

WORKING PAPER · NO. 2026-62

Retrospective Versus Prospective Meritocracy

Steven N. Durlauf

APRIL 2026

Retrospective Versus Prospective Meritocracy
Steven N. Durlauf
April 2026
JEL No. D63, D78, I24, I28, J24

ABSTRACT

This article proposes a way of understanding meritocracy from retrospective versus prospective points of view. Retrospective meritocracy is static or backwards-looking: Merit is based on an individual's characteristics or past achievements as representative of excellence or as desert for a position already obtained. Prospective meritocracy is forward-looking: Merit is functionally defined as the comparative contribution that an individual makes to a specified set of social objectives. I use formal models to show that these alternative conceptions have very different implications for meritocratic assignments of students to schools or workers to jobs and that they involve differing information needs for a policymaker. These different approaches to merit demonstrate how alternate versions of meritocracy may or may not be socially efficient. I discuss implications for the use of meritocracy as a desideratum for various public policy contexts.

Steven N. Durlauf
The University of Chicago
Harris School of Public Policy
Stone Center for Research on Wealth
 Inequality and Mobility
and NBER
durlauf@gmail.com

1. Introduction and general argument

In this essay, I offer some ways of thinking about meritocracy that consider the contexts in which individuals are matched to positions or opportunities—specifically, students to classrooms, applicants to universities, and workers to jobs. These contexts are standard environments for economic analysis, and social scientists and philosophers continue to debate the extent to which assignments in these environments should or should not be meritocratic. My objective is to provide some clarity to these debates by exploring ways to assess whether a given assignment rule is more or less meritocratic. I do so from the vantage point of an economic theorist, studying the rules that are used to assign students to schools and workers to firms and assessing the consequences of the assignments. I do not account for the way in which various assignment rules become policies. Economists call this type of exercise a “social planner” problem, which takes the lofty vantage point of a planner who can implement any policy they wish. This allows for clarity in understanding the internal logic and intrinsic features of various rules. I proceed by studying a social planner who chooses assignment rules for a given socioeconomic environment and then consider how to characterize meritocratic assignments in that environment and assess their consequences.

My claim is that, for two reasons, the “economic perspective” can speak to a number of dimensions in the debates over meritocracy. First, to discuss efficiency, it is necessary to describe the objectives that an assignment rule is supposed to meet. In other words, efficiency requires assumptions about the ways that a social planner assesses the structure and outcomes produced by an assignment rule. Economic models force an explicit specification of objective functions in order to assess efficiency. Second, economic models require specification of the structure of the socioeconomic environment in which assignments are being made. This is necessary in order to understand the implications of a given objective function for how to make the assignments.

In pursuing these objectives, I will propose a way of answering a famous critique of Amartya Sen about debates over meritocracy and justice:

...meritocracy, and more generally the practice of rewarding merit, is essentially underdefined, and we cannot be sure about its content—and thus about the claims regarding its “justice”—until some further specifications are made (concerning, in particular, the objectives to be pursued, in terms of which merit is to be, ultimately, judged). The merit of actions—and (derivatively) that of persons performing actions—cannot be judged independent of the way we understand the nature of a good (or an acceptable) society. (Sen 2000, 5–6)

I attempt to answer Sen’s underdefinition critique by choosing a specific socioeconomic context in which to discuss meritocracy and by providing a range of candidates as definitions of merit, none of which are themselves novel.

My discussion will develop a distinction between *retrospective* versus *prospective* meritocracy, a distinction that is implicit in Durlauf (2008) and made explicit and elaborated in Durlauf (2026). Retrospective meritocracy refers to the idea that assignments should be based on an individual’s achievements and characteristics in a way that these are rewards. (The reasons that certain ones of these constitute merit while others do not are discussed below.) For example, meritocratic college admissions are often understood that standardized test scores and grades should, in isolation, determine priorities for admissions. Retrospective meritocracy, more generally, corresponds to the conventional understanding of meritocracy in public policy debates. This perspective is essentially static and backward-looking. That is, when making allocations in schools or firms, retrospective meritocracy does not involve a nontrivial evaluation of the consequences of the assignment rule for various future outcomes.¹ It can also be interpreted as conceptualizing merit as a reward. Clearly, the reward notion is natural in certain contexts: One’s merit for the title World Chess Champion is determined by performance in the competition.² For such cases, the retrospective notion is, in my judgment, sufficient. Competitions define rewards and thus the merit of participants. In other cases, such as college admissions, the analogy to competition is indirect, in the sense that grades and test scores are regarded as determining who deserves a position.

By prospective meritocracy, I refer to the idea that assignments are made to best achieve the objectives of the decision-maker—the person making the assignments—and that assessment accounts for the effects of the assignments today and in the future. This conception of meritocracy is essentially forward-looking because the trajectories of the

socioeconomic outcomes that follow from the assignments are integral to the assessment of whether a given rule is meritocratic. Prospective meritocracy therefore provides a *functional* definition of merit. The comparative merits of two individuals for a given position are implicitly defined by their contributions to the social planner's objectives. As such, meritocratic decision rules are, by construction, efficient, conditional on these objectives.

These formulations of meritocracy employ social planner objectives and the facts of the socioeconomic context under study in different ways. Retrospective meritocracy uses intrinsic excellence or desert as sufficient to characterize meritocracy. In doing so, retrospective meritocracy reasons from an a priori concept of merit that resolves the specification of objectives and does not involve the socioeconomic context under consideration in the way that prospective meritocracy does. This is because the prospective meritocrat will evaluate the consequences of rule based on how the rule influences that future trajectories of those who would be affected by the rule. This difference implies that a retrospective meritocrat needs less information about the socioeconomic environment under analysis to operationalize meritocracy than does a prospective meritocrat.

