Affirmative action, meritocracy, and efficiency

Steven N. Durlauf

Politics Philosophy Economics 2008 7: 131
DOI: 10.1177/1470594X08088726

politics, philosophy & economics article

London
Thousand Oaks, CA
and New Delhi

1470-594X
200805 7(2) 131–158

Affirmative action, meritocracy, and efficiency

Steven N. Durlauf
University of Wisconsin, USA

abstract This article provides a framework for comparing meritocratic and affirmative action admissions policies. The context of the analysis is admissions to public universities; admission rules are evaluated as part of the public investment problem faced by a state government. Meritocratic and affirmative admissions policies are compared in terms of their effects on the level and distribution of human capital. I argue that (a) meritocratic admissions are not necessarily efficient and (b) affirmative action policies may be efficiency enhancing relative to meritocratic ones. Both these claims, as well as their negations, depend on features of individual behavior for which there is little empirical evidence. The implications of this absence of evidence are then explored, with a focus on policy evaluation when equality and efficiency are both desiderata. I argue that standard statistical decision theoretic approaches do not apply to the affirmative action case, even if equality and efficiency are rendered commensurable based on a scalar payoff function. In this context, I suggest that a presumption for equality-enhancing policies leads to support for affirmative action, but I emphasize the contingent nature of this conclusion.

keywords affirmative action, meritocracy, college admissions, efficiency

1. Introduction

This article describes a framework for meritocratic and affirmative action policies from the perspective of standard economic notions of efficiency. This is done in the context of admissions to public universities. I focus on public college admissions in order to discuss admissions rules as a way of determining public investment – in this case, investment by a state government in the education of its population.

DOI: 10.1177/1470594X08088726
Steven N. Durlauf, Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, W1 53706-1393 [email: sdurlauf@ssc.wisc.edu]

Affirmative action constitutes a redistribution of college slots toward minority students, particularly for more exclusive institutions. The most widely publicized study, due to Derek Bok and William Bowen, finds that for five elite colleges, eliminating affirmative action would reduce the conditional probability of admission for a black applicant from 42 percent to 13 percent.[1] For our purposes, this study is delimited by the absence of public universities in the sample, but it is nevertheless suggestive. Other studies, which include public universities draw similar conclusions. Thomas Kane concludes that for top-quintile schools (defined with respect to the median Scholastic Aptitude Test score), affirmative action for black students is equivalent to adding 400 points to the Scholastic Aptitude Test (SAT) or two-thirds of a point to the grade-point average of an applicant.[2] Mark Long finds for top-decile schools, elimination of affirmative action would reduce minority enrollment from 10.6 percent to 7.8 percent.[3] While one cannot determine how many of these students would simply not have been admitted anywhere, these studies indicate that a large number of minority students are affected by affirmative action. Similar findings occur at the graduate level. Linda Wightman concludes that six-sevenths of the black students admitted to law school in 1990 would not have been admitted to the programs they attended under an admissions policy based entirely on test scores and grades.[4] Most recently, Eric Grodsky argues that previous studies have generally underestimated the effects of affirmative action on admissions because they fail to account adequately for the endogeneity of the application decision by the student and the associated role of universities in influencing the application decision; his empirical evidence suggests that the range of colleges practicing affirmative action is substantially larger than previously believed.[5]

It also appears that efforts simultaneously to eliminate affirmative action and introduce different admissions criteria so as to preserve minority enrollment have not been successful. For example, flagship universities in California, Florida, and Texas have, in response to the elimination of affirmative action, employed admissions policies under which the highest ranked x percent of students from each high school in a state are guaranteed acceptance to the state system; these do not appear to have preserved minority admissions at more elite schools, although Florida's very generous admission standard is a partial exception.[6] Nor is it clear that there are as yet untried policies that might be employed to bolster minority enrollment without explicit attention to race. Maria Cancian shows that one leading alternative, class-based affirmative action, is unlikely to sustain minority enrollments.[7] In fact, Sigal Alon and Marta Tienda argue that more selective colleges have, over time, increased reliance on test scores in admissions, exacerbating the trade-off between merit (so defined) and diversity.[8] For the short run, it therefore appears that substantial minority enrollment in more selective colleges requires the presence of affirmative action policies. It is, further, very possible that minority enrollment rates that reflect the population may require affirmative action in the longer run. As shown by Alan Krueger et al.,[9]

this will be so for the most elite schools even if the convergence rate in black–white test scores were to return to the high levels of the 1980s as opposed to current stagnation. Unless this optimistic scenario is realized, affirmative action will need to persist in order to preserve minority admissions.

The affirmative action debate often proceeds on the basis that, absent equality concerns with respect to historically disadvantaged groups, there is a natural meritocratic basis for determining college admissions, one which amounts to setting higher admissions standards at better schools. Arguments about affirmative action typically focus on the redistribution of college slots and its interpretation in terms of a justice consideration. Advocates of affirmative action often defend the redistribution on equality of opportunity or compensation for past harms grounds; in contrast arguments against affirmative action contend that this redistribution is inconsistent with notions of individual desert and merit.[10] Yet both sides often treat meritocratic admissions rules as a presumption from which deviations need to be justified. My goal is to explore the efficiency side of admissions policies and so address this presumption. In doing so, I will further discuss how efficiency and equity interact when one considers the policymaker's decision problem of choosing among alternative decision rules. In using efficiency and equity language in describing this decision problem, I hope that the loss of richness in conceptualizing justice is compensated for by the insights provided in my formulation of the ways in which trade-offs may be evaluated.

My analysis is conducted from the perspective of a policymaker who sets admissions rules based on an underlying set of preferences which may be summarized by the distribution of human capital induced by the rule. The analysis assumes that the level of state expenditures on education is taken to be fixed in both quantity and allocation across campuses. Admissions are therefore contingent on a fixed set of slots and fixed level of university quality. Further, I assume that the only output of universities is human capital embedded in students; as such I ignore the effects of student body composition on, for example, the quality of research that is produced by the faculty.[11] The use of an admittedly desiccated objective function has the advantage that it allows for clear comparison of different admissions rules. Finally, I assume initial human capital is accurately measured for every student, despite arguments by Claude Steele and James Aronson that minority student test scores are deflated by 'stereotype threat' which leads to underperformance relative to true ability; Mary Fischer and Douglas Massey find evidence of this for the SAT test.[12] These assumptions allow me to bring to bear specific arguments from economic theory as well as to delineate the relevant empirical evidence.

Overall, I argue that there is little basis for regarding either meritocratic or affirmative action admissions as more efficient than the other. Economic theory can provide conditions under which one is more efficient, but these are dependent on certain behavioral parameters for which theory provides no guide in terms of magnitudes. Nor does existing empirical research provide much insight into

the parameters' values. This lack of information has implications for how one considers equity–efficiency trade-offs. Following ideas by Gerald Gaus, I will conclude that the redistributive effects I have described can provide a basis for justifying affirmative action in light of the lack of knowledge of efficiency effects.[13] But this judgment presupposes a stance on the justice issues associated with affirmative action. My contribution is not to these considerations, but to the salience they should have when a policymaker is also concerned with efficiency.

2. College admissions as a policy problem of human capital formation

In this section, I describe a basic structure for understanding college admission rules as a human capital formation problem. The formalism will be useful in understanding some aspects of efficiency considerations in admission rules.

A framework for understanding the question of college admissions as one of choosing human capital levels may be based on what is known as an overlapping generations model.[14] Individuals are modeled as members of family dynasties; dynasties may be denoted by 'i'. Within a dynasty, individuals live two periods, denoted as 0 and 1; these correspond to youth and adulthood; individuals are thus distinguished by their family dynasty and the time in which they are born t. In youth, a student's initial human capital is formed, measured by a scalar $h_{i,0,t}$ which is observable to colleges.[15] Colleges use initial human capital as part of the data employed in admissions decisions.

Adult human capital $h_{i,1,t}$ is determined by college attendance. For student (i,t) who attends college c:

$$h_{i,1,t} = \phi(h_{i,0,t}, h_{-i,0,c,t}, \lambda_{c,t}) \tag{1}$$

Here $h_{-i,0,c,t}$ denotes the human capital of students other than i and $\lambda_{c,t}$ is a measure of 'fundamental' college quality. I will assume that this function exhibits positive spillovers which means (a) the adult human capital is increasing in the initial human capital of other students and in college quality, that is, if $h_{-i,0,c,t} > h'_{-i,0,c,t}$:

$$\phi(h_{i,0,t}, h_{-i,0,c,t}, \lambda_{c,t}) > \phi(h_{i,0,t}, h'_{-i,0,c,t}, \lambda_{c,t}) \tag{2}$$

and (b) adult human capital is increasing in college quality, if $\lambda_{c,t} > \lambda'_{c,t}$, that is:

$$\phi(h_{i,0,t}, h_{-i,0,c,t}, \lambda_{c,t}) > \phi(h_{i,0,t}, h_{-i,0,c,t}, \lambda'_{c,t}) \tag{3}$$

These assumptions may be questioned. There is some evidence that for high school students, weaker students may become demoralized by more successful peers.[16] Similarly, one can imagine a student with a low academic background performing better at a school of lower quality, for example, a community college versus the Ivy League. These assumptions are in fact not necessarily essential in

understanding efficient admissions rules, but matter for understanding equilibrium college choices on the part of applicants.

