks and women. This should not be surprising. In a system with a significant amount of race and sex prejudice, covert and overt, conscious and unconscious, the ‘willingness to pay’ criterion – as it is reflected in purchasing and selling decisions of employers, employees, consumers, and others – will ensure that those subject to discriminatory attitudes will be at a comparative disadvantage in the market. At least this is so when the participation of disfavored groups is highly visible, as indeed it often is. In this context, the most natural initial judgment is that if discrimination is the problem, markets are hardly the solution.

But the judgment, as stated, seems too crude. Suppose a set of companies in a competitive industry – call them “D companies” – decides not to sell to or employ an identifiable segment of the community; suppose too that the D companies must compete with nondiscriminators engaged in the same line of work (“ND companies”). In the long run, the D companies should be at a significant disadvantage. They will be unable to draw on as large a pool of workers; they will be unable to sell to as large a group of purchasers. In practice, their discriminatory behavior will place a tax on their operations. To say the least, businesses that impose on themselves a tax not faced by their competitors are unlikely to fare well.

As patrons and prospective employees drift to ND companies, the D companies will be driven out. Eventually, the consequences should be nothing short of disastrous for discrimination in general. It is this understanding that accounts for the plausible claim that an effective anti-discrimination policy consists of reliance on decentralized markets.

Notwithstanding its plausibility (and even, in some settings, its accuracy), I want in this section to make three arguments against precisely this claim. In setting out these arguments, I make no claim to special originality. All of them can be found, at least in some form, in standard economic discussions of the problem.

A. Third parties. An influential approach to the problem of discrimination, captured in the account just offered, attributes the phenomenon to employers’ “taste” for discrimination. On this view, employers discriminate because they do not like to associate with blacks, women, or others. The engine that drives discrimination is employer preferences – not economic self-interest, narrowly defined. Indeed, economic self-interest is in conflict with employers’ taste for discrimination.

This idea has a degree of truth to it. But sometimes discrimination is caused


\[\text{This view finds its origins in Gary Becker, The Economics of Discrimination (Chicago: University of Chicago, 1957).}\]
WHY MARKETS DON'T STOP DISCRIMINATION

by employer practices that result precisely from economic self-interest, rather than from the employer’s non-economic goals. The reason is that third parties are frequently in a position to impose financial punishments on non-discriminatory employers. Suppose, for example, that purchasers or fellow employees refuse to buy from or work for a company that does not discriminate. Third parties can pressure employers in the direction of discrimination, even if employers would, other things being equal, choose not to discriminate or have no particular view about whether to discriminate or not. Ironically, it is the failure to discriminate that operates as a tax on the employer’s business, rather than vice versa. And when this is so, reliance on competitive pressures will force employers to behave in a discriminatory manner if they wish to survive.

The phenomenon is hardly unusual. Consider, for example, a shopkeeper whose customers do not like dealing with blacks or women; a commercial airline whose patrons react unfavorably to female pilots; a university whose students and alumni prefer a primarily male or white faculty; a law firm whose clients prefer not to have black lawyers; or a hospital whose patients are uncomfortable with female doctors or black nurses. The persistence of private segregation in major league baseball is a familiar example of the phenomenon.

In all of these cases, market pressures create rather than prevent discrimination. An employer who is intent on introducing norms of equality into the workforce will be punished, not rewarded. Indeed, an employer who wants to introduce such norms will be engaging in what is, from his point of view, a form of affirmative action. From his point of view, discrimination is in fact neutral, a quite ordinary form of profit-maximization through catering to consumer demand. From his point of view, a refusal to discriminate is economically harmful to him; if he is a non-discriminator, it is in order to promote the long-range goal of race and sex equality. If discrimination is to be eliminated, it must be a result of legal rules forbidding discrimination from taking root.

Ideas of precisely this sort played a role in the passage of the Civil Rights Act of 1964. Restaurants and hotels sought regulation constraining their own behavior by outlawing discrimination. Their goal was to obtain the force of the law to overcome the effects of private racism (in the form of violence as well as competitive sanctions). They sought to overcome, through law, the discrimination-producing consequences of market pressures.

