

The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action

Author(s): Glenn Ellison and Parag A. Pathak

Source: *The American Economic Review*, MARCH 2021, Vol. 111, No. 3 (MARCH 2021), pp. 943-975

Published by: American Economic Association

Stable URL: <https://www.jstor.org/stable/10.2307/27027277>

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/10.2307/27027277?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



American Economic Association is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*

JSTOR

The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago’s Exam Schools[†]

By GLENN ELLISON AND PARAG A. PATHAK*

Several K-12 and university systems have adopted race-neutral affirmative action in place of race-based alternatives. This paper explores whether these plans are effective substitutes for racial quotas in Chicago Public Schools (CPS), which now employs a race-neutral, place-based affirmative action system at its selective exam high schools. The CPS plan is ineffective compared to plans that explicitly consider race: about three-quarters of the reduction in average entrance scores at the top schools could have been avoided with the same level of racial diversity. Moreover, the CPS plan is less effective at adding low-income students than was the previous system of racial quotas. We develop a theoretical framework that motivates quantifying the inefficiency of race-neutral policies based on the distortion in student preparedness they create for a given level of diversity and use it to evaluate several alternatives. The CPS plan can be improved in several ways, but no race-neutral policy restores minority representation to prior levels without substantially greater distortions, implying significant efficiency costs from prohibitions on the explicit use of race. (JEL H75, I21, I28, J15)

Affirmative action is one of the most contentious issues in American public policy, particularly in school admissions. In the 1970s, race-based plans were widespread in K-12 and university admissions. Since then, such plans have been challenged on multiple fronts. Some states have banned race-based affirmative action.¹ In 2003, the US Supreme Court established a standard of “strict scrutiny” for race-based plans at public institutions, requiring that they must serve a compelling government

* Ellison: Department of Economics, Massachusetts Institute of Technology, and NBER (email: gellison@mit.edu); Pathak: Department of Economics, Massachusetts Institute of Technology, and NBER (email: ppathak@mit.edu). Debraj Ray was the coeditor for this article. We thank Katie Ellis, Susan Ryan, and the staff at Chicago Public Schools for their expertise and help with the data. Vivek Bhattacharya, Adrian Blattner, Jackie Bredenberg, Brandon Enriquez, Raymond Han, Alex Olssen, and Rahul Singh provided excellent research assistance. We’re grateful to seminar participants at CESifo, Harvard Law, MIT, and Tufts for comments. Ellison acknowledges support from the Toulouse Network for Information Technology. Pathak acknowledges support from the William T. Grant Foundation and National Science Foundation grants SES-1056325 and SES-1426566. This study was approved by MIT’s committee on human subjects as IRB 1506128499 on July 15, 2015.

[†] Go to <https://doi.org/10.1257/aer.20161290> to visit the article page for additional materials and author disclosure statements.

¹ California and Michigan passed constitutional amendments by ballot in 1996 and 2006, respectively.

interest that cannot be effectively achieved in a race-neutral manner.² Many public institutions have adopted alternative plans that avoid explicitly using race.

We explore the consequences of the shift from race-based to race-neutral affirmative action at Chicago Public Schools (CPS). From 1980–2009, the assignment process for Chicago’s exam high schools involved racial quotas that restricted White enrollment. In 2010, CPS shifted to a race-neutral system that targets socioeconomic integration via a neighborhood-based approach. In the system, admissions preferences are based on the socioeconomic characteristics of the census tract in which a student resides. Each tract is placed into one of four tiers. Within each tier, the tie-breaker is based on a composite score which combines results of a specialized entrance exam, prior standardized test scores, and grades in prior coursework. Students rank schools and a deferred acceptance algorithm generates placements, factoring in both tier and tie-breaker.³ We focus on how this change has affected the minority and low-income composition, and curriculum matching for the incoming class at the two most selective exam schools, Walter Payton College Prep and Northside College Prep.

Chicago’s race-neutral plan has substantially increased the spread of composite scores at Payton and Northside. This is driven by two forces: the CPS plan admits students who have much lower composite scores than any student who would have been admitted under a race-based plan, and the CPS plan fails to admit many high achieving low-income and minority students. Even though the plan aspires to broaden the definition of disadvantaged students, the CPS plan admits *fewer* low-income students than would a pure racial quota with a comparable effect on the distribution of scores of admitted students. Moreover, within-school achievement gaps between majority and minority students widen compared to racial quotas.