Different efficiency considerations can therefore be understood via the distribution of $h_{i,1,t}$ induced across individuals i and dates t by a given college admissions rule. A college admissions rule is, for my context, simply a mapping of the human capital level of each young person, combined with whatever additional information is available about an individual, denoted as $X_{i,0,t}$, into a binary outcome $z_{i,0,c,t}$ which equals 1 if the student is admitted and 0 otherwise (0 includes a case in which the student did not apply). Formally, this may be expressed as:

$$z_{i,0,c,t} = \Psi_{c,t}(h_{i,0,t}, X_{i,0,t}) \tag{4}$$

The admissions function depends on the identity of the college c and the admissions year t, in that admissions standards may vary across colleges and time in order to address slot constraints. Race-based policies (of any type) mean that this mapping is affected by the ethnic identities of the agents, that is, race is coded into the data $X_{i,0,t}$.

Once these admissions are accepted or rejected by students, the composition of college student bodies is determined. The output of the admissions process is a vector of adult human capital levels $h_{1,t}$ as produced by colleges.

3. Efficiency: theory

In order to discuss efficiency,[17] it is of course necessary to define an objective function for the policymaker. Suppose that the goal of the state government is to maximize the level of human capital among the next generation of adult citizens. This means that the state wishes to maximize:

$$\sum_i h_{i,1,t} \tag{5}$$

The linearity of the equation ensures that the distribution of the aggregate level across individuals is irrelevant to the policymaker. This sort of objective function can be rationalized if a state wishes to maximize aggregate output; Richard Posner's defense of wealth maximization suggests an objective function of this type.[18] Another defense of this objective function may be that distributional concerns should be addressed with different policy mechanisms; that is, educational policy should maximize the pie, while other policies can affect its allocation.

I will not directly solve for the efficient allocation of students across colleges. Rather, I will consider particular admissions rules and evaluate them relative to the objective function (5). Specifically, I consider meritocratic admissions policies and contrast them with alternative rules.

For the college context, I understand a strictly meritocratic admissions policy to mean that each student is admitted to college exclusively on the basis of $h_{i,0,t}$,

with higher human capital students receiving priority over lower human capital students so that when slots are scarce, the higher human capital students are the ones admitted to a given school. Formally, this is expressed as:

$$z_{i,0,c,t} = \Psi_{c,t}(h_{i,0,t});\ \Psi_{c,t}(\cdot) \text{ nondecreasing in } h_{i,0,t} \tag{6}$$

From this perspective, meritocratic decisions are ones based on a limited information set; this is similar to Amartya Sen's idea that welfarist criteria may be understood in terms of the limitations placed on information when ranking social states.[19]

For clarity, I will focus on meritocratic admissions when there are two schools of equal size; while Steven Durlauf and Ananth Seshadri show that the equal size assumption is not innocuous, the force of their finding is diminished for large (in terms of number of members) organizations.[20] The school with the higher fundamental quality $\lambda_{c,t}$ sets an admission standard $\bar{h}$ and the other school sets a lower standard $\underline{h}$. Assume that students prefer attending college to not attending. Under the positive spillover assumptions (2) and (3), students always prefer classmates with higher human capital and colleges with higher fundamental quality, so it is possible for the two colleges to choose admissions standards $\bar{h}$ and $\underline{h}$ such that (a) all slots are filled and (b) there is assortative matching in equilibrium college choices, that is, no student in the college with the lower admission standard has initial human capital higher than the lowest human capital in the school with the higher admissions standard. Students with higher initial human capital prefer to be with each other, and their ability to segregate themselves from weaker students is ensured by the admissions criterion of the higher-quality school. The assumption that the higher-quality school is the one that sets the higher admissions standard ensures that students with relatively high human capital will never be tempted to switch to the other school because of its quality. The equilibrium college choices of students thus translate the meritocratic admissions rule into what is conventionally understood to be a meritocratic outcome: better students attend better schools.

The key idea that I wish to develop from this formal model is that the meritocratic assignment of students is not necessarily equivalent to an efficient allocation of students, as evaluated from the perspective of the objective of maximizing aggregate human capital as expressed in (5). In other words, there do not exist a priori reasons why meritocracy in college admissions, as I believe it is conventionally understood, should be equated with efficiency, at least in terms of aggregate human capital production as I have defined it. To be clear, meritocracy may be efficient, but this will depend on details of the human capital production functions for each college.

Why is it the case that a meritocratic admissions rule may not be efficient? The reasons are best understood at two levels. First, consider the case where two schools have identical quality. The efficient allocation of students across the

schools will in this case entirely depend on the interactions between students. As first recognized by Gary Becker, the general efficiency of assortative matching depends on a particular feature of the college human capital production function: that the benefits of higher human capital peers are greater for a higher human capital student than a lower human capital one.[21] To understand the relevant condition, fixing quality at $\bar{\lambda}$, assortative matching is always efficient if, and only if, for any pair of human capital levels of classmates $h_{-i,0,c}$ and $h'_{-i,0,c}$ such that $h_{-i,0,c} > h'_{-i,0,c}$ (since vectors are being compared, the inequality applies to corresponding elements of the vectors):

$$\phi(h_{i,0,t}, h_{-i,0,c,t}, \bar{\lambda}) - \phi(h_{i,0,t}, h'_{-i,0,c,t}, \bar{\lambda}) > \phi(h'_{i,0,t}, h_{-i,0,c,t}, \bar{\lambda}) - \phi(h'_{i,0,t}, h'_{-i,0,c,t}, \bar{\lambda}) \quad (7)$$

This property is known as complementarity, and formalizes the idea that student human capital levels reinforce each other in their effects on college human capital formation. It is an empirical question whether this condition holds. Hence, one cannot equate meritocracy with efficiency on a priori grounds.

How does heterogeneity of colleges interact with heterogeneity in student quality in determining efficient allocations? Even if student human capital exhibits complementarity for a given school quality, it might be the case that low human capital students benefit more from high-quality schools than high human capital ones because of college heterogeneity. When college quality is heterogeneous, then the efficiency of assortative matching where better students are assigned to the better college requires that, for $\lambda_{c,t} > \lambda'_{c,t}$:

$$\phi(h_{i,0,t}, h_{-i,0,c,t}, \lambda_{c,t}) - \phi(h_{i,0,t}, h_{-i,0,c,t}, \lambda'_{c,t}) > \phi(h'_{i,0,t}, h_{-i,0,c,t}, \lambda_{c,t}) - \phi(h'_{i,0,t}, h_{-i,0,c,t}, \lambda'_{c,t}) \quad (8)$$

This condition may be understood as requiring complementarity between student quality and college quality.

Therefore, under a simple objective function for policymakers, one that is insensitive to any aspect of the human capital distribution outside of its average value, there is no logical reason to think that meritocratic admissions represent a baseline from which deviations need to be justified. To summarize, there are two distinct empirical requirements for the efficiency of admissions rules that assign better students, as measured by initial human capital, to better colleges. The peer group interactions and student quality–college quality interactions must both exhibit complementarity. Without these properties, it may still be the case that meritocratic admissions are efficient, but this will depend on details of the human capital production function as well as details of the distribution of initial human capital.

The differences between complementarity and positive interactions are relevant to conceptualizing what desert means for college admissions. The complementarity conditions (7) and (8) parallel the positive interactions conditions (2) and (3). However, the complementarity conditions relate to how inputs to the

human capital process interact, not just whether higher levels of inputs lead to higher output. Put differently, efficiency concerns the value added to different students. This distinction seems relevant to understanding whether better students deserve admission to better colleges. It is commonly argued that affirmative action is unjust because it violates the notion of justice as merit, where merit is a basis for desert.[22] In my model, this argument amounts to the claim that higher $h_{i,0,t}$ individuals deserve positions at better universities.

To understand the merit claim, it is useful to consider the analogous claim that the most qualified applicant for a job deserves to receive it. A useful analysis of the latter is due to David Miller, who argues that one can derive a claim that one applicant deserves a job versus another from the notion that there exists a particular set of criteria which characterize the appropriate scope of the hiring decision.[23] Miller further argues that these criteria are delimited in certain ways. Specifically, the larger social effects from the decision, for example any role model consequences of hiring, do not come under the purview of a firm as opposed to the greater society and so cannot affect an individual's merit claim.

An analogous argument to Miller's might seem to apply to public education. My analysis of the potential inefficiency of meritocracy was based on a consideration of the aggregate human capital in the population. The units of education, however, are individual colleges. Even if assortative matching is inefficient, if individual colleges wish to maximize the human capital of their own students, then each college will, at an individual level, want the highest human capital for their students. The net effects on aggregate human capital of moving a high human capital student to another college are simply irrelevant for a given college that seeks to maximize the human capital it produces in isolation. So, it would seem that Miller's argument that merit is firm specific can be extrapolated to college.