On this account, it remains necessary to explain why third parties are not themselves hurt by their “taste.” At least part of the reason lies in the fact that third parties often are not market actors in the ordinary sense – that is, they will not suffer competitive injury in markets if they indulge their discriminatory preferences. Consider, for example, people who prefer not to fly in airplanes piloted by blacks or women (or having black or male stewards). At least in many cases, the harm suffered by people who indulge their prejudice – perhaps

---

8 Becker, *ibid.*, in fact makes this point, describing the “taste” for discrimination as including fellow workers and customers as well as employers.

higher prices of some form or another—will not drive them out of any “market,” but instead, and more simply, make them somewhat poorer.

In a competitive market that contains private racism and sexism, then, the existence of third-party pressures will create significant spheres of discrimination. To be sure, sometimes the participation of blacks, women, or the handicapped is invisible; in any case, racism and sexism are currently not so widespread as to make it costly for all employers to act in a non-discriminatory fashion. Because of their decentralized character, markets are less effective than governmental controls in perpetuating discrimination, as the experience in South Africa has (at least sometimes) showed. Moreover, third-party effects will probably produce a degree of occupational segregation instead of, or as well as, discrimination. 10

But discriminatory practices in many areas are likely to persist over time precisely because of the market—which by its very nature registers consumer preferences, including discriminatory ones.

There is an additional problem here. 11 Suppose that society includes a number of traders who want to interact with blacks or women as well as a number of traders who do not. Suppose, too, that it is difficult for traders who are themselves indifferent to race and sex distinctions to tell which traders are discriminators (a frequent phenomenon in competitive markets). If these conditions are met, discrimination is likely to be the rational strategy for traders. New companies will do poorly if they refuse to discriminate. All this suggests that markets will frequently have great difficulty in breaking down discrimination if discriminatory tastes are even somewhat widespread.

B. Statistical or economically rational discrimination. I have referred to the common-sense understanding that employers treat blacks and women differently because of irrational hostility or prejudice on their part—an idea that often accompanies the claim that markets are an enemy of discrimination. But the category of irrational prejudice is an ambiguous one. Perhaps we can understand it to include (a) a belief that members of a group have certain characteristics when in fact they do not, (b) a belief that many or most members of a group have certain characteristics when in fact only a few do, and (c) reliance on fairly accurate group-based generalizations when more accurate classifying devices are available. In all of these cases, it is possible to say that someone is acting on the basis of irrational prejudice.

Suppose instead, however, that discriminatory behavior is a response to generalizations or stereotypes that, although quite overbroad and from one point of view invidious, provide an economically rational basis for employment decisions. 12 Stereotypes and generalizations, of course, are a common ingredient of day-to-day market decisions. There are information costs in making distinctions

10 See Becker on this point.
WHY MARKETS DON’T STOP DISCRIMINATION

within categories, and sometimes people make the category do the work of a more individualized and often far more costly examination of the merits of the particular employee. Such categorical judgments are not only pervasive; they are entirely legitimate in most settings. We all depend on them every day. Employers rely on “proxies” of many sorts, even though those proxies are overbroad generalizations and far from entirely accurate. For example, test scores, employment records, level of education, and prestige of college attended are all part of rational employment decisions. Emphasis on these factors is a common form of stereotyping; at least in ordinary circumstances, it would be odd to say that such factors reflect hostility or prejudice of the sort that the law should ban.

There can be no question that at least in some contexts, race and sex operate as similar proxies. In various ways, blacks differ from whites and women differ from men; there are “real differences” between the two groups. Indeed, in light of past and present discrimination against blacks and women, it would be shocking if group-based differences were not present. Women, for example, are more likely than men to be the primary caretakers of children and more likely to leave the employment market because of that role.13 Along every indicator of social welfare – poverty, education, employment, vulnerability to private (or public) violence, participation in violent crime – blacks are less well-off than whites. And in light of those differences, it is fully possible that in certain settings, race- and sex-based generalizations are economically rational as proxies for relevant characteristics. Indeed, it is fully possible that race or sex is, in some contexts, every bit as accurate a signalling device as, say, test score, education, and previous employment. (Notably, the use of those proxies may itself have discriminatory consequences and indeed might be counted as a form of discrimination. But I put this point to one side for now.)

If race or sex can be a good signalling device, an employer might discriminate not because he hates or devalues blacks or women, or has a general desire not to associate with them, or is “prejudiced” in the ordinary sense, but because he believes (on the basis either of plausible assumptions or actual experience) that the relevant stereotypes have sufficient truth to be a basis for employment decisions. For example, an employer might believe that women are more likely to leave a high-pressure job than men, or that they are less suited for physically demanding employment; he might think that blacks are less likely to have the necessary training. With respect to the handicapped, the same scenario is of course readily predictable: “real differences” call for differential treatment. As I understand it here, statistical discrimination occurs when the employer does not harbor irrational hatred or discriminatory feelings, but instead acts according to stereotypes of the sort that are typically relied on by market actors, and that are no less false than are those ordinary stereotypes.