The distinction I make between prospective and retrospective meritocracy parallels an important distinction in the meritocracy literature between principled and instrumental defenses of meritocracy. Principled defenses argue that merit creates intrinsic claims on the part of an individual to certain outcomes, such as college admissions or jobs, without an assessment of the attendant consequences *per se*. Mulligan's³ is a valuable discussion of intrinsic excellence as warranting assignment to certain positions. His analysis addresses why certain achievements and characteristics warrant serving as measures of merit; for example, mathematical skill is a merit for a position in a physics PhD program, not a football team. Retrospective defenses of meritocracy may also be derived from the notion of desert. While conceptually distinct, desert can provide a powerful basis for meritocracy (see Mulligan [2018] for a strong defense of this view and Mulligan [2023] for how the argument has been made in the philosophy literature). Instrumental defenses of meritocracy, in contrast, are predicated on the extent to which the rules of meritocracy produce a system with desirable features that lie outside the intrinsic claims of excellence or desert.

This retrospective/prospective distinction facilitates understanding the different ways that social scientists have assessed the degree of meritocracy in particular contexts. One approach assesses the extent of meritocracy by determining the extent to which certain variables, for example, education, explain intergenerational persistence in class or occupational status (see Breen and Goldthorpe 1999, 2001). This is a retrospective notion of meritocracy, where education is the intrinsic excellence that warrants an individual's outcomes. Bhattacharya et al. (2017) and Bhattacharya and Shvets (2024) study how different factors in admissions predict academic performance and so engage in a prospective assessment.

To elaborate on these ideas, I will work through a set of simple examples that focus on the assignment of students to schools. Other types of assignment problems may be analyzed with mathematically equivalent models. For example, the assignment of students to classes in a school and the assignment of workers to jobs each follow the identical abstract logic of the college assignment problem.⁴ To conduct my analysis, I employ some mathematical formalism for two reasons. First, the formalization will make precise a range of types of claims about which particular assignment rules are or are not meritocratic. Formal economic models always require deep simplifications relative to the phenomena under study. My hope is that this discussion will illustrate how, despite these simplifications, formal theory can produce substantive insights by revealing common features that underlie different environments. Second, these formal models will help to make clear what types of empirical evidence are needed to determine whether an assignment rule is or is not meritocratic in a given context.

Before proceeding, I should note that I do not address broad assessments of meritocracy, in particular, critical arguments made by Markovits (2019) and Sandel (2020). Markovits emphasizes the harms that competitive pressures can inflict and the links between self-esteem and economic success on all members of a society, successful or not. He and Sandel argue that meritocratic thinking damages social solidarity and induces contempt for the less successful, a view that Williams (2024) demonstrates has been a recurrent theme in moral and political philosophy since Plato. My goal is more modest: to characterize what meritocracy means from the vantage point of a social planner trying to implement it. Whether the social planner should be a meritocrat—and, if

so, how to weigh meritocratic objectives for other social desiderata—is outside the domain of my discussion.

2. Meritocracy and school assignments

University admissions are a standard context in which public policy debates have invoked meritocracy. The history of debates over affirmative action, and associated Supreme Court decisions, ranging from *Regents of the University of California v. Bakke* in 1978 (upholding affirmative action but banning racial quotas in admissions) to *Students for Fair Admissions (SFFA) v. Harvard* and *SFFA v. University of North Carolina* (ending affirmative action in 2023), testify to the salience of meritocracy in major policy controversies.

These debates have tended to presume that there is a natural notion of merit (i.e., past academic performance) and treat rules that do not prioritize those with higher grades and test scores as unmeritocratic. I argue that the assessment of meritocracy is far more complex. To explore general ideas underlying retrospective versus prospective meritocracy, I consider a sequence of stylized college admissions decision problems and characterize how retrospective and prospective meritocrats would choose.⁵

i. A baseline model

Imagine two colleges and two students. One college is high-quality, measured by c^H , and the other low-quality, measured as c^L . In parallel, one student has an initial high skill level e^H and the other has an initial low skill level e^L . Attendance at a college means that a student will, upon graduation, possess an educational skill level measured as a scalar g .⁶ The skill level g is determined by an education production function,

$$g = f(e, s).$$

Assume that the level of skill at graduation is increasing in the student's initial skill e and in the college's quality s .⁷ Neither assumption is empirically problematic.

Suppose that a social planner sets the admission rules for the two colleges.⁸ This simple environment has only two possibilities: Students are assigned to pair high skills and high quality or the inverse holds. The matching of like with like is often called assortative matching, while the inverse is called negative assortative matching. Conventional discussion of meritocracy assumes that it entails assortative matching of student skills and college quality. But is this so? To rigorously link meritocracy to college assignments, it is necessary to specify whether the social planner is a retrospective or a prospective meritocrat and then to determine how merit is defined for each type of social planner.

For the retrospective meritocrat, the solution to the student assignment problem requires specifying the intrinsic excellence associated with merit or the characteristics of a student that create a desert claim on an admission. For this example, initial educational skill is a natural definition for merit from the retrospective vantage point. Educational skill represents the intrinsic excellence that Mulligan (2018)⁹ discusses and is consistent with the idea that a student deserves admission because of the skill they possess at the time of admission. This view of merit is, I believe, the conventional one in public discourse.¹⁰

If an applicant's skill level represents their merit, then the meritocratic assignment rule is straightforward: The high-skilled student is assigned to the high-quality college, and the low-skilled student is assigned to the low-quality college. In this model, the excellence of a student is measured by their skill, since (by assumption) they will be able to enjoy better outcomes at the high-quality college than the low-skilled student will at the high-quality college. This is an unobjectionable assumption but warrants highlighting since it means that retrospective meritocratic arguments involve empirical claims, even for this context. The assumption also captures what seems a requirement that the high-skilled student deserves the opportunity to attend the high-quality college; if a student's education would be *diminished* by attending the high-quality institution rather than the low-quality school, it is not clear how they could reasonably be said to deserve that slot over the other student. Further, if grades and test scores are used to uncover a student's educational skill level, then the meritocratic assignment may be operationalized by setting

different thresholds for the two schools. A meritocratic admissions rule, therefore, is one that sets combinations of grades and test scores such that applicants who exceed them are admitted while those who fall short are not.

