Mine the Gap: Using Racial Disparities to Expose and Eradicate Racism

James S. Liebman
*Columbia Law School*, jliebman@law.columbia.edu

Kayla C. Butler
*Cravath, Swaine & Moore*

Ian Buksunski
*Morrison & Foerster*

Follow this and additional works at: https://scholarship.law.columbia.edu/faculty_scholarship

Part of the Civil Rights and Discrimination Commons, Law and Race Commons, and the Race and Ethnicity Commons

**Recommended Citation**


Available at: https://scholarship.law.columbia.edu/faculty_scholarship/2751
MINING THE GAP: USING RACIAL DISPARITIES TO EXPOSE AND ERADICATE RACISM

JAMES S. LIEBMAN*, KAYLA C. BUTLER** & IAN BUكسUNSKI***

ABSTRACT

For decades, lawyers and legal scholars have disagreed over how much resource redistribution to expect from federal courts and Congress in satisfaction of the Fourteenth Amendment’s promise of equal protection. Of particular importance to this debate and to the nation given its kaleidoscopic history of inequality, is the question of racial redistribution of resources. A key dimension of that question is whether to accept the Supreme Court’s limitation of equal protection to public actors’ disparate treatment of members of different races or instead demand constitutional remedies for the racially disparate impact of public action.

For a substantial segment of the nation’s population as well as its judiciary and legal culture, governmentally mandated redistribution, and particularly racial redistribution, of resources to remedy the disparate results of public action is anathema to our constitutional order—so much so that such redistribution may provoke violence that horribly magnifies inequality. Avoiding that prospect leads us to propose a new constitutional understanding of the relationship between disparate impact and treatment to serve as an alternative to racial redistribution—or, should our legal

** Associate, Cravath, Swaine and Moore; B.S. 2013, Ohio University; J.D. 2020, Columbia Law School.
*** Associate, Morrison & Foerster; B.A. 2015, Vassar College; J.D. 2020, Columbia Law School. The authors are grateful for the outstanding research assistance of Tessa Baizer, Cheng Jun Koh, Julia J. Levitan, and Ira Rosenberg.
culture change sufficiently in reaction to current events, as a necessary supplement to redistribution.

While acknowledging the need to mind the racial and other gaps that public action persistently creates and tolerates, our strategy calls upon public actors and oversight bodies to mine the gaps for dispositive evidence of disparate treatment. Compared with how federal courts and our legal culture currently understand disparate treatment, our approach is more honest about the existence and meaning of centuries of unrelenting racial disparities and more insistent on transparency about why disparities keep occurring and whether they are innocent. Yet, the proposal also is moderated by its continuing prioritization of disparate treatment over disparate impact per se; by the extent to which it remains constitutionally and culturally precedent; by its objective of reform but not necessarily outright racial redistribution; and by its effort to avoid rowing upstream against the nation’s individualistic current or being swept by it over treacherous and violent falls.

In offering this approach, we recognize the need constantly to calibrate the breadth of the concession being made to liberty over equality and community, in order to keep the voracious appetite of the nation’s individualism from consuming all hope of equity and social solidarity among diverse populations.

* * * * *

[O]n the big unfinished goals in this country... broad majorities agree on the ends. That’s why folks with power will keep trying to divide you over the means. That’s how nothing changes. You get a system that looks out for the rich and powerful and nobody else. So expand your moral imaginations, build bridges, and grow your allies in the process of bringing about a better world.
Barack Obama
Address at the Show Me Your Walk H.B.C.U. Edition
# TABLE OF CONTENTS

**I. INTRODUCTION: RACIAL REFORM, REDISTRIBUTION, RAGNARÖK**

- 3

**II. DISPARITY-DRIVEN DUTY TO INQUIRE**

- 10
  - A. DUTY TO LEARN
    - 10
  - B. RESPONSIBLE LEARNING
    - 13
  - C. WHAT'S NEW HERE
    - 16
      1. Taking Seriously the Insight that Persistent Disparate Outcomes Indicate Bias that Is Deeply and Ineffably Built In
        - 17
        - 17
      3. Using Disparities Among Traditionally Disadvantaged Individuals as Well as Between Them and Traditionally Privileged Individuals as Evidence of Inequity
        - 21
      4. Bringing Multiple Perspectives to Bear on Whether Suspicious Disparities Are Purposeful or Recklessly Indifferent to Black Lives
        - 22
      5. Promoting Improved Outcomes for All, While Keeping General Benefits from Sufficing as a Response to Persistent Gaps
        - 23

**III. DISPARITY-DRIVEN LEARNING IN ACTION**

- 29
  - A. DISABILITY AND EMPLOYMENT DISCRIMINATION
    - 30
  - B. JUVENILE JUSTICE
    - 33
  - C. PORTFOLIO SCHOOL REFORMS
    - 39
      1. Improved Graduation and College-readiness Rates
      - 43
      2. Increased Proficiency on State Math and English Tests
      - 47
      3. Gap Closure on National Tests
      - 52
      4. Impact on Racial Disparities
      - 54
  - D. SCHOOL DESEGREGATION 1970–1985
    - 59

**IV. OBJECTIONS**

- 73

**V. CONCLUSION AND POSTSCRIPT**

- 83

**VI. APPENDIX**

- 85
For decades, lawyers and legal scholars have disagreed over how much resource redistribution to expect from federal courts and Congress in satisfaction of the Fourteenth Amendment’s promise of equal protection.1 Of particular importance to this debate and to the nation, given its kaleidoscopic history of inequality is the question of racial

---

redistribution of resources through affirmative action,\(^2\) reparations,\(^3\) modern-day abolitionism,\(^4\) or the complex of actions needed so that Black lives matter.\(^5\) A key dimension of that question is whether to accept the Supreme Court’s limitation of equal protection to public actors’ disparate
treatment of members of different races\(^6\) or instead demand constitutional remedies for the racially disparate impact of public action.\(^7\)

At base, this is a debate over how U.S. law and legal thinking should prioritize and align liberty, equality, and social solidarity. Accepting liberty’s primacy might imply that only an individual’s exercises of liberty for the invidious purpose of withdrawing the liberty of another—for example, through personal or property crimes or purposeful racial discrimination—can justify state intervention to withdraw an actor’s liberty by redistributing her or his property in service of equality or social stability.\(^8\) A number of left liberal thinkers have argued to the contrary that a constitutional commitment to the primacy of each individual’s liberty to define and pursue her or his own life plan implies an affirmative communal duty—including through mandated redistribution of resources—to assure all individuals equitable access to the foundational prerequisites of liberty, such as an adequate education and some measure of bodily, economic, medical, and social security.\(^9\) Disciplining public action that generates disparities in access to those prerequisites by Black Americans or others who historically have been

\(^{6}\) See, e.g., McCleskey v. Kemp, 481 U.S. 279, 292 (1987); City of Mobile v. Bolden, 446 U.S. 55, 66 (1980); Pers. Adm’r v. Feeney, 442 U.S. 256, 279 (1979); Washington v. Davis, 426 U.S. 229, 239 (1976); see also Richard A. Primus, Equal Protection and Disparate Impact: Round Three, 117 Harv. L. Rev. 493, 496 (2003) (“Equal protection has become hostile to government action that aims to allocate goods among racial groups, even when intended to redress past discrimination.”); The Supreme Court—Leading Cases, 123 Harv. L. Rev. 282, 283 (2009) (“Equal protection theory has long been troubled by the conflict inherent in requiring unequal treatment in order to avoid or remedy unequal result.”); Thomas Sowell, Discrimination and Disparities 23–29 (2019) (arguing that racial disparities in economic outcomes do not primarily result from discrimination and that conflating disparities with discrimination generates ineffective social policies); infra notes 237–266 and accompanying text (discussing potential equal-treatment objections to a constitutional focus on disparities and equal results).


\(^{8}\) See, e.g., Epstein, supra note 1, at 543; Nozick, supra note 1, at 167–73.

denied them is one way to enforce that duty.\textsuperscript{10}

The proposition that liberty implies some degree of resource equality, however, has not penetrated federal constitutional law.\textsuperscript{11} On the contrary, for a substantial segment of the nation’s population as well as its judiciary and legal culture, overt redistribution of any sort, and particularly racial redistribution, is fundamentally anathema to our constitutional order\textsuperscript{12} —so much so that it may provoke violence that horribly magnifies inequality.\textsuperscript{13}

\textsuperscript{10} See Epstein, supra note 1, at 543; Nozick, supra note 1, at 167–73.

\textsuperscript{11} See, e.g., San Antonio Indep. Sch. Dist. v. Rodriguez, 411 U.S. 1, 13–14 (1973) (holding that poverty is not a suspect classification for federal equal protection purposes and that, as important as public education is to the exercise of liberty, a state is not required to distribute educational funding equally to all students or in a manner that compensates for some students’ economic disadvantage). As Frank Michelman put the point more generally in 1973, “the mainstream of our legal tradition” and our “legal order” generally are “noticeably lacking in norms, principles, and categories of analysis directly applicable to the evaluation of distributional outcomes.” Michelman, supra note 9, at 963; see also Andrias, supra note 1, at 11–23.


That prospect led us to conceive a new constitutional understanding of the relationship between disparate impact and treatment to serve as an alternative to racial redistribution. We acknowledge from the start that compromising in this fashion with the nation’s settler vision of liberty is painful, given the links between that vision and antebellum Black slavery and postbellumpeonage, apartheid, and prisons; and given the Faustian compromises that put slavery in the Constitution, spread it westward, ended Reconstruction, and contrived “separate but equal” (among so many others). As we eventually came to see,

---


15 See, e.g., U.S. CONST. art. I, § 2, cl. 3 (“Representatives and direct Taxes shall be apportioned among the several States . . . according to their respective Numbers, which shall be determined by adding to the whole Number of free Persons . . . three fifths of all other Persons.”); Missouri Enabling Act of Mar. 6, 1820, Ch. 22, § 8, 3 Stat. 545, 548 (admitting Missouri as a slave state and Maine as a free state in what became known as the “Missouri Compromise”); Plessy v. Ferguson, 163 U.S. 537, 544 (1896) (holding that enforced racial segregation, as long as it was “separate, but equal,” does not offend the Constitution), overruled by Brown v. Bd. of Educ., 347 U.S. 483, 494–495 (1954); Michael W. McConnell, The Forgotten Constitutional Moment, 11 CONST. COMMENT. 115, 122–29 (1994) (describing the Hayes–Tilden Compromise of 1877—

In hopes of fending off the cataclysm these competing impulses portend, this Article suggests a new middle position. That position is different from how federal courts and our culture currently understand disparate treatment. It is more honest about the fact and meaning of centuries of unrelenting racial disparities, and more insistent on transparency about why disparities keep occurring and whether they are innocent. Our proposal, however, still prioritizes disparate treatment over disparate impact per se; still hopes to stay somewhat constitutionally and culturally preceded; and aims for reform not, necessarily, outright racial redistribution; and is still chary of rowing upstream against the nation’s individualistic current or being swept by it over treacherous and violent falls. While acknowledging the need to mind the racial resource and outcome gaps that public action persistently creates and tolerates, we offer a new strategy for mining those gaps for dispositive evidence of disparate treatment. In offering this approach, we recognize the need constantly to calibrate the breadth of the concessions being made to liberty, in order to keep the voracious appetite of the nation’s individualism from consuming all hope of equity and social solidarity among diverse
populations. Part II describes our proposal for disparity-driven learning and five ways it more practically and powerfully roots out racism. Part III provides examples of disparity-driven learning in action. Part IV responds to objections.

II. DISPARITY-DRIVEN DUTY TO INQUIRE

A. DUTY TO LEARN

Our starting point is a deep suspicion that disparate treatment is at work whenever an identifiable group of individuals and communities almost always finds itself negatively affected by disparate impacts across the full spectrum of social activity in which public institutions play an important part. As in the case of Black Americans, on whom we focus, the persistence of this pattern across time, space, and subject matter is constitutionally suspicious because it raises two strong inferences: (1) that public actors willingly accept the harmful disparities affecting the target category of people or of the social conditions that persistently generate those disparities; and (2) that these chronically disadvantageous disparities or their causal conditions are intended or at least reveal such extreme indifference to the lives of the affected people that they may as


[18] Much of what we say here applies as well to public action brought under suspicion by its chronically disparate impact on Native American, Latinx, economically disadvantaged, disabled, immigrant, and other historically oppressed groups. We focus on Black individuals and communities as the paradigm U.S. case.
Not every disparity that fits this pattern constitutes actionable disparate treatment, however, or requires judicial scrutiny that is strict in theory and usually fatal in fact. But as long as the disparities persist, the suspicions just noted justify a practical inquiry by the responsible public actors that either proves that neither of the two inferences applies or, instead, reveals to the world—and to the actors themselves if they previously were unconscious of it—that the disparities are purposeful or recklessly indifferent to the welfare of Black Americans. To be clear, the goal is not a world in which all individuals—or that members of privileged and suspect groups on average—have exactly the same amount of every condition and resource on which happiness can depend. The goal, instead, is a world in which there is no reason to expect that the population-wide distribution of those conditions and resources will systematically skew in a direction that correlates with attributes historically associated with bias, oppression, and deprivation.

In other words, the Equal Protection Clause’s long-accepted proscription against public actors’ disparate treatment of members of suspect groups, and the suspicion that such treatment accounts in significant part for historically enduring and socially pervasive patterns of disparate outcomes, should oblige public actors to be on the lookout for outcomes they generate that fit this pattern. When such results appear, those actors should be obliged to take responsible steps to convince themselves and the public that an inference of disparate treatment is unwarranted. We propose that they do this by experimenting with strategies for diminishing the disparities at no great expense beyond that incurred by the steps that generate the disparities. If experiments

19 Cf. MODEL PENAL CODE § 210.2 (defining the most serious criminal offense—murder—by treating as equivalent the taking of life that is purposeful, that is knowing even if not intended, and that is grossly reckless in that the actor proceeded with an awareness of and “extreme indifference” to the high likelihood that death would result).


22 See Primus, supra note 6, at 532–36 (noting situations in which prominent racial and other disparities can signal the presence of decisional racial bias).

23 Among other questions this Article raises without answering is the relationship between the public cost of disparity-reduction steps and the steps’ status as “responsible” or not. Our initial thoughts on the matter are that:
succeed in diminishing harmful disparities affecting the target group, the resulting lower disparity levels establish a new, lower constitutional ceiling on disparities below which public action thereafter must stay in similar circumstances. These obligations continue as long as the public action in question, to any considerable degree, extends or leaves in place disparities that fit the pattern. Importantly, the object is not, as the Supreme Court suggested might be unconstitutional in *Ricci v. DeStefano*, for public actors to discontinue the disparity through intentionally race-based counteractive action regardless of the disparity’s causes or contributions to other values, but instead to explore—to mine—the disparity for evidence one way or the other on those questions.

- The duty to act responsibly does not end when new costs arise. The duty is, however, sensitive to the extent of and likely disparity-reduction return on investment in terms of time, money, and administrative capacity expended relative to the overall size and budget of the activity generating the disparities and the extent and number of people negatively affected by the disparities.
- It may be responsible to explore less rather than more expensive fixes first, but the costs of the initial efforts should be discounted by reasonable estimates of the probability that broader measures will later be needed.
- The cost of monitoring gaps and experimentation must be incurred. It can be reduced in a number of ways including by leveraging existing administrative data collection and infrastructure, building the steps into new or existing continuous improvement routines, attending to disparity reduction possibilities in devising strategies in the first place, and participating in networks of similar actors or agencies engaged in a parallel search for productive ways to reduce disparities. See, e.g., Charles F. Sabel & William H. Simon, *Contextualizing Regimes: Institutionalization as a Response to the Limits of Interpretation and Policy Engineering*, 110 MICH. L. REV. 1265, 1287–91 (2012).
- Given the potential of experimentation to generate systemic efficiencies that offset short-term costs, long-term costs should count more than short-term ones, and estimates of long-term costs should be offset by conscientiously estimated long-term gains.
- Apart from other gains, the probable extent and presumptive importance of equity gains count heavily in the calculus.
- These considerations may define “responsible action” equivalently to the balancing implied by middle-level equal protection scrutiny. See, e.g., ERWIN CHEMERINSKY, *CONSTITUTIONAL LAW: PRINCIPLES AND POLICIES* 671–74 (6th ed. 2019) (“Under intermediate scrutiny, a law is upheld if it is substantially related to an important government purpose. . . . The means used need not be necessary, but must have a ‘substantial relationship’ to the end being sought.”).

24 In other words, there is no going back, only forward, in the race to end chronic disparities.

25 *Ricci v. DeStefano*, 557 U.S. 557, 593 (2009) (invalidating under Title VII while precluding the constitutionality of official race-based action taken solely to keep a racial disparity from arising (discussed infra notes 270–272 and accompanying text)). Disparities begin, but do not end, the analysis, which then aims to understand why disparities arise. See, e.g., *infra* notes 51–67 and accompanying text (noting the attention due public action affecting multiple similar sites where disparities only sometimes arise or are of different magnitude, exposing hidden influences unrelated to public need).
B. RESPONSIBLE LEARNING

What, though, if remedial steps benefiting members of the previously harmed group also benefit members of more privileged groups so that some part of the harmful gap remains? Can such steps nonetheless qualify as responsible efforts to reduce gaps? For reasons we explain in more detail below, changes that benefit others besides Black Americans are expected and welcome as long as those steps do not compound the suspicion of disparate treatment by disproportionately benefiting members of the previously privileged group. In particular, steps that turn out to benefit others in addition to the target group may well qualify as responsible gap-reducing experiments if their disparate effects (if any) favor the target group, thus diminishing the preexisting gap. Again, however, remaining gaps continue the duty to experiment.

Systematic and predictable outcome disparities for which racial bias is likely to be at least partly to blame remain the starting point for efforts to achieve equal protection. And eradication of those disparities is those efforts’ long-term aim. But the immediate goal is a progression of forward-looking steps that demonstrably reduce suspicion of invidious motivation by disparately benefiting the people previously on the wrong end of disparities. Given the nation’s history of public action with disparate effects unfavorable to Black Americans, taking actions that disparate, even if they do not solely, favor them, is a sufficiently strong proxy for progress that we are prepared to treat them as presumptively constitutional.

We ground this conclusion, as well, in the practical welfare both of the longstanding victims of bias and deprivation and of the population at

26 See infra notes 57–73 and accompanying text.
27 Experiments that end up benefiting others equally or more than the target group are not for that reason alone ineffective. Indeed, innovations with positive general welfare effects as well as target equity effects are an intended byproduct of this approach. Such actions, however, will not exhaust and may add to the actor’s duty to experiment responsibly with steps to diminish chronic disparities. The continuing duty, along with traditional subjective and objective methods of proving racial animus, moreover, should discourage strategic behavior, such as inefficiently magnifying racial disparities at the outset to facilitate the search for ameliorative steps later, or knowingly pursuing a less effective option in hopes of diverting attention from more effective ameliorative steps.
29 The validation is only presumptive, however, because conventional backward-looking constraints on disparate treatment would continue to apply.
large. If over the long haul, we want to improve the lot of people chronically harmed by bias, we must at times be willing in the short and medium term to replace gap eradication with progressive outcome gains plus disparities in welfare gains favoring traditionally underserved populations. Treating as constitutionally sufficient steps that over the middle term in fact benefit members of chronically disadvantaged groups and evidence a good-faith effort to diminish harmful disparities against those members acknowledges that those individuals' lot depends on two considerations: the quantity and quality of the resources and conditions associated with happiness to which they have access, as well as how closely those resources and conditions approximate those of those of members of more privileged groups. Our constitutional baseline acknowledges, as well, that actionable bias can be past as well as present, unconscious as well as conscious, structural as well as episodic, and objectively as well as subjectively manifested.

As long as disparities persist, officials remain obliged to explore further gap-closing measures—an obligation that is heightened when experiments enlarge rather than close gaps. Over time, therefore, gaps can be expected to close appreciably. But not every experiment that fails to close gaps is, for that reason alone, irresponsible. If failed experiments were always insufficient, actors would have little incentive to experiment. Nor would it be irresponsible for public actors to continue or extend an experimental treatment after discovering beneficial effects other than gap closure—as long as they keep looking for ways to close remaining gaps.

An important indicator that steps to diminish disparities are “responsible” is public transparency—the extent to which the actors in question acknowledge the disparities their actions generate or preserve and document the rationale and results of experiments they pursue to lessen those disparities. In turn, wide publication of such results justifies treating officials’ failure to consider adopting or adapting proven disparity-reducing steps undertaken elsewhere under similar circumstances as likely irresponsible. Beyond that, responsible experimentation requires that the actors conscientiously work to identify and test disparity-reducing steps through flexible cycles of identifying gaps, analyzing possible causes, hypothesizing and experimenting with solutions, and adjusting policies and operations based on results.  

---

30 “Publicly attending” to evidently effective steps others take to reduce disparities does not require other public actors blindly to replicate the steps, or to accept the validity of the results or their applicability in different circumstances. Responsible consideration of the potential bearing of such steps on the disparities at issue is required, however.