A race-neutral policy remains far away from the optimal admissions plan for two reasons. First, the policy may not have been chosen well. We show that CPS’s plan can be substantially improved with some relatively minor changes. By reweighting and transforming the CPS tract-level socioeconomic measures, we can improve on the performance of the plan. Further improvements are possible if we use applicant’s free or reduced price lunch (FRPL) status.⁴

Second, the policy is bounded away from the optimal frontier because of the constraint that it is race-neutral. Our most important conclusion is that a substantial portion of the minority inefficiency of the CPS plan appears to be an inevitable consequence of a restriction on using race. We draw this conclusion from an exploration of more complex plans constructed using machine learning techniques. Admissions policies which may be beyond a legally permissible use of race cannot achieve much more than what our simple improvements to CPS’s tract-level variables achieve. Possibilities for race-neutral plans differ by school and are even more limited at Northside.

We anchor our empirical investigation with a model of elite schools. In the model, students are assumed to benefit from a curriculum tailored to their needs and from learning within a diverse student body. Affirmative action plans can

²A June 2007 Supreme Court decision applied the earlier decision to strike down race-based admissions plans for public schools in Seattle and Jefferson County, Kentucky.

³Dur, Pathak, and Sönmez (2020) describe the assignment mechanism in more detail.

⁴CPS chose not to use individual-level data in part due to concerns about whether the data would be reliable if used.

benefit students by increasing diversity, although this comes at a cost of reducing curriculum-matching benefits. We propose that the efficiency of plans be assessed by measuring how close a plan is to minimizing the adverse effects on the distribution of student-preparedness conditional on the level of diversity. Our overall efficiency measure compares a given plan with other plans that simultaneously achieve the same level of minority and low-income representation. Our minority and low-income efficiency measures reflect how efficiently a plan accomplishes diversity on each dimension separately. Our efficiency measures involve subjective decisions about which adverse effects to consider and how to cardinalize them. We have tried to be conservative in labeling plans as inefficient, both relative to a utility-based model that motivates the measure and in ignoring additional disadvantages of the CPS plan, e.g., its contribution to stereotype reinforcement.

Under purely score-based admissions, Payton and Northside would have few low-income and minority students. Race-based or free or reduced price lunch (FRPL)-based affirmative action plans can substantially increase minority and/or low-income representation with a modest effect on the average composite score of admitted students. However, according to our metric, the CPS race-neutral plan is 28 percent and 22 percent efficient at Payton and Northside compared to the race- and FRPL-based benchmark. Focusing more narrowly on the CPS plan as a means to boost minority enrollment, the CPS plan is only about 26 percent and 19 percent efficient at Payton and Northside. Focusing narrowly on the plan as a means to boost low-income enrollment, it is even more inefficient. By making fairly straightforward modifications to how CPS measures socioeconomic status and how SES is used in the assignment mechanism, it is possible to improve overall efficiency to 44 percent and 39 percent at Payton and Northside. But this is about all that can be done: a more involved change (that may not survive legal challenges) results in a modest additional efficiency improvement.

The model we use to motivate our approach builds on several ideas in the previous literature. In particular, our model of curriculum-matching and school system design generates a trade-off between score maximization and diversity studied in other papers. The effect of CPS's plan on the distribution of admitted students verifies the practical relevance of Chan and Eyster's (2003) observation that optimal race-neutral plans may expand the lower-tail of the admitted student distribution by rejecting some very highly qualified minorities and admitting some majority candidates who would have not been admitted under a race-based plan. Our finding that within-school achievement gaps are larger under the CPS plan than under a race-based plan empirically confirms a phenomenon that Ray and Sethi (2010) noted was a necessary feature of the constrained optimal plan, but which might have been regarded as counterintuitive. Finally, our approach to developing an upper bound on what may be achieved, exploring what can be done with plans that probably go beyond what would be legally permissible, follows Fryer and Loury's (2005) admonition that too-efficient forecasts of race are not race-blind in a meaningful sense.