On closer examination, this extrapolation is not warranted. College admissions do not admit a clean division of labor in that individual public colleges are part of a common education process: they are funded and supervised by the state government as a whole. This seems sufficient to require that college admissions consider the effects of the system as a whole, not just each college in isolation. By analogy, a corporation owning several firms in a market will allocate workers across the firms in order to maximize overall profitability.[24]

My argument does not require that the state government is internalizing any sort of externalities. Just as that worker who will be most productive in a job, subject to Miller's caveat, can be said to deserve the slot as he maximizes corporate profits, a college student deserves a slot to the extent his placement in it maximizes human capital. Firms do not hire workers as a reward for past achievement, but rather on the basis of prospective achievement, in the sense of making the firm more profitable. This move from merit as reward to merit as effectiveness is not innocuous as it means that merit needs to be assessed relative to the proper-

ties of the human capital production functions. A summa cum laude graduate has a meritocratic claim relative to a high school dropout with respect to an associate position at a law firm, but not for a job as a fast-food clerk.

Of course, none of this proves that better students do not deserve admissions to better colleges, but rather means that if maximization of total human capital is the goal of the policymaker, meritocratic admissions are not deserved from the perspective of this objective.

4. Efficiency and race

The arguments made so far demonstrate that there are not good reasons to believe that meritocratic admissions are required by efficiency. However, these arguments do not provide a basis for affirmative action per se. Rather, they call into question the common assumption that meritocratic standards are the natural default standard for admissions. Here I consider why an efficient college admissions rule might include the race of an applicant.

First, it may be the case that the composition of the student body with respect to race is a direct argument of the human capital production function. Arguments of this type alter the terms of the baseline model by assuming that the human capital production function depends on student characteristics beyond human capital; denoting individual i's human capital as $X_{i,0,c,t}$ and the additional characteristics of others as $X_{-i,0,c,t}$, the associated human capital production function may be written:

$$h_{i,1,t} = \phi(h_{i,0,t}, h_{-i,0,c,t}, \lambda_{c,t}, X_{i,0,t}, X_{-i,0,c,t}) \tag{9}$$

If the only relevant nonhuman capital characteristic is the racial composition of the student body, letting r_c measure racial heterogeneity produces the function:

$$h_{i,1,t} = \phi(h_{i,0,t}, h_{-i,0,c,t}, r_{c,t}, \lambda_{c,t}) \tag{10}$$

For this function, even if, conditional on equal racial heterogeneity, human capital formation is maximized across two schools by assortative matching on initial human capital, it still may not be efficient because of its effects on heterogeneity. The claim that diversity improves education for all students is the basis of much of the Grutter defense made by the University of Michigan in the most recent Supreme Court decision on affirmative action.

A different way of introducing race into the human capital production process is to argue that education is affected by the extent to which a student feels isolated. Charles Moskos and John Butler argue that an effect of this type is present in the US military, in that the percentage of blacks is large enough that individuals do not feel obliged to act as the representative black.[25] Letting $h_{i,0,w,c,t}$ denote initial human capital among white students and $h_{i,0,b,c,t}$ initial human capital

among black students, this suggests that human capital should be understood as produced by:

$$h_{i,1,b,t} = \phi_b(h_{i,0,t}, h_{-i,0,b,c,t}, h_{-i,0,w,c,t})$$
$$h_{i,1,w,t} = \phi_w(h_{i,0,t}, h_{-i,0,w,c,t}, h_{-i,0,b,c,t}) \quad (11)$$

Even if each of these functions exhibit complementarities, affirmative action may be justified to maximize overall human capital. That said, there is no logical connection between diversity and aggregate human capital, or, for that matter, aggregate human capital in either group taken in isolation. It might even be the case that efficiency is promoted by segregation; put less negatively, this sort of specification can perhaps argue for the continuing value of historically black colleges and women's colleges.

When one considers more general objective functions than average human capital, the efficiency case for race-based admissions seems stronger. In other words, rather than generalize the human capital production function, one might argue that the objective function should explicitly account for the overall social state produced by a policy – to put it another way, that the objective function depends on the full distribution of adult human capital levels in the population as well as the distribution of individual characteristics:

$$W = \pi(h_{i,t}, X_{1,t}, X_{0,t+1}) \quad (12)$$

where $X_{1,t}$ and $X_{0,t+1}$ denote the full sets of adult and child characteristics, respectively. To link this generalization to affirmative action, suppose that the policymaker's objective function depends on the distribution of human capital across the workforce. This can occur if workers tend to concentrate in certain regions or tend to offer services to members of their own groups. In this case, the distribution of college admissions will affect the distribution of services available to the population. In essence, this modifies the objective function from one that only depends on aggregate human capital to one based on:

$$s\left(\sum_i (h_{i,1,b,t}), p_b\right) + s\left(\sum_i (h_{i,1,w,t}), p_w\right) \quad (13)$$

where $s(.,.)$ transforms the level of race-specific human capital into a measure of the associated services; here I allow the measure to depend on the percentage of each race in the population. The idea of this division is to allow for differences in the marginal value of additional human capital in each group. In particular, the natural assumption in this regard is that $s(.,.)$ is concave in human capital. This means that the marginal value of human capital is decreasing within each group. This assumption means that equal human capital across groups is preferred, all other things being equal. The modified efficiency condition can motivate race-based admissions. This follows immediately from the form of (13) since the

marginal benefit of an increase in black human capital can differ from the marginal benefit of an increase in white human capital, something that is ruled out for the linear function described by (5).

Two main conclusions may be drawn from this theoretical discussion. First, with respect to standard efficiency arguments, meritocracy (as conventionally understood) does not deserve a presumption to be an admissions criterion for colleges. Second, if the goal of a policymaker is to maximize the benefits of human capital in a population, race-based admissions may be justified because of their distributional effects with respect to the provision of services. Notice this is not a direct argument for equity; rather, it follows from concavity in the marginal value of human capital within a racial group.[26]

5. Evidence

Sections 3 and 4 have described how the theoretical efficiency case for meritocratic, color-blind admissions depends on various properties of the human capital formation process. I now turn to what is known about these properties.

5.1. Complementarity

As far as I know, the body of social science studies on education provides no strong basis for concluding that human capital formation exhibits either of the types of complementarity needed to justify meritocratic admissions. This statement should be understood as largely negative: the question of complementarity has in fact received little empirical analysis for college education; the standard survey of the consequences of affirmative action by Harry Holzer and David Neumark reports few studies that even indirectly address the question.[27] In particular, there does not exist good empirical evidence on the value added by college education for different types of students.

Some indirect evidence may be ascertained from the vast literature on tracking in elementary and secondary schools. Julian Betts and Jamie Shkolnick summarize this literature as of 2000 by concluding: 'past studies which compare students from different ability groups to heterogeneously grouped students find evidence that the top students are helped by ability grouping and that the bottom students are harmed, resulting in a net effect that may be positive or negative, but which is usually close to zero'.[28] This suggests that complementarities, even if present, are not large.[29] There is also a difficulty in extrapolating from pre-college to college-level education. Adam Gamoran has made important arguments that the effects of tracking cannot be assessed without considering how tracking affects the educational process; following his reasoning, assessing classroom tracking requires evaluation of the influence of tracking on teacher expectations of weaker students and the ability of teachers to tailor instruction.[30] It is unclear to what extent such factors come into play at the college level.

Moving beyond specific studies, there are good reasons to question whether

statistical analyses can uncover the sorts of factors that produce complementarities. Within economics, effects of the characteristics and behaviors of others on individuals are referred to as social interactions. The identification of social interactions is rendered difficult by three distinct problems.[31] First, the effects of characteristics and behaviors of one student on another are difficult to disentangle; this occurs because the logic of complementarity models implies these should move together. For the college context, this general problem means, for example, that it is hard to disentangle the effects of high initial human capital peers from high effort peers. Charles Manski refers to this difficulty as the reflection problem.[32] Second, in contexts such as colleges, it is necessary to account for self-selection. As emphasized in the work of James Heckman,[33] self-selection can occur with respect to individual characteristics that are not observable to an analyst. For example, individual ambition presumably matters for the value produced by college attendance. More ambitious students presumably will self-select into higher-quality schools. Third, the identification of social interactions requires accounting for the presence of group-level unobservables. College quality is neither measurable nor necessarily constant across time. Hence evidence that student performance covaries with the observed characteristics of classmates may be spurious if there are unmeasured features of college quality. There is an active econometric literature that has proposed ways to address these problems; these methods have not been widely employed. So while there is an extensive literature that contains claims that social interactions are empirically important, it is generally straightforward to question whether the findings are based on untenable assumptions.