From the perspective of prospective meritocracy, the assignment problem is more complicated. First, it is necessary to specify an objective function. Suppose that the goal of the social planner is to produce the most outstanding student possible. In other words, the objective is extreme excellence. Metaphorically, one can analogize the question of assigning students to music conservatories, where the goal is to create the most extraordinary composer. In this case, the prospective meritocratic assignment rule is straightforward. The high-skilled student is assigned to the high-quality college: Mozart attends Juilliard over anyone else. For this example, the retrospective and prospective rules coincide although the rationales behind them differ since, given the logic of the prospective meritocrat, Mozart does not have an intrinsic claim to be accepted by Juilliard.

The prospective meritocratic decision rule will not coincide with the retrospective rule for alternative objective functions. In contrast, suppose that the social planner wishes to maximize the outcome for the lower-skilled student. This objective function is often motivated by Rawls' (1971) difference principle; the qualitative features of this case would be preserved for more general prioritarian principles (Parfit [1991] 2002) for giving greater weight to the outcomes of lower-skilled students when aggregating across larger student populations. Here, I would ask what a prospective meritocrat would assign students when the objective functions embody explicit egalitarian considerations.

In this example, the egalitarian meritocrat's assignment rule is again simple: The lower-skilled student is assigned to the high-quality college. Relative to the objectives of the social planner, this outcome is efficient. While it may seem strange to call this assignment rule meritocratic, that is only because merit is *functionally* defined here; it is not intrinsic or justified by desert.¹¹

This solution and the way that it contrasts with the meritocratic decision rule maximizing extreme excellence illustrate two things. First, once one moves away from notions of intrinsic excellence and desert, the objective function of a social planner can produce very different implied levels of relative merit and, in the case of the egalitarian planner, orderings of merit that appear counterintuitive. Second, egalitarian objectives

can influence comparative merit and be efficient. Efficiency is defined relative to a set of objectives, and there is no reason an equality dimension must be excluded from them. And it is for this reason that facile discussions of equality/efficiency trade-offs can be misleading.

To complete the discussion of the policy choices of the prospective meritocrat in this stylized example, I consider a third objective function. Suppose that the social planner wishes to maximize the sum of the educational skill levels of the two students upon graduation. This type of objective function can be justified if the social planner is concerned about the average skill level of the labor force for a state. This objective function does not assign any intrinsic value to egalitarian considerations (see Durlauf [2026] for more discussion). This objective function is the baseline for assessing efficient assignments in many economic contexts.

The important feature of this objective function is that one cannot, given the information I have provided, determine the prospective meritocratic decision rule. The prospective meritocratic rule will assign the high-skilled student to the high-quality college and the low-skilled student to the low-quality college if

$$f(e^H, c^H) + f(e^L, c^L) > f(e^H, c^L) + f(e^L, c^H).$$

However, if the inequality is reversed, then the high-skilled student is assigned to the low-quality college and the low-skilled student is assigned to the high-quality college, even though the social planner does not have a preference that weights the low-skilled student differentially from the high-skilled one.

How can this be the case? To understand why this is so, rewrite the inequality that justifies matching the high-skilled student and the high-quality college:

$$f(e^H, c^H) - f(e^H, c^L) > f(e^L, c^H) - f(e^L, c^L).$$

What does this inequality say in words? The left-hand side of the expression is the differential in educational attainment from attending the high-quality college by the high-

skilled student while the right-hand side is the differential benefit to the low-skilled student from attending the high- versus the low-quality college. If the high-skilled student benefits relatively more, then they are assigned to the high-quality school. When the differential benefits of a high-quality match are increasing this way, student skill and college quality are said to be complements; if the inequality is reversed, student skill and college are said to be substitutes. The idea that the efficiency of matching like with like, known as assortative matching, is efficient when complementarity is present, but inefficient when substitutability is present, was originally recognized in a classic paper by Gary Becker (1973). The finding is robust to many richer specifications than this example (Durlauf and Seshadri 2003). In the case of Mozart, his assignment to Juilliard is efficient if the marginal value to him of attending it versus another school is greater than that of a less innately gifted student. I use this example as it illustrates a context where it might reasonably be the case that the less gifted student benefits more. The upshot of this argument is that the equivalence between the retrospective and prospective meritocratic assignment rules is determined by a feature of the socioeconomic environment—complementarity versus substitutability—that is an empirical question.

Therefore, the conventional economic notion of efficiency—maximizing average output—does not entail the conventional retrospective notion of meritocratic assignments. At a minimum, the conventional notion may be defended as producing efficient outcomes, but it need not do so. Now compare this to the distinctions between intrinsic and instrumental meritocracy. This example illustrates how they may lead to different assignment rules. Another way to say this is that intrinsic excellence and desert considerations need not be socially efficient, unless adherence to these principles is defined as the sole social objective of the social planner.