The most closely related empirical antecedent to our paper is Fryer, Loury, and Yuret (2008). We follow their main approach of comparing the effects that race-based and race-blind affirmative action procedures would have on class composition via counterfactual simulations. Importantly, however, we overcome the

limitations of their dataset, which only contains information on admitted students and not unsuccessful applicants, making it challenging to assess how policies would work in practice. Cestau, Epple, and Sieg (2017) develop an econometric model of the referral process for taking the admissions tests for selective elementary schools. They report that profiling by race and income together with affirmative action based on FRPL status can achieve 80 percent of level of Black enrollment as a race-based affirmative action plan. Corcoran and Baker-Smith (2018) study admissions policies at New York's exam schools, focusing on a descriptive account of application decisions. Though their main interest is not in affirmative action, they simulate top 10 percent rules based on seventh grade scores and find that such policies lead to an increase in Black and Hispanic representation at the schools examined. Epple, Romano, and Sieg (2008) develop an equilibrium model of affirmative action and tuition policies to show that a ban on affirmative action leads to a decline in minority students at top-tier colleges.

The rest of this paper is organized as follows. Section I describes Chicago's exam schools and the new race-neutral assignment policy. Section II computes the trade-off between diversity and selectivity at Payton and Northside. It also reports on racial, income, and achievement gaps under the CPS policy compared to a race and FRPL-based benchmark. Section III develops a model of optimal admissions with affirmative action to motivate our measure of the relative efficiency admissions plans. Section IV computes the relative efficiency of Chicago's plan. Section V studies the extent that simple or larger-scale changes to Chicago's plan could increase its efficiency. Section VI concludes.

I. Chicago's Exam Schools

A. Schools and Admissions Process

In 2014, Chicago Public Schools operated ten "selective enrollment" or exam high schools, listed in Table 1.⁵ Each school offers students advanced curricula and a high-achieving peer group compared to most CPS schools. Payton and Northside are the most selective and highest-performing. Table 1 shows that Payton and Northside are also fairly small: 233 and 265 students enroll respectively, compared to total eighth grade CPS enrollment of about 28,000. Jones and Whitney Young are the next most selective; their average ACT scores are at the eighty-sixth and eighty-ninth percentile, respectively. Performance at the other exam schools is substantially lower.

To obtain an exam school seat, applicants must take an admissions test and submit a ranking of up to six schools. Students who are competitive for Northside or Payton will be sufficiently far above the cutoff for the bottom half of the schools so that they can be certain that they will be admitted to any of them, and should therefore list their top six choices in their true preference order.⁶ This fact motivates our

⁵We omit South Shore College Prep because it only became a selective enrollment school in 2013–2014 so statistics on 2014 graduates would not reflect the current admissions procedure and curriculum.

⁶Haeringer and Klijn (2009) and Pathak and Sönmez (2013) also show that for students who prefer less than six schools, it is weakly dominant strategy to rank schools truthfully.

TABLE 1—CHICAGO'S EXAM HIGH SCHOOLS

	Mean ACT score (percentile) (1)	Average number of AP exams passed (2)	Average AP score (3)	Enrolled students (4)
Northside College Prep	30.1 (94)	4.68	3.83	265
Walter Payton College Prep	30.2 (94)	5.20	3.74	233
Whitney Young Magnet	28.3 (89)	3.04	3.20	468
Jones College Prep	26.8 (86)	2.25	3.18	427
Lane Technical	24.7 (78)	2.13	2.91	925
Brooks College Prep	23.5 (74)	0.81	1.97	199
Lindblom Math and Sci Acad	22.8 (69)	1.34	2.42	228
Westinghouse	21.4 (57)	0.79	2.01	278
King College Prep	20.5 (57)	0.25	1.80	148

Notes: ACT Percentiles are for rounded average ACT scores by school, and based on national distributions for high school graduates from 2015, 2016, and 2017 (ACT 2020). ACT data was received from the Prairie State Achievement Examination (ISBE 2014). Average number of AP exams passed refers to the average number of AP exams ever passed (score 3 or higher) for the cohort of seniors in 2014. Average AP score refers to the average of the mean score achieved by the same cohort. Enrolled students refers to the number of ninth grade students in 2014.

assumption that rankings for these schools for students correspond to true preferences and would not change under the alternate admissions policies we consider.⁷