It is important to be careful with this point. I do not mean to deny that

13 This phenomenon, of course, does not mean that differential treatment is permissible. Whether such treatment should be allowed need hardly be made to depend solely on the fact that it is economically rational.
ordinary, irrational prejudice accounts for a significant amount of discrimination, or even that it is hard both in practice and in principle to distinguish between ordinary prejudice and economically rational stereotyping. Sometimes employers who refuse to follow a stereotype will find people who will move them far ahead in the marketplace. Moreover, the decision about what kinds of economically rational stereotypes should be adopted might itself reflect prejudice. People are highly selective in their creation and selection of the very categories with which they view people—as college graduates? as children of successful parents? as nonreligious? as white males? as tall? as good-looking? People are also selective in their decisions about which categories ought to count in various sectors. Thus, for example, the use of sex as a proxy for flexibility with respect to overtime work may be fully rational; however, there may be other proxies that are just as good or better. Further, selectivity in the choice of the proxy may be a product of prejudice. Finally, people tend to notice events that are consistent with previously-held stereotypes and to disregard events that are inconsistent with them. This is a built-in obstacle to changes in social norms, even when those norms increasingly fail to mirror reality.

One might agree with all this and still acknowledge what seems undoubtedly true: discrimination sometimes persists because it is economically rational to rely on a race- or sex-based generalization. And if this is so—if discrimination of this sort is a significant part of modern practice—then markets will be the furthest thing from a corrective. If discrimination is rational, then discriminatory behavior is rewarded in the marketplace. Efforts to view people through more finely-tuned (and, hence, more expensive) devices than stereotypes will be punished.

Moreover, if discrimination is rooted in economically rational stereotypes, it remains necessary to ask what is wrong with it. The (apparent) social consensus that prejudice is irrational tends to downplay the difficulty. Contemporary American law consistently condemns discrimination even if it is fully rational—that is, if it is as accurate as other generalizations typically used in the labor market. It is no defense to a claim of race or sex discrimination that the practice under attack is reasonable stereotyping. And because contemporary law singles out one kind of rational stereotyping and bans it, the distinction between affirmative action and anti-discrimination norms is extremely thin and perhaps invisible in principle. In some settings, an anti-discrimination norm itself operates to bar economically rational behavior in the interest of long-term social goals. An anti-discrimination norm requires innocent victims—for example, customers who must pay higher prices—to be sacrificed in the interest of that goal.

Indeed, the distinction between affirmative action and anti-discrimination is crisp only to those who see discrimination as always grounded in hostility and
irrationality, which it clearly is not. To say this is hardly to say that there are no differences between affirmative action and an anti-discrimination principle that forbids economically rational behavior. But it is to say that a great failure of the assault on affirmative action is its failure to account for the ways in which a requirement of non-discrimination involves very much the same considerations.

However that may be, the central point remains. Markets will not drive out discrimination to precisely the extent that discrimination consists of economically rational stereotyping.

C. The effects of discrimination on human capital. Suppose that there is widespread discrimination for any of a number of reasons: employers have a taste for discrimination; third parties impose pressures in discriminatory directions; race or sex is used as a proxy for productivity (whether or not productivity includes the losses produced by the reactions of third parties). If any of these things is so, markets will perpetuate discrimination for yet another reason, one having to do with harmful effects on human capital.\(^{17}\)

The central point here is that the productivity of blacks or whites and women or men is endogenous to, or a product of, discrimination in the labor market. Since this is so, it is a mistake to see productivity within any social group as static or independent of decisions by employers. Such decisions will have important dynamic effects on the choice of blacks and whites or women and men to invest in human capital. Decisions about education, training, drug use, tradeoffs between work and leisure, and employment programs will be affected by existing patterns of discrimination. In a market that contains such discrimination, blacks and women will invest relatively less in such programs. Indeed, lower investments on their part are perfectly rational. As market actors, women should invest less than men in training to be (say) pilots, economists, politicians, or lawyers if these professions discriminate against women and thus reward their investment less than that of men.