This third example has an important general implication. Prospective meritocracy focuses on the value added from matches, not on who will benefit the most from them per se. If the context under study were the assignment of workers to jobs, this would be completely natural. As discussed in Miller (1992), the natural notion of a desert claim for a job is that one candidate would be more productive than another. My discussion here employs that same claim for students, with the exception that “being better at a job” is

replaced with “will have a larger relative change in skills from the assignment” and converts the output of a firm to the total educational attainment of a school’s graduates.

In elaborating my discussion, I will focus, for prospective meritocrats, on the case where the social planner wishes to maximize average outcomes, since this involves the most interesting implications of the generalizations I examine.

ii. **Dynamic environments**

I now consider a two-stage generalization of this problem¹². Suppose that a social planner is determining assignment rules for high schools and colleges. Middle school graduates are assigned to high schools given their skill levels at middle school graduation. High schools determine the skills that are then used to assign students to colleges. In parallel to the earlier discussion, suppose that the two students graduate middle school with skill levels y^H and y^L and are assigned to high schools. One high school has quality level r^H and the other r^L . Assume that the skill level of a student at the time they are assigned to a college is determined by their initial skills and their high school match via:

$$s = g(y, r).$$

Assume high school skills are increasing in the student’s initial skills and the quality of the high school.¹³

What makes this environment qualitatively different from the first one I discussed? The answer is that the two stages may not naturally decompose so that one determines the assignment rule at one stage independently from the assignment rule for the other. The static assignment problem I have described treated the skills of the students as predetermined, i.e., the social planner takes them as given when determining the assignments of students. However, in this dynamic environment, the student skills of high school graduates are not predetermined; the social planner can influence them by the assignments of the students to high schools. The implications for the retrospective and the prospective meritocrats are very different.

If skills measure intrinsic excellence or determine merit, the retrospective meritocrat could argue that the assignments can be decomposed. A higher-skilled middle school graduate will have higher intrinsic excellence for high school and may deserve the slot at the high-quality high school because of this. This scenario corresponds to my claim that retrospective meritocracy thinks about assignments as a reward for past achievement.

For the prospective meritocrat, assume, as before, that their objective is to maximize the total educational quality of the college graduates. Further, I will focus on the case where the high school education production $g(y,r)$ and the college education production function $f(e,s)$ both exhibit complementarities. This assumption is the best-case scenario, so to speak, for the prospective and retrospective assignment rules to coincide given the objective function.

Under complementarity of the college production function, the prospective meritocrat will assign the high-skilled high school graduate to the high-quality college and the low-skilled graduate to the low-quality college. The reasoning here is that once the students have completed high school, the college assignment decision can take their skills as given, so original assignment analysis is unchanged.

What should be done in terms of assigning middle school graduates to high schools? Different assignment rules will produce different values for e^H and e^L . What do we know about these values as a function of the assignment rules? We know that assigning the high-skilled middle school student to the high-quality high school will maximize the sum of the skills of the two students when they graduate from high school, $e^H + e^L$, because middle school skills and high school quality are complements. However, the prospective meritocrat does not intrinsically value the average skills of high school students. Rather, the distribution of skills across the two students is a means to an end—in this case, the average skills of college graduates. There is no distinct meritocratic evaluation made for high school assignments since high school is an intermediate step in the process of producing educational quality after college.

The challenge for the prospective meritocrat is to choose between two different pairs of high school skills. Assortative matching of high school students and schools will

produce a pair of high school graduation skill levels $e^H(AM), e^L(AM)$, while the negatively assortative matching of the students produces skill levels $e^H(NAM), e^L(NAM)$. Assortative matching in colleges will occur under the assumption of complementarity. Therefore, assortative matching in high school will be justified if

$$f(e^H(AM), c^H) + f(e^L(AM), c^L) > f(e^H(NAM), c^H) + f(e^L(NAM), c^L).$$

Complementarity in the high school production function does not ensure this condition holds. Why? Complementarity ensures that the average skill level upon entering high school is maximized. However, it also produces the most variation between the skill levels of different students. The effects of this variation on skills can be negative for the overall education produced. This will depend on another property of the high school education production function: the concavity or convexity of $f(e, s)$ with respect to e .¹⁴ If the function is concave, the marginal effects of a given increase in skills on high school attainment are lower for a student with higher skills than an equivalent change for a student with lower skills. When this is so, then for a given average skill level, the greater the dispersion of skills between the two students, the lower the total skills the students will have upon graduation. It is possible that assortative matching at the high school level maximizes the total college education in a population. But it need not do so.

The lessons of this stylized example apply if one integrates education systems and labor markets. As suggested earlier, a prospective meritocrat may be concerned about economic performance and the skills of the workforce. If this is so, then an additional layer of dynamics needs to be considered. Suppose that the educational system produces a set of workers who are then employed by firms. In this case, the educational system produces a distribution of workers who will be matched to various firms. Suppose firm i has a production function

$$p(m_i, g_i),$$

where m_i is the skill level of the firm's manager and g_i the skill of the firm's employee. Assume the production function exhibits constant returns to scale and exhibits complementarity. This means that the high-skilled manager and the high-skilled worker will be matched, when managers and workers are paid their marginal products, so the equilibrium labor market outcomes produce total output:

$$p(m^H, g^H) + p(m^L, g^L).$$

Given the skills of managers and workers, assortative matching maximizes total output. But in parallel to my earlier discussion, it may not be the case that the assortative matching assignment rules in education are efficient, even if they maximize the sum of college skills across the two workers. Concavity of the firm production function works in a way analogous to concavity of the educational production functions, so a lower value of total college skills will be more productive in labor markets because of lower skill dispersion. If the objective of the social planner is to maximize aggregate output, the distribution of skills matters, not just the average.