31 For other explications of this standard, see Olatunde C.A. Johnson, Disparity Rules, 107 COLUM. L. REV. 372, 407–16 (2007); James S. Liebman, Perpetual Evolution: A Schools-Focused A
In some cases, irresponsibility will be clear, as when officials do not transparently monitor disparities or seek to mitigate ones they identify or fail to spread the fruits of successful experiments. In other cases, assuring that steps taken were responsible will be the work of multiple evaluators operating from different perspectives. Not least among these perspectives is officials’ own sense of responsibility. Providing the transparency that is the easiest aspect of our proposal to assess will force officials to confront the racial disparities they cause and to explain and defend the ameliorative steps they undertake. If their efforts succeed, belying previously held assumptions about why disparities arise, their own cognitive and moral insights will expose the irresponsibility of tolerating disparities going forward, without outside supervision.

Courts, regulators, and public and private funders can continue their longstanding backward-looking search for subjective, causal racial animus, supplemented by the “conscientious inquiry” analysis described above. To add yet another perspective—that of the public at large—we ultimately would apply a test the Supreme Court uses to assess religious discrimination under the Establishment Clause. The question that test asks is whether the extent of the disparities together with the quality of the officials’ past and responsive actions and inaction conveys a public message that chronic disparities of this sort are the acceptable norm. If so, the actions are irresponsible, whether or not the public officials in question believe their tolerance of existing disparities is responsible, and whether or not direct or inferential evidence of racial animus is present. If instead, the message conveyed is that chronic disparities are presumptively unacceptable and require conscientious steps to eliminate them, then the public actors are absolved from liability for now, though


Although the contemplated causal analysis may strive to be deductive, “abductive” analysis will likely be the norm. See, e.g., DAN SIMON, IN DOUBT: THE PSYCHOLOGY OF THE CRIMINAL JUSTICE PROCESS 22 (2012) (“Abductive reasoning[—]a recursive process of generating and testing hypotheses, geared toward eliminating invalid hypotheses and substantiating the correct one[—]involves] a search for information, followed by its evaluation [and] drawing of correct inferences from that information. While the evaluation of the information entails logical inference, the generation of hypotheses . . . require[s] intuitive and conjectural thinking.”).

32 See, e.g., Santa Fe Indep. Sch. Dist. v. Doe, 530 U.S. 290, 310 (2000) (“[State] sponsorship of a religious message is impermissible [if] it sends the ancillary message to members of the audience who are nonadherents ‘that they are outsiders, not full members of the political community, and an accompanying message to adherents that they are insiders, favored members of the political community.’” (quoting Lynch v. Donnelly, 465 U.S. 668, 688 (1984) (O’Connor, J., concurring)).
their obligation to continue searching for ways to reduce disparities remains.

C. WHAT'S NEW HERE

We realize that our definition of constitutionally suspicious disparities differs more in principle than in effect from conventional definitions of suspect classifications. Nor is there anything novel about reformist proposals to factor in the hidden influence of historical racism, search for objective evidence of invidious motivation in actors' responses to disparities they may or do create, and acknowledge the frequency with which such motivations operate unconsciously. We also realize that these backward-looking reforms have not kept our culture and legal system from vastly underestimating and over-tolerating disparities that, however suspicious, are hard to disentangle from "neutral" explanations and to redress without arousing resentment and risking violent resistance to suspected racial redistribution. Our proposal, however, has five features that distinguish it from other reform proposals and increase

33 See supra notes 17–25 and accompanying text.
our confidence in its capacity to achieve equal protection.

1. Taking Seriously the Insight that Persistent Disparate Outcomes Indicate Bias that Is Deeply and Ineffably Built In

Advocates of heightened efforts to achieve racial equity rightly point out that racism is not an idiosyncratic and episodic deviation from a neutral or equitable norm. It is the structured-in norm. Making this point to support radically less tolerance for disparities, however, can obscure an important implication of the insight—that the more deeply racist structures penetrate our psyches and social systems, the harder they are to prove and purge. Purging discrimination thus will take time and grindingly hard work. It likely will require a mix of external incentives and penalties, and positively experienced internal changes in (1) mindsets and beliefs—what we are attuned to seeing and inferring, and how effectively we recognize and correct stereotyped judgments—and (2) how we organize social and productive activity.

Consistently with this insight, our proposal advances an understanding of equal protection as a long-haul project of habituating public actors to notice disparities and treat them as favorable opportunities to inquire into the beliefs, evidentiary and inferential judgments and misjudgments, and social arrangements and systems of all kinds, small and large, that generate disparities. In some cases, the obligation to identify and ameliorate disparities will quickly reveal effective alternatives. Other disparities will require patient inquiry across multiple experiments, with success marked by gradual decreases in disparities’ frequency and magnitude.

2. Making the Mandated Inquiry Practical, Forward-looking, and Self-executing

The problem with a doctrine dependent on proof of disparate treatment is that the inquiry it imposes upon litigants and courts is

37 See, e.g., sources cited supra notes 4, 28, 34; infra note 38.
41 See, e.g., infra notes 106–117 and accompanying text.
etiological, looking backwards in time in search of two things that are notoriously hard to find:

- A public decisionmaker’s desire to treat the races differently, proved either directly through the actor’s admission or indirectly through “objective” evidence that racial animus is the only reasonable explanation for what the actor claims to have done for race-neutral reasons;42
- A causal chain linking disparate outcomes to that official’s racial motivation rather than to an infinite array of other possibly causal race-neutral motivations, intervening personal and sociological forces, and discriminatory acts by private parties.43

Our proposal seeks instead to explore possible links between treatments and outcomes in the same practical manner as researchers, businesses, and regulatory and social-service agencies do in their everyday work: through forward-looking experimentation in search of ways to maximize desirable and minimize untoward outcomes. Rather than looking backwards to expose and punish illicit behavior connected to harmful results, our proposal calls for the forward-looking exploration of new ways of conducting business to achieve better results. The requisite inquiry thus is not a court’s or regulatory agency’s retrospective, judgment-laden examination of the etiology of public actions affecting different groups differently that occasionally and haphazardly come to their attention. The proposal instead entails each public actor’s ongoing, self-reflective steps to acknowledge and explore ways to diminish troubling disparities that arise in the regular course of daily routines.

This approach is more practical in multiple ways. First, it is more likely to be informative because it places the focal conditions—in this case, disparities and ways of diminishing them—under self-conscious scrutiny as they occur. Regularity and simultaneity make the approach more effective than intermittently trying to ferret out easily disguisable and morally abhorrent mental states and psychologically and

42 See supra notes 6 and 35 and accompanying text.
43 See, e.g., Parents Involved v. Seattle Sch. Dist., 551 U.S. 701, 750 (2007) (Thomas, J., concurring) (“Although presently observed racial imbalance might result from past de jure segregation, racial imbalance can also result from any number of innocent private decisions, including voluntary housing choices.”); McCleskey v. Kemp, 481 U.S. 279, 298–99 (1987) (affirming Black defendant’s capital sentence despite sophisticated regression analysis with hundreds of control variables showing killings of white victims are substantially more likely to generate death sentences than otherwise-identical killings of Black victims because McCleskey could not identify a particular actor in the criminal justice process who intentionally discriminated against him).
sociologically complex causal factors after the fact. Second, it puts the responsible public actors themselves on guard for racial disparities and habituates these actors to considering ways to lessen those results. As such, it invites those same public actors to be on the lookout for new and better ways to accomplish the legitimate, socially beneficial aims they do or claim to want to achieve. Third, unlike the traditional approach to disparate treatment, which creates a huge risk of angering everyone—public officials not wanting to be accused of racism and victims who rarely get redress—our approach has the capacity to change minds in positive ways.

The last-mentioned practicality is the most important. By exposing inaccurate stereotypes and other forms of unconscious racism, successful experiments to diminish disparities can help public actors better understand (1) the relevant background conditions and how they previously misapprehended them, (2) weaknesses in their decisionmaking processes, (3) why their past actions generated disparities, (4) how they can do their jobs more effectively when misapprehensions no longer divert them, and (5) how they can widen the reach of beneficial outcomes—all with positive spillover effects on their later official and personal actions. Even if responsible experimentation does not result in fewer disparities, or only does so at significant cost to legitimate goals, public actors still are better informed about the new experiments they remain obliged to run. And in the meantime, actors have a response to claims of racism associated with the disparity in question. Additionally, persistent racists can be identified more easily by their failure to track disparities, responsibly experiment with ways to diminish them, and scale experiments that work. Under our proposal, therefore, a lack of transparency as to the extent of the disparities, how officials monitor them, steps they do and do not take to mitigate them,

44 See, e.g., Pedro A. Noguera, Ties That Bind, Forces That Divide: Berkeley High School and the Challenge of Integration, 29 U.S.F. L. Rev. 719, 731 (1995) ("Fear of being charged with racism, whether deliberate or unconscious, is such a terrifying prospect that most whites go out of their way to avoid dealing directly with race.").

45 See, e.g., Stephen Rynkiewicz, Workplace Plaintiffs Face Long Odds at Trial, Analytics Data Indicates, A.B.A. J. 17, 17 (July 17, 2017), https://www.abajournal.com/news/article/workplace trial analytics lex machina ("Only 1 percent of plaintiffs who file federal job discrimination, harassment and retaliation claims win on the merits at trial, according to an analysis by the Lex Machina legal analytics firm."); infra note 283.

46 See infra notes 106–117, 130–141 and accompanying text (providing examples of experimentation exploding long-held myths that juveniles not appearing for court dates were guilty or absconding and that no more than 50 percent of New York City’s mainly Black and Latinx students could be expected to finish high school).
and the rationale for and rigor and results of responsive steps taken will expose racism—not obscure it as often happens now.47

Illustrating these points, the table below compares the Supreme Court's treatment of commonly offered objective and historical evidence that disparities are racially discriminatory with the treatment such evidence would receive under our proposal.

<table>
<thead>
<tr>
<th>The evidence</th>
<th>Response to the evidence by . . .</th>
<th>our proposal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Officials rejected reasonable alternatives that might have reduced disparities.</td>
<td>Officials' belief at the time or now that alternatives would have cost too much, had other harmful effects, or failed to diminish disparities deserves deference.48</td>
<td>Officials' failure to try alternatives violates the duty to take responsible steps.</td>
</tr>
<tr>
<td>Officials failed to look for, discover, or consider reasonable alternatives that might have reduced disparities.</td>
<td>Officials had no reason to conduct such a search given the reasonableness of the action they took or its consistency with common practice. Any alternatives they might have considered would have had the same problems as noted in the cell above.49</td>
<td>Officials must try out reasonable alternatives, the results and costs of which speak for themselves. If common practice produces disparities, that is a reason to explore alternatives.</td>
</tr>
<tr>
<td>The disparity-tolerating action is one of a series stretching back years, including explicitly racist ones.</td>
<td>Absent a clear link between prior actions and current results, prior actions do not count, or the inference of current racism is weak.50 (Implicitly: the longer the suspicious history, the greater the social engineering backwards-looking analysis would require reluctant courts to undertake.)</td>
<td>The longer the history of disparities and thus the more chronic they are, the clearer is the duty to explore less disparate possibilities going forward.</td>
</tr>
</tbody>
</table>

---

47 See, e.g., McCleskey, 481 U.S. at 293–97 (discussed supra note 43).
49 See, e.g., id. at 267–70.
3. Using Disparities Among Traditionally Disadvantaged Individuals as Well as Between Them and Traditionally Privileged Individuals as Evidence of Inequity

Public actors and institutions often operate across many sites, affecting multiple individuals and communities. Many, as well, engage in actions paralleling those taken by other officials and institutions under similar circumstances. Think, for example, of a state department of education interacting with multiple school districts, with each district encompassing multiple schools, each school operating multiple classrooms, and each classroom teacher interacting with multiple students over multiple years. In these situations, variation in outcomes across multiple similar sites provides naturally occurring opportunities to compare treatments based on their outcomes. Our proposal encourages public actors to use these naturally occurring conditions as experiments with steps that may aggravate or mitigate racial disparities.51 The opportunity to do so adds to our proposal’s practicality, given its reliance on actions already occurring in the regular course of public activity.

Whether naturally occurring or designed in response to the obligations we propose, this kind of experimentation often will reveal outcome disparities not only between Black and White individuals within and across sites, but also between groups of Black individuals across sites.52 For example, it is clear that similarly situated Black and Latinx children systematically learn more and less at different schools in the same district and state, and in different districts in the state.53 Under current equal protection doctrine, these differences either are irrelevant to whether

51 See supra note 31 (emphasizing the “abductive,” not “deductive,” nature of most such experiments, given the need to account for more than a single potentially disparity-altering difference between sites or time periods, thus making it difficult to be certain—even as much is learned—about causation).

52 See, e.g., 1, 153, 154, 155, 156, 157, 158.

53 See, e.g., 147, 150, 153, 154, 155, 156, 157, 158.

See, e.g., CTR. FOR THE FUTURE OF Ariz., WHY SOME SCHOOLS WITH LATINO CHILDREN BEAT THE ODDS . . . AND OTHERS DON’T 14, 18–21 (Mar. 2006), https://morrisoninstitute.asu.edu/sites/default/files/beattheodds-someschoolswithlatinochildrenwell-somedont.pdf [https://perma.cc/UKA2-EVJ3]; Average Share of Students in “Beating the Odds” Schools for 3 Most Recent Years of Data in Math, CTR. ON REINVENTING PUB. EDUC. [CRPE], https://www.crpe.org/examples/four [https://perma.cc/9YFQ-M4DP] (documenting outcome disparities between “beat the odds” and other public schools with large proportions of students of color and in poverty in fifty school districts nationwide; also noting disparities among those fifty school districts in the percent of students in superior schools, ranging from almost none in Raleigh, North Carolina to nearly 40 percent in Newark, New Jersey); Sean F. Reardon & Demetra Kalogrides, The Geography of Racial/Ethnic Score Gaps, 124 AM. J. SOC’Y 1164, 1165–66, 1181–87, 1194–96 (2019); infra notes 175–240 and accompanying text (discussing schools’ divergent outcomes with similar children depending on their adoption of portfolio reforms and of racial integration).
discrimination accounts for Black/Latinx-White disparities or may imply that because some Black or Latinx students have thrived on the defendants' watch, the others or their families must themselves be at fault. Again, uncertainty—in this case, about why Black-Black disparities occur—counts against those alleging racial bias.

Under our proposal, the onus of the ambiguity again reverses. Assume, for example, that Black children learn less on average than White children in the district, but that the Black children at School A learn considerably more than those at School B. Under our proposal, district officials are obliged to extend the positive results of their naturally occurring experiment at School A to the Black children at School B, or take other steps to diminish Black-White disparities at School B to at least the lesser level present at School A. If the Black children at both schools are situated roughly the same, and if school officials fail to take action, their bias or indifference toward the Black students at School B is manifest. If instead school officials discover that the Black students at the two schools are not similarly situated because of a condition at School A that is not present at School B, the officials have now isolated a condition they know is worthy of further inquiry. They remain obliged to pursue this inquiry at both schools as long as Black-White disparities persist. A further distinguishing feature of our proposal, therefore, is its treatment of disparities among members of underserved populations, as well as between them and more privileged populations, as actionable occasions and laboratories for improving equity.

4. Bringing Multiple Perspectives to Bear on Whether Suspicious Disparities are Purposeful or Recklessly Indifferent to Black Lives

Our approach mobilizes the learning capacities of multiple actors applying multiple standards to provide layers of internal and external accountability for conscientiously examining the defensibility or indefensibility of ongoing disparities. These include (1) the public officials themselves, based on their understanding of their job and constitutional responsibilities, informed by realities their experiments reveal; (2) the wider public’s perceptions via the “message conveyance” measure of discrimination; and regulators applying both (3) traditional disparate treatment and (4) our new “conscientious inquiry” standards to determine whether “responsible steps” overcome suspicions generated by

---

54 See, e.g., McCleskey v. Kemp, 481 U.S. 279, 297–98 (1987) (limiting objective proof of equal protection violations to actions taken that, given their sharply monochromatic impact on—and only on—and identifiable group, manifest animus against members of that group).
chronic disparities.55

5. Promoting Improved Outcomes for All, while Keeping General Benefits from Sufficing as a Response to Persistent Gaps

Current constitutional doctrine invites discrimination defendants to obscure racial motivation in a welter of race-neutral considerations that also may have influenced policy.56 Our proposal again makes a virtue of what otherwise impedes the search for the truth by encouraging public actors to experiment with disparity-reduction techniques that serve multiple public purposes. The difference is that, if pursuing other purposes does not close or instead widens gaps, it will not discharge and may reinforce officials’ obligation to continue exploring new gap-closing possibilities.

This feature might seem to undermine our proposal’s ability to improve the lives of people who are chronically underserved. Why not instead insist upon actions with the single purpose of gap closure or at least no other purpose that might limit gap closure by simultaneously aiming to improve the lives of others? In fact, however, this feature is central to our proposal’s equity and welfare effects on the chronically underserved.

For starters, we have no objection in principle to “rising tides that raise all boats,”57 and we certainly encourage public action generating welfare gains for the chronically underserved. What we do find objectionable, however, is defending such steps on equity grounds when in fact many such efforts have the opposite, gap-increasing effect of raising the boats of the privileged vastly more than those of the

55 See supra notes 17–36 and accompanying text.
chronically underserved. That is why we limit our invitation to all-boats-raised strategies to ones that benefit chronically underserved more than privileged populations. With that caveat, we encourage such strategies, because they are more likely to be effective over time than the single-purpose gap-closing strategies that equity proponents often prioritize.

We make this strong claim because, as we document below, some of the most beneficial recent reforms from the perspective of chronically underserved communities have had this quality. It makes sense that they would. If as gap-closure advocates have shown, racially disparate outcomes often are motivated by unexamined and inaccurate assumptions, exposing the inaccuracies should lead individuals and systems long held back by their own biases to improve overall efficiency, benefiting themselves and other client populations along with those they previously underserved. This indeed is true of the “rising tide” reforms discussed below.

Going further, if as equity advocates also have shown, racial


59 On civil rights advocates’ preference for remedies “targeted” on improving the lot of particular underserved communities, see Samuel R. Bagenstos, Universalism and Civil Rights (with Notes on Voting Rights after Shelby), 123 YALE L.J. 2838, 2852–55, 2859–62, 2865 (2014) (summarizing civil rights proponents’ concerns that pursuing “universalist” goals in service of anti-discrimination aims may weaken the practical or moral force of challenges to or be a screen for preserving discriminatory structures).

60 See infra notes 74–253 and accompanying text.

61 See, e.g., Daniel Kiel, No Caste Here? Toward a Structural Critique of American Education, 119 PENN. ST. L. REV. 611, 624 (2015) (describing racial bias inherent in measures used to group children in tracks in schools, “[s]tructurally maintain[ing], even widen[ing], disparities in educational opportunities”); William M. Wiecek & Judy L. Hamilton, Beyond the Civil Rights Act of 1964: Confronting Structural Racism in the Workplace, 74 LA. L. REV. 1095, 1122 (2014) (identifying faulty assumptions about “dependability, presentability, communication skills, and work ethic” resulting in “significant racial disparities” in workplaces); Erik J. Girvan, On Using the Psychological Science of Implicit Bias to Advance Anti-Discrimination Law, 26 GEO. MASON U. CIV. RIGHTS L.J. 1, 35 (2015) (study revealing implicit bias at work in “teachers’ expectations for the performance of ethnic minority compared to non-minority children” in racial performance gaps on standardized tests and in other racial disparities).

62 See e.g., VIVIAN HUNT ET AL., MCKINSEY & CO., DELIVERING THROUGH DIVERSITY (Jan. 18, 2017), https://www.mckinsey.com/business-functions/organization/our-insights/delivering-through-diversity (finding that companies in the United States and five other countries studied that have the most racially and culturally diverse executive teams “are 33 percent more likely to outperform their [less diverse] peers on profitability”).

63 See infra notes 74–253 and accompanying text.
subordination is deeply built into public structures,\(^6^4\) rooting it out will require deep modification of systems and structures. If undertaken wisely, modifications of this scope are likely to benefit everyone associated with the endeavor—not just those previously discriminated against—for the reasons given above and because people motivated to make changes of this magnitude are likely motivated (legally, politically, and ethically) to pursue broader benefits as well. This “renaissance effect” was central to the “systemwide” school desegregation plans of the 1970s and 1980s that most substantially closed gaps between Black and White students.\(^6^5\) Examples abound of organizations using “disruptions” caused by untoward effects of their actions as an opportunity to rethink entire systems, discarding obsolete path dependencies and other time-honored inefficiencies.\(^6^6\)

---


\(^{6^5}\) See infra notes 236–242 and accompanying text.