The result will often be a vicious circle or even spiral. Because of existing discrimination, the relevant groups will invest less in human capital; because of this lower investment, the discrimination will persist or perhaps increase because its statistical rationality itself increases; because of this effect, investments will decrease still further; and so on. Markets are the problem, not the solution.

Undoubtedly, this picture is too bleak in many settings. Proxies not rooted in sex and race have frequently evolved and been used; they are ordinarily far more accurate, and employers who use them sometimes prosper. More generally, the extraordinary persistence of blacks and women in attempting to enter professions dominated by whites and men is one of the most striking phenomena of the post-World War II period. Blacks and women frequently appear to invest huge amounts of human capital even in sectors that treat them inhospitably. In

fact, some people in some circumstances respond to discrimination by increasing rather than decreasing their investments in human capital. In Western history, some racial and ethnic groups have prospered as a result of such investments, notwithstanding discriminatory tastes on the part of employers, customers, and fellow employees. But there can be little question that discrimination does have a large effect on human capital in many contexts, and that the discouraging reception known to be accorded to blacks and women perpetuates the exclusion of both groups from certain sectors of the economy. And when this is so, reliance on markets as an anti-discrimination policy is badly misconceived.

D. In general: reinforcing effects. The three arguments thus far—coupled with the existence of race- and sex-based hostility or devaluation on the part of employers—build on standard economic accounts of discrimination. It is notable that none of the various effects reflects a market failure in the conventional sense. The incorporation of racist or sexist preferences is efficient, if the efficiency criterion is based on private willingness to pay, so too with profit-maximizing reactions to the desires of third parties and with statistical discrimination. In all of these cases, governmental interference will probably produce an efficiency loss, at least in the short term.

The only possible exception is the effect of discrimination on human capital. In that case, it is possible to conceive of discrimination as producing a serious externality, in the form of harmful effects on outsiders—prospective labor market entrants—with corresponding efficiency losses. But even here there is an optimal level of investment in human capital, and it is not entirely clear why and to what extent existing investments that are adaptive to employer behavior in the labor market should be seen as suboptimal. Any generalization that is relied on in the labor market will have a signalling effect that will shift investments in human capital, and it is not easy to develop a model revealing which signals produce suboptimal investments. In general, then, markets will perpetuate discrimination for reasons that are unrelated to the market failures that provide traditional economic grounds for legal intervention.

Moreover, and crucially, each of these effects reinforces the others in potentially powerful ways. If there is ordinary prejudice, it will interact with statistical discrimination so as to produce more of both. People notice events consistent with their prejudice and disregard events that are inconsistent with them, and the result will be more in the way of both prejudice and statistical discrimination. If third parties attempt to promote discrimination, they will increase the existence of both prejudice and statistical discrimination. Employers will hire fewer blacks and women, who will in turn appear less frequently in desirable positions, with consequent reinforcing effects on both prejudice and statistical discrimination.

19But see Schwab, “Is Statistical Discrimination Efficient?”; Donohue, “Is Title VII Efficient?”
In addition, if there is prejudice and statistical discrimination, and if third parties promote discrimination, there will be decreased investments in human capital. Such decreased investments will be a perfectly reasonable response to the real world. And if there are decreased investments in human capital, then prejudice, statistical discrimination, and third-party effects will also increase. Statistical discrimination will become all the more rational; prejudice will hardly be broken down; consumers and employers will be more likely to be discriminators. Of course, it would be necessary to compile detailed factual evidence to assess the magnitude of the relevant effects; in the abstract, it is hard to know whether they will be large or small.

In any case, the fact that markets often perpetuate discrimination should present no puzzle at all. The actual effect of markets on discrimination will depend on the particular circumstances. The likelihood that markets will break down discrimination should decrease if third-party prejudice is widespread, if stereotypes or generalizations are accurate, and if ordinary racism and sexism are not sharply limited. What may be surprising is not that markets produce discrimination, but that discrimination has decreased as much as it has. And all this follows from quite conventional economic approaches to the problem.

### III. Markets and Discrimination: Noneconomic Accounts

In this section, I provide two additional arguments against the view that markets will prevent discrimination. Neither of them is standard within the economic tradition, but both of them have considerable explanatory power.