Further, efforts to identify social effects, including many in the tracking literature, generally focus on the identification of positive spillover effects rather than complementarities and so are not necessarily informative on the question of efficiency. In fact, complementarities are a priori ruled out by certain functional form assumptions that are used in models of social interactions. For example, it is common to use linear models to identify social interactions. Suppose one uses the model:

$$h_{i,1,t} = \alpha + \beta h_{i,0,c,t} + \gamma \bar{h}_{-i,0,c,t} + \delta Z_{i,0,t} + \varepsilon_{i,0,t} \tag{14}$$

where $\bar{h}_{-i,0,c,t}$ denotes the average human capital of a student's peers and $Z_{i,0,t}$ denotes individual heterogeneity. This functional form for adult human capital production rules out any complementarities between students, since the effects of changes in peer characteristics are independent of the student's initial human capital. This problem is ameliorated in some studies by allowing for nonlinearity; for example, Maureen Hallinan[29] does this implicitly by allowing the parameters of the effects of a tracked classroom on a student to depend on the type of classroom to which the student was originally assigned. But there is no reason to believe this is an appropriate functional form to permit complementarities. Put

differently, complementarity involves interactions between an individual's characteristics and others', but the nature of this interaction is something that an analyst needs to evaluate, not assume.

Of course, statistical studies do not represent the only basis on which to draw conclusions on the presence of complementarities. Hence, the most appropriate assessment of the existing empirical evidence is that one's prior beliefs ought not to be strongly affected by the existing empirical evidence.

5.2. Diversity

Evidence that diversity represents an input into the educational process is also weak.[34] I am focusing on the effects of diversity on human capital, not on attitudes concerning race; the effects of diversity in reducing prejudice might be an argument in favor of affirmative action, but is not relevant here. Given the frequency with which the claim is made, it is remarkable how poor the evidentiary support is for the proposition. A standard form of evidence of this type derives from experimental settings in which some problem is assigned to groups; the beneficial effects of diversity are adduced that heterogeneity in group composition facilitates the problem solving; a recent study by Samuel Sommers of the effects of racial composition on the deliberations of mock juries is a good example of this type of research.[35] The value of such studies for the question I am studying is, however, limited in that they bear little relation to college interactions. The interactions studied in diverse groups presuppose that these groups have been formed exogenously for specific tasks. In contrast, intra-college interactions emerge through repeated interactions and via the endogenous formation of social relationships.

Unlike the complementarity case, the claims that diversity has some general effect on education strike me as much less plausible on *ex ante* grounds. If the 'diversity improves education' argument is that classroom teaching is affected, say via differences in ethnically based perspectives, then it would seem that for Mathematics and Sciences, the claim is a non-starter. Moreover, it would seem to suggest the argument is irrelevant to large lecture classes. It is interesting to note that in the expert testimony provided by the University of Michigan arguing in favor of affirmative action the only direct studies of the link between education and diversity 'found very few direct relationships between diverse student enrollments and educational outcomes after four years of college' and consist of two unpublished analyses.[36]

5.3. Allocation of resources across groups

In assessing the ethnic-group-specific efficiency effects, I will focus on medical services given the rich literature that is available on the subject. A number of studies have argued that the medical access of African Americans and Hispanics is sensitive to the local supply of physicians, and that these populations are generally served by physicians of similar ethnicity; a prominent study of this type is

that of Miriam Komaromy et al.[37] Claims of this type, which operationally involve finding, after controlling for various factors, correlations between doctor and patient ethnicity, appear to suffer from serious statistical problems. For example, Komaromy et al. do not deal with the general equilibrium nature of doctor–patient matching, that is, the study does not provide an explicit model of the supply of and demand for physicians of different types in different localities under different assumptions about the ethnic composition of physicians in the aggregate. Hence the resultant changes in the relationship between physician ethnicity and patient ethnicity that would occur if the ethnic composition of physicians were to change cannot be inferred.[38] So it is hard to see how to infer efficiency benefits from different policies based on the sorts of correlations they identify. Interestingly, Martin Stinson and Norman Thurston find that the correlation between doctor and patient race is substantially diminished once one controls for factors such as location and physician specialty.[39] It is perhaps noteworthy that the 2000 survey by Holzer and Neumark reports no studies on this question which have appeared in social science as opposed to medical journals.

Statistical problems also exist with studies arguing that white doctors may provide worse services to black patients than black doctors; this represents a different source for race-specific efficiency from matching per se. Claims that diagnoses and treatments are affected by the ethnicity of a patient need to account for the self-selection of patients to doctors and vice versa, as well as differences in the nature of unobserved heterogeneity between blacks and whites. While I cannot claim to have exhaustively read the literature on doctor and patient ethnicity and treatment, what I have read does not meet the evidentiary standard expected in economics for evidence, say, of discrimination. Heckman provides a nice overview of difficulties involved in identifying evidence of discrimination for both nonexperimental and experimental data.[40]

Claims that affirmative action is efficient on access grounds contain an implicit assumption, namely that the effects of affirmative action on the supply of minority group professionals is positive in the sense of equations (2) and (3). This implicit assumption has been challenged by some authors. It is easy to imagine a student with a low academic background struggling in the presence of peers with stronger backgrounds. In the affirmative action context, this possibility has been referred to as ‘mismatch’. This mismatch argument matters, relative to the baseline model I described, as it suggests a defect in the individual. As a result, there is a body of work that has tried to identify whether affirmative action has had an adverse consequence on those admitted.

The most widely publicized study of the effects of affirmative action on minority students is due to Bok and Bowen, who conclude that for very elite schools, those admitted through affirmative action have benefited from the experience. This study is useful in summarizing data on college performance, but does not employ formal statistical methods to account for self-selection problems induced by the endogeneity of the admission process; Alon and Tienda consider the same

question with particular attention to self-selection problems and conclude that mismatch is not occurring, when the outcome of interest is graduation.[41] Linda Datcher-Loury and David Garman examine the effects of college on labor market earnings and conclude that admission to more elite colleges has significant positive effects on future earnings and that the effect is greater for blacks than whites.[42] At the same time, they find that these benefits are offset for weaker students in the sense that the likelihood of graduation is lower for students with lower (below classmate median) SAT scores. They conclude that mismatch is a significant offset for these students.

A number of analyses of mismatch have been conducted in the context of law schools. In a widely publicized study, Richard Sander argues that affirmative action admissions to law school hurt the recipients as the recipients tend to receive poor grades, thereby hurting their subsequent placement in law firms.[43] There appear to be general conceptual problems with the study; Jesse Rothstein and Albert Yoon are especially persuasive that Sander fails to set the correct counterfactuals.[44] Sander assumes that affirmative action in law school admissions has no effect on overall law school attendance by African Americans, that is, the policy only affects where attendance occurs; Rothstein and Yoon show this is not the case. Further, they demonstrate that any evidence of adverse effects is concentrated among black students with very poor academic backgrounds. This suggests a flaw in the structure of the given affirmative action program as opposed to an intrinsic defect.

Indirect evidence for efficiency gains in a race-sensitive distribution of human capital may be derived if one considers the effects of diversity in an occupation. Here, the most studied example concerns affirmative action in police departments. The best statistical study of this question, by Justin McCrary, finds little evidence that affirmative action policies affect police quality, and indicates why claims to this effect by John Lott are dubious.[45] David Sklansky provides a useful complement to the statistical findings in describing why African American officers are of value in African American communities and how they can promote efficiency. His arguments suggest that the statistical analyses using aggregated data may mask certain types of efficiency gains.[46]

In my judgment, the issue of matching human capital to the needs of the entire population is, relatively speaking, the most empirically plausible, efficiency-based argument for affirmative action, when efficiency is understood to refer to (13). This plausibility stems from the persistence of residential segregation combined with my view that claims on mismatch are incorrect. It does not stem from firm statistical evidence and so my own views may be understood as nothing more than beliefs I have prior to reading the statistical literature.

6. Extensions of the theoretical model

In this section, I consider two extensions of the basic theoretical model. My objective is to illustrate how these extensions introduce additional complications in the assessment of efficiency in admissions.

6.1. Effort

The analysis so far has abstracted from the question of effort; the vector of human capital levels has been treated as an object that evolves without any actions on the part of individuals. A natural extension of the sort of analysis I have described is the introduction of educational effort. One would expect that different admissions policies will affect student effort levels. Operationally, effort is most naturally incorporated in the analysis by considering its role in producing initial human capital. Formally, initial human capital may be modeled as the outcome of effort in youth $e_{i,0,t}$ and a vector of background variables $Z_{i,0,t}$ as follows:

$$h_{i,0,t} = \phi(e_{i,0,t}, Z_{i,0,t}) \tag{15}$$

In economics, it is standard to treat individual decisions as rational, in the sense that choices such as effort are purposeful and are optimal given an individual's preferences, beliefs, and constraints.[47] While the theoretical arguments I will describe presuppose rationality, they retain interest even if one rejects the rationality, since the question for us is whether and how admissions rules affect effort, thereby producing distinct effects on human capital formation from those so far described.