\(^{6^6}\) See, e.g., Christopher K. Ansell, Pragmatist Democracy: Evolutionary Learning as Public Philosophy 67–9, 14, 47–54 (2011) (describing disruptive strategies for transforming how police departments interact with communities and how the World Health Organization handles the occasionally adverse consequences of generally beneficial public health responses to disease); Tristin K. Green, Targeting Workplace Context: Title VII as a Tool for Institutional Reform, 72 FORDHAM L. REV. 658, 664–68 (2003) (describing NASA’s use of the Challenger explosion to identify multiple flaws baked into its rocket-building process, improving system’s overall productivity); Arvind Singhal & Karen Greiner, Using the Positive Deviance Approach to Reduce MRSA at the Veterans Administration Healthcare System in Pittsburgh, in Leading Change in Healthcare: Transforming Organizations Using Complexity, Positive Psychology, and Relationship-Centered-Care 177, 181–92 (2011) (describing efforts to reduce MRSA infections at Western Pennsylvania veterans hospitals by identifying and spreading usually productive habits vis-à-vis such infections displayed by employees at all levels of the health care process, fundamentally altering the hospitals’ cultures and capacity to identify and address other health care risks); Steven J. Spear, The High Velocity Edge: How Market Leaders Leverage Operational Excellence to Beat the Competition 87–107 (2010) (describing how Alcoa’s mandated inquiry into reasons and solutions for injuries from production accidents revealed cures for many small defects in production processes, improving company’s safety, productivity, profitability, and capacity to address other problems such as CO2 emissions).

Our proposal’s aspiration to use efforts to reduce chronic racial disparities as an occasion to consider how to improve overall welfare has affinities with what Professor Bagenstos calls “universalism” in civil rights enforcement. See Bagenstos, supra note 59, at 2842–47. The
Of course, gap-closing reforms designed to keep privileged constituencies from benefiting might violate prevailing equal protection norms and invite the zero-sum competition between social groups we hope to avoid. But even simply leaving privileged communities out of the calculus is problematic, given the likely ill-effects on the chronically underserved. Just as exposing the false assumptions that drive bias and innovatively reorganizing systems structured to subordinate groups are promising ways to improve systems’ overall functioning, designing systems to help some, but not all, is a recipe for reducing efficiency and innovation while building in new constraints to success. Such strategies risk diminishing the welfare of everyone the operation touches, with the worst effects on members of chronically underserved groups, if history is any guide. For like reasons, we would not prioritize reforms designed to achieve equity, but not welfare gains for Black individuals; or only welfare gains arising from more equitably distributed “positional” goods, as when White graduates’ educational attainments are leveled down to position Black graduates more favorably in the competition for jobs. The goal of increased equity for long-subordinated populations should encompass welfare gains as well, which we fear may not be the result of gap-closing-

difference is that Bagenstos advocates a search for across-the-board legislative protections for everyone that in the process aid victims of racial discrimination, see id. at 2875, while we advocate a search for practical, context-specific responses to chronic racial disparities that in the process benefit everyone affected by the actors in question. A fixation on gap-closing also might be criticized for privileging the values of advantaged populations as to what counts as merit or success, reinforcing racial biases. This criticism applies to our proposal, and to any that gives propulsive force to outcome disparities on traditional welfare metrics. If the implication of this claim is that Black Americans should content themselves with what they have that White Americans do not have and not worry about the reverse, it is troubling on its face. The more compelling implication is that the key disparity is not how much White Americans have that Black Americans do not have, but how little opportunity Black Americans have to define what they want and to work to achieve it, given how broadly White definitions dominate. Attacking that disparity will take even more effort than diminishing those under discussion here. Still, the effort required by our proposal affords White officials everyday experiential access to the needs and wants of Black Americans, which could make way for the deeper change.

A more subtle problem is that focusing on disparities limits analysis to a small subset of the considerations that determine individuals’ access to success (as when we focus on students’ academic achievement, ignoring practical problem-solving ability, planning, administrative functioning, grit, and the like). Our proposal is better suited than most to respond to this problem, however, given its emphasis on context-specific inquiry into why disparities exist, which may reveal other, previously hidden disadvantages—and advantages—that traditional measures obscure.

See William S. Koski & Rob Reich, When “Adequate” Isn’t: The Retreat from Equity in Education Law and Policy, and Why it Matters, 56 EMORY L.J. 545, 549 (noting the “‘positional good’ aspects” of education: “one person’s possession of more education necessarily decreases the value of another’s education,” raising concerns that improving all students’ achievement without diminishing existing inequality will “mak[e] some people educationally worse off”).
only remedies. Although we cannot prove it, we wonder whether the non-systematicity in this sense of single-purpose efforts to root out the effects of systemic racism accounts for those efforts’ limited history of success.

These reflections may seem obvious, but they are not to some aspiring reformers. Over the past decade, this Article’s Senior Author has asked business, education, law, and policy graduate students in an interdisciplinary class focused on public education reform to identify their preferences among four hypothetical reforms with differing distributions of positive and negative results across populations (see Figure 1 below). By hypothesis, each reform is adopted when K–12 academic outcomes for Black students are substantially lower than for White students. Reform A equalizes all academic outcomes at the initial level of White students. Reform B equalizes all outcomes at the initial level of Black students. Reform C raises the outcomes of all students by an equal amount, moving Black educational attainment above the initial level of White attainment but leaving the achievement gap the same in absolute (though not proportional) terms. Reform D achieves the same increases as Reform C for Black students, while raising the outcomes of White students by a lesser absolute amount, diminishing the achievement gap in absolute and proportional terms. In Reforms C and D, therefore, welfare measured (imperfectly) by academic achievement is higher for all students as a whole, and for Black students in particular, than in Reforms A and B—and the welfare of all students measured both by academic achievement and by the absolute reduction in outcome disparities is higher than initially for all students in Reform D.

---

69 For an example, see Eliza Shapiro, $773 Million Later, de Blasio Ends Signature Initiative to Improve Failing Schools, N.Y. TIMES, Feb. 26, 2019, at A20 (discussed infra note 179).
Year in and year out, many or most students have preferred Reform A—and some have preferred Reform B—to Reforms C and D. (All prefer Reform A to Reform B, and Reform E to all others if it is attainable.) Preferences remain about the same when it is stipulated that the jurisdiction in question has defined the academic level reached by Black students in Reforms C and D—but not the level reached in Reforms A or B—as at or above the level needed for success in college and careers.

We acknowledge that Reforms A and B have equity gains—and welfare gains from a “positional good” standpoint—and that Reforms C and D do not. But particularly in a knowledge economy in which the issue increasingly is not how well graduates compete with each other for jobs, but how well they compete with machines that can do the sorts of low-skill and low-analytic-capacity work that once predominated in industrial economies, we expect that the net result of choosing Reform A or B over Reforms C and D is a substantial welfare loss for Black students.

The greater worry, however—within the education context—is that focusing on gap closure alone will leave Black students with no achievable

70 See supra note 68 (discussing positional goods).
72 First, Black graduates in Reform D have acquired more knowledge and skills than those in Reform A, giving them access to higher education opportunities and jobs for which Black Reform A graduates do not qualify. With regard to those benefits, Black graduates’ gains are not positional; they achieve the same additional access to benefits as White graduates with the same qualifications. Second, the value of efforts at Reform A have to be discounted by the probability of outcomes more like Reforms C or F. White families are unlikely to sit still even if school officials are paying little attention to their children, meaning the cost of diminished welfare gains for Black graduates may not be more, and could well be less, equity. Better, then, to combine all families’ and students’ motivations to improve rather than pitting them against each other. Additionally, the microeconomic importance of competitive advantages between individual workers is overstated given the macroeconomic advantages of increasing the skill level of individuals of both races. Jobs are attracted to locations with large numbers of workers with a variety of skills. As such, a Reform D situation in which everyone on average is better educated than in the Reform A situation is likely to place considerably more jobs at the disposal of Black (and White) workers than is Reform A, further undermining the welfare impact of the equality advantage Black students have relative to White students in Reform A compared to Reform D.
possibilities other than leveling-down à la Reform B, with all its attendant welfare, political, and social-conflict costs. Indeed, over decades in that context, there are no sustained examples of successful equity-plus-welfare gains of anything like the Reform A sort.73 Since 1970, however, as we develop below, the nation has seen two roughly fifteen-year periods of equity-plus-welfare gains of the Reform D sort. Both suggest that systemic changes pulling all students up can improve the lot of Black students on both equity and welfare dimensions. Both also reveal the capacity of such reforms to hold the line distinguishing equity-positive Reform D from equity-neutral Reform C in ways that our proposed equal protection standard could substantially reinforce.

III. DISPARITY-DRIVEN LEARNING IN ACTION

But how is Reform D possible? Just as leveling-up Reform A may end up as leveling-down Reform B, might rising-tide Reform D end up, as rising-tide efforts so often do, as a Reform F, with White students gaining more than Black students, making escalating inequity the cost of significant welfare gains? Below, we offer several Reform D examples. We begin by briefly gesturing to two examples previously developed by others: efforts by retailers to diminish service disparities affecting sight-impaired customers and employment disparities affecting aspiring women managers (Section (A)) and efforts by Congress and local agencies to diminish racial disparities in the juvenile justice system (Section (B)). We end with two public education case studies that we more fully develop and tie to experimentation ourselves: large-city portfolio reforms from the turn of the

---

73 The closest courts have come to requiring absolute school equity are orders by a few courts, starting in California in Serrano v. Priest, 487 P.2d 1241, 1266 (Cal. 1971), that equalized per-capita funding across school districts. The result in California, however, according to some experts, was that the state leveled-down funding, shrinking the spending gaps between districts but not the achievement gaps. See Patricia F. First & Louis F. Miron, The Social Construction of Adequacy, 20 J.L. & Educ. 421, 428 (1991) (listing "undesirable consequences" of equalized resources distribution without also improving adequacy of resources, including "leveling-down" of the acceptable minimum of educational offerings); W. Norton Grubb, The Money Myth: School Resources, Outcomes, and Equity 258 (2009) (tracing history of Serrano reforms, which were intended to "level up" funding by poor school districts but instead "level[ed] down" funding, achieving greater cross-district funding equality but no similar convergence of student outcomes). During this period, California went from being one of the nation's top-performing to one of its lowest-performing states educationally. See Lisa Trei, Documentary Examines How California Public Schools Fell 'From First to Worst', STANFORD REP., Jan. 21, 2004 (describing Merrow Report documentary detailing "how California’s school system went from the nation’s best in the 1950s to its worst by 1994").
current century to roughly 2015 (Section (C)) and school desegregation from 1970 to the mid-1980s (Section (D)).

A. DISABILITY AND EMPLOYMENT DISCRIMINATION

Starting in the mid-1990s, Lainey Feingold and the law firm of Goldstein, Borgen, Dardarian & Ho initiated a series of structured negotiations with the nation’s largest banks to develop “talking” automatic teller machines (ATMs) to overcome disparities in banking services between sighted and sight-impaired customers. At a time when ATM machines and the Americans with Disabilities Act were both new, Feingold and colleagues worried that “[i]f we brought the issue to a judge, maybe the judge would say, ‘What the hell? Blind people shouldn’t go to an ATM.’”

So in lieu of a lawsuit, they began discussions, first with Citibank, then with Bank of America and Wells Fargo, to convince them that enhancing ATMs with text-to-speech screen readers would be good for the companies as well as for their blind customers—discussions that ultimately prompted technology that set the industry standard. Subsequent negotiations with Dollar General, Target, Trader Joe’s, and other retailers led them to add braille keypads to touch-screen point-of-sale checkout devices that blind customers otherwise could not use. Realizing that “[w]hen technology becomes a compliance issue, creativity is lost, our enthusiasm is lost, and things get stuck in the law office[,]” Feingold directed the negotiations beyond adherence to minimum legal requirements toward ways of “integrating technology, web development and usability” that benefited businesses as well as her disabled clients. Over time, Feingold convinced other companies, including Microsoft, Google, and Apple, to “embrace[ ] accessibility, not just to include the disabled public in products and services, but because of the many benefits of accessibility in addition to protecting civil rights,” including enhanced search engine optimization and market...

74 In presenting these reforms, we intend not to endorse them all but to illustrate how gap mining can reveal many biases and misimpressions that generate—and many ways to close—the gaps.
76 Id.
77 Id.
78 Id.
79 Id.
80 Id.
share, corporate goodwill, and other innovations in the use of digital space.\textsuperscript{81} Feingold and colleagues have also applied the approach successfully in pre- and post-litigation settlement negotiations with the Houston and San Francisco public transit agencies, Major League Baseball, Denny’s, Charles Schwab, Kaiser, Anthem, and other disability-rights defendants.\textsuperscript{82}

Goldstein firm lawyers used similar techniques in workplace racial and gender discrimination lawsuits by developing expertise in designing settlements that comprehensively restructured hiring, promotion, and training practices to improve defendant companies’ overall efficiency and profitability while increasing the number of women and people of color hired and promoted.\textsuperscript{83} An example is a settlement with Home Depot of a lawsuit alleging gender discrimination in filling managerial positions.\textsuperscript{84} Aiming to “link equity and effectiveness,” lawyers on both sides designed “a new system for hiring, promotion and training” that extended to the company’s personnel practices its longstanding innovative approach to “marketing practices, opening stores, [and] figuring out how to make products accessible to people.”\textsuperscript{85} Home Depot’s human resources (“HR”) middle managers appreciated the new system because it helped them “tap [the company’s] pool of workers in a rigorous way,” “fill[] jobs with the right people,” streamline procedures, and reduce turnover.\textsuperscript{86} Company leaders were so pleased with the system that they voluntarily extended it


\textsuperscript{82} Ward, \textit{supra} note 75; \textit{see also} Matthew Hirsch, \textit{Meet Early and Settle Often}, \textit{LAW.COM} (July 21, 2007), \url{https://www.law.com/alm1D/l l8233035 l923/} [https://perma.cc/7V6L-G5VU]. It may seem a shame that defendants do the morally and legally right thing only upon seeing “there’s something in it for us.” Were that all there was to it, we still might accept that “hard reality” as a regrettable human given that we have to work around. Added to self-interest, however, are these defendants’ newly acquired self-awareness of their susceptibility to false assumptions and self-disciplining habit of inquiring when disparities raise that possibility anew. This combination of self-interest, awareness, and discipline is just what morality and law are expected to generate in order to guide future behavior in public-regarding directions.


\textsuperscript{85} Sturm, \textit{Second Generation}, \textit{supra} note 83, at 510–12.

\textsuperscript{86} \textit{Id.} at 511, 518–19.
beyond women in the geographic division governed by the consent decree, to people of color as well as women company-wide. These steps helped generate 28 and 30 percent increases in the number of Home Depot minority and female managers, respectively, during the system's first few years of operation. Defendant-side lawyers in the suit, likewise, offered other companies their expertise in restructuring HR practices to improve organizational efficiency and profitability while diminishing disparities in hiring and promoting women, Black, and other categories of workers.

Although Feingold, Goldstein, and their plaintiff- and defense-side colleagues might not describe themselves this way, they serve as disparity-reduction experimentation entrepreneurs in support of strategies like the one we propose to constitutionalize. They began by exposing disparities in access to technological conveniences and jobs affecting members of groups that are chronically disfavored by such disparities, then promoted and provided expertise in experimenting with new technologies and HR structures for overcoming the disparities in ways that enhanced overall efficiency. The experiments' success led defendant firms to extend the approach companywide, reducing disparities affecting members of other undervalued groups (as in the case of Home Depot) and led competitors to fall in line (as in the ATM example).

The retail firms behaved this way because the successful experiments exposed built-in structural conditions that generated disability-related, gender, and racial disparities for which there turned out to be no legitimate business need. This realization in turn created legal, as well as ethical and business, incentives to scale the experiments to other chronically underserved groups. By exposing false assumptions about sight-impaired individuals' inability to use technology and women's inability to serve as homebuilding experts, these structural and mindset changes spawned system-wide advances that benefited multiple constituencies—in the case of Home Depot, for example, all qualified job and promotion candidates, HR middle managers, and shareholders. Nor did extending the benefits of settlement agreements beyond the covered plaintiffs to other chronically underserved individuals keep the reforms from achieving real gap closure—at Home Depot, for example, 28 to 30 percent increases in minority and women managers in a few years. Instead, it assured that the beneficiaries of the company-wide rising tide were disproportionately women and minority managers.

---

87 Id. at 517–19.
88 Id.
89 See id. at 478 n.59 (discussing presentations by Home Depot's mediation counsel "recommending Home Depot's system to [other] companies as a prototype").
To be sure, not all of the people benefited by the Home Depot settlement were women and minority promotion candidates, nor did the remedy close all of Home Depot's gender and racial employment gaps. In theory, therefore, an exclusively gap-closing remedy might have diminished chronic employment disparities more than the actual remedy did. In practice, however, it seems unlikely that limiting the settlement benefits to the plaintiff class would have generated nearly as comprehensive a set of both equity and welfare returns for the chronically subordinated groups as the actual outcomes did—not to mention the less biased and more equity-oriented mindsets the changes generated among other Home Depot constituents.

B. JUVENILE JUSTICE

A number of statutes condition federal funding for state and local public agencies on their commitment—enforceable when funding renewals are sought—to (1) track racial disparities or other questionable outcomes at each step or in each category of the agency's operations; (2) develop a plan to reduce disparities or other untoward results for each step or category; and (3) report periodically on the success of ameliorative actions taken and adjustments made when actions falter, including by comparison to reductions achieved by other funded agencies under similar circumstances. Questionable outcomes regulated by such legislation include racial disparities in spending roughly $800 billion in federal stimulus funds in the midst of the 2009 recession, racial disparities in juvenile justice contact and confinement, and prison rapes.

---


93 Prison Rape Elimination Act of 2003, id. §§ 15605(a)–(e).
The most widely studied of these funding regimes is the Juvenile Justice and Delinquency Prevention Act ("JJDPA"). First adopted in 1974, the JJDPA allocates funds to state, local, and private juvenile justice agencies nationwide; evaluates and provides technical assistance and training to federally assisted programs; "establish[es] a centralized research effort on the problems of juvenile delinquency;" and coordinates "the collection, preparation, and dissemination of useful data regarding the treatment and control of juvenile offenders." Extending this broad research and experimentation function from the national to the state and local levels, federal funding is predicated on the submission of plans for achieving the Act’s statutory purposes and for expanding local "research, training, and evaluation capacity."

Initially, the Act directed national and local experimentation at juvenile delinquency prevention and diminishing juvenile confinement through innovative diversion programs, community-based alternatives, and school retention. In successive amendments in 1988, 1992, 2002, and 2018, however, Congress defined reducing disproportionate minority contact and confinement ("DMC") as another "core" requirement of the Act and put funded agencies at risk of losing a quarter of their funding if they fail to address in good faith racial disparities at all decision points in the juvenile justice continuum (for example, arrest, diversion, adjudication, and court disposition). Plans must include methods of tracking racial disparities at each decision point and disparity-reduction strategies for all points where racial disproportion is found. Funded agencies have broad discretion over the content of their plans, guided by federal training and the "back and forth between [funded agencies], policymakers, and advocates" that helps benchmark plan quality nationally.

---

96 See id. § 102(a) (codified at 42 U.S.C. § 5602) (listing the evaluation, technical assistance, training, centralized research and data dissemination, standards-generating, funding, and other purposes of the Act).
97 Id.; see Juvenile Justice Reform Act of 2018 § 102, 34 U.S.C. § 11102 (updating the purposes of the JJDPA).
99 Id. § 223(a) (codified at 42 U.S.C. § 5633).
100 Id. §§ 223(a)(11)–(13) (codified at 42 U.S.C. §§ 5633, 5651).
102 See id. at 409.
103 See id.
Since 1992, the Annie E. Casey Foundation’s Juvenile Detention Alternatives Initiative (JDAI) has supported the DMC regime and facilitated cross-agency benchmarking by connecting federally funded juvenile justice agencies in a nationwide learning network for developing and spreading effective disparity-reduction practices.\(^{104}\) Through “peer to peer learning,” networked jurisdictions share resources, receive expert advice and data, and take guidance from five “network learning labs”—four cities and one state—that the Foundation funds to disseminate findings to other jurisdictions.\(^{105}\)

Experimentation mandated and supported by JJDPA has generated many effective disparity-reduction practices.\(^{106}\) These include small but effective changes, such as telephonic reminders of court hearings and in-community reporting centers to avoid detention orders based on juveniles’ failure to appear in court or to attend meetings with probation officers.\(^{107}\)
They also include the replacement of racially biased standards for assessing flight and re-offense risks (for example, the minor “is a danger to others;” "has gang associations;" "exhibits “negative attitudinal and personality traits;" or lacks “a parent to monitor behavior" or “good family structure") with equally effective but less loaded standards (for example, severity of the offense; prior delinquency adjudications or failures to appear; or no “adult willing to take responsibility"). Changes of this sort “may lead officials to see that assumptions that . . . formerly disposed them to detain certain categories of children more often than others—for example, that a failure to appear in court evidenced guilt, flight proneness, or irresponsibility—are wrong.”