Additional complications occur if one extends the educational process to include student effort. To do this, one would have to model educational outcomes as a function of initial student skills, student effort, and school quality. The complication relative to our baseline model is that one needs to have a model of student effort. The simplest way to do so is to model students as trading effort for educational attainment. In static models, the natural generalization does not change the qualitative claims I have made. In dynamic models, however, it can because student efforts today will depend on the effect of their efforts on future school enrollments. We can therefore introduce complicated interactions between future school quality and current effort levels. It is possible that if assortative matching in later stages of education ensures that earlier efforts have low payoffs, then disincentives for effort may prove to be dynamically inefficient. This observation is about more than mathematical models. There are reasons to believe that affirmative action enhances aspirations and efforts of students because opportunities are enhanced; see

Bleemer (2022) and Akhtari et al. (2024) for evidence that affirmative action enhances efficiency in this sense.

iii. Matching with interactions

The lessons for meritocracy for dynamic models may be amplified when we consider interactions that result from matching. High schools and colleges admit classes, not single students. A natural generalization of the high school education production functions so far employed is to model the educational attainment of a high school student as

$$s = g(y, r, y_{-i}),$$

where y_{-i} denotes the skills of high school classmates. Similarly, suppose that college educational attainment is determined by

$$g = f(e, s, e_{-i}),$$

where e_{-i} denotes the skills of a student's classmates. Suppose that there are complementarities between the skills of different students and the quality of the school for both high school and college.

This richer model of the school environment again exhibits the differences between retrospective and prospective meritocrats. The retrospective meritocrat sorts the students for high school and college according to entering skills. The prospective meritocrat does not have a straightforward decision rule. Again, focusing on an objective of maximizing average college graduate skills, the efficient assignments of students in high school need to account for interstudent influences and trade off complementarity and concavity.

Similar considerations hold for richer models of firms. Suppose that the educational system produces a set of workers who are then employed by firms. In this case, the educational system produces a distribution of workers who will be matched to various

firms. Suppose firm i has a production function in which output is determined by the skill of the manager and the skills of the workers at the firm, i.e.,

$$p_i(m_i, g_{1i} \dots g_{iN}),$$

where $g_{1i} \dots g_{iN}$ is the vector of educational skill levels for the N workers employed by the firm. For this environment, the quality of each person's coworkers will determine their own marginal productivity. This creates the possibility that the highest-skilled workers will have higher productivity if coworkers have higher skill levels, even if the associated educational assignment rules did not maximize those workers' skills.¹⁵

There are some broad messages that may be taken from the stylized examples I have described. From the dynamic perspective, a sequence of conditionally meritocratic decision rules (i.e., they are meritocratic when applied at a point in time without consideration of their intertemporal consequences) may be dynamically inefficient and therefore unconditionally not meritocratic. The inefficiency derives from the excess segregation that occurs when considering the assignments outside their dynamic consequences. A school with year-by-year classroom tracking or even an education system that places adolescents on college or vocational tracks can exhibit such possibilities. This possibility can be amplified by the same complementarities that made segregation efficient in static cases. See Durlauf (2026) for elaboration.¹⁶

Prospective meritocracy becomes even more complex when one considers links between generations. The educational admissions policies at a given level of school help determine the overall distributions of educational outcomes for parents, along with differences in occupations and incomes. These different inequalities affect inequalities in the next generation that will matter for social objectives that apply to all generations. For example, income inequalities today help determine heterogeneities in neighborhood and school quality that determine the differences in skills involved in high school and college assignments. The relationship between contemporary parental inequalities and persistence of socioeconomic status between parents and children is sometimes called the Great Gatsby Curve. See DiPrete (2020) and Durlauf et al. (2022) for surveys of evidence and mechanisms.

3. Limits to Knowledge

One significant implication of this discussion about retrospective and prospective meritocratic rules is that they have very different information requirements for their operationalization. My discussion emphasizes that, if a social planner is to implement dynamically efficient assignment rules, they need information on properties of the educational production functions for schools and colleges and the output production functions of firms. This is qualitatively different from the knowledge that is necessary to be able to characterize factors that make up intrinsic excellence. For the examples I have outlined, a prospective meritocrat requires knowledge of educational production functions of schools and firms. The information requirements become an order of magnitude greater if intergenerational considerations come into play.

Currently, empirical social science has very limited information to address these information requirements. Consider education. In a major survey of the classroom tracking literature, Betts (2011) concluded there were few general lessons to be taken on whether tracking raised or lowered average student performance. Similarly, Bleemer and Rothstein (forthcoming) conclude that there is little evidence on the complementarity versus substitutability between college students and schools. Studies of the effects of different assignment rules on educational attainment need to account for the interactions that occur among students, school officials, and peers and are known to face complicated identification problems. Blume, Brock, Durlauf, and Ioannides (Benhabib et al. 2011, ch. 18), Epple and Romano (Benhabib et al. 2011, ch. 20), and Graham (Benhabib et al. 2011, ch. 19) together provide a comprehensive discussion of the host of issues that arise. This is not to say that there are any impossibility theorems in the econometrics and statistics literature that preclude generating the types of knowledge needed by the prospective meritocrat, but rather that great care is needed in empirical work to make analyses credible. Further, this does not mean that individual programs have not yielded information about the effects of different assignment rules. Card and Giuliano (2016) and Nomi et al. (2021) provide evidence of the benefits of classroom tracking for specific

programs and specific school districts. Their findings suggest how detailed knowledge of specific contexts is necessary to draw lessons across evidence for particular programs.

4. General Lessons for Meritocracy

What messages may be taken from the stylized models that I have reviewed?