#### A. Preference and belief formation.

In economic theory, preferences and beliefs are usually taken as given—an approach that is often helpful in building models for understanding social life. But in the context of discrimination, and perhaps elsewhere as well, this approach will cause both descriptive and normative problems. The central point here is that preferences are endogenous to current laws and practices. Once those laws and practices are entrenched, there are special obstacles to bringing about change through market ordering.

In the setting at hand, the problem is that private preferences, on the part of both discriminators and their victims, tend to adapt to the status quo, and to do so in a way that makes significant change hard to achieve. The reduction of cognitive dissonance is a powerful motivating force: people attempt to bring their beliefs and perceptions in line with existing practice. The victims of inequality tend to reduce dissonance by adapting their preferences to the available opportunities. Consider the story of the fox who concludes that he does not want unavailable grapes because he considers them sour; the reason that he considers them sour is that they are unavailable. Or people may adapt their aspirations to the persistent and often irrationally-held belief that the world is


just. The beneficiaries of the status quo tend to do the same, concluding that
the fate of the victims is deserved, or is something for which victims are
responsible, or is part of an intractable, given, or natural order. All of these
claims have played an enormous role in the history of discrimination on the
basis of race and sex.

There is of course extremely powerful evidence in the psychological literature
for the thesis that human beings try to reduce dissonance. Some work here
reveals that people who engage in cruel behavior change their attitudes toward
the objects of their cruelty and thus devalue them. Observers tend to do the
same. The phenomenon of blaming the victim has clear cognitive and motivational
foundations. The notion that the world is just, and that existing inequalities are
deserved or desired, plays a large role in the formation of preferences and
beliefs. The reduction of cognitive dissonance thus operates as a significant
obstacle to the recognition that discrimination is a problem, or even that it
exists.

This problem can be tied quite tightly to the operation of a market economy,
whose participants are self-consciously involved in catering to existing tastes
and in perpetuating and reinforcing them. Markets, partly because of this effect,
are an engine of productivity and respect for individual autonomy in most
circumstances; but in this setting, they sometimes perpetuate inequality. This
is especially true in the context of sex discrimination, where advertising and
consumption powerfully reinforce existing stereotypes, with consequences for
the development of preferences and beliefs. Consider, for example, the multiple
ways in which the beauty industry, broadly understood, attempts to define and
to commodify femininity, in efforts to reach men and women alike.

The inevitable effect of a discriminatory status quo on preferences and beliefs
is related to the phenomenon, noted above, of decreased investment in human
capital. In both cases, the response to discrimination is endogenous rather
than exogenous and has the consequence of perpetuating existing inequalities. But
with respect to preferences and beliefs, the effect is especially pernicious. Here
the consequence is not merely to shift investments in self-development but,
instead, to make people believe that the existing regime, including discrimination,
presents no problem at all.

Once preferences and beliefs are affected, the likelihood of social change
through markets diminishes dramatically. Here, as well, there is a mutually
reinforcing effect among the various sources of market discrimination. The
effect of discrimination on preferences and beliefs fortifies existing prejudice,
produces a decrease in investment in human capital, leads to more in the way
of statistical discrimination, and increases the unwillingness of third parties to
deal with blacks and women.

---

WHY MARKETS DON'T STOP DISCRIMINATION

B. Baselines and discrimination law. Almost by definition, markets incorporate the norms and practices of advantaged groups. Conspicuous examples include the many ways in which employment settings, requirements, and expectations are structured for the able-bodied and for traditional male career patterns. Here the market, dependent as it is on the criterion of private willingness to pay, is extremely unlikely to eliminate discrimination. What the market values will revolve around is in significant part the norms and practices of certain social groups.25 To say that the refusal to provide flexible time, child care, or building access for the disabled does not count as "discrimination" is to rely on an exceptionally narrow conception of what discrimination is — a conception that is repudiated in much of current American law, that is inconsistent with an approach that is alert to discriminatory purposes26 or effects, and that will in any case do very little about existing inequalities.

In these circumstances, a legal system committed to an anti-discrimination principle might in some cases restructure market arrangements so as to put members of disadvantaged groups on a plane of greater equality — not by allowing them to be "like" members of advantaged groups, but by changing the criteria themselves, at least when those criteria do not have a firm independent justification. Consider here the conventional test of American discrimination law: is the member of the disadvantaged group "similarly situated" to the member of the advantaged group? That is: have women who are otherwise the same as men, or disabled people who are otherwise the same as the able-bodied, been treated differently from men or the able-bodied? The problem with this test is that it itself reflects inequality, since it takes the norms and practices of the advantaged groups as the baseline against which to measure equality.27 And here market ordering, dependent as it is on the criterion of private willingness to pay, will inevitably present a problem.