.....

from almost 40 percent in 1994 to 11 percent in 2003); Sabel & Simon, Contextualizing Regimes, supra note 23, at 1289–90 (discussing evening reporting center and other neighborhood-based services).


Patrick McCarthy et al., Nat’l Inst. of Just., The Future of Youth Justice: A Community Based Alternative to the Youth Justice Model, at 16 (Oct. 21, 2016), https://www.aecf.org/m/resourcedoc/NIJ-The_Future_of_Youth_Justice-1021.16.pdf [https://perma.cc/RF7P-2P6E]; see Annie E. Casey Found., No Place for Kids, supra note 104, at 15 (noting that, although “[m]any youth without serious offending histories are placed into custody for repeatedly violating rules and/or behaving disrespectfully toward judges, probation officers, and other authorities,” institutional noncompliance of this sort is “not very predictive of the risk of recidivism”).

Wolfson & Hubner, supra note 107, at 48–49 (emphasis added).

Id. at 49.

See Steinhardt, supra note 108, at 11 fig.1 (emphasizing importance of developing risk assessment instruments with a point scale based only on proven high-risk factors associated with non-appearance and dangerous behavior).

Id. at 11.


Wolfson & Hubner, supra note 107, at 48–49 (emphasis added) (acknowledging that prior terminology “worked as a bias against minority youth” and that updated terminology takes advantage of “the strengths of extended families that often exist in minority communities”).

Liebman, Perpetual Evolution, supra note 31, at 2028. One of the most important assumptions—with severely disproportionate effects on Black and Brown children—that JDAI-supported experimentation has disproven is that substituting nonincarcerative for incarcerative responses to behaviors by youth that enmesh them in the juvenile justice system will increase rates of juvenile crime. See McCarthy et al., supra note 110, at 1–4 (describing misguided fears of predator juveniles that triggered vast expansion of juvenile prisons and incarcerations). Numerous disparity-reduction experiments have demonstrated the opposite to be true. See, e.g., McCarthy et al., supra note 110, at 15, 19 (reporting that as juvenile incarceration rates decreased, public safety outcomes improved in 90 percent of JDAI sites; in New York State, for example, a 65
Juvenile justice agencies’ systematic experimentation with racial disparity reduction helped generate a broader result: an astonishing 61 percent decline from 1997 to 2017 in the proportion of youths in the country who were detained as a result of interactions with the juvenile justice system (dropping from 356 to 138 per 100,000 youths). Had the 1997 detention rate persisted, 68,444 more youths would have been detained in 2017. During this period, juvenile crime and recidivism rates also dropped substantially, entailing overall a vast welfare gain for youth and society at large.

What about the program’s equity effects? In 1997, Black youths accounted for 40 percent of all juvenile detainees; in 2017, they were 41 percent. The comparable numbers are 18 and 21 percent for Hispanic youths but 37 and 33 percent for White youths. From this perspective, there was no equity gain, as critics on both the left and right have noted. But
when respective declines in the detention rate for each racial category of youths are considered, the case for positive equity effects is stronger. For Black youth, the detention rate declined by 60 percent, for Hispanics by 75 percent, and for Asians by 90 percent, compared to 59 percent for Whites. Most importantly, due to the vast overrepresentation of children of color in the juvenile justice system, the 68,844 youths who would have been in custody in 2017 had 1997 detention rates persisted would likely have consisted of many more youth of color than White youth. From this latter perspective, the beneficiaries of the reform are disproportionately Black, Hispanic, and other youths of color, and the reform is a rare public initiative with beneficiaries who are disproportionately members of chronically disfavored groups.122

Equally significant are the equity dividends that may be paid in the future by sensitization of adults running the system to their unconscious biases.123 Indeed, lessons learned and dispositions changed through these efforts demonstrate the workability of radically decarcerative proposals for juvenile offenders,124 and likely deserve some of the credit for juvenile

---

122 This distribution of JJDPA’s welfare effects on Black and Latinx youth, compared to White youth, arises only because of horrendous racial disparities in historical Black and Latinx compared to White youth detention rates. Even so, we count this as a positive equity effect, especially given that the progress achieved now points the way to a reform—decarceration of youth generally—that would end disparities entirely, albeit still via a “universalist,” not single (racial)-purpose, policy. See infra notes 123–125 and accompanying text (discussing decarceration proposal); see also supra note 59 (discussing “universal” reforms).

123 See supra note 36 and accompanying text.

justice agencies’ ability to achieve dramatic decreases in juvenile detention early in the COVID-19 crisis.\textsuperscript{125}

Even without factoring in the benefits to White youth and society at large, the combined equity and associated welfare gains for communities of color from JJDPA’s and JDAl’s contributions to the 20-year decline in detentions of mainly Black and Latinx youth are a net positive, if incomplete, increase in racial justice. From an equity and welfare perspective, the ongoing duty to experiment with disparity-reduction strategies is worth continuing, albeit with added pressure to diminish remaining racial disparities—exactly as Congress has applied through its 2018 revisions to the JJDPA.\textsuperscript{126}

C. PORTFOLIO SCHOOL REFORMS

Another example are gains in student achievement following big-city “portfolio” reforms prompted in part by experimentation requirements in the No Child Left Behind Act of 2001.\textsuperscript{127} Among others, the school districts in Baltimore, Boston, Camden, Chicago, Denver, Houston, Indianapolis, Los Angeles, Memphis, New Haven, New Orleans, New York, Newark, Oakland, and Washington, D.C. engaged in these reforms.\textsuperscript{128}


\textsuperscript{126} See Five Things to Know about the New Juvenile Justice Act, ANNIE E. CASE FOUND. (Feb. 8, 2019), https://www.aecf.org/blog/five-things-to-know-about-the-new-juvenile-justice-act (summarizing the Juvenile Justice Reform Act of 2018, Pub. Law No. 115-385, 132 Stat. 5123, that amended the JJDPA to require funded agencies to experiment with systems-improvement strategies to reduce disparities, "such as establishing coordinating bodies to advise states; using data to identify bias in decision making; and creating and implementing a work plan with measurable objectives to address the needs identified by the data").


\textsuperscript{128} James S. Liebman, Elizabeth Cruikshank & Christina Ma, Governance of Steel and Kryptonite Politics in Contemporary Public Education Reform, 69 FLA. L. REV. 365, 367–68, 387–401 (2017); PAUL HILL ET AL., CRPE, PORTFOLIO SCHOOL DISTRICTS FOR BIG CITIES: AN INTERIM REPORT (Univ. of Wash. Oct. 2009); PAUL HILL & CHRISTINE CAMPBELL, CRPE, GROWING NUMBER OF DISTRICTS SEEK BOLD CHANGE WITH PORTFOLIO STRATEGY, at 1 (Univ. of Wash. June 2011) (“New York City has become an incubator for portfolio district leadership. Today New
In many cases, change began with legislation giving mayors greater control over districts previously governed diffusely by elected school boards, central bureaucrats, community boards, area supervisors, and collective bargaining agreements. Other common features included rigorous learning standards, greater autonomy for schools in deciding how to meet standards, more accountability for whether they did, closure of poor-performing schools and intentional “portfolios” of new schools, educator evaluation tied to student outcomes, and rich data and tools for diagnosing student needs and developing strategies to meet them.

The leading-edge reforms in New York City (the “City”) are the most widely studied. Starting in late 2002, after the state legislature gave the city’s mayor full control of the City school district, newly elected Mayor Michael Bloomberg and his Department of Education Chancellor Joel Klein implemented several key changes. First, they gave principals control over hiring, budget, class schedules, curriculum, and professional development, while holding them accountable for using that flexibility to find ways to increase their students’ (particularly low-performing students’) learning gains to levels equal to or better than the highest levels achieved at other schools with similar populations. Second, they used student Orleans, Newark, Chicago, and Baltimore are all led by people mentored by former NYC Schools Chancellor Joel Klein.

Orleans, Newark, Chicago, and Baltimore are all led by people mentored by former NYC Schools Chancellor Joel Klein.

See Liebman, Cruikshank & Ma, supra note 128, at 387–88 (citing sources).

See id. at 393–94 (citing sources).


See James C. McKinley, Jr., State Senate Passes Bill Giving Mayor Control of Schools, N.Y. TIMES, June 12, 2002, at B6.

See Hill, supra note 131 (noting reforms’ diminution of pre-existing “concentration of the lowest skilled teachers in schools serving poor and disadvantaged children” and per-pupil funding increase of $5,273 for high-poverty schools versus $3,962 for low-poverty schools): N.Y.C. INDEP. BUDGET OFF., EXAMINING FIVE YEARS OF SCHOOL ALLOCATIONS UNDER FAIR STUDENT
outcomes, student/parent/teacher surveys, and expert "quality reviews" of schools to (1) identify and close chronically poor-performing schools; (2) replace them with district-led or independent charter schools implementing innovative instructional models; and (3) facilitate family choice among schools.134 Third, they increased per-pupil funding, calibrating it to differential student need and shifting hundreds of millions of dollars from the central bureaucracy to schools.135 Fourth, they recruited more qualified teachers, placing them in lower-performing schools,136 raising salaries substantially, and jettisoning seniority rules that previously concentrated less qualified teachers in schools with more disadvantaged students. Finally, they built educator capacity via (1) a Leadership Academy that recruited the system’s best teachers to be principals and trained them to lead high-need schools and use newly available qualitative and quantitative data to identify and respond to each student’s learning needs; and (2) a cadre of facilitators for school-based “inquiry teams” of educators who hypothesized causes of instructional failure and implemented solutions by closely observing students’ responses to different modes of instruction.137

Each aspect of the reform promoted experimentation. Whether at the student, classroom, school, or district level, the focus was on new ways to overcome disparities between each student’s potential and actual results, between state proficiency standards and underperforming students’ results, and among racial and other categories of students.138


134 Liebman, Cruikshank & Ma, supra note 128, at 394–95.
135 See N.Y.C. INDEP. BUDGET OFF., supra note 133, at 6–12.
136 Lindsey Christ, City Teachers Outperform Teachers in Rest of State, State Data Finds, NY1 (Sept. 5, 2013), https://www.ny1.com/nyc/all-boroughs/archives/2013/09/05/city-teachers-outperform-teachers-in-rest-of-state--state-data-finds.NYC_188311 [https://perma.cc/B62C-4X4P] (reporting state teacher-quality ratings in last year of Bloomberg reforms; 11 percent of City teachers received highest ranking versus 5 percent in rest of State; 11 percent of City teachers were rated less than “effective” versus 20 percent elsewhere).

137 See Stacey Childress et al., Managing for Results at the New York City Department of Education, in Education Reform in New York City, in EDUCATION REFORM IN NEW YORK CITY:
measures, as a variety of rigorous studies reveals, these reforms improved student outcomes in New York City and nationwide as other cities followed suit.

AMBITIONS CHANGE IN THE NATION’S MOST COMPLEX SCHOOL SYSTEM, 87, 90 (Jennifer A. O’Day et al. eds., 2011) (quoting Chancellor Klein’s explanation of reforms as “shift[ing] the locus of power from central office to the schools,” “shift[ing] the organizational culture to a focus on results,” and “build[ing] into the equation some variability in terms of problem solving at the school level and learn[ing] from it”); KELLEHER, supra note 131, at 1–2 (“[N]ew York City’s most successful reforms created conditions that permitted school-level innovation . . . ”); Hill, supra note 131 (describing reforms as enabling schools to “assemble a group of teachers who are united around a school mission” to “work closely with one another and with individual students to meet high academic expectations”); Liebman, Cruikshank & Ma, supra note 128, at 389 (describing the goal of transforming district leaders from “deciders and implementation micromanagers” into “assemblers of a ‘portfolio’ of differentiated schools” with flexibility, resources, data, and facilitation needed to explore effective strategies).

139 See, e.g., Will Dobbie & Roland G. Fryer, Are High-Quality Schools Enough to Increase Achievement among the Poor?, 3 AM. ECON. J. APP. ECON. 158 (2011) (estimating effects on predominantly Black students of attending a set of New York City charter elementary schools as large enough to close the racial achievement gap in math and English); Roland G. Fryer & Will Dobbie, The Medium-Term Impacts of High-Achieving Charter Schools, 125 J. POL. ECON. 985 (2015) (finding, by comparing academic results six years later of students who entered lottery to attend a set of New York City charter schools, that students who participated in the lottery and attended charter schools had higher academic achievement, educational attainment, and on-time-to-graduate metrics and lower rates of teenage pregnancy and incarceration than students who participated in the lottery but attended other schools); James S. Liebman & Jonah E. Rockoff, Moving Mountains in New York City, EDUC. WEEK, Nov. 30, 2010, at http://www.edweek.org/ew/articles/2010/11/30/liebman.h30.html [https://perma.cc/6B9M-ENVY] (“[T]he best studies show that each of three key policies—grading schools based on student outcomes, supporting high-quality charter schools, and replacing failing schools with new ones run by empowered principals and teachers—lifted student performance.”); James S. Liebman, Columbia L. Sch., Presentation at Harvard Law School Advocates for Education 2013 Annual Conference: Reforming the Governance and Accountability of the Nation’s Public Schools 14, 23 (Mar. 26, 2013) (on file with authors) (showing that students on whom inquiry teams focused gained, on average, 11 percent more proficiency in math and 6 percent in English over a year than other students with similar starting points); infra note 143 (small high schools); infra note 139 (charter schools); Jonah E. Rockoff & Lesley J. Turner, Short Run Impacts of Accountability on School Quality (Nat’l Bureau of Econ. Rsch., Working Paper No. 14564, 2008) (using quasi-experimental analysis to find that a school’s “receipt of a low grade” from New York City’s school-evaluation system “significantly increased student achievement” in English and math the next year); Marcus A. Winters & Joshua M. Cowen, Grading New York: Accountability and Student Proficiency in America’s Largest School District, 34 EDUC. EVAL. & POL’Y ANALYSIS 313 (2012) (similar); see also Howard S. Bloom et al., Lessons for New York City’s Small Schools of Choice about High School Features that Promote Graduation for Disadvantaged Students, 39 J. POL’Y ANALYSIS & MGMT. 740 (2020) (analyzing “rich dataset based on naturally-occurring lotteries for 68 new small non-selective high schools in New York City [‘SSCs’]” and identifying as predictors of SSC’s higher graduation rates for disadvantaged students “school leadership quality, teacher empowerment, teacher mutual support, teacher evaluation and feedback, teacher professional development, data-driven instruction, teacher/parent communication, academic rigor, personalized learning, and teacher/student respect”); JAMES J. KEMPLE, RSCH. ALL. FOR N.Y.C. SCHOOLS, THE CONDITION OF NEW YORK CITY HIGH


1. Improved Graduation and College-readiness Rates

During the early 2000s, multiple agencies tracked New York City’s graduation rate using different measures. Between 1986 and 2002, across the terms of nine education chancellors, graduation rates on the least rigorous measure never exceeded 51 percent, flat-lining around 46 percent until 1991 and around 50 percent until 2002. As Figure 2 illustrates, however, between 2003 and 2013, that rate climbed 33 percent to 68. On the most rigorous, federal measure (first calculated in 1999), the pre-reform rate hovered around 37 percent. Halfway through the reform (2007), when the federal government last calculated the rate, it had risen 49 percent to 55.

On New York State’s (the “State’s”) two graduation rates (first calculated in 2005), the City’s rate rose 42 percent to 66 during the remainder of the reforms; on the state’s more rigorous, college-ready measure, the rate jumped 108 percent to 63.5 percent. A major cause of these increases—

SCHOOLS, at 4 (NYU Steinhardt Mar. 2013) [hereinafter KEMPLE, NYC HIGH SCHOOLS], https://research.steinhardt.nyu.edu/scmsAdmin/media/users/sgl58/PDFs/conditionNYCHS.pdf [https://perma.cc/9CAX-UDV9] (attributing rising graduation and college-readiness rates to reforms’ “systematic method for identifying and phasing out high schools that exhibit persistently low levels of performance” and replacing them with “portfolio of new smaller, mission-driven high schools,” “high school choice,” and “accountability system that measures and holds schools responsible for improving key outcomes”).


141 Eliza Shapiro, Bloomberg Announces Graduation Rates Early, POLITICONEWYORK, Dec. 4, 2013, https://www.politico.com/states/new-york/city-hall/story/2013/12/bloomberg-announces-graduation-rate-increase-early-009912 [https://perma.cc/WLY7-JZ2]; see KEMPLE, NYC HIGH SCHOOLS, supra note 139, at 1 (comparing New York City high schools in 1980s and 1990s, when graduation rates “hovered at or below 50 percent, nearly 20 percentage points lower than state and national averages” and when “[s]uccessive waves of ineffective reforms . . . offered little hope of improvement” to changed “high school landscape” as of 2011, with “steady improvement in student outcomes, across all groups of students” in graduation (“rates increased from 51 percent of those who entered high school in 1999 (scheduled to graduate in 2003) to 69 percent of those entering in 2007 (scheduled to graduate in 2011.”)); “more students earning a Regents diploma (rather than the less demanding Local diploma)”; declines in dropout and transfer rates; and “improved” attendance, credit accumulation, rates of being on-track to graduate, and taking and passing college-reading examination); see also Hill, supra note 131 (“Every year [as of a decade into the reform] 18,000 [more New York City] young people graduate high school than would have been expected in 2002.”); Aaron M. Pallas, Policy Directions for K–12 Public Education in New York City, in Toward a 21st Century City for All: Progressive Policies for New York City in 2013 and Beyond 1, 5 (John Mollenkopf et al. ed., 2013) [hereinafter Pallas, Policy Directions], https://www.gc.cuny.edu/CUNY_GC/media/CUNY-Graduate-Center/PDF/Centers/Center%20for%20Urban%20Research/Resources/21cfonall_education.pdf [https://perma.cc/9YE4-8RYU] (“High school graduation rates have risen substantially over the past six years across all groups of students.”).
which coincided with the State’s adoption of tougher graduation requirements—was the city’s closure of 140 failing high schools with graduation rates perennially below 50 and their replacement with new small high schools with above-average graduate rates."

142 See Kemple, NYC High Schools, supra note 139 at 2, 12 & n.4 (discussing state’s abandonment during reforms of lesser “local diploma” and requirement that graduating students qualify for more demanding “Regents” diploma).

143 See Liebman, Cruikshank & Ma, supra note 128, at 397; see Hill, supra note 131 (“On campuses where new small schools replaced large underperforming high schools, the overall graduation rate increased from 37.9 percent to 67.7 percent . . . [while] serving [students] in poverty, with disabilities, [and] learning English”; “students who entered the new small schools with the lowest test scores benefited the most”); Leanna Stiefel, Amy Ellen Schwartz & Matthew Wiswall, Does Small High School Reform Lift Urban Districts? Evidence from New York City, 20 EDUC. RES. 1, 2–3, 6–10, 18 (2015), https://wagner.nyu.edu/files/faculty/publications/EDUCATIONAL_RESEARCHER-2015-Stiefel-Sch-Wis_Sm_Schools.pdf (contrasting uneven results of small high schools elsewhere in nation and in New York City before Bloomberg reforms with substantial graduation gains across all types of high schools during New York City’s reforms; highest gains were in high schools created by the reforms, with “graduation rates roughly 13 percentage points higher than” in other schools; new small high schools’ “salutary effects on continuously operating large schools” validate City’s “portfolio” strategy combining large and small schools); Rebecca Untermann & Zeest Haider, MDRC, New York City’s Small Schools: A First Look at Effects on Postsecondary Persistence and Labor Market Outcomes, at 1 (Apr. 2019), https://www.mdrc.org/sites/default/files//SSC-First_Look%20Brief.pdf (using “naturally occurring lotteries within New York City’s high school assignment process to rigorously study” small high schools of choice (SSCs) and finding that all types of city high schools improved during the reforms, but SSCs had larger “positive impacts on students’ secondary school and college outcomes . . . at no additional cost per graduate”; students “of all backgrounds” “who won a lottery and enrolled in an SSC . . . were 9.5 percentage points more likely to graduate from high school than those who lost a lottery and did not enroll in an SSC”). Demographic change cannot account for the City’s rising graduation rate. See Stiefel, Schwartz & Wiswall, supra (noting “remarkab[el]” demographic stability in City’s student body during reform years).
After remaining flat for 16 years, NYC's graduation rate has increased by 33% since 2002.

New York City's graduation gains outpaced those in other cities. In all reform years when calculated rates allow apples-to-apples comparisons, New York City's rates began lower, ended higher, and maintained more consistent growth when compared to rates in other cities for which data are available:

| Federal graduation rate, percentage-point increase 2001–2007¹⁴⁴ |
|-----------------------|-----------------|----------------|-----------------|-----------------|
| New York City         | Charleston      | Chicago        | Denver          | Los Angeles     |
| +16.6                 | -10.1           | +6.6           | +12.3           | -5.8            |

| New York State large city graduation rates, 2005–2013¹⁴⁵ |
|------------------------|-----------------|-----------------|-----------------|
| Start / end rate       | Point increase  | Percent increase| Change in gap with state |
| New York City          | 43.5 / 61.3     | +17.8           | 41%             | -34%            |
| Buffalo, Rochester, Syracuse, Yonkers | 45.3 / 52.4 | +7.1            | 16%             | +20%            |

However defined, high school graduation enhances job prospects and other

opportunities. Still, one may wonder whether New York City’s graduation rate increases reflected greater knowledge and skills. Suggesting they did are concurrent increases in the rates at which:

- Ninth graders entering high school enrolled in college four years later and persisted in college;  
- High schoolers took at least one College Board Advanced Placement Exam and passed at least one;  
- High school graduates met the City University of New York (CUNY) college-readiness standard, enrolled at CUNY, and entered CUNY without needing to take remedial courses.