First, the divergence between retrospective and prospective meritocratic perspectives shows that Sen was correct in claiming that “meritocracy” underdetermines the assessment of different socioeconomic environments. Most important, different stances on meritocracy as a social desideratum carry different implications if one uses principled versus instrumental justifications. On the other hand, the examples I have discussed show how Sen’s challenge can be met by an explicit characterization of socioeconomic environments and of the social planner’s objectives. These characterizations are not unique. By implication, any claims that a given assignment rule is or is not meritocratic requires prior specification of the normative and positive claims about a given socioeconomic environment. Meritocracy does not speak for itself.

Second, standard claims about the relationship between meritocracy and efficiency are flawed for two reasons. First, since each of the prospectively meritocratic assignment rules I have described is efficient, it is not meaningful to claim there is a trade-off. Equality and efficiency require specific definitions with respect to a socioeconomic environment. Once one has defined them, the extent to which each is fulfilled will be determined by the assignment rule (or other policy) that has been implemented. In other words, the equality and efficiency properties of an environment are jointly determined. By analogy, one never speaks of prices and quantities exhibiting a positive or a negative trade-off in a market; they are jointly determined by forces that affect supply and demand. In discussing meritocracy, one can go further to say that a blanket claim that meritocracy is inefficient is false in the sense that a principled meritocrat who chooses retrospective assignment rules for students might produce inefficient outcomes, if efficiency is measured by average educational attainment or total output.

Third, I believe this discussion leads to a general problem in thinking about meritocracy: its inappropriate application to each of the stages of a dynamic process. There exists a plethora of ways in which students are assigned to distinct educational environments at different ages. As described by Diaz and Rivera (2025), some sorting starts as early as preschool. Schools can track students either explicitly or implicitly via differences in course offerings. There is a range of elite public high schools in America with competitive admissions. Colleges, of course, set admissions rules for undergraduates and distinct rules for graduate programs and professional schools. The upshot of my discussion of prospective meritocracy is that these assignment rules can, even if conditionally efficient at each stage, be collectively inefficient. Durlauf (2026) gives an example where sequential static efficiency is dynamically inefficient. This example has the interesting property that the efficient assignment rule starts by, first, integrating students of different abilities and, second, segregating them. The example also demonstrates how the retrospective meritocrat can choose rules that are inefficient in terms of overall educational quality. This is perhaps no surprise since the efficiency argument for meritocracy does not entail particular implications with respect to the reasons for principled meritocracy.

Why is integration in earlier years followed by later segregation efficient? Sequential assortative matching at every stage of education ensures that some students will have low skills by the time that they enter college, which will reduce average educational outcomes regardless of how the students are matched in the college admissions process. The appropriate way to think about education, in other words, is that the assignment rules across the education life course are a means to an end: the distribution of educational attainment at the end of the process. Before the final period, assignment rules shape the distribution of skills, which will be used for assortative matching at the end. The dynamically efficient assignment rules can initially integrate students of various skills with various schools and then segregate them in the ways I have described.

Fourth, this discussion has implications for understanding the relationship between meritocracy and equality of opportunity. A disagreement exists between scholars of meritocracy over whether meritocracy entails equality of opportunity. Miller (this volume)

treats equality of opportunity as a requirement of meritocracy since, in his view, there must be fair access to the skill and ability acquisition that applies to school and firm positions, while Markovits (2019) argues that meritocracy can reduce equality of opportunity as elites hoard opportunities for their children. Prospective meritocracy provides a way of understanding their disagreements. Dynamic efficiency considerations can require that the ability of parents to hoard opportunities is limited by policy interventions. However, because of complementarities and convexities in the educational and production processes, there is no reason to conclude that dynamically efficient meritocracy would lead to equality-of-opportunity conditions that would meet the ethical demands of responsibility-sensitive egalitarianism, for example, that of Roemer (1993, 1998).¹⁷

5. Conclusions

Michael Young (1958) created the term “meritocracy” to characterize the rise of a self-perpetuating, educated elite in the United Kingdom. It was meant to be a warning. It is an irony, then, that Young’s idea has proven to be so protean in its conception and wide-ranging in its reception. My contribution here is an attempt to provide some contour to the many ways that meritocracy has been thought of and evaluated. The notions of retrospective and prospective meritocracy provide a way of operationalizing the ideas of intrinsic versus instrumental justifications for meritocracy. In turn, these different views of meritocracy lend themselves to analysis through formal economic models, which can provide links between general concepts and specific policies.

The important conclusions of this article are twofold. First, retrospective and prospective meritocrats can produce very different policies that reflect their different conceptions of the meaning of merit. Prospective meritocrats want to address social and economic efficiencies in ways that are not available to retrospective meritocrats. In other words, meritocracy does not speak for itself. Second, the informational requirements for achieving prospective meritocracy are greater if merit is understood to be a function of a policymaker’s preferences and the structure of the socioeconomic environment. Hence,

there are differences in the potential for complete implementation of meritocracy between the two conceptions.

The differing information requirements of retrospective and prospective meritocracy, and their association with principled and instrumental justifications for meritocracy, suggest an interesting parallel with arguments made by Hayek (1960) and later extended and deepened by Gaus (1998, 2007) that policy choices should be made on a priori principles since consequentialist claims for the policies are undermined by the unpredictability of their consequences. One might conclude that, in other words, the information requirements of the prospective meritocrat are so great as to create an argument for acting as a retrospective one. In my judgment this is too harsh. Robust decision-making is an active area of research; see Hansen and Sargent (2008) for one influential approach. The work develops ways to design policy in information-thin environments. Further, there are cases where various qualitative features of a socioeconomic environment can be independent of details of the environment (Durlauf 2012). Exploring the feasibility of different meritocratic approaches under different degrees of information about socioeconomic structures is therefore a natural next step in the types of analysis I have described.