The CPS assignment procedure starts by assigning each student a “composite score,” which we rescale from 0 to 100. The composite score places equal weight on three factors: the entrance exam score, the student’s seventh grade score on Illinois’ statewide standardized test, and the student’s seventh-grade grade-point average. The composite score can roughly be thought of as corresponding to a student’s national percentile. For example, an applicant would receive a score of 98.9 if she achieved a score CPS deemed to be at the national ninety-eighth percentile on their admissions test, had ninety-seventh percentile scores in both English and Math on the ISAT, and had a perfect middle school GPA.⁸ Each student is also assigned to one of four SES tiers. The SES tiers are determined by a place-based affirmative action scheme. Student addresses are mapped to each of about 800 census tracts. Six characteristics of each census tract are then used to construct the SES score for the tract: (i) median family income, (ii) percentage of single-parent households, (iii) percentage of households where English is not the first language, (iv) percentage of homes occupied by the homeowner, (v) adult educational attainment, and (vi) average Illinois Standards Achievement Test (ISAT) scores for attendance-area schools.⁹ Tracts are ranked by these scores, and then divided into four tiers, each with approximately the same number of school-age children. Tier 1 tracts are the most disadvantaged, while Tier 4 tracts are the most advantaged. Panel A of Figure 1 illustrates the distribution of tiers

⁷ Student preferences over schools could change if changes to the admissions procedure made large enough changes to the achievement levels or demographics of the admitted students. There is also evidence from lab experiments that participants do not submit their preferences truthfully in strategy-proof mechanisms (see, e.g., Chen and Sönmez 2006 and Li 2017). Students might also move in response to a change in the admissions process. Cullen, Long, and Reback (2013) document that this has occurred in Texas after Texas adopted a school-based admissions procedure for college. The online Appendix shows that such behavior seems to be quite rare in our context.

⁸ The percentiles are not calibrated to the same population, and a full 100 GPA points are given to any student with straight As in seventh grade so the interpretation of scores as percentiles is not precise.

⁹ The SES score is the average of a tract’s percentile rank within the city of Chicago on the six dimensions.

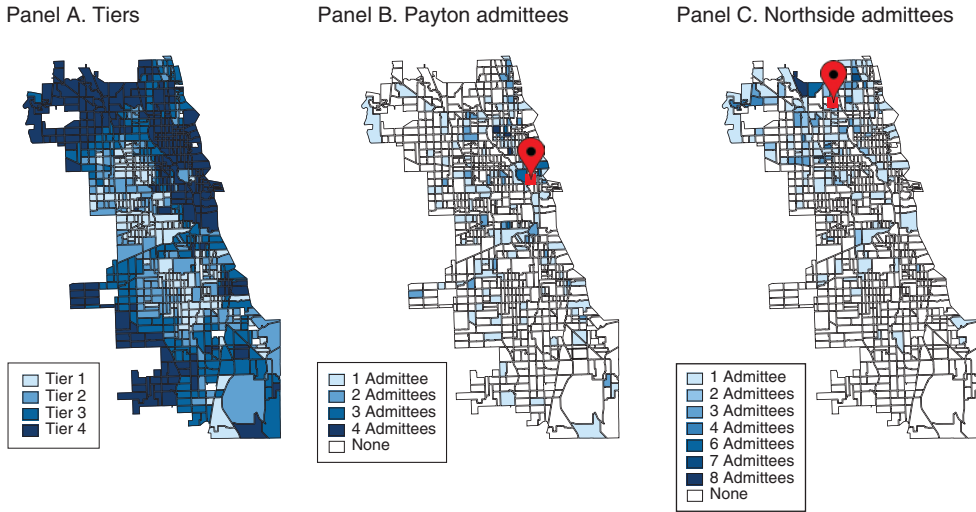


FIGURE 1. CHICAGO TIERS AND ADMITTED STUDENTS AT PAYTON AND NORTHSIDE

Notes: Panel A shows the geographical distribution of tiers across tracts. Panels B and C show the number of applicants admitted to Payton and Northside, respectively, by tract.

throughout the city. Many of the highest-tier census tracts are at the northern or western edges of the city or along Lake Michigan north of downtown.

The CPS assignment mechanism is based on deferred acceptance (DA). In the system, each school is divided into five sub-schools. Merit seats, representing 30 percent of the seats, are open to all students regardless of their SES tier. Tier seats, each with 17.5 percent of the available seats, prioritize applicants from the corresponding tier, using the composite score to break ties. In the algorithm, applicants who rank a school are first considered for merit seats and then for the seats reserved for their tier. Tier seats are almost exclusively assigned to applicants from the corresponding tier.¹⁰ An unmatched student is defaulted to their neighborhood school.¹¹ An unmatched student also could matriculate at other choice options including career and technical academies, magnet schools, military academies, and charter schools.