Graduates of New York City’s new small high schools had even greater gains. Although a diploma still eludes too many City high school

---

146 See, e.g., James J. Heckman et al., Earnings Functions and Rate of Return, 2 J. HUM. CAP. 1 (2008) (finding that high school graduates are employed at higher rates and earn more than students who did not graduate).

147 KRISTIN BLACK & VANESSA COCA, RSCH. ALL. FOR N.Y.C. SCHOOLS, NEW YORK CITY GOES TO COLLEGE, at ES-iii (June 2017) (documenting rise in college-enrollment rates during reforms, from 55 percent of students starting high school in 2003 to 61 percent of those starting in 2008, and modestly higher college persistence rates).

148 KELLEHER, supra note 131, at 6 (“Between 2002 and 2012, the number of New York City public high school students taking one or more Advanced Placement examinations has grown from 17,165 to 32,471, an increase of 89 percent [while] pass rate held steady at 56 percent . . . .”); N.Y.C. Dep’t of Educ., Division of Academics, Performance, and Support, at 9 (July 2012) (on file with authors) (documenting 9,736 AP test takers in 2002: \((.56(32,471) - 9,736)/9,736 = 87\%\) increase in students passing at least one test); see also Philissa Cramer, Gains in City’s AP Exam Pass Rate Outpaced Participation Growth, CHALKBEAT (Sept. 25, 2012, 4:20 PM), https://ny.chalkbeat.org/2012/9/25/21089590/gains-in-city-s-ap-exam-pass-rate-outpaced-participation-growth [https://perma.cc/8EV2-GNKL] (“[B]ucking a common trend in standardized tests [that] as testing pools grow and become more diverse, average scores . . . fall. . . . [9.1 percent] more [New York City] students took [AP] exams [than year before, but] the number of students passing the exams rose by even more—12.7 percent . . . .”); KELLEHER, supra note 131, at 6 (noting that between 2007, when first calculated, and 2013, the number of SAT test takers in New York City public schools “increased by nearly 14 percent,” while average score “held steady as participation increased”).

149 Hill, supra note 131 (“The percentage of [New York City] graduates who enter college [as of 2013] without needing to take remedial courses has doubled since 2001.”); N.Y.C. Dep’t of Educ., Division of Academics, Performance, and Support, at 5 (July 2012) (on file with authors) (documenting 56 percent increase in students meeting CUNY college-readiness standard from 2005 to 2011; 54 percent rise in CUNY enrollment from 2002 to 2010; and 32 percent increase in proportion of CUNY enrollees who matriculated to 4-year colleges from 2002 to 2010).

150 UNTERMAN & HAIDER, supra note 143, at 1–2, 4 (finding that students selected by lottery to attend new small high schools of choice (SSCs) were more likely to enroll in postsecondary education, persist in college for four years, and be enrolled in postsecondary education and/or employed four years later than identical students who applied to attend but were lotteried-out of SSCs); see HOWARD S. BLOOM & REBECCA UNTERMAN, MDRC, SUSTAINED PROGRESS: NEW
seniors and does not assure others success in college, advances during the reform years were substantial.

2. Increased Proficiency on State Math and English Tests

City students’ gains on New York State standardized math and English assessments are another example of improvements resulting from the school reforms, despite upgrades in the rigor of state “proficiency” cut scores and in the tests themselves. From 2003 through 2010, the story is straightforward: the reforms moved the district’s entire curve, as Figure 3’s analysis of City students’ fourth-grade math results in those two years shows. The same is true for eighth grade math as well as fourth and eighth grade English—encompassing all of the grades and subjects in which annual tests were given every year during the reforms.  

FINDINGS ABOUT THE EFFECTIVENESS AND OPERATION OF SMALL PUBLIC HIGH SCHOOLS OF CHOICE IN NEW YORK CITY, at 7-8 & Tbl.2 (Aug. 2013), https://www.mdrc.org/sites/default/files/sustained_progress_FR_0.pdf (enrolling in an SSC increased college readiness on CUNY scale by 20 percent). 

See BLACK & COCA, supra note 147, at 6-10.

Liebman & Rockoff, supra note 139 (“Between 2002 and 2010, New York City educators shifted the entire bell curve of state-test results decisively to the right, toward systematically greater skill in [fourth grade] math,” and “[m]uch the same is true for 6th grade math and 4th and 8th grade English.”); see Panel Presentation, supra note 144, at 6-8 (presenting similar graphs for other grade-subject combinations); see also KELLEHER, supra note 131, at 6 (finding that “[b]etween 2006 and 2009, New York City’s students showed substantial performance increases on state standardized tests”). The district confirmed that these changes reflected more learning—not easier tests—by examining whether over time, given scores on the exam in eighth grade predicted lower high school graduation rates for same students years later, and finding no appreciable diminution in the graduation probability a given eight-grade score predicted. N.Y.C. Dep’t of Educ., Overview of New York City Accountability, Results 2002–2012 (2012) (on file with authors) (finding minor score inflation for top-performing students in math and no evidence of score inflation for other students in math or any students in English).
In 2010, New York State raised the scores needed to demonstrate proficiency (for example, from the 637 to the 674 fourth grade math cut-score lines on Figure 3). Then, in the last year of the Bloomberg reforms (2013), the State adopted more rigorous learning standards and exams. These changes confound year-to-year district-specific comparisons after 2010. The best way to assess progress after 2010, therefore, is to compare rates of change in test scores in New York City with those in other districts in the state—all of which administered the same exams—while accounting for demographic differences between districts. Demographically, New York City and State students are mirror images: City schools are about 67 percent Black and Latino and 30 percent white and Asian; schools elsewhere in the State are about 30 percent Black and Latino, and 68 percent white and Asian.

City students are more likely to be disabled (1.5x), economically disadvantaged (1.67x), English Learners (2.33x), and homeless (8x)—all factors that predict lower average student outcomes in the City compared to the rest of the State, especially as tests become more difficult.\(^\text{156}\) That, indeed, was the case before the reforms, when the City’s average reading and math proficiency rates across all tested grades were, respectively, 22 and 25 percentage points lower than in the rest of the State.\(^\text{157}\)

In 2009, six years into the reforms, however, the City’s less privileged student body was rapidly catching up to peers elsewhere in the State, cutting their proficiency gap (all tested grades) 72 percent in math (from 25 to 7 points) and 40 percent in English (from 22 to 13 percentage points).\(^\text{158}\) Although new cut scores and tests after 2010 caused proficiency rates and scores to plummet statewide,\(^\text{160}\) they dropped less and recovered faster in the City than elsewhere in the state—counter to what socioeconomic and education-status conditions would predict.\(^\text{161}\) By 2013, City
students had narrowed their 2003 math and English proficiency gaps with other students in the State by, respectively, 94 percent (remaining gap 1.4 points) and 78 percent (remaining gap 4.7 points). Boosting the City’s test score gains was its expanding cohort of charter schools, which on average substantially outscored traditional public schools both in the City and statewide. Since 2013, scores have climbed statewide but New York City’s scores have climbed faster, propelling its students’ average proficiency rates above those of other students in the State in both subjects.

County-, district-, and school-level comparisons confirm the City’s gains relative to the rest of the State and reveal the strength of those gains across all five City boroughs (county equivalents), including the Bronx, the

---


163 CTR. FOR RsCH. ON EDUC. OUTCOMES, CHARTER SCHOOL PERFORMANCE IN NEW YORK CITY, at 7–9 (Feb. 20, 2013), https://www.nyccharterschools.org/sites/default/files/resources/CredoReport2013.pdf [https://perma.cc/HU4L-JEWU] (comparing annual student learning growth in 2007-2011 reading and math by students in New York City charter schools and in traditional public schools (“TPS”) the charter students otherwise would have attended; on average, charter students “gain[ed] an additional one month of learning in reading over their TPS counterparts per year of schooling” and “about five months of additional learning” in math); CTR. FOR RsCH. ON EDUC. OUTCOMES, NATIONAL CHARTER SCHOOL STUDY, at 29 (2013), https://credo.stanford.edu/sites/g/files/sbiybj6481/f/ncss_2013_final_draft.pdf [https://perma.cc/NT2F-CFDF] (finding that nationally during same period, charter students in reading barely outperformed, and in math barely underperformed, matched TPS students); CTR. FOR RsCH. ON EDUC. OUTCOMES, CHARTER SCHOOL PERFORMANCE IN NEW YORK, at 42–44 (Sept. 2017), https://credo.stanford.edu/sites/g/files/sbiybj6481/f/ny_state_report_2017_06_01_final.pdf [https://perma.cc/SN1Q-3PN] (finding in a later study of New York State charter schools during the last two years of the Bloomberg reforms and following three years, that annual “benefits for charter students,” “mostly driven by the performance of charter schools in New York City,” were equivalent to “34 days of additional learning in reading and 63 additional days in math” compared to learning by matched TPS students); see Caroline M. Hoxby & Sonali Murarka, Charter Schools in New York City: Who Enrolls and How They Affect Their Students’ Achievement 30–32 (Nat’l Bureau of Econ. Rsch., Working Paper No. 14852, 2009), https://www.nber.org/papers/w14852.pdf [https://perma.cc/YPS5-6KV4] (finding causal evidence that students in City charter schools outperformed statistically identical students in regular public schools in math and English).

State’s poorest county.\textsuperscript{165} Comparing New York City to the four other densely populated cities in the State—all of which had similar score gaps relative to the State average as of 2003—tells a similar story.\textsuperscript{166}

<table>
<thead>
<tr>
<th>Fourth Grade Scale-Score Gap with Rest of State</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>New York City</td>
</tr>
<tr>
<td>Buffalo, Rochester, Syracuse, Yonkers</td>
</tr>
</tbody>
</table>

\textsuperscript{165} New York State’s sixty-two counties range from one of the nation’s wealthiest (Westchester) to one of its poorest (Bronx). Bronx, Kings (Brooklyn), New York (Manhattan), Queens, and Richmond (Staten Island) counties comprise New York City, with median family income lower than the state average in four of the five counties (with 81 percent of the City’s students). Top Staten Island Public Schools, PUB. SCH. REV., https://www.publicschoolreview.com/new-york/staten-island [https://perma.cc/6K3U-C84X]; QuickFacts, U.S. CENSUS BUREAU, https://www.census.gov/quickfacts/fact/table/newyorkcitynewyork,bronxcountybronxboroughnewyork,kingcountybrooklynboroughnewyork,newyorkcountymanhattanboroughnewyork,queenscountyqueensboroughnewyork,richmondcountrystatenislandboroughnewyork/PST045219 [https://perma.cc/36AZ-KCRA]. In 2002, the City’s five boroughs had the bottom four, and five of the seven lowest, average fourth-grade reading proficiency levels among the sixty-two counties statewide. Jesse Margolis, A New View of New York City School Performance, 2002–2009, at 1 (on file with the authors). By 2009, however, Queens, Staten Island, Manhattan, and Brooklyn had leap-frogged dozens of other counties, ranking fourth, sixth, tenth, and fifteenth out of sixty-two. Bronx outperformed several counties and substantially closed the gap with other City boroughs. Id. The study compared counties’ actual five-year gains to gains their different starting points predicted. Outside the City, most counties’ expected and actual gains were within two points of each other, and none beat its expected average proficiency level by more than five points. In contrast, all five New York City counties “surpassed their expected gains by more than five points,” and together did so by almost nine points. Id. at 7–8. Controlling for lower starting positions, “New York City’s five boroughs made more progress in English and math than any other county in the state.” Id. at 1–8.

The State divides New York City public schools into 32 geographic school “districts” to compare them to the State’s other 656 districts. At the district level, Margolis found that, “[a]cross all grades in English and math, 27 of [the State’s] 30 most improved districts between 2002 and 2009, including the top 17, were in New York City.” Id. at 8. Comparing the City’s 32 districts to the 32 districts closest in size elsewhere in the State, and dividing each group into lowest, middle, and highest thirds by starting performance level, Margolis found that the City’s lower, middle, and highest-third districts beat their expected gains by, respectively, 8, 7, and 2 percentage points, while their statewide counterparts did so by only 2, 0, and .5 points. Id. at 8, 10.

A similar pattern characterized the school level. See KELLEHER, supra note 131, at 6 ("[B]y 2013, New York City was home to 22 of the 25 highest-performing schools [on state tests, compared to] none [in 2002, and] New York City’s share of the state’s lowest-performing schools shrank from 62 percent to 30 percent.").

\textsuperscript{166} Panel Presentation, supra note 144, at 5; see Liebman & Rockoff, supra note 139 ("[I]f the city’s gains were the result of easier statewide tests, we ought to have seen kids improve as much in Buffalo, Yonkers, and the rest of the state. We didn’t"); N.Y.S. Educ. Dep’t, 2013 Grades 3–8 English Language Arts and Mathematics, at 48, 152 (Aug. 6, 2013) (on file with authors).
3. Gap Closure on National Tests

Every two years since 2003, New York City, along with all fifty states and a subset of other large U.S. cities (many of which adopted similar reforms), have participated in the National Assessment of Educational Progress ("NAEP"). The NAEP's rigorous reading and math tests, administered to representative samples of fourth and eighth graders nationwide, permit reliable cross-jurisdiction comparisons. Figure 4 tracks average NAEP scores in fourth-grade math from 2003 (when New York City’s reforms began and it first took part in NAEP) to 2019, comparing New York City and State, all participating large city ("LC") districts nationally, and the national public-school aggregate.
During the reform’s first seven years through 2009, NAEP scores rose nationwide. But they rose faster in New York City, shrinking score gaps with New York State in all four grade-subject combinations, and with the nation in three of the four.\textsuperscript{169}

<table>
<thead>
<tr>
<th>Change in New York City’s NAEP Score Gaps, 2003–2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourth Grade</td>
</tr>
<tr>
<td>Math</td>
</tr>
<tr>
<td>Gap with Nation</td>
</tr>
<tr>
<td>Gap with New York State</td>
</tr>
</tbody>
</table>

New York City also increased its initially modest proficiency advantage over other large cities in the fourth-grade subjects.

In the reforms’ last four years, the City’s NAEP scores leveled off some, while national scores rose. The national increases were propelled in part by rapidly rising results in other large cities as they adopted reforms similar to New York City, often led by City transplants.\textsuperscript{170} During this

\textsuperscript{169} Between 2003 and 2009, New York City closed 64 percent of its gap with the nation in fourth grade math, 67 percent of its fourth grade reading gap, and 11 percent of its eighth grade math gap, with a 10 percent increase in its eighth grade reading gap. In the same period, the New York City/State gap closed 58 percent in fourth grade math, 30 percent in fourth grade reading, 33 percent in eighth grade math, and 8 percent in eighth grade reading. See Liebman & Rockoff, supra note 139 (“English and math gains by [New York City] 4th graders [from 2002 to 2009]—most of whom are poor, black, and Hispanic—are equivalent to a third of the achievement gap between minority and white students that has persisted for decades in this country.”)

period, however, the City gained ground on the State in the previously lagging eighth-grade subjects as students who benefited from earlier reforms reached the higher grade. 171

<table>
<thead>
<tr>
<th>Change in New York City’s NAEP Score Gaps, 2003–2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourth Grade</td>
</tr>
<tr>
<td>Math</td>
</tr>
<tr>
<td>Gap with Nation</td>
</tr>
<tr>
<td>Gap with New York State</td>
</tr>
</tbody>
</table>

4. Impact on Racial Disparities

In 2013, tracking similar changes in other cities, New York City voters elected Bill de Blasio mayor on a platform of ending the Bloomberg-Klein school reforms.172 Why, if the reforms were as successful as we have suggested, were they ended? Several aspects of the reforms—closing neighborhood schools, opening charter schools, using standardized-test scores to evaluate schools and teachers—were unpopular among teachers and many, especially middle-class, voters.173 Aggravating matters, the reformers treated policy experimentation as a technocratic exercise that engaged teachers, students, parents, and communities only as the instruments and subjects of the experiments rather than as partners in the innovations’ development and evaluation—with particularly calamitous results in the case of school closures.174 Overlaying these critiques, however, was another that broadened their appeal: the systemic reforms had

reforms “beg[un in New York City”); HILL & CAMPBELL, supra note 128, at 1; WONG & SHEN, supra note 167, at 44 (noting 2003–2011 NAEP gains across mayoral-control districts). The surge in other cities’ NAEP scores relative to New York City’s was most notable in 2011. Compare Kemple, Student Outcomes, supra note 159, 271–74 (finding higher rates of NAEP gains in New York City than in other large cities and nationwide from 2003 to 2009), with Pallas, Policy Directions, supra note 141, at 3 (noting New York City students’ “substantial increase” in NAEP 4th and 8th reading and math performance between 2003 and 2011, and that by 2011, other cities were “catching up”). Cf. Hill, supra note 131 (“NAEP gains [may be] more meaningful in New York than in other cities, because New York experiences more population growth among disadvantaged students than any other U.S. city and therefore must constantly make up a score deficit.”).

172 See Lieberman, Cruikshank & Ma, supra note 128, at 440–41 (citing sources).
173 Id. at 368–71, 438–41.
174 Id. at 371–72, 436–41.
not solely targeted and did not fully close, the systems’ racial and other outcome disparities.\textsuperscript{175}

The critics were right that substantial racial and other disparities remained as of 2013.\textsuperscript{176} What they missed, however, is the extent to which system-wide experimentation benefited children of color and lessened previously grotesque disparities in academic achievement. Consider that two-thirds of New York State’s Black and Latinx students benefited from the City’s substantial reduction of its NAEP and graduation gaps with the nation, and its remarkable erasure of its achievement gap with the rest of the State, compared to only 15 percent of the State’s White students.\textsuperscript{177} Further, in most cases, Black and Latinx students’ achievement rose more steeply than that of White students.\textsuperscript{178}

\textsuperscript{175} See, e.g., NORM FRUCHTER ET AL., ANNENBERG INST. FOR SCH. REFORM, IS DEMOGRAPHY STILL DESTINY?, at 1 (2012), http://www.aqeny.org/wp-content/uploads/2018/04/2012_10_24-1s-demography-still-destiny.pdf [https://perma.cc/MZQ3-SR5V] (criticizing New York City reforms because “racial composition and average income of a student’s home neighborhood are very strong predictors of a student’s chance of graduating high school ready for college, [and] gaps between neighborhoods are enormous”); Javier Hernandez, In Mayoral Race, Looking for Substance in Schools Conversation, N.Y. TIMES, Aug. 8, 2013, at A13 (quoting claims by candidate to replace Mayor Bloomberg that his education reforms had “dismal” effects on students of color); Luz Yadira Herrera & Pedro Noguera, Children First and its Impact on Latino Students in New York City, INMOTION MAG., Sept. 7, 2013, https://inmotionmagazine.com/er13/pn_nylatinos.html [https://perma.cc/EJK8-PG2T] (“[D]espite the bold assertions by Mayor Bloomberg . . . , there have been no significant reduction in the achievement gap separating New York City’s African American and Latino students from White students . . . ”); Hill, supra note 131 (“Critics point out the remaining deficiencies of . . . schools in poor and minority areas [as] many children are still stuck in weak schools.”); KEMPLE, N.Y.C. HIGH SCHOOLS, supra note 139, at 19 (“[A]mong young men who entered high school in 2007, more than 70 percent of Asian and white youth graduated with a Regents Diploma, compared to 48 percent of black and Hispanic young men.”); Ramos, supra note 133 (noting claims of only “modest” declines in achievement gaps under Bloomberg); Shapiro, supra note 141 (“[C]ritics . . . have accused [Bloomberg] of overseeing a vastly unequal system, with minority students losing out on the resources their white and Asian counterparts have . . . ”); ELAINE WEISS & DON LONG, BROADER BOLDER APPROACH TO EDUC., MARKET-ORIENTED EDUCATION REFORMS’ Rhetoric Trumps Reality, at 5 (Apr. 2013) https://files.epi.org/2013/bba-rhetoric-trumps-reality.pdf [https://perma.cc/KJM3-K4P6] (“[R]eform gains slightly narrowed the black-white gap in the city, but not enough to counter growing gaps in other subjects and grades.”).

\textsuperscript{176} See, e.g., KEMPLE, N.Y.C. HIGH SCHOOLS, supra note 139, at 19 (quoted supra note 175).

\textsuperscript{177} See supra note 175 and accompanying text; N.Y. State Public School Enrollment (2017–2018), supra note 155.