References

- Akhtari, Mitra, Natalie Bau, and Jean-William Laliberté. 2024. Affirmative action and precollege human capital. *American Economic Journal: Applied Economics* 16 (1): 1–32.
- Anderson, Axel. 2015. A dynamic generalization of Becker's assortative matching result. *Journal of Economic Theory* 159 (A): 290–310.
- Becker, Gary S. 1973. A theory of marriage: Part I. *Journal of Political Economy* 81 (4): 813–846.
- Bénabou, Roland. 1996. Heterogeneity, stratification, and growth: Macroeconomic implications of community structure and school finance. *American Economic Review* 86 (3): 584–609.
- Benhabib, Jess, Alberto Bisin, and Matthew O. Jackson, eds. 2011. *Handbook of social economics, volume 1B*. Handbooks in Economics. North-Holland.
- Bhattacharya, Debopam, Shin Kanaya, and Margaret Stevens. 2017. Are university admissions academically fair? *Review of Economics and Statistics* 99 (3): 449–464.
- Bhattacharya, Debopam, and Julia Shvets. 2024. Inferring trade-offs in university admissions: Evidence from Cambridge. *Journal of Political Economy* 132 (11): 3737–3784.
- Bleemer, Zachary. 2022. Affirmative action, mismatch, and economic mobility after California's Proposition 209. *Quarterly Journal of Economics* 137 (1): 115–160.
- Bleemer, Zachary, and Jesse Rothstein. Forthcoming. The meritocratic consensus and stratification in higher education. *Journal of Economic Perspectives*. Available from https://zacharybleemer.com/wp-content/uploads/Working-Papers/Meritocracy_Paper.pdf.
- Betts, Julian R. 2011. The economics of tracking in education. In *Handbook of the economics of education*, volume 3, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann. Elsevier.
- Breen, Richard, and John H. Goldthorpe. 1999. Class inequality and meritocracy: A critique of Saunders and an alternative analysis. *British Journal of Sociology* 50 (1): 1–27.
- Breen, Richard, and John H. Goldthorpe. 2001. Class, mobility and merit: The experience of two British birth cohorts. *European Sociological Review* 17 (2): 81–101.

- Card, David, and Laura Giuliano. 2016. Can tracking raise the test scores of high-ability minority students? *American Economic Review* 106 (10): 2783–2816.
- Cooper, Suzanne J. 1998. A positive theory of income redistribution. *Journal of Economic Growth* 3 (2): 171–195.
- Diaz, Estela B., and Lauren A. Rivera. 2025. Essentializing merit: Disability and exclusion in elite private school admissions. *American Sociological Review* 90 (3): 455–492.
- DiPrete, Thomas A. 2020. The impact of inequality on intergenerational mobility. *Annual Review of Sociology* 46:379–398.
- Durlauf, Steven N. 1996a. Associational redistribution: A defense. *Politics & Society* 24 (4): 391–410.
- Durlauf, Steven N. 1996b. Neighborhood feedbacks, endogenous stratification, and income inequality. In *Dynamic disequilibrium modelling: Theory and applications; Proceedings of the ninth international symposium on economic theory and econometrics*, edited by William A. Barnett, Giancarlo Gandolfo, and Claude Hillinger. International Symposia in Economic Theory and Econometrics. Cambridge University Press.
- Durlauf, Steven N. 1996c. A theory of persistent income inequality. *Journal of Economic Growth* 1 (1): 75–93.
- Durlauf, Steven N. 2008. Affirmative action, meritocracy, and efficiency. *Politics, Philosophy & Economics* 7 (2): 131–158.
- Durlauf, Steven N. 2012. Complexity, economics, and public policy. *Politics, Philosophy & Economics* 11 (1): 45–75.
- Durlauf, Steven N. 2026. Meritocracy and its discontents. *Public Choice* 206 (1–2): 129–147.
- Durlauf, Steven N., Andros Kourtellis, and Chih Ming Tan. 2022. The Great Gatsby curve. *Annual Review of Economics* 14:571–605.
- Durlauf, Steven N., and Ananth Seshadri. 2003. Is assortative matching efficient? *Economic Theory* 21 (2–3): 475–493.
- Gaus, Gerald F. 1998. Why all welfare states (including laissez-faire ones) are unreasonable. *Social Philosophy and Policy* 15 (2): 1–33.
- Gaus, Gerald F. 2007. Social complexity and evolved moral principles. In *Liberalism, conservatism, and Hayek's idea of spontaneous order*, edited by Louis Hunt and Peter McNamara. Palgrave Macmillan.