In some contexts, effort effects can render assortative matching inefficient. Durlauf and Seshadri give an example in which assortative matching can be inefficient because of the incentive efforts.[48] Suppose, for example, that human capital depends on effort and initial human capital. It is possible, if effort is easier for high human capital students, that mixing will lead to more human capital in equilibrium. The trick to the argument is that high human capital students choose high effort levels regardless of the quality of their peers, whereas lower human capital students only have an incentive to do so when they are surrounded by high human capital peers. What this does is create a case where, *ex post*, assortative matching is efficient, but *ex ante* it is not because it depresses effort levels. Hence, meritocratic admissions can, because of incentive effects, fail to maximize human capital.

The Durlauf and Seshadri example is developed in a very simple environment, and more sophisticated models can produce different conclusions. Stephen Coate and Glen Loury, in one of the most important theoretical contributions to the study of affirmative action, argue that affirmative action (for the purposes of this discussion, non-meritocratic affirmative action) policies may, as a theoretical matter, create adverse incentives to effort by students; in the environment studied, they show it is possible to produce a 'patronizing equilibrium' in which

minority students choose low initial human capital levels (via effort decisions), so that those minorities who are admitted to college *ex post* have lower qualifications than those in the majority group.[49] This argument is important, but suffers from the limitation that it does not address the question of the optimal design of an affirmative action policy. What I mean by this is that Coate and Loury compare two regimes, one without affirmative action and one with affirmative action which consists of explicit quotas. This begs the question as to whether one can design policies which depend on race and preserve incentives. Kim Sau Chung shows that this is in fact possible, but as is common in such demonstrations, the implied admissions rules are quite complicated, so much so that it is unclear whether they are realistic.[50]

I am not aware of any empirical evidence that allows one to evaluate directly the effects of affirmative action on effort, as opposed to infer effort from conclusions about phenomena such as mismatch. For reasons discussed in the evaluation of evidence on complementarities, my conjecture is that such effects would be extremely difficult to identify.

6.2. Inter-temporal considerations

The efficiency perspective I have developed has considered admission policies from the perspective of a one time decision problem. Many of the affirmative action debates focus on the evolution of the human capital distribution across time. For example, one important issue in affirmative action debates is whether the policy is justified as compensation for past discrimination.[51] Here, I focus on a different issue, namely the persistence of inequality in initial human capital. In particular, I am concerned with understanding the implications of affirmative action when human capital inequality is persistent. The limiting form of persistence is sometimes known as a 'trap', meaning that if a family or group of families starts with low socioeconomic status, the descendants of the family will not be able to escape that status. Of course, such traps are idealizations; what substantively matters is the possibility that inequality is sufficiently persistent that it renders socioeconomic status highly correlated across generations.

In assessing educational outcomes among blacks, it seems that a significant source of disparities lies in the area of social influences. One manifestation of this is the question of attitudes toward education. John Ogbu offers a critical study in this regard. His analysis focuses on the Shaker Heights school district in Ohio.[52] Shaker Heights commissioned Ogbu (among others) to study the sources of racial differences in educational outcomes. Ogbu concluded that one important source was a difference in attitudes toward education. For example, it was common for black students to equate interest in education with 'acting white' so that a component of black identity excluded educational achievement. The idea that individual behaviors may be shaped by a conscious or unconscious desire to fulfill a perceived group identity has been explored by George Akerlof and Rachel Kranton; Hanming Fang and Glenn Loury suggest this can be a source of per-

sistent interracial inequality.[53] To be clear, none of these explanations in any way 'blame the victim'.

To complete the picture of human capital formation, it is therefore necessary to consider the process by which $h_{i,0}$ evolves. A general form of this process is:

$$h_{i,0,t} = \gamma(h_{i,1,t-1}, h_{-i,t-1}, X_{i,1,t}, X_{-i,0,t}, X_{i,0,t}, X_{-i,1,t}) \tag{16}$$

which allows initial human capital for each individual to be affected both by family-specific factors as well as the characteristics of the population as a whole. Notice that this form incorporates the effects of the population-wide characteristics on neighborhood formation and hence allows for social influences at various levels.

From this perspective, a college admissions rule has intergenerational effects because it not only influences the human capital of the next generation of adults, but also affects the initial human capital of the generation after next. This specification differentiates between the influences of one parent and the rest of the population. When one moves to this dynamic perspective, one also needs to consider an inter-temporal objective function; in parallel to my simplest specification, which is based on the level of human capital at one point in time, it is appropriate to consider a weighted average of the levels across time. A standard way to make this move involves working with:

$$W = \sum_t \beta^t \sum_i h_{i,1,t} \tag{17}$$

that is, a discounted sum of adult human capital across two generations. While the choice of the discount rate β is problematic, that does not affect the argument here.

Efficiency in student allocation is far more complicated than before. I am unaware of any simple way of describing efficiency conditions for college assignment rules analogous to (4). For my purposes, what matters is not the explicit description of the optimal allocation, but the recognition that this optimal allocation hinges on the nature of (16) and as such depends on a very broad range of often not well-understood processes. For example, the process in (16) includes a role for neighborhood influences on individuals. Different distributions of adult human capital will produce different distributions of neighborhood quality.[54] In turn, differences in neighborhoods produce differences in initial human capital formation. Some neighborhood influences are a consequence of the rules for school finance: the major role that local taxes play in school revenue creates a dependence of one child's education on all the adults in the community, both in terms of the tax base and the collective choices made on tax and spending levels. Other influences include role model effects, which relate the educational aspirations and expectations of the young to the distribution of occupations in a community. These social influences aggregate to what Durlauf calls the 'mem-

berships theory of inequality', and are relevant when considering equity issues in college admissions.[55]

A number of affirmative action arguments may be interpreted as claims about (16). The most obvious example is role models; the observation of adults who have benefited from education plausibly influences the effort levels of the young. Alternatively, if black teachers are better able to educate black students, this provides an intergenerational analog to the services argument discussed above.

Arguments that social influences matter, so that human capital formation in children is affected by the human capital of adults in their relevant reference groups (residential neighborhood, ethnicity, and so on) represent questions of positive spillovers, rather than complementarities. In my view, the overall evidence for positive spillovers is relatively strong.[56] But this conclusion primarily derives from my reading of the ethnographic literature in sociology and the experimental literatures in psychology and economics; the identification problems I have described for uncovering social interactions apply to statistical studies of intergenerational effects.[57] Yet there is an additional problem: the body of evidence on social effects does not speak to the quantitative significance of affirmative action as a way of altering these influences. The evidence can be construed as, at best, providing some presumption for the claim that these effects exist. So for both intergenerational channels, I suspect one cannot do more than make reasoned guesses.

7. Policy and the limits to knowledge

I now turn to the question of constructing a conclusion about meritocratic versus affirmative action admissions rules from the perspective of the policymaker. My focus has been on the lack of clear theoretical and empirical reasons to regard meritocratic admission policies as more efficient than policies that incorporate affirmative action. By themselves none of these arguments are sufficient to determine whether the policy should or should not be adopted. In particular, for policy analysis my assumptions on the policymaker's objective function are no longer adequate; it is necessary to reintroduce equity considerations. Abstractly, one way to formalize the comparison of an affirmative action policy and a meritocratic policy is to assign a payoff to each policy. The affirmative action policy (AA) produces a payoff to the policymaker of:

$$V_{AA} = C_{AA,EFF} + C_{AA,EQ} + \varepsilon_{AA,EFF} + \varepsilon_{AA,EQ} \tag{18}$$

where $C_{AA,EFF}$ is the predictable efficiency payoff produced by the policy, $C_{AA,EQ}$ is the predictable equity payoff produced by the policy, $\varepsilon_{AA,EFF}$ is the unpredictable efficiency payoff produced by the policy, and $\varepsilon_{AA,EQ}$ is the unpredictable equity payoff produced by the policy.[58] By construction, the unpredictable pieces have expected value of 0. The analogous payoff function for a meritocratic policy is:

$$V_M = C_{M,EFF} + C_{M,EQ} + \varepsilon_{M,EFF} + \varepsilon_{M,EQ} \tag{19}$$

This formulation ignores issues of commensurability between equity and efficiency considerations; commensurability is not germane to the arguments I want to develop; rather, my goal is to explore how the expected and unpredictable components interact in a policy evaluation.