\textsuperscript{178} Graduation: N.Y.C. DEPT’OF EDUC., PRELIMINARY NEW YORK CITY GRADUATION RATES CLASS OF 2013 (Dec. 4, 2013), at 2, 5 (reporting increases in four-year graduation rates from first reform cohort (2005-graduates) to last (2013 graduates): all students, +42 percent; Asian students, +22 percent; Black students, +53 percent; Black males, +75 percent; Latinx students, +58 percent; Latinx males, +76 percent; students with disabilities, +119 percent (from 17 percent to 37 percent graduating); by 2013, the 2003 Black-White and Latinx-White gaps had narrowed 11 and 6
percentage points, respectively); see Kemple, N.Y.C. HIGH SCHOOLS, supra note 139, at 1–4, 15–16, 18 (finding that from 2005 to 2011, graduation rates improved at a moderately faster-than-average pace for Black and Latinx students and improved more for male students of color, economically disadvantaged students, English learners, and students with special needs; “largest improvements” were for students entering high school with “lowest relative attendance and achievement test scores from middle school”; +176 percent for students who were chronic absentees in the 8th grade, compared to +39 percent for 8th graders with 95 percent or greater attendance; and +400 percent for students with the lowest 8th grade ELA and math scores compared to +125 percent for middle-range students); see also Bloom & Untermann, supra note 150, at 1–3, 5, 10–11 (enrolling in Bloomberg-era new small high school increased four-year graduation rates by 9.5 percentage points for all students (70.4 percent compared to 60.9 percent), with largest gains by Black males (13.5 percent more likely to graduate) and Hispanic females (10.4 percent more likely).

College readiness: see Kelleher, supra note 131, at 6 (finding that AP pass rate for Black students rose 49 percent between 2002 and 2011, while overall pass rates were static; City’s 14 percent increase in SAT takers “was driven by more African American and Latino students taking the test”); Black & Coca, supra note 147, at ES-iii (“The largest percent increases in both high school graduation and college enrollment have been among . . . students in the poorest neighborhoods, Black and Latino students, and young men.”); Panel Presentation, supra note 144, at 20 (reporting 2002 to 2009 increases in city graduates’ enrollment in CUNY: Asian students, +88 percent; Black students, +44 percent; Latinx students +83 percent; White students +16 percent).

State achievement tests: Scale-score comparisons by race are available only before and after tests changed. From 2006 to 2010, Black and Latinx average math scores (all grades) rose 25 and 26 points, compared to 16 and 19 points for Asian and White students; in English, Black and Latinx scores rose 14 points, compared to 3 and 5 for Asian and White students—with declines in score gaps (all tested grades) of: 11.1 points (37 percent) Black-White English; 10.5 points (26 percent) Latinx-White English; 5.6 points (18 percent) Black-White math; and 6.5 points (22 percent) Latinx math. Panel Presentation, supra note 144, at 9–10. Between 2013 and 2014, the first comparison available under the new test, Black and Latinx gains in English (8.1 percent and 8.7 percent, respectively) outpaced Asian and White gains (6.4 percent and 5.9 percent), as did Black and Latinx gains in math proficiency levels (14.9 percent and 21.5 percent) versus Asian and White gains (11.8 percent, 11.1 percent)—albeit against much lower starting points for Black and Latinx than for Asian and White students. ENGAGENY, MEASURING STUDENT PROGRESS IN GRADES 3–8 ENG. LANG. ARTS & MATH, at 22, 34 (Aug. 2014), http://www.p12.nysed.gov/irs/ela-math/2014/2014Grades3-8ELAMath-final8-13-14.pdf [https://perma.cc/BT66-K9N5]: see Hill, supra note 131 (summarizing research on reforms showing that “New York City’s black and Hispanic students were 10 percentage points more likely, while white and Asian students were 5 points more likely, to be in above-average schools [in 2013] than in 2006”); Kemple, Student Outcomes, supra note 159 (showing that reforms had stronger beneficial effects on students with lower starting scores); see also CTR. ON EDUC. OUTCOMES, supra note 163, at 42–44 (finding that New York charter school students who gained most compared to matched students attending traditional public schools from 2012 to 2016 were Black students in poverty (receiving equivalent of “an additional 38 days of learning in reading and 74 days of learning in math”) and “Hispanic charter students in poverty” (receiving “an additional 68 days of learning in math”).

NAEP outcomes: Between 2003 and 2013, average fourth grade and eighth grade math and reading scores of the City’s Asian, Black, Latinx, and White students increased in all grade/subject combinations for all populations with the exception of White students’ fourth grade reading scores, which declined a point. Asian, Black, and Latinx scores rose more than White scores in three-quarters of the grade/subject combinations. NAEP Data Explorer, NAT'L CTR.
Conversely, ending system-wide experimentation in favor of (since-abandoned) resource infusions targeting only the City’s lowest-performing schools has widened most gaps and slowed the shrinkage of others. Figure 5 compares the first six years of the reforms, when the disparity between the City’s mainly Black and Hispanic student body and public school students nationwide on fourth-grade NAEP scores shrunk by almost two-thirds, with the gap’s 80 percent increase in the six years since the reforms ended. Similarly, Figure 6 illustrates that in the reform’s first six years, the City’s Latinx, English Learner, and disabled students closed a quarter to over a third of their gaps with White students, only to lose that ground in the six years since the reforms ended. Students in all three categories scored the same as they did seventeen years ago, while the scores of White students and students nationally have risen. The story is worse for Black students, whose modest gains on White students during the reforms have given way to losses—meaning Black students’ scores are now well below their 2003 averages. Likewise, after steadily shrinking during the reforms, Black-White and Hispanic-White gaps on state achievement tests have widened since the reforms ended, as Figure 7 illustrates. The Black-White and Hispanic-White graduation gaps have continued to decline since the reforms ended, but the pace of improvement has declined. As


179 Early in his tenure, Mayor de Blasio ended the systems’ “portfolio” policy of tracking learning growth at all schools and replacing weaker schools with newer stronger ones and chose instead to pour resources into a small number of the City’s lowest performing schools. Three years later, he abandoned the plan as a failure and began closing the schools. See Eliza Shapiro, $773 Million Later, de Blasio Ends Signature Initiative to Improve Failing Schools, N.Y. TIMES, Feb. 26, 2019, at A20 (“Mayor Bill de Blasio is canceling one of his signature education initiatives, acknowledging that despite spending $773 million he was unable to turn around many long-struggling public schools in three years after decades of previous interventions had also failed.”).

180 NAEP Data Explorer, supra note 178. The data underlying Figures 5–7 are from id. and are further displayed in Tables 2 and 3 in the Appendix.

181 Id.

182 Id.

183 Id.

184 N.Y.C. Dep’t of Educ., Test Results, ACADEMICS, https://infohub.nyced.org/reports/academics/test-results [https://perma.cc/4KCB-7N2D]; Panel Presentation, supra note 144, at 10 and accompanying text. Figure 7 covers 2004 to 2010 and 2013 to 2019. Changes in tests and how their scores were scaled between those two periods confound longitudinal comparisons. See supra notes 153–158 and accompanying text.

185 See Rsch. All. for N.Y.C. Schools, How Have NYC’s High School Graduation and College Enrollment Rates Changed Over Time, SPOTLIGHT ON N.Y.C. SCHOOLS (June 28, 2019),
tempting as it was for the City to drop middle class schools out of the reforms and focus only on a small number of failing schools attended mainly by poor students of color, the results have been harmful to them.

**Figure 5**

![Percent Change in NYC Score Gaps with US Public Students: NAEP 4th Grade Math](https://research.steinhardt.nyu.edu/site/research_alliance/2019/06/28/how-have-nycs-high-school-graduation-and-college-enrollment-rates-changed-over-time)

**Figure 6**

![Percent Change in NYC Score Gaps with NYC White Students: NAEP 4th Grade Math](https://perma.cc/TC4L-6PHL)

https://research.steinhardt.nyu.edu/site/research_alliance/2019/06/28/how-have-nycs-high-school-graduation-and-college-enrollment-rates-changed-over-time

<table>
<thead>
<tr>
<th>Change in N.Y.C. Graduation Gap with White Students</th>
<th>Percentage Point Change</th>
<th>(Percent Change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004–2010</td>
<td>-8 [-1.33] (-33%)</td>
<td>-8 [-1.33] (-30%)</td>
</tr>
<tr>
<td>2004–2013</td>
<td>-9 [-1.00] (-36%)</td>
<td>-10 [-1.11] (-37%)</td>
</tr>
<tr>
<td>2013-2019</td>
<td>-3 [-0.60] (-20%)</td>
<td>-3 [-0.60] (-18%)</td>
</tr>
</tbody>
</table>
In 1973, in refusing to recognize a constitutional right to education, the Supreme Court expressed uncertainty whether courts fairly could expect educators to close Black and Latinx achievement gaps with White students. In roughly the first fifteen years of this century, countless classroom, school, and district level experiments by New York City and other large districts disproved that suggestion and validated the counter-suspicion the Court said it had no way to test: that officials perpetuate racial disparities based on the false ascription of racial limitations. There is, however, a way to test these competing suspicions by interpreting the Equal Protection Clause to require steps to plumb the uncertainty, exposing racial disparities supportable only based on false assumptions, benefiting Black communities more than targeted steps, and aiding the broader public as well.

D. SCHOOL DESSEGREGATION 1970–1985

Observers often characterize the educational achievement gap as “persistent” and “fail[ing] to close,” but as Figure 8 illustrates, this is not

---


true. Twice over the past fifty years, the Black/White and Latinx/White NAEP gaps shrunk substantially in nearly all tested grades and subjects: from 1975 to 1985, and again from 2003 to 2011. As a result, “black and Hispanic students today are roughly three years ahead of their parents’ generation in math” and “roughly two to three years ahead” in reading. “White students’ scores have also improved, but not by as much.”

Figure 8

[Diagram showing racial and ethnic achievement gaps on NAEP, 1975–2012]

Stanford Ctr. for Educ. Pol’y Analysis, supra note 188, at fig.1.

White Achievement Gap, Phi Delta Kappan, Oct. 2004, at 105, 106 (arguing that “persistent achievement gaps” show that school policies cannot solve problems stemming from broad social forces).

188 Stanford Ctr. for Educ. Pol’y Analysis, Racial and Ethnic Achievement Gaps, EDUC. OPPORTUNITY MONITORING PROJECT, https://cepa.stanford.edu/educational-opportunity-monitoring-project/achievement-gaps/race [https://perma.cc/UF5X-8PQ9] (observing, based on data underlying Figure 8, that “white-black and white-Hispanic achievement gaps have been declining, albeit unsteadily,” “narrow[ing] sharply in the 1970s and first half of the 1980s,” than “stall[ing]” and even widening until the new century when they declined again through 2012). We adapted Figure 8 from id. by adding the two labeled boxes.

189 Id.
Causal explanations for national variations in gap change are hard to prove, but meaningful gap closure in “experimentalist” urban school districts in the first fifteen years of the century suggests an explanation for the later period of declines. The earlier declines also coincided with a national experiment in education reform: racial integration of vast swaths of the nation’s schools, which began in earnest only in 1970, then began collapsing in the mid-1980s due to retrenchment by the Supreme Court and successive Republican administrations. Considerable social scientific evidence links 1970s-to-1980s racial integration to enhanced academic and life outcomes for Black, Latinx, and economically disadvantaged students. More recently, Economist Rucker C. Johnson has used the “natural policy lab” provided by longitudinal data on the life experiences of Black and White adults who took part in “significant efforts to integrate schools that occurred only for about 15 years, and peaked in 1988” to draw a causal link between that integration and substantial improvements in Black Americans’ “achievement gaps, educational attainment, earnings, and health status.”

---


191 As Figures 4 (above), and 9–11 (Appendix) reveal, national, state, and local NAEP scores and gap closure tailed off after 2013, when many cities abandoned their education reforms. See Liebman, Cruikshank & Ma, supra note 128, at 440–41.

192 See infra notes 203–206 and accompanying text.

193 Asked why “during the height of desegregation in the 1970s and 80s, the [NAEP] achievement gap between black and white students narrowed” more rapidly than ever before or since, Stanford sociologist Sean Reardon, Figure 8’s creator, said “‘[t]here was a lot going on [in the 1970s and 1980s]. But clearly desegregation improved outcomes for blacks, and didn’t harm them for whites.’” Sarah Garland, Was ‘Brown v. Board’ a Failure, THE ATLANTIC, Dec. 8, 2012, https://www.theatlantic.com/national/archive/2012/12/was-brown-v-board-a-failure/265939 [https://perma.cc/LVV-9P8G]; see The Benefits of Socioeconomically and Racially Integrated Schools and Classrooms, CENTURY FOUND. (Apr. 29, 2019) [hereinafter CENTURY FOUND.], https://tcf.org/content/facts/the-benefits-of-socioeconomically-and-racially-integrated-schools-and-classrooms [https://perma.cc/HHM8-UDM5] (“[T]he racial achievement gap in K–12 education closed more rapidly during the peak years of school desegregation in the 1970s and 1980s than it has overall in the decades that followed—when many desegregation policies were dismantled.”).

strong school funding [integration generated], the better their [gap-closing] outcomes in adulthood." \(^{195}\)

Although the school desegregation era formally began with *Brown v. Board of Education* in 1954, \(^{196}\) the Supreme Court’s “all deliberate speed” \(^{197}\) temporizing permitted massive then passive resistance across the South, delaying the onset of systematic racial integration of schools there until the 1969–70 school year. \(^{198}\) Doubts about how *Brown* applied in the North and West delayed the start of sustained court-ordered desegregation

historical “accident” that some Black and White children in K–12 schools during the 1960s to 1980s experienced, while others did not experience, court-ordered desegregation and did so for different amounts of time; (2) compares the educational, economic, and health outcomes through 2013 of those children as adults, as reported on a long-running longitudinal household survey of 5000 households; (3) controls for other potentially causal forces such as Title I funding, Head Start, introduction of kindergarten, AFDC, Medicaid, Food Stamps, Community Health Centers, and hospital desegregation; and (4) concludes that prior, less sophisticated studies typically found, but substantially understated, the positive effects of integration on Black students. See also Wills 7 Hawley & Mark A. Smylie, *The Contribution of School Desegregation to Academic Achievement and Racial Integration*, in Eliminating Racism: Profiles in Controversy 281, 289 (P. Katz & D. Taylor eds., 1988) (“[E]ven though both massive and passive resistance have been more common than genuine efforts to make it work, school desegregation has benefited most of those who have experienced it”); AMY STUART WELLS ET AL., CENTURY FOUND., *How Racially Diverse Schools and Classrooms Can Benefit All Students*, at 11-17 (Feb. 9, 2016), https://tcf.org/content/report/how-racially-diverse-schools-and-classrooms-can-benefit-all-students [https://perma.cc/87M8-LEWH] (noting that researchers were just starting to understand how to implement racial integration with the best academic and social-emotional results when desegregation efforts were reversed in the 1980s, suggesting that its benefits would have increased had integration continued).

195 Johnson, *Why School Integration Works*, supra note 194: see Johnson, *Long-run Impacts*, supra note 194, at 2 (“[F]or blacks, school desegregation significantly increased both educational and occupational attainments, college quality and adult earnings, reduced the probability of incarceration, and improved adult health status [with] no [ill] effects on whites[, and those] adult attainments increased significantly with . . . the duration of desegregation exposure.”); see also id at 3, 25 (“[B]lack children who were exposed to court-ordered school desegregation for the majority of their school-age years experienced significantly improved education, economic, and health outcomes in adulthood, compared with their older siblings who grew up in segregated school environments . . .”).


in those parts of the country until 1973.\textsuperscript{199} In the early 1970s, however, school desegregation exploded nationwide.\textsuperscript{200} As is documented by a federal study of 125 randomly selected school districts nationwide, by 1987, federal court or administrative desegregation orders were in place in an astonishing 87 percent of school districts, and the proportion of Black students attending essentially all-minority schools had fallen from 63 to 30 percent since 1970, while the proportion of Black students in schools that were 25 to 75 percent White had grown from 17 to 44 percent.\textsuperscript{201} As it happened, the 1987 report coincided with racial integration’s apogee. Thereafter, a series of mid-1970s,\textsuperscript{202} then early 1990s,\textsuperscript{203} Supreme Court decisions limiting the availability and reach of desegregation orders, coupled with the withdrawal of federal enforcement by the Reagan and Bush administrations from 1981 to 1992, spelled the end of the social experiment.\textsuperscript{204} As Orfield and Yun report, “[t]he percent of black students

\textsuperscript{199} Keyes v. Sch. Dist. No. 1, 413 U.S. 189, 214 (1973) (approving “all-out” school desegregation order in city where segregation had not previously been statutorily required); Columbus Bd. of Educ. v. Penick, 443 U.S. 449 (1979) (similar).

\textsuperscript{200} See Liebman, Desegregating Politics, supra note 198, at 1465–70.

\textsuperscript{201} See Finis Welch & Audrey Light, U.S. Comm’n on Civ. Rts., Ed-293-936, New Evidence on School Desegregation, at 4–8, 16–21, 40–41, 66–67 (1987); Liebman, Desegregating Politics, supra note 198, at 1466–67 nn.6, 10–11 (citing similar studies and noting criticism of the 1987 federal study for under-sampling southern rural districts, where integration rates were especially high); see also Johnson, Long-run Impacts, supra note 194, at 7 (finding that 88 percent of the Black respondents in a large national representative survey of individuals born between 1945 and 1968 that the survey “followed into adulthood grew up in a school district that was subject to a desegregation court order sometime between 1954 and 1990”).

\textsuperscript{202} See Dayton Bd. of Educ. v. Brinkman, 433 U.S. 406, 417 (1977) (limiting desegregation remedies to school segregation causally linked to particular, identifiable acts of intentional segregation); Miliken v. Bradley, 418 U.S. 717, 733, 745, 753 (1974) (barring inclusion of all-White suburbs in desegregation orders remedying statewide and local segregation of urban schools, even if necessary to provide a modicum of racial integration, if suburban districts did not themselves segregate schools).

\textsuperscript{203} Early 1990s decisions making it progressively easier for school districts under desegregation orders to be declared “unitary” and return to segregative neighborhood assignment policies include Missouri v. Jenkins, 515 U.S. 70 (1995); Freeman v. Pitts, 503 U.S. 467 (1992); Bd. of Educ. v. Dowell, 498 U.S. 237 (1991); see Sean F. Reardon et al., Brown Fades: The End of Court-Ordered School Desegregation and the Resegregation of American Public Schools, 31 J. Pol’y Analysis & Mgmt. 876, 899–901 (2012) (finding increased school segregation in over 200 districts released from court-ordered desegregation pursuant to the cited Supreme Court decisions).

in majority white schools peaked in the early 1980s and declined to the levels of the 1960s by the 1996-97 school year.  

The evidence linking racial and economic integration to better academic and other life outcomes for Black, Latinx, and economically disadvantaged students is robust and multifaceted. Rigorously studied input and outcome effects of K-12 racial integration on Black students and of economic integration on low-income students, with no adverse effects on White students, include:

- Higher per pupil funding, lower student-teacher ratios, better teachers, facilities, and curricula, higher academic expectations for students by educators.

---


206 See Brief for 553 Social Scientists as Amicus Curiae Supporting Respondents at App.1–28, Parents Involved v. Seattle Sch. Dist. 551 U.S. 701 (2007) (No. 05-908) [hereinafter Social Scientists' Brief] (comprehensively cataloguing studies); CENTURY FOUND., supra note 193 (reviewing later studies); Johnson, Long-run Impacts, supra note 194, at 1, 31–37 (reviewing recent studies); Liebman, Desegregating Politics, supra note 198, at 1486, 1621, 1624–30 (reviewing research as of 1990); WELLS ET AL., supra note 194 (reviewing research). See Liebman, Desegregating Politics, supra note 198, at 1621 ("[N]o respectable study has ever claimed that desegregation . . . decreases the academic accomplishments of white children.").

207 See Johnson, Long-run Impacts, supra note 194, at 2, 16–18 ("[A]mong blacks with 12 years of exposure to school desegregation, average school-age per-pupil spending was 22.5 percent greater . . . than that experienced among unexposed black cohorts[, with the] large increase . . . driven solely by the infusion of state funds following the timing of court-ordered school desegregation.").

208 See id. at 2, 17–18 (finding that class-size for Black students fell sharply following desegregation orders, with average reductions of three to four students per class two years into the process and no significant effects on class size for White students).