One feature of Chicago that makes it attractive for studying affirmative action is that the two most selective exam schools are differently situated, so we can study how race-neutral affirmative action plans fare in different environments. Figure 1 shows the schools' locations and where their students live. Payton is closer to the center of Chicago and enrolls students from across the city. Northside is a relatively attractive location for students from the northern parts of the city which consist of primarily tier 3 and tier 4 neighborhoods. For this reason, there may be greater

¹⁰Dur, Pathak, and Sönmez (2020) provide more details on Chicago's DA implementation with merit and tier seats.

¹¹A small number of students may also be admitted to each school under special education and "principal's discretion" programs. We have no way of knowing if any students would have been admitted via these programs under alternative admissions policies, so in all of our comparisons we assume that under the current policies the school consists only of those students selected by the main deferred acceptance algorithm. All statistics on the admitted classes at Payton and Northside refer to students admitted through the main admissions process, and therefore will not exactly match other measures of the schools' demographics computed directly from the set of enrolled students.

efficiency costs at Northside if an affirmative action plan fails to offer admission to the low-income and/or minority students living in these neighborhoods.

B. Data

Our primary data consist of application files from the Chicago Public Schools for the 2013–2014 year and a separate file containing the factors underlying the tier formula for each census tract (CPS 2014). Each record includes students' composite scores and underlying components, school choices, tier, and final assignment.¹² We augment this file with data on the census tract factors used to construct tiers from Eder and Gregg (2014) and additional information on tract-level characteristics from IPUMS (Manson et al. 2010). Our analysis sample has 16,818 students and 77,051 student choices. The average applicant ranks about 4.6 schools. Applicants who rank all six choices are more likely to be Black, to qualify for a subsidized lunch, and have lower average composite scores.¹³

Applicants to exam schools are positively selected compared to CPS eighth graders. Table 2 shows that the standardized test scores of applicants (measured as percentiles) are well above average. Exam applicants are also less likely to qualify for a subsidized lunch than a typical CPS eighth grader. Compared to the broader pool of exam applicants, applicants who rank Payton and Northside are somewhat more positively selected as shown in columns 3 and 6 of Table 2. Northside has fewer Black and more Hispanic applicants compared to Payton.

Race-neutral admissions plans will by necessity either admit fewer minority students or reduce average composite scores at the most selective schools compared to race-based plans. Column 4 reports characteristics of students admitted under the current CPS procedure, while column 5 reports on students who would have been admitted had CPS used racial quotas to maintain the same level of Black and Hispanic representation at Payton and Northside as under the previous race-based process.¹⁴ At Payton, the shift to a race-neutral plan has cut the fraction of Black admits nearly in half, from 28 percent to 15 percent. Average composite scores are unchanged, and there have been only small reductions in the fraction of Hispanic students and the fraction of students qualifying for free or reduced price lunch (FRPL). At Northside, the CPS plan increased minority and FRPL enrollment and decreased average composite scores.¹⁵ Across both schools, 124 of the students

¹²The application files also indicate whether a student qualifies for special education programming; because those applicants are assigned through a different process, we exclude them from our analysis sample.

¹³The fact that an applicant does not rank all schools does not necessarily imply that she has not revealed her preferences truthfully. Rather it could indicate that any unlisted choice is less preferred to the applicant's outside option. For high-scoring applicants, the fraction of applicants who do not rank all six choices increases with the composite score. Many such high-scoring applicants would likely opt for schools outside of the public district if not admitted to the most selective exam schools.

¹⁴We implemented racial quotas at Payton and Northside by setting separate Black, Hispanic, and White/Asian quotas to be equal to the average share of Black, Hispanic, and White/Asian students from the three years prior to the adoption of race-neutral admissions. The number of offered seats is not exactly the same as the number of students who enroll in Table 1 because of new offers made after the initial round.

¹⁵The share of Black and Hispanic students was much lower at Northside than at Payton under the former race-based policy. This was a side effect of the fact that the 1980 consent decree required that no more than 35 percent of a school be White. Northside had a substantially larger Asian enrollment than Payton and by 2008, Asian students took up many of the non-White seats at Northside.