Ideas of this sort underlie recent efforts to ban discrimination on the basis of pregnancy and disability. As in the case of statistical discrimination, these forms of discrimination are perfectly rational from the standpoint of employers and others. The impetus for legal controls is that, in these settings, the criteria used in private markets are hardly prepolitical or natural, have neither good moral status nor powerful independent justification, and at the least should be overridden when their predictable consequences are exclusion or second-class citizenship for certain social groups.

In the case of disability, for example, the handicapped face a wide range of obstacles to participation in both public and private spheres. These obstacles are not a function of "handicap" itself (a most ambiguous concept) but are instead the necessary consequence of humanly created barriers, including stairs,

25 If it is to be made plausible, this is not a conspiracy theory. There has been no self-conscious decision by whites or men to do something of this sort. The phenomenon is instead a consequence of a collection of small decisions without global or world-transforming aspirations — in short, of a market.


doors, and standards in general made by and for the able-bodied. The market in these circumstances is wholly nonresponsive. The market itself will not value the contributions of handicapped people, which are by hypothesis relatively low if measured by normal productivity standards, understood as these are through the lens of willingness to pay. Discrimination of this sort requires a legal remedy. And though the specification of the content of that remedy is a difficult question, markets will not respond, even in the long run.

The point, then, is that the valuation of the market will be a reflection of prevailing norms and practices, and those norms and practices sometimes are what an anti-discrimination principle is designed to eliminate or reduce. When this is so, reliance on markets will be unsuccessful. The rationale for anti-discrimination laws in these cases becomes more controversial, complex, and difficult. But however that may be, private markets will not stop discrimination.

IV. WHAT’S WRONG WITH DISCRIMINATION

Thus far, the discussion has been entirely descriptive and explanatory. I have sought to show why markets are unlikely to eliminate discrimination on the basis of race, sex, and disability. But it will not have escaped notice that I have only provisionally defined discrimination—a highly protean concept, in light of the foregoing remarks—or said what is wrong with it. Indeed, one of the largest gaps in present civil rights theory consists of the absence of a full-fledged explanation of just what discrimination is and why we should eliminate it. In this highly tentative section, I want to say something about the appropriate characterization of the claim of discrimination and about how that claim should be treated.

Discrimination should, I suggest, be understood to include any decision that treats an otherwise similarly qualified black, woman, or handicapped person less favorably than a white, male, or able-bodied person, whether the reason for the decision lies in malice, taste, selective empathy and indifference, economic self-interest, or rational stereotyping. This understanding of discrimination picks up not merely covert unequal treatment, but also requirements that are neutral “on their face” but that would not have been adopted if the burdened and benefited groups had been reversed.28 It does not pick up measures merely having discriminatory effects unless those effects are, in the sense indicated, tied up with racial, sexual, or other bias.

It follows that the claim of discrimination, best understood, is not for prevention of certain irrational acts, or of “prejudice,” but instead for the elimination, in places large and small, of something in the nature of a caste system. The concept of caste is hard to define, and I will have to be tentative and somewhat vague about it. I do not mean at all to suggest that the caste-like features of current practices are precisely the same, in nature or extent, as those features of genuine caste societies. I do mean to say that the similarities are what make the current practices a reason for collective concern.

The motivating idea behind an anticaste principle is Rawlsian in character: without very good reason, legal and social structures should not turn differences that are irrelevant from the moral point of view into social disadvantages. They certainly should not be permitted to do so if the disadvantage is systemic. A systemic disadvantage is one that operates along standard and predictable lines in multiple important spheres of life, and that applies in realms—like education, freedom from private and public violence, wealth, political representation, and political influence—that are deeply implicated in basic participation as a citizen in a democratic society.29

In the areas of race and sex discrimination, and of disability as well, the problem is precisely this sort of systemic disadvantage. A social or biological difference has the effect of systematically subordinating the relevant group and of doing so in multiple spheres and along multiple indices of social welfare: poverty, education, political power, employment, susceptibility to violence and crime, and so forth. That is the caste system to which the legal system is attempting to respond.30