209 See CENTURY FOUND., supra note 193, at 3 n.19 (citing studies); Social Scientists' Brief, supra note 206, at App. 31–33 & nn.101–06 (documenting that schools with concentrations of Black and Latinx students tend to have higher-than-average teacher turnover; teachers who are less qualified, experienced, and effective; and less "viable learning communities"); WELLS ET AL., supra note 194, at 12 & n.46, 14 n.54 ("[I]ntegrating schools leads to more equitable access to important resources such as structural facilities, highly qualified teachers, challenging courses, private and public funding, and social and cultural capital.") (citing sources).
and students themselves, increased parental involvement, and lower suspension rates; Enhanced academic motivation, creativity, critical thinking, and problem-solving skill (also true of White students in integrated settings); Higher IQ scores; Higher academic achievement; Lower probability of teenage pregnancy; Higher graduation rates;

211 See Liebman, Desegregating Politics, supra note 198, at 1625 n.675 (citing studies); WELLS ET AL., supra note 194 (citing studies, footnotes 48–49).
212 See Social Scientists’ Brief, supra note 206, at App.27–28 & nn.88–91 (citing studies linking higher levels of racial integration of schools to higher levels of parental involvement by families of Black students, including those attending school outside their neighborhood).
213 See WELLS ET AL., supra note 194, at 15 nn.52–53 (citing studies).
214 See Social Scientists’ Brief, supra note 206, at App.12–13 & nn.34–37 (“[S]tudents experiencing classroom diversity—specifically racial and ethnic diversity—showed the greatest engagement in active thinking processes, growth in intellectual engagement and motivation, and growth in intellectual and academic skills.”).
215 See Liebman, Desegregating Politics, supra note 198, at 1625 n.675 (citing authority).
216 See id. at 1624–25 & n.675; Reardon & Kalogrides, supra note 53, at 1203–09 (analyzing 200 million standardized math and reading tests in thousands of school districts from 2009 to 2013, finds that school segregation is associated with wider achievement gaps and explains more of the variation in gaps than any other policy correlate); Social Scientists’ Brief, supra note 206, at App.13–17, 37–38 & nn.38–49, 120–23 (reviewing (1) earlier studies revealing consensus that “school desegregation has a modest positive impact on the achievement of African Americans” in contrast to “[m]ost school reforms [which had] little or no effect on improving students’ outcomes” and (2) later studies with superior methodology associating racial segregation with lower achievement by Black students, which, if cumulated over years of integration could “substantially” increase Black achievement); WELLS ET AL., supra note 194, at 12 & nn.41–44 (“Attending racially diverse schools . . . is associated with smaller test score gaps between students of different racial backgrounds [as a result of] black and/or Hispanic student achievement increase[s].”); see also CENTURY FOUND., supra note 193, at 1 nn.1–2, 2 nn.7–9 (citing studies finding higher NAEP and other achievement levels and smaller achievement gaps for low-income students attending economically integrated schools than matched students attending “schools with concentrated poverty”).
217 See Liebman, Desegregating Politics, supra note 198, at 1625–26 & n.676 (citing studies).
218 Professor Johnson makes multiple findings of relevance here:

• “For Black students [whose graduation gap with white students, absent desegregation, was 15 percentage points], there is an immediate jump in the likelihood of graduating from high school with exposure to court-ordered desegregation, and each additional year of exposure leads to a 1.8 percentage-point increase in the likelihood of high school graduation with an additional jump for those exposed throughout their school-age years.”
• “[E]ffect size is comparable to the influence of having college-educated parents.”
• There was no significant change in White graduation rates in same schools, ruling out the influence of trends other than desegregation on rising Black attainment.

Johnson, Long-run Impacts, supra note 194, at 18–19; see Social Scientists’ Brief, supra note 206, at App.20 & n.59 (“Studies find that attending desegregated schools is related to [higher likelihood of] completing high school for nonwhite students.”); see also CENTURY
More years of completed secondary and higher education;\footnote{219}

- Expanded college access, attendance, and success;\footnote{220}
- Higher quality of college attended;\footnote{221}
- Higher adult earnings, occupational status, and family economic status;\footnote{222}

\footnote{219} See Johnson, *Long-run Impacts*, supra note 194, at 18–19 ("For black children, exposure to court-ordered desegregation in all 12 school-age years increases [K–12] educational attainment by roughly one year [which is] large enough to eliminate the black-white educational attainment gap on this measure.").

\footnote{220} See CENTURY FOUND., supra note 193, at 1 & n.3 ("[S]tudents with similar socioeconomic backgrounds . . . at more affluent schools are 68 percent more likely to enroll at a four-year college than their peers at high-poverty schools."); Liebman, *Desegregating Politics*, supra note 198, at 1626 & nn.677–82 ("[B]lack students who graduate from desegregated schools (controlling for [relevant] criteria) attend college, attend four-year colleges, get high marks from their professors, graduate from college, and enter predominantly white employment settings in greater numbers than do [those] who attended segregated public schools."); Social Scientists' Brief, supra note 206, at App.20 & n.60, App.21–22 & nn.61, 65–69 (summarizing the results of two bodies of relevant research: (1) "Nonwhite students who graduate from integrated schools are more likely to have access to the social and professional networks that . . . increase[] access to professional jobs and information about college-going opportunities [and] to complete more years of education, earn higher degrees, and major in more varied occupations than graduates of all-black schools."); (2) "Students who experienced integration prior to attending an undergraduate institution are more likely to connect positively with diverse peers in college and to exploit the academic opportunities that a diverse university can offer."); WELLS ET AL., supra note 194, at 12 & n.47 ("The gap in SAT scores between black and white students is larger in segregated districts; change from complete segregation to complete integration in a district would reduce as much as one quarter of the SAT score disparity").

\footnote{221} See Johnson, *Long-run Impacts*, supra note 194, at 19–20 (finding that exposure to desegregation has "large estimated effects"—increasing as years of desegregation exposure increase—on the quality of colleges Black students attend, as defined by average SAT scores of entering freshmen).

\footnote{222} Professor Johnson makes multiple findings of relevance here:

- "Adult outcomes for blacks generally improved monotonically with the number of school-age years of exposure to desegregation."
- "[A]verage effects of a 5-year exposure to court-ordered school desegregation [include] about a 15 percent increase in wages and an increase in annual work hours of roughly 165, which combined to result in a 30 percent increase in annual earnings."
- Exposure to integration is associated with a 5.2 point increase on the occupational prestige index (closing 17 percent of the Black-White job prestige gap), and "about [a] 25 percent increase in annual family income."

*Id.* at 20–22; see Social Scientists' Brief, supra note 206, at App.22–23 & nn.66–69 (discussing rigorous studies linking attendance at desegregated schools with higher wages and income and probability of a white-collar job as an adult).
• Lower poverty rates;\textsuperscript{223}
• Lower incarceration rates;\textsuperscript{224}
• Better health as an adult and longer life expectancy;\textsuperscript{225}
• Increased likelihood of working, residing, and educating one's children in integrated settings (also true of White adults formerly attending integrated schools or living in neighborhoods with racially integrated schools);\textsuperscript{226} and
• More favorable views of and more positive relationships with members of other races (also true of White adults formerly attending integrated schools).\textsuperscript{227}

\textsuperscript{223} See Johnson, \textit{Long-run Impacts}, supra note 194, at 22 (associating 5-year exposure to integrated schools with "11 percentage-point decline in the annual incidence of poverty in adulthood").
\textsuperscript{224} See \textit{id.} at 22–24 (attending desegregated schools diminished probability of adult incarceration by 10, 15, and 22 percentage points for Black students in desegregated schools starting in, respectively, high, middle, and elementary school).
\textsuperscript{225} See \textit{id.} at 22–23 ("[F]or blacks, adult health status improves monotonically with duration of exposure to court-ordered desegregation[, with] a 5-year exposure ... yield[ing] an 11 percentage-point increase in the annual incidence of being in excellent/very good health"; "impact of each additional year of desegregation exposure [is] equivalent (on average) to blacks reaching a level of health deterioration about 1 year later than if that year were spent in segregated schools.").
\textsuperscript{226} See \textit{Century Found.}, supra note 193, at 2 n.12; \textit{Jennifer Hochschild, The New American Dilemma: Liberal Democracy and School Desegregation} 179–85 (1984) (finding that, as of 1984, strongest support for desegregation among white Americans was among southerners whose children participated in mandatory desegregation plans); Liebman, \textit{Desegregating Politics}, supra note 198, at 1622 & nn.670–71 (citing studies linking certain kinds of desegregation plans with increases in housing integration); Social Scientists' Brief, \textit{supra} note 206, at App.23, 25–27 & nn.70, 78–87 (citing numerous studies linking metropolitan-wide school desegregation plans to greater and more stable housing integration than in areas with little or no racial integration of schools and linking prior attendance at racially desegregated schools with increased disposition to live in racially integrated neighborhoods).
\textsuperscript{227} See \textit{Century Found.}, supra note 193, at 2 nn.11–12 (citing studies); Liebman, \textit{Desegregating Politics}, supra note 198, at 1621, 1626–27 & nn.667–68, 683–95 ("Black and white adults who previously attended desegregated elementary and secondary schools (after proper controls) seem to be less likely to express negative views about members of the other race[,] are significantly more comfortable in integrated work and social settings[, and are more likely to] live in integrated neighborhoods and [to] report having personal relationships with persons of the other race ... than blacks and whites who went to segregated schools."); Social Scientists' Brief, \textit{supra} note 207, at App.23–25 & nn.70–77 (citing studies showing that compared to Black and White individuals who attended racially segregated schools, those "who attended desegregated schools are more likely to have friends from other races, work in desegregated workplaces, live in desegregated neighborhoods, and favor desegregated schools for their own children, beyond the influence of [other relevant] characteristics"; feel "comfortable in racially diverse or predominantly white environments"; "bring fewer racial and ethnic stereotypes into the workplace"; and work productively with members of other races); \textit{Wells et al.}, \textit{supra} note
Given racial integration’s 1970s-to-1980s reversal of manifold racial resource and outcome disparities in public education that made “separate but equal” inevitably unequal, court-ordered racial desegregation counts as both a grandly successful and tragically curtailed national experiment in disparity reduction. Partially to blame for its demise, court-ordered desegregation often was an old-fashioned, bureaucratic experiment in top-down, one-size-fits-all policymaking, in which—making matters worse—the “top” consisted not of elected officials or expert administrators but of appointed generalist judges lacking democratic legitimacy and policy competence. One-size remedies worked well in southern rural settings where the blatantly irrational and inequitable operation of dual school systems easily gave way to unification once the Supreme Court finally ordered desegregation “now.” Desegregation faltered, however, in southern cities and across the North and West, when courts encountered disparities complicated by the difficulty of disentangling the effects of public (legally exposed) and private (legally insulated) decisions about where families live and school their children. And debates ensued—admittedly beyond judges’ ken—over the educational efficacy and equity of policies like neighborhood schools, busing, academic tracking, teacher certification, and didactic versus flexible curricula. Uniform edicts issued from on high by non-expert judges turned out to be a poor way to conduct context-sensitive experiments in input and outcome.

194, at 15 & nn. 65–67 (citing studies). Qualitative reflections by Black and White adults on the experience of attending desegregating schools and its effects on them as adults confirm the last two listed benefits of integration, see SUSAN EATON, THE OTHER BOSTON STORY (2001) and AMY STUART WELLS, BOTH SIDES NOW (2009).


231 See supra notes 196–200 and accompanying text.


233 See Liebman & Sabel, supra note 230, at 195–201 (attributing nation’s retreat from Brown in part to inability of bureaucratic judicial remedies to provide the legitimacy, policy expertise, and flexibility needed to address local variations in racial segregation and its causes); Liebman, Perpetual Evolution, supra note 31, at 2013–24 (discussing failings of old-style—and calling for new, more flexible, context-sensitive forms of—public-law litigation).
disparity reduction across a myriad of local contexts, prompting severe backlash against court-ordered remedies.\footnote{234}{See Liebman, \textit{Perpetual Evolution}, supra note 31, at 2016–17 \& n.60 (citing sources).}

In one important respect, however, desegregation orders anticipated the localized experimentation this Article proposes. A subset of metropolitan-wide desegregation orders “redistribute[d] children among schools and redefined districts not as [a judicially mandated] end in itself but instead to renew the processes through which local, county, and state officials and educators interacted to administer and design schools . . . .”\footnote{235}{Liebman & Sabel, supra note 230, at 201 n.76} These orders—in places like Charlotte-Mecklenburg County, North Carolina; Louisville-Jefferson County, Kentucky; Nashville-Davidson County, Tennessee; Tampa-St. Petersburg (Pinellas and Hillsborough Counties) in Florida; and Wilmington-New Castle County, Delaware—“energize[d] surprising and effective” interracial “coalitions of actors, both inside and outside the schools,” which “in turn, revitalized entire regional educational systems and even the cities they straddled.”\footnote{236}{Liebman & Sabel, supra note 230, at 201 n.76} Reconfiguring segregated metropolitan school districts into new sets of racially mixed districts prompted the reorganized units “under the guidance of newly interlocked local, regional, and state officials and a variety of actors from the private sector” to experiment. Reconstituted districts explored new ways “to reorganize the governance, administrative, and pedagogical structures of the newly reconfigured schools and districts” into more racially integrated polities, rather than simply shoehorning integrated classrooms into otherwise fundamentally segregated school systems and communities.\footnote{237}{Id.} The result was what Political Scientist Jennifer Hochschild calls an “educational renaissance,” characterized by a revived administrative bureaucracy; funding increases; and “organizational, curricular, and pedagogical improvements for all students.”\footnote{238}{Hochschild, supra note 226, at 80–82 \& nn.140–44.}

In part because many of the best desegregation plans aimed to benefit all students and the nation as a whole—diluting its focus on enhanced educational outcomes for Black children—influential Black critics advocated ending court-ordered integration in favor of creating
“model black schools.” This view culminated in a lawsuit by Black parents in Louisville-Jefferson County, Kentucky demanding an end to that metropolitan district’s voluntary race-based school-enrollment policy after a court declared the district free of the vestiges of legally mandated segregation. The Supreme Court decision the lawsuit triggered—the final nail in desegregation’s coffin—bars school districts not under order to remedy past discrimination from basing school assignments on race as a way to maximize integration.

Two especially powerful exponents of this view were Judge Robert Carter and Professor Derrick Bell, both former counsel for Black families in Brown and following cases. In Judge Carter’s formulation, “[w]hile we fashioned Brown on the theory that equal education and integrated education were one and the same, the goal was not integration but equal educational opportunity. . . . If that can be achieved without integration, Brown has been satisfied.”

Going further, Bell argued that white-


240 See generally SARAH GARLAND, DIVIDED WE FAIL: THE STORY OF AN AFRICAN-AMERICAN COMMUNITY THAT ENDED THE ERA OF SCHOOL DESEGREGATION (2013) (describing Black Louisville families’ opposition to voluntary integration because it did not assure equal outcomes or control over their children’s education).


242 Carter, supra note 239, at 27.
dominated courts had intentionally elided Brown's promise of improved educational opportunity for Black children into a requirement of racially integrated schools because the latter more fully converged with White interests. Bell criticized orders “balancing the student and teacher populations by race in each school, eliminating one-race schools, redrawing school attendance lines, and transporting students to achieve racial balance” for having only “debatable” educational benefits and for not “guarantee[ing] black children better schooling than they received in the pre-Brown era” or protecting them from “discriminatory policies... the loss of black faculty and administrators, suspensions and expulsions at much higher rates than white students, and varying forms of racial harassment.” Bell believed that generously funding model all-Black schools would more effectively eliminate those disparities than funneling resources into schools serving White students as well.

With the exception of charter and similarly empowered schools in cities that included them in their early 2000s portfolios of new educational models, there is no systematic evidence that predominantly Black and Latino schools have diminished racial disparities or improved Black educational and life outcomes at anything like the rates at which 1970-1985 racial integration did. Instead, there is tragically uniform evidence that the large and growing set of all-minority public schools that the demise of both fifteen-year experiments left in their wake are disastrous drags on the life chances of Black and Latino youth.

The connection between the two reforms has not gone unnoticed, but ironically it has been drawn by opponents of both who criticize both for forsaking Black communities in service of more broad-gauged improvements. At the tail end of the second reform period, for example,
Journalist Sarah Garland, author of a sympathetic account of the Black Louisville parents whose lawsuit brought desegregation down, said she was struck, as [she] listened to [Louisville parents’] criticisms of busing, at how similar their complaints were to the frustrations with the current set of education reforms: the charter schools and accountability systems that replaced desegregation. As the era of desegregation ended, black communities across the nation were once again facing unilateral school closings and mass firings of black teachers. Many felt disenfranchised, and wondered whether reformers cared about their own vision for their children’s education.248

We, too, have criticized the cavalier ways in which advocates and courts in the earlier period and reform technocrats in the later period conducted their experimentation—typified by the dismissal of Black teachers in the earlier case, and the closure of valued neighborhood schools by later reformers.249 Participatory voids in the experimental processes at work violated fundamental democratic norms and sapped the power of the experiments for lack of guidance from the Black and Brown communities that know best what their children need and how these systems currently fall short.250

But it is a tragic irony that both reforms were shut down just as practitioners had begun to fill the reforms’ democratic gaps,251 developing what by now is a robust body of knowledge about why broad participation and Black leadership matter and how to achieve them.252 Most importantly, this crucial and correctable weakness is not a function of failing to focus on


249 See Klein, supra note 131, at 197–212 (describing debates with the senior author of this Article over whether and how to engage communities in designing and implementing school reforms); Liebman, Cruikshank & Ma, supra note 128, at 438–43 (criticizing New York City reforms for leaving communities of color and teachers out of the experimentation process); Liebman, Perpetual Evolution, supra note 31, at 2103–18 (criticizing bureaucratic public-law education-reform litigation of the 1960s to 1980s).

250 See WELLS ET AL., supra note 194, at 11–15; Liebman, Cruikshank & Ma, supra note 128, at 441–43.

251 See Liebman, Cruikshank & Ma, supra note 128, at 441–43; WELLS ET AL., supra note 194, at 11–15.

252 See Liebman, Cruikshank & Ma, supra note 128, at 446–63.
disparities alone in favor of systemic experimentation—a strategy that, against all odds, has moderated disparities more than any others. On the contrary, disparities have given way most substantially when court orders place White and Black families in “ethical situations,” through which the coincident wellbeing of their children has motivated collaborative experimentation to improve the learning and life chances of both groups of children.253

IV. OBJECTIONS

This Section briefly responds to anticipated objections to our proposal, starting with the most obvious: the proposal is legally unprecedented and not something the current Supreme Court is likely to substitute for its longstanding prioritization of racial motivation over disparity in equal protection analysis.254 Our immediate audience, however, is not the Supreme Court but the burgeoning set of public and private organizations nationwide expressing readiness to address rampant, long-ignored racial disparities in their operations.255 In different ways, they too are obliged by the Fourteenth Amendment to pursue equal protection, to acknowledge the unfinished racial struggle animating the Amendment, and to recognize the proven power of evolving social practices and norms to compel the Supreme Court to abandon other tragically misbegotten interpretations of the Amendment.256

---

That said, our examples above illustrate the proposal’s many statutory precedents for requiring or incentivizing experimentation with disparity or other kinds of harm-reduction. Those precedents demonstrate our proposal’s consistency with Congress’s exercise of its responsibility to apply the Fourteenth Amendment and Delinquency Prevention Act,259 Prison Rape Elimination Act of 2003,260 American Recovery and Reinvestment Act of 2009,261 Americans with Disabilities Act of 1990,262 Individuals with Disabilities Education Act (IDEA),263 No Child Left Behind Act and superseding Every Student Succeeds Act,265 Title VII of the Civil Rights Act of 1964,266 Age


257 U.S. Const., amend. XIV, § 5.


260 Id. §§ 15605(a)–(c); see Charles F. Sabel & William H. Simon, Democratic Experimentalism, in SEARCHING FOR CONTEMPORARY LEGAL THOUGHT 477, 484–85 (Justin Desautels-Stein & Christopher Tomlins eds., 2017) (discussing Prison Rape Elimination Act of 2003); supra note 95 and accompanying text.


264 See, e.g., Liebman & Sabel, supra note 230, at 213–66 & n.214 (discussing effective experimentation in the shadow of NCLB by Kentucky, North Carolina, and Texas).

265 See, e.g., Redesigning High Schools to Support Their Communities in the 21st Century, CROSS STATE HIGH SCH. REDESIGN COLLABORATIVE, https://www.hsredesign.org [https://perma.cc/M7SM-AAJT] (describing joint steps by Illinois, Louisiana, Massachusetts, Mississippi, New Mexico, New York, Ohio, and Johns Hopkins University “to use The Every Student Succeeds Act (ESSA) to enable struggling high schools in high-needs communities to reinvent themselves”).