TABLE 2—DESCRIPTIVE STATISTICS ON APPLICANTS

	CPS 8th graders (1)	Exam applicants (2)	Applicants Payton (3)	Offered Payton		Applicants Northside (6)	Offered Northside	
				Current CPS (4)	Racial Quota (5)		Current CPS (7)	Racial quota (8)
Black	0.42	0.39	0.37	0.15	0.28	0.26	0.06	0.05
Hispanic	0.45	0.42	0.44	0.22	0.24	0.50	0.29	0.23
Female	0.50	0.55	0.55	0.65	0.67	0.57	0.65	0.65
FRPL	0.86	0.74	0.71	0.25	0.26	0.72	0.34	0.24
Composite score		62.17	65.89	97.99	98.0	66.87	97.75	98.93
GPA		61.82	65.67	99.36	99.3	67.39	99.16	99.90
Standard tests		75.86	78.30	97.61	97.5	78.68	97.49	98.24
Admissions test		48.82	53.70	97.01	97.1	54.56	96.58	98.65
Observations	27,944	16,818	10,549	220		8,274	259	259

Notes: This table shows descriptive statistics for the CPS students enrolled in eighth grade and applicants to exam schools in our analysis sample. Current CPS refers to the current tier-based admission scheme and GPA to seventh grade average grades.

admitted under the new policy, or roughly one-quarter of admitted students, would not have been admitted to the same school under the racial quota policy we simulate.

C. The CPS Composite Score

The CPS composite score is closely related to curricular offerings and educational outcomes at schools. The relationship between an applicant's composite score and an exam school's distribution therefore provides a measure of whether the student is suited for the exam school's curriculum.

Advanced Placement (AP) courses are advanced college-level curricula offered to high school students who may obtain placement or credit for high enough scores. Exam schools offer AP courses, and more selective schools are able to offer more such courses and teach them at a more advanced level. Panel A of Figure 2 examines the relationship between entering students' composite scores and AP test performance, where AP test performance data is from CPS (2014). To measure performance, we define AP Performance as the number of points in excess of two that students achieve on all of the AP tests that they take in 11th or 12th grade.¹⁶ We then plot average AP Performance against student composite score, for students with scores greater than 85. Recall that we have scaled the CPS composite score to run from 0 to 100, so this represents roughly the top 15 percent of students. AP Performance measures both the relationship between student ability and AP test performance and course offerings at any given exam school. At less selective exam schools, students have both lower baseline scores and tend to have less access to AP courses. Students with composite scores below 90 typically pass few AP tests. Expected AP performance increases sharply among students with composite scores greater than 95. In particular, students

¹⁶Specifically, we give them no credit for tests on which they score 1 or 2, one point for each test on which they score 3, two points for each score of 4, and three points for each score of 5. The motivation for this transformation rather than simply adding all AP scores is that scores of 1 or 2 are generally regarded as failing scores. We do not want results to be affected by differences in whether students did or did not take tests on which they would receive such scores.