One way to formulate the policy evaluation problem is to ask, what is the probability that a given policy is the better one given the preferences of and information available to a policymaker? I use this formulation as it provides a way of thinking about the strength of the case for one of the policies versus the other. Formally, given values of $C_{AA,EFF}$, $C_{AA,EQ}$, $C_{M,EFF}$, and $C_{M,EQ}$, what is the conditional probability that $V_{AA} > V_M$? From the linear functional form I have employed:

$$\begin{aligned}&\text{Prob}\,(V_{AA} > V_M | C_{AA,EFF}, C_{AA,EQ}, C_{M,EFF}, \text{and } C_{M,EQ}) = \\ &\text{Prob}\,(C_{AA,EFF} - C_{M,EFF} + C_{AA,EQ} - C_{M,EQ} > \varepsilon_{M,EFF} - \varepsilon_{AA,EFF} + \varepsilon_{M,EQ} - \varepsilon_{AA,EQ})\end{aligned} \tag{20}$$

This formulation makes clear that the strength of the case for affirmative action depends not only on $C_{AA,EFF} - C_{M,EFF} + C_{AA,EQ} - C_{M,EQ}$, which is simply the net predictable payoff to the affirmative action policy, but also on the probability measure that describes the unpredictable net payoff $\varepsilon_{M,EFF} - \varepsilon_{AA,EFF} + \varepsilon_{M,EQ} - \varepsilon_{AA,EQ}$. For example, suppose that:

$$\text{Prob}\,(\varepsilon_{M,EFF} - \varepsilon_{AA,EFF} + \varepsilon_{M,EQ} - \varepsilon_{AA,EQ} \leq z) = \frac{1}{1 - \exp(-\beta z)}\,,\ \beta \geq 0 \tag{21}$$

This is the logistic probability density; it is frequently employed in the discrete choice literature.[59] The parameter β indexes the spread of the probability measure of the unpredictable terms; the smaller β, the more diffuse the probability density of $\varepsilon_{M,EFF} - \varepsilon_{AA,EFF} + \varepsilon_{M,EQ} - \varepsilon_{AA,EQ}$. For the logistic formulation, it is easy to verify that if $\beta = \infty$, then Prob $(V_{AA} > V_M | C_{AA,EFF}, C_{AA,EQ}, C_{M,EFF}, \text{and } C_{M,EQ})$ equals either 0 or 1, depending on whether $C_{AA,EFF} - C_{M,EFF} + C_{AA,EQ} - C_{M,EQ}$ is positive or negative. The reason for this is that there is no unpredictable component to the payoffs from the policies; one either performs better or worse than the other. (The logistic density means that the probability the payoffs are equal is 0.) In contrast, if $\beta = 0$, then Prob $(V_{AA} > V_M | C_{AA,EFF}, C_{AA,EQ}, C_{M,EFF}, \text{and } C_{M,EQ}) = 1/2$ regardless of the value of $C_{AA,EFF} - C_{M,EFF} + C_{AA,EQ} - C_{M,EQ}$. The reason for this is that it is a probability of 1 that the realized value of $\varepsilon_{M,EFF} - \varepsilon_{AA,EFF} + \varepsilon_{M,EQ} - \varepsilon_{AA,EQ}$ swamps the predictable value $C_{AA,EFF} - C_{M,EFF} + C_{AA,EQ} - C_{M,EQ}$; by symmetry, the unpredictable part is equally likely to produce a large positive or large negative value.

This formulation of the policy problem indicates how the uncertainties about complementarities and other features of my efficiency analysis matter. Suppose

that the equity effects of the policies are known, that is, $\varepsilon_{M,EQ} - \varepsilon_{AA,EQ} = 0$. This may be justified if equity considerations are developed on a basis that is not contingent on empirical claims. My discussion on efficiency amounts to arguing that there are not good reasons to assume that $C_{AA,EFF} - C_{M,EFF} \neq 0$. Finally, assume that $C_{AA,EQ} - C_{M,EQ} > 0$, so that from the equity perspective, affirmative action is justified. For these assumptions, there is a certain equity payoff from affirmative action and an unknown efficiency effect, one that is centered at 0. Under these assumptions, the greater the uncertainty about efficiency effects $\varepsilon_{M,EFF} - \varepsilon_{AA,EFF}$, in the sense illustrated by the logit model, the weaker the case for one policy versus another. This is so even though it might appear that I have 'rigged' the model so as to provide a case in favor of affirmative action; the case will be strong only if one can argue that the uncertainty about the efficiency effects is not too large.

This type of analysis presupposes that one can assign probabilities to alternative values of the unpredictable components in the payoff functions. Yet it is not clear that one can do this for affirmative action. Specifically, the available empirical evidence does not even permit the assignment of probabilities. Such environments are said to be characterized by ambiguity as opposed to risk, to use current economic parlance. Focusing on complementarities between student types, whose presence is sufficient to imply the efficiency of assortative matching, the empirical literature does not allow one to assign probabilities to the presence of complementarities at the college level let alone to magnitudes. To take another example, while Coate and Loury give a theoretical reason to believe that disincentives are produced, it is far from clear that these are empirically salient. Ellis Cose's conclusion regarding attitudes by blacks toward affirmative action is that it levels the playing field by reducing the effects of discrimination, not that it ensures success regardless of effort.[60] It is plausible to think that effort among the disadvantaged would be reduced by eliminating affirmative action programs. As noted above, this is not to say that affirmative action policies cannot be structured in ways to ameliorate the Coate and Loury problem. Rather, on such a basic consideration, it is difficult to talk coherently about the likelihood that incentive effects work in one direction versus another, let alone how large they are.

The ambiguity associated with the effects of affirmative action on college outcomes represents a problem. Statistical decision theory incorporates a rich set of results on optimal decision-making in the presence of uncertainty when this uncertainty may be characterized by a probability measure over a set of possible outcomes. The deep problem for assessing affirmative action is, in my judgment, that we do not have good bases for assigning probabilities. Nor is it clear that decision-making methods that have been proposed to address ambiguity can help. Within statistical decision theory, criteria such as minimax or minimax regret have been proposed to address ambiguity and have received recent attention in economics.[61] These methods assume that the policymaker has knowledge of the range of the uncertain outcomes even though he does not know the probabilities

to assign to particular values. For affirmative action, one does not even know the bounds for the magnitude of incentive effects, role model influences, and the like. It is not clear that one even knows the support for factors such as the levels of effort students may choose.[62]

This level of ignorance on the part of the policymaker provides a motivation for a way to conceptualize policy evaluation as suggested by Gaus in a recent analysis. Gaus argues that in weighting principled and expedient reasons for a policy, information limits on expedient reasons should lead them to be down-weighted relative to principled ones.[63] This seems to be correct; there should be a burden of proof if one wishes to override a principled reason with an expedient one. Gaus's idea provides a way of proceeding in evaluating our candidate admissions policies. When one down-weights $C_{AA,EFF} + \varepsilon_{AA,EFF} - C_{M,EFF} - \varepsilon_{M,EFF}$ in the payoff function, one operationally reduces the variance of $\varepsilon_{M,EFF} - \varepsilon_{AA,EFF}$. This down-weighting allows for $C_{AA,EQ} - C_{M,EQ} > 0$ to produce a value of $\text{Prob}(V_{AA} > V_M)$ that is substantially greater than 1/2, when $C_{AA,EFF} - C_{M,EFF} = 0$ and the density of $\varepsilon_{M,EQ} - \varepsilon_{AA,EQ}$ is concentrated near 0.

This line of reasoning provides a basis from which affirmative action may be justified. My theoretical and empirical discussions of efficiency suggest $C_{AA,EFF} - C_{M,EFF} = 0$ is a reasonable claim. With respect to equity, current policies produce substantial direct redistribution toward a group that is, in the aggregate, substantially disadvantaged. In the non-ideal society of 21st-century America, persistent inequality is highly associated with racial inequality. I therefore regard the redistributive effects of affirmative action as meaning that the policies contribute to equity in the sense that $C_{AA,EQ} - C_{M,EQ} > 0$; notice that I am in essence folding all aspects of affirmative action and equity not captured by the fact of redistribution into the unpredictable term. Further, the lack of a firm basis for regarding meritocratic admissions as efficient undermines what I regard as the best argument that affirmative action violates individual claims to merit. In contrast, I see good reasons to regard affirmative action policies as promoting equality of opportunity as developed by John Roemer.[64] These reasons lead me to believe that neither the non-negative support of $\varepsilon_{M,EQ} - \varepsilon_{AA,EQ}$ nor the probability that $\varepsilon_{M,EQ} - \varepsilon_{AA,EQ} > 0$ is large. Together, I conclude that $\text{Prob}(V_{AA} > V_M)$ is large. I must emphasize, however, that the philosophical literature has addressed the relationship between affirmative action and justice in ways that are far deeper than my assumption that redistribution of educational opportunities toward minorities means that the policy is equity enhancing, or, in Gaus's terminology, principled. Hence my claim that $C_{AA,EQ} - C_{M,EQ} > 0$ may certainly be challenged.[65]

Put differently, my personal conclusion is contingent on the existence of a presumption in favor of policies that are equality enhancing, that is, it assumes that, exempting countervailing reasons, a policy whose direct effects are equality enhancing is to be preferred. I believe such a presumption may be derived from a range of welfarist and non-welfarist perspectives.[66] To the extent one rejects the view that equality has a primary claim in evaluating government policies, or

believes that my assumption that the redistribution associated with affirmative action implies that $C_{AA,EQ} - C_{M,EQ} > 0$, my conclusion should be rejected. To be clear, I do not claim that efficiency considerations cannot trump the equality claim for affirmative action, merely that there is currently an insufficient basis to do so.