Discrimination in Employment Act of 1967,\textsuperscript{267} and Fair Housing Act of 1968.\textsuperscript{268}

In fact, the Supreme Court itself adopted an experimentalist interpretation of the Equal Protection Clause in \textit{Fisher v University of Texas}.\textsuperscript{269} There, the Court upheld the University of Texas's ("UT's") race-conscious college admissions policy precisely because UT had engaged in responsible policy experimentation to show both that the forms of educational diversity it was pursuing were of compelling importance, and that its race-conscious steps were as limited as they could be and still serve that goal.\textsuperscript{270} Analogous to our proposal, the Court's ruling made clear that, as long as UT continued to take constitutionally questionable action—in that case, factoring race into admissions—it had a duty to continue experimenting with ways to increase the benefits and limit the constitutional risks incurred.\textsuperscript{271} As we have summarized elsewhere, Justice Kennedy's majority opinion "provides a case study" of public actors' constitutionally sufficient experimentation with ways to minimize constitutional risks and of how a constitutional court can review such action to assure that experimentation was responsible.\textsuperscript{272}

notes 83–90 and accompanying text (providing an example of Title VII’s experimentalist application).\textsuperscript{267} Age Discrimination in Employment Act § 4(a), 29 U.S.C. § 623(a) (2018); see Smith v. City of Jackson, 544 U.S. 228, 235 (2005) (plurality opinion) (interpreting id.); Smith, 544 U.S. at 243–47 (Scalia, J., concurring in part and concurring in the judgment); Liebman & Mbikiwa, supra note 266 (discussing Age Discrimination in Employment Act).


\textsuperscript{269} See, e.g., Fisher v. Univ. of Tex., 136 S. Ct. 2198, 2209–10 (2016).

\textsuperscript{270} \textit{Id.} at 2205, 2209–13.

\textsuperscript{271} \textit{Id.} at 2210:

Through regular evaluation of data and consideration of student experience, the University must tailor its approach in light of changing circumstances, ensuring that race plays no greater role than is necessary to meet its compelling interest. . . . The type of data collected, and manner in which it is considered, will have a significant bearing on how the University must shape its admissions policy to satisfy strict scrutiny in years to come.

\textsuperscript{272} Liebman, \textit{Perpetual Evolution}, supra note 31, at 2033 (quoting \textit{id.} at 2205, 2209–13) (footnotes omitted):

As summarized by the Court, UT's process provides a case study in evolutionary learning. UT began by testing race-neutral admissions . . . . Only after "[n]one of these efforts succeeded" and "a reasonable determination was made that the University had not yet attained its goals" did UT embark on a "significant evolution" of a race-conscious policy. To learn iteratively what was possible, UT experimented with different types of
That said, even apart from Fisher’s 5-4 instability on a changing Court, it by no means compels the changes in equal protection jurisprudence we advocate. Its focus on subjective racial motivation, side-lining of racial disparities, and treatment of affirmative action as the constitutional risk that Fisher calls for experimentation to mitigate all reinforce doctrinal dispositions our proposal questions. Indeed, affirming prior caselaw, Fisher expressly rejects racial-disparity reduction as a valid reason for race-conscious college admissions. Still, given Fisher’s manifest effect of facilitating disparity reduction in college-going in Texas, and its instantiation of legally mandated and reviewable experimentation as a way to allay forbidden discrimination, Fisher could yet help motivate the Court to catch up with realities on the ground demanding steps to reduce racial disparity.

At the ready—once the Court is—are a variety of established constitutional doctrines with affinities to our proposal. These include:

- The Court’s acceptance of “objective evidence” of racial motivation, including officials’ refusal to consider less racially disparate alternatives or to implement them once they prove themselves;
- Burden-shifting devices requiring defendants charged with discrimination to reveal what they know about racial disparities they generated at places and times not too distant from those plaintiffs have shown to be discriminatory;
- Middle-level equal-protection scrutiny imposing modest duties of inquiry in the face of suspicious diversity the university might try to achieve and different steps for achieving them, using “the experience the school has accumulated and the data it has gathered since the adoption of its admissions plan” to make “periodic reassessment[s] of the constitutionality, and efficacy, of its admissions program” and provide a “reasoned, principled explanation” of actions it took at each step.

273 Fisher, 136 S. Ct. at 2210–11 (citing prior cases).
274 See Fisher, 136 S. Ct. at 2225–28 (Alito, J., dissenting) (arguing that disparity reduction, not diversity, motivated UT’s plan and the majority opinion).
275 See supra notes 48–50.
276 In school desegregation cases, proof of an intent to segregate children by race at a particular school or point in time generates a presumption that racial segregation arising elsewhere or at another time in the same district is also racially motivated absent proof to the contrary by the defendant. See, e.g., Dayton Bd. of Educ. v. Brinkman, 443 U.S. 526, 537–40 (1979); Keyes v. School Dist. No. 1, 413 U.S. 189, 211–12 (1973); Swann v. Charlotte-Mecklenburg Bd. of Educ, 402 U.S. 1, 28 (1971).
mistratment of disfavored minorities;277
• “Reckless indifference” and “message conveyance” standards calling into question officials’ ongoing tolerance of constitutionally suspicious harms;278
• Penalty default rules incentivizing officials to search for better ways to reduce constitutional harms than those the Court itself has identified;279 and
• Numerous lower-court institutional-reform decrees, some affirmed by the Supreme Court, adopting experimentalist remedies for equal protection and other constitutional violations by police, prisons, mental-health facilities, and child-welfare agencies.280

A second objection arises from a brewing debate over Title VII of the 1964 Civil Rights Act, one of our proposal’s statutory analogies.281

Consider three crude visions of how Title VII’s disparate-impact provisions operate. A neutral layperson might argue that Title VII bars employers from using practices to hire, promote, or fire people that generate racially disparate outcomes unless they had a good business reason for the practice.282 A left-leaning skeptic might say Title VII validates any disparity-creating employment practice for which someone can dream up a non-racial explanation after the fact that is not facially false or irrational,

---

278 See Liebman, Perpetual Evolution, supra note 31, at 2040 & n.183 (citing Supreme Court cases).
281 See Kenneth L. Marcus, The War between Disparate Impact and Equal Protection, 2008–09 CATO SUP. CT. REV. 53, 55–56, 72–73 (arguing that Title VII’s disparate impact rules and other facially neutral but potentially racially redistributive actions such as Texas’ “ten percent plan” for admitting students to public colleges in order, by indirection, to increase minority enrollment are unconstitutional); Mark Tushnet, How the Supreme Court’s Ruling on Ricci v. DeStefano Hints at Trouble Ahead, DISSERT, July 12, 3009, https://www.dissentmagazine.org/online_articles/how-the-supreme-courts-ruling-on-ricci-v-destefano-hints-at-trouble-ahead [https://perma.cc/BF8D-EKKH] (noting potential “constitutional problem” with Title VII identified by legal academics and conservative publicists).
rendering the law a nullity. Right-leaning skeptics might describe it as the federal government “mandating” employers to do what the Equal Protection Clause forbids the government to do directly—“discriminate on the basis of race” by forcing employers to give people jobs because of their race not their qualifications.

The quotations in the last-mentioned view are from Justice Scalia’s concurrence in Ricci v. DeStefano, which in his view prophesied “the evil day on which the Court will have to confront the question” whether Title VII’s disparate-impact provisions are consistent with equal protection. Justice Scalia was clearly attracted to the third view of Title VII, describing it as “a racial thumb on the scales, often requiring employers to evaluate the racial outcomes of their policies, and to make decisions based on (because of) those racial outcomes.” Justice Scalia acknowledged, however, a fourth, less crude interpretation of the statute: “as simply an evidentiary tool used to identify genuine, intentional discrimination—to ‘smoke out,’ as it were, disparate treatment”—which, he said, could “defend the law” against constitutional attack if it actually worked that way (which he doubted).

By “requiring” officials in every case to “evaluate the racial outcomes of their policies,” our proposal might seem to step directly into the constitutional trap Scalia has set. Our proposal, however, escapes the trap’s second prong. Rather than requiring officials “to make decisions based on (because of) . . . racial outcomes,” our approach uses mandated evaluation of disparities and experiments to “smoke out” and avoid the stereotypes and conscious or unconscious biases that otherwise would lead to decisions “based on (because of)” false beliefs about race. As between decisions about which chronic racial disparities create real doubt as to their racial motivation, and decisions using inquiry and learning to replace doubt and reveal race-neutral—often generally beneficial—practices, there is no reason why the Equal Protection Clause should opt for ignorance and constitutional doubt.

A third objection is a pragmatic version of the second. The proposal is impractical, given the high cost of disparity-reducing experimentation, legal exposure when experiments fail or are found to be insufficient, and

---

285 Id.
286 Id. at 595 (quoting Primus, supra note 6, at 498–99, 520).
steps to avoid those ills by bestowing benefits on undeserving people so disparities disappear. Of course, if racial gaps do close, and if the agency and its other stakeholders are not harmed or are benefited, net cost is not a problem. Nor would compelling officials to take known steps to diminish disparities at no cost to the public interest fall afoul of traditional equal protection analysis, because doing so “smokes out” false assumptions or biases that drove preexisting disparities. Additionally, given the breadth of existing gaps, taking truly damaging steps to close them completely, with no real hope of improving the public good—at the expense of constituents previously benefited by disparities—will not likely be a price that officials are willing or able to pay in response to the incentives our proposal would create.

The real question, then, is whether the equity and welfare benefits to Black Americans of efforts to “smoke out” any biases underlying the latest iteration of chronic disparities, plus the public benefits that conscientious experimentation typically seeks and often discovers, plus greater certainty as to whether racial bias is at play, outweigh the cost of experiments that do not diminish racial disparities or improve the public good. Given our limitation on the obligation to experiment with steps that do not unduly burden the public fisc, and given our preference for doing something to try to diminish or better understand suspicious disparities over legal rules that justify doing nothing as such disparities mount, we assess the balance as favoring benefits over costs. Given our hope to change action on the ground before changing the Supreme Court’s mind, we have no doubt that our proposal is a worthy experiment in its own right, which we commend287 to actors interested in minimizing racial disparities. Overtime, we expect our proposal will generate more examples of equity and welfare benefits exceeding harms that dispel practical worries.

Still, one might worry about the difficulty of determining whether experimentation that fails for now to close gaps entirely—as often will occur—was sufficiently conscientious to pass muster. Even if it is sensible for large federal departments with performance management expertise in setting and measuring the effectiveness of strategies, effective disparity-reduction experimentation may not seem realistic for many modest local agencies—for example, the nation’s thousands of school districts with only a few administrators. Consider, though, the classic justification for so thoroughly decentralizing U.S. public education, juvenile and criminal

justice, and many other public activities: the ability it gives localities to address problems in local context. Rather than imposing novel requirements, therefore, our proposal calls on local public actors to apply their ingenuity to closing the many local opportunities for racial discrimination. As our examples reveal, moreover, local education, juvenile justice, and other agencies have willingly accepted similar statutory obligations in order to qualify for coveted federal funding. And by comparing experimentation steps and success levels of comparably situated local actors, federal and state funds-dispensing and oversight agencies, support organizations, and practitioners’ networks have learned to recognize more and less conscientious experimentation in widely varying contexts. Courts can do the same—as some already do in enforcing bans on workplace sexual harassment, disability discrimination, and racial and gender-based bias. Here, again, constitutional courts trailing behind longstanding practice in analogous contexts have plenty of experience on which to draw in developing disparity-reduction standards and reviewing experiments by a multiplicity of public actors.

These responses to objections from the right side of the political spectrum invite more troubling ones from the left. Braiding them into one, the objection is that, “we have tried all this before at a variety of subconstitutional levels, and we are still awash in racial disparity.” Experimentation or not, chronically underserved populations are likely to remain in that position because of a poisonous mix of (1) malevolently


289 Discussing government and private-sector practices in various contexts that gradually improve public safety and service delivery by comparing and benchmarking the quality and effectiveness of plans or other steps local entities take under similar circumstances to reduce harms or achieve public objectives are, e.g., Cary Coglianese & David Lazer, Management-Based Regulation: Prescribing Private Management to Achieve Public Goals, 37 L. & SOC. REV. 691, 693–700 (2003) (food and workplace safety; pollution prevention); Bradley Karkkainen, After Backyard Environmentalism, 44 AM. BEHAV. SCI. 690, 694–703 (2000) (toxic chemicals reduction, Chesapeake Bay pollution control, protection of endangered species’ habitats); Helen F. Ladd, Education Inspectorate System in the Netherlands and New Zealand: A Policy Note, 5 ED. FIN. & POL’Y 378, 378–83 (2010) (public school improvement); Joseph V. Rees, Hostages of Each Other 94–127 (1994) (regulation of nuclear power generation); Sabel & Simon, Contextualizing Regimes, supra note 23, at 1290, 1278–97 (avoiding leafy green contamination, racial discrimination in juvenile justice administration, dolphins kills in commercial tuna fishing).

chosen, assiduously concealed, discrimination; (2) unconscious bias perpetrators are unlikely to surrender even when publicly exposed; (3) sprawling wealth effects of 400 years of slavery, segregation, and economic subjugation that exponentially multiply disparities in a capitalist environment in which wealth-determined choices are the principal distributional mechanism;\textsuperscript{291} and (4) the “positional” fact as to some White privileges and widespread beliefs as to other such privileges that the perquisites of happiness are finite, so your gain is my loss.\textsuperscript{292} Under these circumstances, the objection goes, it is both morally imperative and practically necessary to risk lost generalized benefits, social conflict, and a probability of cataclysmic failure by insisting on reparative, racially redistributive gap closure now. Black communities have no time left to waste.

All of us must make our own decisions about how much risk we, our families, and communities can reasonably be asked to take—and how much our uniquely, fragilely diverse polity can stand. But even under the most favorable scenario, racially reparative justice will get us only so far. The chances of full financial racial repair are uncertain at best, given reasonable estimates that we would have to multiply Black Americans’ existing wealth by a factor of from ten to one hundred to pull them even with White Americans\textsuperscript{293}—not to mention the large additional amounts needed to account for administrative imperfection in distributing large amounts of money to millions of individuals in ways that benefit them over the long haul. And even if all preexisting material gaps could be back-filled, discriminatory muscle memory aggravated by animosities provoked by a bitter reparations fight will likely cause new disparities to arise. Because disparities will continue to arise, any sensible reparative regime will have to be backstopped by a duty on officials’ part not to take back with the right hand what the left hand has given. Either way, therefore—pre-, post-, or in lieu of, reparations—we will need to keep doing what we can to identify and alter the racist habits, institutional muscle memory, embedded wealth inequities, and shortsightedness that are the legacy of centuries of racial subjugation. Perhaps the best solution is a blend of the two, offering

\textsuperscript{291} See supra notes 2–5, 17, 157, 217 (citing sources).
\textsuperscript{292} See supra note 67 (discussing zero-sum positional inequity).
\textsuperscript{293} See Hannah-Jones, supra note 3 ("The average black family with children holds just one cent [worth] of wealth for every dollar that the average white family with children holds."); Michael W. Kraus et al., The Misperception of Racial Economic Inequality, 14 J. PERSPECTIVES PSYCH. SCI., Sept. 2019, at 1, 5–6 (finding that Black households hold $10 in wealth for every $100 held by White households and that disparity increases as Black income (as opposed to wealth) increases).
reparations in the form of preferences for participation in or subsidies to expand official disparity-reduction experiments. 294

As for past experimentation failing to end disparities, the key questions are how fully, quickly, and instead of (or alongside) what other steps we should expect experimentation to work? As incomplete as the results of disparity-reduction experimentation have been even in the favorable examples we feature, such steps have outperformed more direct efforts to the substantial benefit of Black and other chronically underserved populations and the public at large. 295 Results would have been even better, but for the premature truncation of some of the efforts in the wake of initially ragged implementation (the downfall of both school improvement strategies we highlight 296); weak enforcement mechanisms (a problem in the juvenile justice context that a recent statutory revision aims to cure 297); and a tendency to forsake ongoing experimentation in favor of mandates to spread initially effective strategies in context-insensitive ways.

Illustrating the last-mentioned stumbling block to effective experimentation and an important way around it is New York City’s 1990s experimentalist policing reforms, which gave way in the 2000s to a disastrous, disparity-ignoring obsession with one strategy—stop and frisk—until an experimentalist judge ordered police to include disparity-reduction as a goal of future crime-reduction experimentation. 298 Following the lead of that judge’s decision and recognizing a generalized constitutional duty of disparity-reduction experimentation that is privately enforceable could bolster enforcement rigor, as long as it retained context-sensitive flexibility. A best-case scenario, then, might be disparity-reducing experimentation as a backstop or implementation mechanism for reparations; robust public participation in experimentation to increase public patience and acceptance as experimentation improves, 299 and improved enforceability through the proposal’s adoption as an appropriate interpretation of the Equal Protection Clause with a Section 1983 private right of action. 300

295 See supra notes 83–254 and accompanying text.
297 See supra note 126 and accompanying text.
299 See Liebman, Cruikshank & Ma, supra note 128, at 446–47, 450–58 (discussing ways to engage the public and build confidence in structured policy experimentation).
V. CONCLUSION AND POSTSCRIPT

We began drafting this Article in late 2019. At the time, we imagined Armageddon coming from a White revanchist right inflamed by calls for racial redistribution, and we embarked on a last-ditch search for a reformist middle ground.

We finished drafting the Article in early summer 2020, soon after Minneapolis police officers—enacting another of countless acts of racially disparate official violence against the nation’s Black people—killed George Floyd. By then, we could as easily see the reckoning coming from a Black-mobilized left, out of impatience with a 400-year wait for justice, and we began looking to and beyond racial reparations. As puzzled readers may have discerned, our perspectives changed along with tectonic shifts that Floyd’s death provoked in the nation’s struggle to understand racial (in)justice.

As the Article was being edited, a racist and revanchist mob assaulted the nation’s Capitol, elevating still more our fears for a nation in need of deep and inevitably disorienting change and for the safety of the intended beneficiaries of such change. President Biden’s effort


soon after to restore a semblance of racial equanimity provides only modest solace.\(^{303}\)

Uncertain of the nation’s evolving, or dissolving, center of gravity on race, we have left in place our fluctuating guesses about the social realities that will determine whether our proposal for a constitutional regime of determined, public-regarding learning from the chronic racial disparities to which George Floyd’s death testifies goes too far or not nearly far enough.

Table 1: Juvenile Detention Rates and Numbers 1997–2017

<table>
<thead>
<tr>
<th></th>
<th>1997</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>detention rate</td>
<td>detained (#)</td>
<td>detained (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>356</td>
<td>105,055</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>201</td>
<td>39,445</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>968</td>
<td>41,896</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>468</td>
<td>19,322</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Am. Ind.</td>
<td>490</td>
<td>1,615</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian, P.I.</td>
<td>195</td>
<td>2,215</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Average NAEP Scale Scores – Fourth Grade Math – N.Y.C. and National Public Students

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>White</th>
<th>Black</th>
<th>Hispanic</th>
<th>FRPL</th>
<th>SWD</th>
<th>ELL</th>
<th>Nat'l Public</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>226</td>
<td>244</td>
<td>219</td>
<td>220</td>
<td>224</td>
<td>205</td>
<td>208</td>
<td>234</td>
</tr>
<tr>
<td>2009</td>
<td>236</td>
<td>248</td>
<td>227</td>
<td>230</td>
<td>235</td>
<td>223</td>
<td>220</td>
<td>239</td>
</tr>
<tr>
<td>2013</td>
<td>236</td>
<td>251</td>
<td>225</td>
<td>228</td>
<td>231</td>
<td>219</td>
<td>215</td>
<td>241</td>
</tr>
<tr>
<td>2019</td>
<td>231</td>
<td>248</td>
<td>216</td>
<td>220</td>
<td>224</td>
<td>204</td>
<td>208</td>
<td>240</td>
</tr>
</tbody>
</table>
Table 3: Percent Change in N.Y.C. Achievement Score Gaps on NAEP Fourth Grade Math; Reform Phases 2003–09 & 2003–13 vs. Post-Reform Period 2013–19

<table>
<thead>
<tr>
<th></th>
<th>All Students</th>
<th>Black Students</th>
<th>Hispanic Students</th>
<th>Economically Disadvantaged</th>
<th>Students with Disabilities</th>
<th>English Learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gap with N.Y.C. White Students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003–09</td>
<td>-33%</td>
<td>-16%</td>
<td>-25%</td>
<td>-35%</td>
<td>-36%</td>
<td>-22%</td>
</tr>
<tr>
<td>2003–13</td>
<td>-17%</td>
<td>+4%</td>
<td>-4%</td>
<td>-</td>
<td>-18%</td>
<td>-</td>
</tr>
<tr>
<td>2013–19</td>
<td>+13%</td>
<td>+23%</td>
<td>+22%</td>
<td>+17%</td>
<td>+38%</td>
<td>+11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gap with All U.S. Public School Students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003–09</td>
<td>-63%</td>
<td>-20%</td>
<td>-36%</td>
<td>-60%</td>
<td>-33%</td>
<td>-19%</td>
</tr>
<tr>
<td>2003–13</td>
<td>-17%</td>
<td>+7%</td>
<td>-7%</td>
<td>-</td>
<td>-18%</td>
<td>-</td>
</tr>
<tr>
<td>2013–19</td>
<td>+80%</td>
<td>+50%</td>
<td>+54%</td>
<td>+60%</td>
<td>+44%</td>
<td>+23%</td>
</tr>
</tbody>
</table>

Figure 9: NAEP Scores 2003-2019
Figure 10: NAEP Scores 2003-2019

Figure 11: NAEP Scores 2003-2019