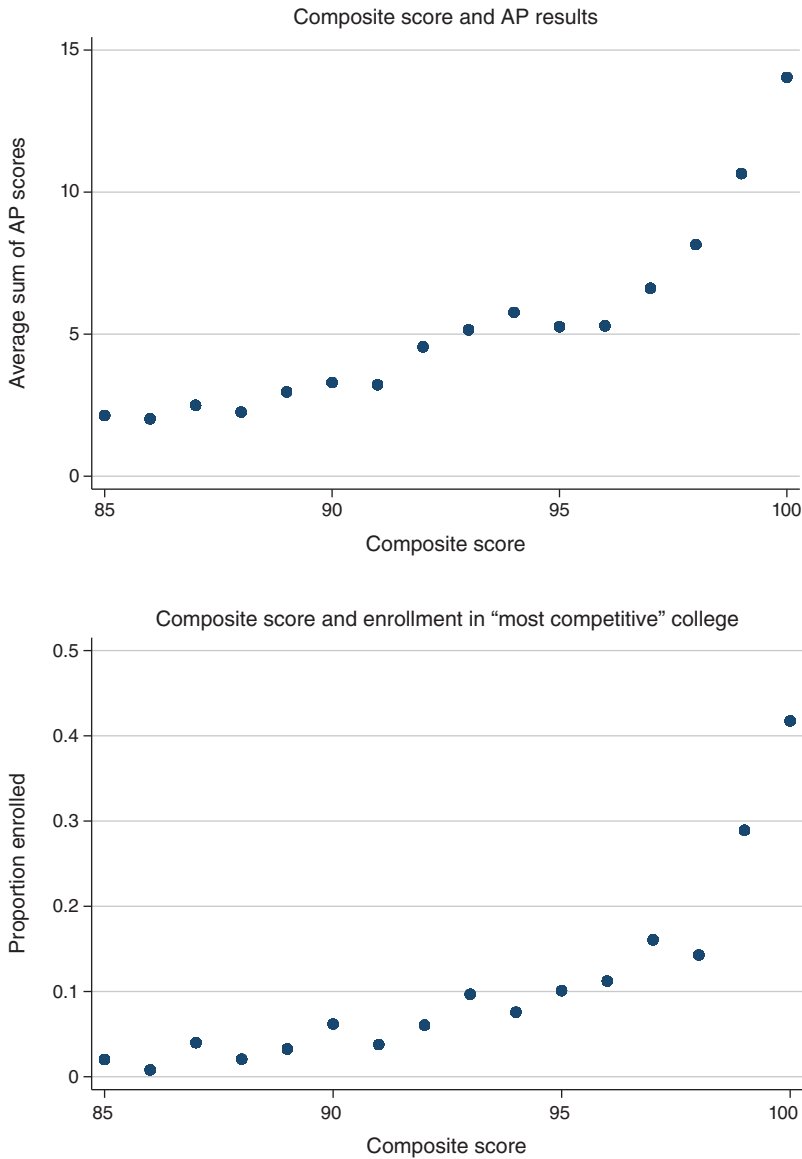


FIGURE 2. RELATIONSHIP BETWEEN CPS COMPOSITE SCORES AND AP PERFORMANCE AND SCORES AND ADMISSION TO THE MOST COMPETITIVE COLLEGES

Notes: The sum of AP Scores is the sum of the AP scores ever achieved by a student, where only scores in the range of 3–5 are considered, and rescaled to 1–3. Most competitive colleges are defined as in the 2009 *Barron’s Profile of American Colleges*.

with the highest composite scores have expected AP Performance of 14, those with composite scores of 99 have expected AP Performance of about 11, and those at 98 have about 8. Therefore, if a student with a composite score in the low 90s or below replaces one with a composite score in very high 90s, we anticipate a substantial efficiency loss due to curriculum matching.

Suggestive evidence of curriculum mismatch from assigning lower scoring students to Payton or Northside is also present when we compare composite scores to college-going. Panel B of Figure 2 plots the relationship between the composite scores and whether students go on to attend a “most competitive” college after graduation.¹⁷ As with AP Performance, the predictive power of composite score in the upper tail is striking: more than 40 percent of students with the highest possible composite score attend a most competitive college, while only about 5 percent of students with composite scores between 85 and 90 do.

The strength of the relationship between composite score and AP Performance and college-going motivates treating the composite score as a proxy for the curriculum to which students are matched. When we examine how affirmative action plans affect within-school standard deviation in the composite score, Figure 2 can serve as a reference to translate the cardinal scale of composite scores into units of AP and college outcomes.

II. Trade-Offs between Diversity and Selectivity

A. Chicago's Frontiers

We begin our analysis by computing the trade-off between diversity and selectivity at Payton and Northside. Figure 3 plots the frontier at both schools, as well as a single point (marked with a triangle) corresponding to the class composition if CPS were to adopt a purely score-based admissions procedure. Under this policy, the average composite score of admitted students would be 99.1 at Payton and 99.0 at Northside. Since these means are near the upper bound score of 100, there is obviously little within-school variation in the composite scores. The majority of students would have composite scores in the 99 to 100 range and all students would have scores of at least 97. The position of these points relative to the x - and y -axes show that the schools would also not be very diverse under purely score-based admissions. The x -axis position reflects that only 21 percent of the students at Payton and 19 percent of the students at Northside would be Black or Hispanic, hereafter underrepresented minorities. The y -axis positions reflect that only 15 percent of the students at Payton and 23 percent of the students at Northside would be sufficiently low income to qualify for FRPL.

The three curves in the figure correspond to admissions policies where some admitted under score-based admissions are replaced with lower-scoring minority and/or low-income students.¹⁸ The innermost curve contemplates a small change in class composition, where the average composite score is 98.8. Such plans can have a nontrivial effect on minority and low-income representation even though only a limited number of offers change. The other two curves illustrate admissions policies which result in larger changes in the composite score, reducing

¹⁷ We define “most competitive” using the 2009 *Barron's Profile of American Colleges*, provided to us by David Deming in electronic form.