- Hansen, Lars Peter, and Thomas J. Sargent. 2008. *Robustness*. Princeton University Press.
- Hayek, Friedrich A. 1960. *The constitution of liberty*. University of Chicago Press.
- Hong, Lu, and Scott E. Page. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences* 101 (46): 16385–16389.
- Huff, Connor, Eric Min, and Robert Schub. 2026. Racial inequality in war. *American Political Science Review* 120 (2): 662–680.
- Markovits, Daniel. 2019. *The meritocracy trap: How America's foundational myth feeds inequality, dismantles the middle class, and devours the elite*. Penguin.
- Miller, David. 1992. Deserving jobs. *Philosophical Quarterly* 42 (167): 161–181.
- Miller, David. 2025. Meritocracy: A qualified defense. *The ANNALS of the American Academy of Political and Social Science* (this volume).
- Mulligan, Thomas. 2018. *Justice and the meritocratic state*. Political Philosophy for the Real World. Routledge.
- Mulligan, Thomas. 2023. Meritocracy. In *Stanford encyclopedia of philosophy*, fall 2023 edition, edited by Edward N. Zalta and Uri Nodelman. Stanford University. Available from <https://plato.stanford.edu/archives/fall2023/entries/meritocracy>.
- Nomi, Takako, Stephen W. Raudenbush, and Jake J. Smith. 2021. Effects of double-dose algebra on college persistence and degree attainment. *Proceedings of the National Academy of Sciences* 118 (27): e2019030118.
- Page, Scott E. 2017. *The diversity bonus: How great teams pay off in the knowledge economy*. Our Compelling Interests. Princeton University Press.
- Parfit, Derek. (1991) 2002. Equality or priority? In *The ideal of equality*, paperback edition, edited by Matthew Clayton and Andrew Williams. Palgrave MacMillan.
- Rawls, John. 1971. *A theory of justice*. Belknap Press.
- Roemer, John E. 1993. A pragmatic theory of responsibility for the egalitarian planner. *Philosophy and Public Affairs* 22 (2): 146–166.
- Roemer, John E. 1998. *Equality of opportunity*. Harvard University Press.

Sandel, Michael J. 2020. *The tyranny of merit: What's become of the common good?* Farrar, Straus and Giroux.

Sen, Amartya. 1979. Utilitarianism and welfarism. *Journal of Philosophy* 76 (9): 463–489.

Sen, Amartya. 2000. Merit and justice. In *Meritocracy and economic inequality*, edited by Kenneth Arrow, Samuel Bowles, and Steven Durlauf. Princeton University Press.

Williams, David Lay. 2024. *The greatest of all plagues: How economic inequality shaped political thought from Plato to Marx*. Princeton University Press.

Young, Michael. 1958. *The rise of the meritocracy, 1870–2033: An essay on education and equality*. Thames & Hudson.

Notes

¹ It is possible for a reward-based assignment rule to generate incentives for those to whom it applies. But that consideration is not a retrospective one per se.

² This example originates with Daniel Markovits.

³ Thomas Mulligan. 2025. "On merit." Unpublished working paper shared with the author.

⁴ I will consider how labor market outcomes matter for assessing meritocracy in educational assignments, but otherwise I do not consider these other contexts.

⁵ My discussion omits any intrinsic benefit to diversity in educational settings. There are good theoretical reasons to believe that diversity can enhance learning (Hong and Page 2004) as well as a range of contexts of empirical evidence of these effects (Page 2017). This omission does not affect the arguments I make here.

⁶ The assumption of a single dimension for skills and qualities is of course unrealistic. Treating either student skills or college quality as multidimensional would not qualitatively affect the claims I make. However, doing so would make the discussion of this environment more cumbersome.

⁷Formally,

$$\frac{\partial f(e,s)}{\partial e} > 0 \frac{\partial f(e,s)}{\partial s} > 0.$$

These conditions mean that the higher entry-level skills lead to higher graduate skills and higher-quality colleges lead to higher graduation skills. Neither is objectionable here, but without these assumptions, better students would prefer worse colleges, etc., so they avoid uninteresting technical complications.

⁸ I ignore the fact that college attendance is a two-sided matching problem, where a college makes a choice to admit students and students then make a choice on whether to accept.

⁹ Thomas Mulligan. 2025. "On merit."

¹⁰ This retrospective argument assumes that the higher-skill student would prefer the higher-quality college. This is unobjectionable in this context. The assumption is less obvious if the subject of analysis was meritocratic assignments of soldiers to tasks of different levels of danger. Huff et al. (2026) discuss how prejudices against Black soldiers in World War I affected combat assignments, as an example.

¹¹ Operationalizing the assignments by grade and test score thresholds would require what appear to be strange rules here (admissions would be driven by upper bounds not lower bounds) because of the simplicity of the environment and the absence of information about the students under consideration. Admissions policies that promote ethnic or socioeconomic diversity are, independent of current legality, not strange and de facto are designed specifically to address admissions outcomes that would be generated by retrospective meritocratic grade and test score rules.

¹² Anderson (2015) gives an analysis of optimal matching in infinite horizon environments and identifies the roles of concavity and convexity as key to efficiency. That said, his results do not fully apply to the contexts I discuss. Schools and firms are finite horizon environments. For example, one matches students to classrooms each year from

kindergarten to the senior year of high school. Finite horizons allow for switches in assignment rules at different points in time while infinite horizon models focus on assignment rules that do not change. Finite horizon environments have not been systematically examined. The treatment here follows ideas in Durlauf (2026), which studies specific cases.

¹³ Formally,

$$\frac{\partial g(y,r)}{\partial y} > 0; \frac{\partial g(y,r)}{\partial r} > 0.$$

As before, these are natural assumptions and avoid technical problems.

¹⁴ Concavity means $\frac{\partial^2 f(\mathbf{e}, \mathbf{s})}{\partial \mathbf{e}^2} < 0$ while convexity means $\frac{\partial^2 f(\mathbf{e}, \mathbf{s})}{\partial \mathbf{e}^2} > 0$.

¹⁵ Bénabou (1996) and Cooper (1998) demonstrate this generally in much richer formal models.

¹⁶ Socially inefficient segregation may separately occur when individuals do not account for the social consequences of their choices. Durlauf (1996b, 1996c) describe models of the dynamic evolution of neighborhoods, in which income segregation emerges from parental neighborhood choices made to maximize their children's income, without considering the social consequences that are produced collectively.

¹⁷ Assignment rules involve inequalities of location in social structure. Durlauf (1996a) discusses the relationship between these types of inequalities and responsibility-sensitive egalitarianism.