Differences are usually invoked as the justification for disadvantage. The question is not, however, whether there is a difference—often there certainly is—but whether the legal and social treatment of that difference can be adequately justified. Difference need not imply inequality, and only some differences have that implication. When differences do have that implication, it is a result of legal and social practices, and these might be altered. The problems faced by the handicapped, for example, are not a function of handicap “alone” (an almost impenetrable idea), but instead of the interaction between physical or mental capacities on the one hand and a set of human obstacles made by and for the able-bodied on the other. It is those obstacles, rather than the capacities taken as brute facts, that create a large part of what it means to be handicapped. For these purposes, it is not decisive whether a present disadvantage can be tightly connected to an intentionally discriminatory act by a public “tortfeasor.” This understanding, I suggest, provides the principal impetus behind antidiscrimination efforts.

A few disclaimers are necessary here. The anticaste principle, if taken seriously, would call for significant restructuring of social practices. The principle is better set out and implemented by legislative and executive bodies, with their superior democratic pedigree and fact-finding capacities, than by constitutional courts. Moreover, it is important to acknowledge that for good Rawlsian reasons a wide range of differences are morally arbitrary, and in a market economy those differences are quite frequently and for fully legitimate reasons translated into social disadvantages. Consider educational background, intelligence, strength, existing supply and demand curves for various products and services, height, beauty, and perhaps even willingness to work hard.


An anticaste principle that would attempt, through law, to counteract all of these factors would be difficult indeed to sustain. In general, the recognition of such factors is inseparable from the operation of a market economy, and a market economy is an important source of freedom, prosperity, and respect for different conceptions of the good. The use of such factors is at least sometimes in the interest of the less or even least well-off, and when this is so an anticaste principle that would bar their use seems perverse. Moreover, an anticaste principle that would override all morally irrelevant factors would impose extraordinary costs on society, both in its implementation and administrative expense and in its infliction of losses on a wide range of people. Those costs are high enough to make such a global principle immensely unappealing.

The principle therefore has greatest weight in discrete contexts, most notably including race, sex, and disability, in which gains from current practice to the least well-off are hard to imagine; in which second-class citizenship is systemic and occurs in multiple spheres and along easily identifiable and sharply defined lines; in which there will be no global threat to a market economy; and in which the costs of implementation are not terribly high.

If an anticaste principle is accepted, the most general question is whether members of the disadvantaged group, or people who might find themselves in that category, could be persuaded that there are good reasons for the practice under attack. That question will pick up unconscious discrimination as well as selective empathy and indifference; here, the answer will frequently be negative. Of course, there will be severe difficulties in administering a system built on a question of that sort. But in this context, the issues should be discussed precisely in terms of possible answers to that question.

V. Conclusion

Except in limited contexts and under limited conditions, competitive markets will not prevent discrimination. Sometimes discrimination is an economically rational response to the desires of third parties; sometimes it reflects rational stereotyping; sometimes it fuels limited investments in human capital, which create a cycle of discrimination. All of these cases are standard. They account for much of the territory of race and sex discrimination in the United States and elsewhere. They explain why, in many contexts, reliance on markets will ensure that discrimination will continue.

Somewhat less conventional explanations of why markets perpetuate discrimination point to the effects of the discriminatory status quo on beliefs and aspirations and the embodiment, in markets, of the norms and practices of advantaged groups — norms and practices that perpetuate the second-class citizenship of the disadvantaged. These accounts interact with those offered above. Taken together, they provide an explanation for the phenomenon, puzzling to some, of simultaneous market ordering and discrimination extending over long periods and in many places.

In order to evaluate these conclusions — offered as simple explanations — one needs a theory of what discrimination is and why it ought to be eliminated. It
is here that the theory of civil rights remains in a surprisingly primitive state. Discrimination is a protean concept. It covers hatred or devaluation, selective empathy and indifference, false or excessive generalization, rational responses to the desires of third parties, the use of plausible stereotypes, the adoption of criteria that would not be used if the burdened and benefited groups were reversed, and other forms of behavior as well. I have suggested that the problem with discrimination is that it turns a morally irrelevant characteristic into a pervasive source of social disadvantage. Thus understood, an anti-discrimination principle turns into an anticaste principle. And if it is so understood, the distinction between anti-discrimination and “affirmative action” is thin in practice and perhaps nonexistent in principle.

These points are of course controversial. But if my argument is persuasive, one point should not be: markets will not cure discrimination. And it is on that point, above all, that I am insisting here.

*Law and Political Science, The University of Chicago*