8. Conclusions

Affirmative action is an exemplar of a complex policy question. Its evaluation requires the adjudication of a range of factors, ranging from economic effects to its interaction with questions of equality of opportunity and desert. In this article, I have focused on the complexities involved in assessing the policy in terms of efficiency. In doing so, I have endeavored to draw explicit comparisons with meritocratic admissions policies. My main conclusion is that there does not exist a strong basis for regarding affirmative action as more or less efficient than meritocratic admissions. I have furthered suggested some ways to think about assessing affirmative action policies in light of this uncertainty. The redistributive effects of affirmative action, in my judgment, provide a defense of the policy in light of the lack of a clear efficiency case, but for someone who makes a different judgment of the justice of the policy, my conclusions on efficiency would lead to the opposite conclusion. If nothing else, this analysis shows that social science research, by failing to provide adequate evidence on the effects of affirmative action or for that matter meritocratic admissions, has rendered the efficiency side of the problem one in which the policymaker faces deep uncertainty.

Notes

1. Derek Bok and William Bowen, *The Shape of the River* (Princeton, NJ: Princeton University Press, 1998).
2. Thomas Kane, 'Racial and Ethnic Preferences in College Admissions', in *The Black-White Test Score Gap*, edited by C. Jencks and M. Philips (Washington, DC: Brookings Institution Press, 1998), p. 432.
3. Mark Long, 'College Applications and the Effects of Affirmative Action', *Journal of Econometrics* 121 (2004): 319–42.
4. Linda Wightman, 'The Threat to Diversity in Legal Education: An Empirical Analysis of the Consequences of Abandoning Race as a Factor in Law School Admission Decisions', *New York Law Review* 72 (1997): 1–53.
5. Eric Grodsky, 'Compensatory Sponsorship in Higher Education', *American Journal of Sociology* 112 (2007): 1662–712.
6. See Mark Long, 'Race and College Admissions: An Alternative to Affirmative Action?' *Review of Economics and Statistics* 86 (2004): 1020–33.
7. Maria Cancian, 'Race Versus Class-Based Affirmative Action in College Admissions', *Journal of Policy Analysis and Management* 17 (1998): 94–105.

8. Sigal Alon and Marta Tienda, 'Diversity, Opportunity, and the Shifting Meritocracy of Higher Education', *American Sociological Review* 72 (August 2007): 487–511.
9. Alan Krueger, Jess Rothstein, and Sarah Turner, 'Race, Income and College in 25 Years: Evaluating Justice O'Connor's Conjecture', *American Law and Economics Review* 8 (2006): 282–305.
10. The philosophy literature of course contains defenses and criticisms of affirmative action on deontological grounds. I have found the following collections of essays to be especially helpful: Steven Cahn (editor), *The Affirmative Action Debate* (London: Routledge, 1995) and Marshall Cohen, Thomas Nagel, and Thomas Scanlon (editors), *Equality and Preferential Treatment* (Princeton, NJ: Princeton University Press, 1977). See also Michel Rosenfeld, *Affirmative Action: A Philosophical and Constitutional Inquiry* (New Haven, CT: Yale University Press, 1993) and the recent, book-length debate by Carl Cohen and James Sterba, *Affirmative Action and Racial Preference* (Oxford: Oxford University Press, 2003).
11. I thank Gerald Gaus for pointing out the importance of this assumption.
12. See Claude M. Steele and James Aronson, 'Stereotype Threat and the Intellectual Test Performance of African Americans', *Journal of Personality and Social Psychology* 69 (1995): 797–811 and Mary J. Fischer and Douglas Massey, 'The Effects of Affirmative Action in Higher Education', *Social Science Research* 36 (2007): 531–49.
13. Gerald Gaus, 'Social Complexity and Evolved Moral Principles', in *Liberalism, Conservatism, and Hayek's Idea of Spontaneous Order*, edited by P. McNamara (London: Palgrave Macmillan, forthcoming).
14. This framework has been a workhorse for the study of inequality since Gary Becker and Nigel Tomes, 'An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility', *Journal of Political Economy* 87 (1979): 1153–89 and Glenn Loury, 'Intergenerational Transfers and the Distribution of Earnings', *Econometrica* 49 (1981): 843–67.
15. This is not a necessary assumption. Test scores and grades may be used to produce an estimate of a student's actual human capital, for example.
16. See Christopher Jencks and Susan Mayer, 'The Social Consequences of Growing Up in a Poor Neighborhood', in *Inner-City Poverty in the United States*, edited by L. Lynn and M. McGreary (Washington, DC: National Academy Press, 1990), pp. 111–186.
17. There is some ambiguity in differentiating equity and efficiency since at one level, efficiency refers to the choice of an admissions policy which is optimal relative to some objective function of a policymaker; this objective function can in turn include measures of what is conventionally understood to mean equity. I believe public policy discussions tend to distinguish between policies which maximize aggregate output in some sense and those that achieve some notion of a just allocation of slots, and so will avoid this ambiguity by equating efficiency with meeting objective functions that do not incorporate such considerations.
18. Richard Posner, 'Utilitarianism, Economics, and Legal Theory', *Journal of Legal Studies* 8 (1979): 103–40. Peter Hammond, 'The Economics of Justice and the Criterion of Wealth Maximization', *Yale Law Journal* 91 (1982): 1493–507 provides a critique of Posner's claim to have developed a perspective that improves

on standard criticisms of utilitarianism.

19. Amartya Sen, 'Personal Utilities and Public Judgements: or What's Wrong with Welfare Economics?' *Economic Journal* 89 (September 1979): 537–58.
20. Steven Durlauf and Ananth Seshadri, 'Is Assortative Matching Efficient?' *Economic Theory* 21 (2003): 475–93.
21. See Gary Becker, 'A Theory of Marriage: Part I', *Journal of Political Economy* 81 (1973): 813–46. Andreas Prat, 'Should a Team be Homogeneous?' *European Economic Review* 40 (2002): 1187–207 and Durlauf and Seshadri, 'Is Assortative Matching Efficient?' develop additional results on the relationship between complementarity and the efficiency of assortative matching.
22. Here, I rely on George Sher, *Desert* (Princeton, NJ: Princeton University Press, 1987).
23. See David Miller, 'Deserving Jobs', *The Philosophical Quarterly* 42(167): 161–81. For Miller, these factors amount to the effect an employee has on the future profitability of a firm, subject to the constraint that this effect is itself measured with respect to legitimate behavior on the part of others. So, the fact that racial prejudice by other workers will affect firm productivity does not affect the desert of a black applicant; presumably, one would also rule out prejudice on the part of customers.
24. The fact that the corporation may have market power in this example is not relevant to the argument. Once one considers profit maximization by a large organization, issues of imperfect information can arise. Gerald Gaus (personal communication) has made the interesting suggestion that competition among campuses for students via meritocratic admissions rules might promote efficiency when policymakers cannot perfectly monitor resource use by the individual campuses. As far as I know, this idea has not been explored either theoretically or empirically.
25. Charles Moskos and John Sibley Butler, *All That We Can Be* (New York: Basic Books, 1996).
26. By analogy, suppose that one is evaluating different distributions of income based on a utilitarian social welfare function $\sum_i u(y_i)$. As is well known, if $u(\cdot)$ is concave, then some preference by the policymaker for egalitarian distributions is built into the social welfare function even though social welfare places no intrinsic value on equality.
27. Harry Holzer and David Neumark, 'Assessing Affirmative Action', *Journal of Economic Literature* 38, 3 (2000): 483–568.
28. Julian R. Betts and Jamie L. Shkolnick, 'The Effects of Ability Grouping on Student Achievement and Resource Allocation in Secondary Schools', *Economics of Education Review* 19 (2000): 1–15.
29. Individual studies may be identified in which classroom complementarities appear to be present. See, for example, Maureen Hallinan, 'Ability Grouping and Student Learning', *Brookings Papers on Education Policy* (2003): 95–124. While Hallinan does not focus on complementarities, she provides evidence that they are present in English but not Mathematics. But my reading of the literature suggests that the Betts and Shkolnick summary of the state of empirical evidence is reasonable. See Betts and Shkolnick, 'The Effects of Ability Grouping on Student Achievement and Resource Allocation in Secondary Schools'.

30. Adam Gamoran, 'Instructional and Institutional Effects of Ability Grouping', *Sociology of Education* 59 (1986): 185–198.
31. See William Brock and Steven Durlauf, 'Interactions-Based Models', in *Handbook of Econometrics*, *Vol. 5*, edited by James Heckman and Edward Leamer (Amsterdam: North Holland, 2001) and Lawrence Blume and Steven Durlauf, 'Identifying Social Interactions: A Review', in *Methods in Social Epidemiology*, edited by J.M. Oakes and J. Kaufman (San Francisco, CA: Jossey-Bass, 2006) for overviews.
32. Charles Manski, 'Identification of Endogenous Social Effects: The Reflection Problem', *Review of Economic Studies* 60 (1993): 531–42.
33. James Heckman, 'Micro Data, Heterogeneity, and the Evaluation of Public Policy', *Journal of Political Economy* 109 (2001): 673–748, provides a brilliant overview of his work.
34. E. Anderson makes similar criticisms of the claim that diversity inherently adds to educational quality. See Elizabeth Anderson, 'Integration, Affirmative Action, and Strict Scrutiny', *NYU Law Review* 77 (2002): 1195–271 and Elizabeth Anderson, 'Racial Integration as a Compelling Interest', *Constitutional Commentary* 21 (2004): 101–27.
35. Samuel Sommers, 'On Racial Diversity and Group Decisionmaking: Identifying Multiple Effects of Racial Composition on Jury Deliberation', *Journal of Personality and Social Psychology* 90 (2006): 597–612.
36. P. Gurin explains these findings as 'primarily because the quality of interracial contacts is a key determinant of many of these educational outcomes', with no explanation as to why this is sufficient to explain the absence of evidence. See Patricia Gurin 'The Compelling Need for Diversity in Higher Education', Testimony in *Gratz, et al. v. Bollinger, and Grutter, et al. v. Bollinger, et al.*, http://www.vpcomm.umich.edu/admissions/research/expert/studies.html. A recent semipopular book, Scott Page, *The Difference* (Princeton, NJ: Princeton University Press, 2007) contains a large number of citations of studies purporting to find efficiency effects, but this compilation of evidence is unfortunately done in a polemical fashion, with no effort seriously to evaluate the studies either on their own terms or in terms of how they may be extrapolated to justify his grandiose claims about the benefits of diversity. Diversity effects are a type of social interaction and are subject to the statistical difficulties I have described.
37. For example, see Miriam Komaromy, Kevin Grumbach, Michael Drake, Karen Vranizan, Nicole Lurie, Dennis Keane and Andrew Bindman, 'The Role of Black and Hispanic Physicians in Providing Health Care for Underserved Populations', *New England Journal of Medicine* 334 (1996): 1305–10.
38. This importance of evaluating the effects of a factor such as the ethnic composition of doctors from an equilibrium-matching perspective has long been known in economics. Gary Becker, *The Economics of Discrimination* (Chicago, IL: University of Chicago Press, 1957) is the classic reference.
39. Martin Stinson and Norman Thurston, 'Racial Matching Among African-American Physicians and Hispanic Patients', *Journal of Human Resources*, 37 (2002): 410–28.
40. James Heckman and Peter Siegelman, 'The Urban Institute Audit Studies: Their Methods and Findings', in *Clear and Convincing Evidence*, edited by M. Fix and

R. Struyk (Washington DC: Urban Institute Press, 1993); James Heckman, 'Detecting Discrimination', *Journal of Economic Perspectives* 12 (1998): 101–16; and Robert Bohnholz and James Heckman, 'Measuring the Disparate Impacts and Extending Disparate Impact Doctrine to Organ Transplantation', *Perspectives in Biology and Medicine* 48 (2005): S95–S122.

41. See Bok and Bowen, *The Shape of the River*, and Sigal Alon and Marta Tienda, 'Assessing the "Mismatch" Hypothesis: Differences in College Graduation Rates by Institutional Selectivity', *Sociology in Education* 78 (October 2005): 294–315.
42. Linda Datcher-Loury and David Garman, 'College Selectivity and Earnings', *Journal of Labor Economics* 31 (1995): 289–308.
43. Richard Sander, 'A Systematic Analysis of Affirmative Action in Law Schools', *Stanford Law Review* 57 (2004): 367–484.
44. See Ian Ayres and Richard Brooks, 'Does Affirmative Action Reduce the Number of Black Lawyers?', *Stanford Law Review* 57 (2004): 1807–54; Jesse Rothstein and Albert Yoon, 'Mismatch in Law School', mimeo (Princeton University, 2006); and Jesse Rothstein and Albert Yoon, 'Affirmative Action in Law School Admissions: What Do Racial Preferences Do?' mimeo (Princeton University, 2007).
45. Justin McCrary, 'The Effect of Court-Ordered Quotas on the Composition and Quality of Police', *American Economic Review* 97 (2007): 318–53; John Lott, 'Does a Helping Hand Put Others At Risk? Affirmative Action, Police Officers, and Crime', *Economic Inquiry* 38 (2000): 239–76.
46. David Sklansky, 'Not Your Father's Police Department: Making Sense of the New Demographics of Law Enforcement', *Journal of Criminal Law and Criminology* 96 (2006): 1209–43.
47. Notice that rational behavior, as I have defined it, can be influenced by factors such as peer group and role model effects, limited information about the future, and so on.
48. Durlauf and Seshadri, 'Is Assortative Matching Efficient?'
49. Stephen Coate and Glenn Loury, 'Will Affirmative-Action Policies Eliminate Negative Stereotypes?', *American Economic Review* 83 (1993): 1220–40.
50. Kim Sau Chung, 'Affirmative Action as an Implementation Problem', mimeo (Northwestern University, 1999).
51. A well known defense is Judith Thomson, 'Preferential Hiring', *Philosophy and Public Affairs* 2 (1973): 364–38. Robert Amdur, 'Compensatory Justice: The Question of Costs', *Political Theory* 7 (1979): 229–44 is a valuable discussion of the merits of such claims.
52. John Ogbu, *Black Students in an Affluent Suburb* (Mahwah, NJ: Lawrence Erlbaum Associates, 2003).
53. George Akerlof and Rachel Kranton, 'Economics and Identity', *Quarterly Journal of Economics* 115 (2000): 715–53; Hanming Fang and Glenn Loury, 'Toward an Economic Theory of Dysfunctional Identity', in *The Social Economics of Poverty: On Identity, Communities, Groups and Networks*, edited by Chris Barrett (London: Routledge, 2005).
54. Race is also a separate and significant factor. See Douglas Massey and Nancy Denton, *American Apartheid* (Cambridge, MA: Harvard University Press, 1993).
55. Steven Durlauf, 'The Memberships Theory of Inequality: Ideas and Implications', in *Elites, Minorities, and Economic Growth*, edited by E. Brezis and P. Temin

(Amsterdam: North Holland, 1999), pp. 161–78; Steven Durlauf, 'The Memberships Theory of Poverty: The Role of Group Affiliations in Determining Socioeconomic Outcomes', in *Understanding Poverty in America*, edited by S. Danziger and R. Haveman (Cambridge, MA: Harvard University Press, 2001), pp. 392–416.

56. I provide a defense of this broad claim in Steven Durlauf, 'Assessing Racial Profiling', *Economic Journal* 116 (2006): F402–F426.
57. I defend this claim about the state of empirical evidence in Steven Durlauf, 'Groups, Social Influences, and Inequality: A Memberships Theory Perspective on Poverty Traps', in *Poverty Traps*, edited by S. Bowles, S. Durlauf, and K. Hoff (Princeton, NJ: Princeton University Press, 2006), pp. 141–75. Steven Durlauf, 'Neighborhood Effects', in *Handbook of Regional and Urban Economics, Vol. 4*, edited by J.V. Henderson and J.-F. Thisse (Amsterdam: North Holland, 2004) surveys the empirical literature.
58. Formally, this model may be thought of as treating both equity and efficiency payoffs as the sum of observable and unobservable components; without loss of generality, I think of the observables and unobservables as independent of each other in the probabilistic sense.
59. A useful reference on discrete choice analysis is Simon Anderson, Andrea de Palma, and Jacques-Francois Thisse, *Discrete Choice Theory of Product Differentiation* (Cambridge, MA: MIT Press, 1992).
60. Ellis Cose, *Rage of a Privileged Class* (New York: Harper Collins, 1993).
61. Lars P. Hansen and Thomas Sargent, *Robustness* (Princeton, NJ: Princeton University Press, 2007) provides a deep reanalysis of many macroeconomic questions from the perspective of ambiguity using minimax ideas. Charles Manski, *Identification for Prediction and Decision* (Cambridge, MA: Harvard University Press, 2007) summarizes his important work in microeconomics using minimax regret.
62. Durlauf, 'Groups, Social Influences, and Inequality: A Memberships Theory Perspective on Poverty Traps', pp. 141–75.
63. Gerald Gaus, 'Social Complexity and Evolved Moral Principles'.
64. John Roemer, 'Three Egalitarian Views and American Law', *Law and Philosophy* 20: (2001), 433–460 specifically discusses affirmative action. See also John Roemer, *Equality of Opportunity* (Cambridge, MA: Harvard University Press, 1998) for a lovely summary of his thought.
65. The division between predictable and unpredictable equity payoffs is artificial in the sense that one can interpret violations of my assumptions as meaning either that the predictable parts have a different sign or that the unpredictable parts do not have a mean of 0. My arguments do not depend on which way one thinks about the violations.
66. My own attempts to justify this presumption may be found in Durlauf, 'Groups, Social Influences, and Inequality: A Memberships Theory Perspective on Poverty Traps' and Durlauf, 'Assessing Racial Profiling'.