¹⁸ To compute these curves, we add bonus points to an applicant's score if she is minority or low-income. With these adjusted scores, we re-compute the outcome of the deferred acceptance algorithm for a grid of bonus points. The plotted curves correspond to minority and low-income percentages which correspond to the given average composite score.

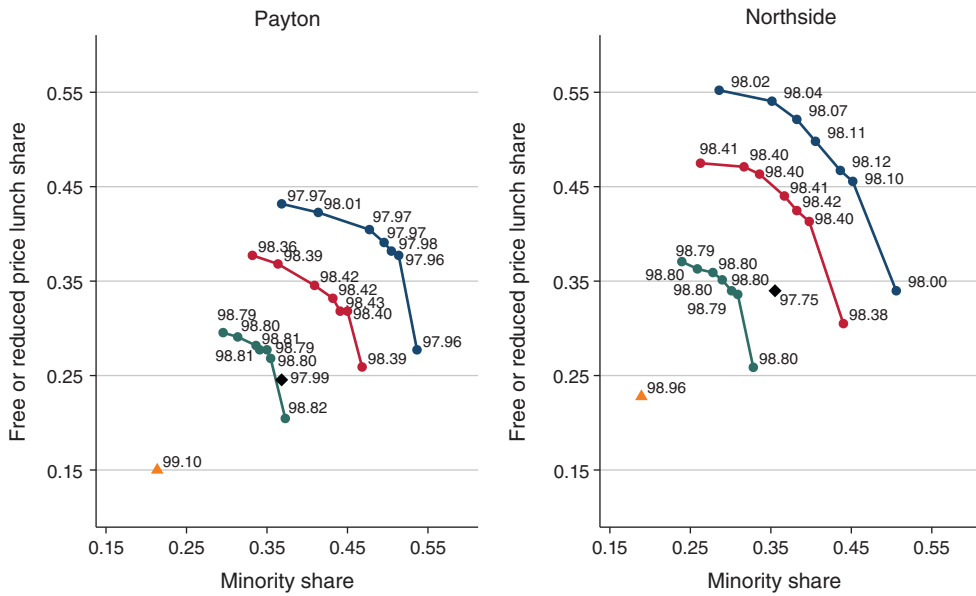


FIGURE 3. LEVELS OF MINORITY AND LOW-INCOME REPRESENTATION THAT CAN BE SIMULTANEOUSLY ACHIEVED AT VARIOUS SCORE LEVELS

Notes: This figure plots the frontiers for Payton and Northside for several levels of the composite score. The triangle corresponds to a purely score-based admission scheme. The black diamond corresponds to the CPS scheme. The green, red and blue curves represent several levels of minority and FRPL representation that can be achieved, while holding the composite score fixed at 98.8, 98.4, and 98, respectively.

the average to 98.4 and 98.0, respectively. These policies make larger changes to student demographics feasible.

The shape of the frontiers demonstrates three main findings. First, absent restrictions on the form of affirmative action, it is possible to make substantial changes in school demographics with only modest changes in the distribution of admitted students' scores. Looking at the endpoints of the innermost curve for Payton, we see that changes that keep the average score at least 98.8 can increase minority representation from 29 percent to 37 percent or increase the share of FRPL students from 20 percent to 30 percent. The pattern is similar at Northside: it is possible to increase minority representation from 23 percent to 33 percent or increase the FRPL share from 26 percent to 37 percent, while maintaining average scores of at least 98.8. Unrestricted affirmative action would therefore be a powerful tool in Chicago.

Second, under admissions policies that change composite scores as much as Chicago's race-neutral system, there is substantially greater scope to alter school demographics. The outermost curves in the figure describe assignments which maintain an average composite score of 98.0, and hence can be thought of as making changes to the composite score distribution which are about as large as CPS has made. (Recall that the average composite score under the CPS policy is 98.0 at Payton and 97.7 at Northside.) With changes of this magnitude, the minority share at both Payton and Northside could be increased to over 50 percent. Moreover, the FRPL shares could be increased to about 43 percent at Payton and 55 percent at Northside.

Third, all of the frontiers are highly concave: they are quite flat near their left endpoints and quite steep near the right endpoints. The flatness means that a modification of a pure income-preference that replaces a portion of the income preference with explicit consideration of race can substantially increase minority representation while having negligible effect on average scores and almost no effect on low-income representation. For example, the leftmost and third from left points on the outmost Payton frontier indicate that one could raise minority representation at Payton from 37 percent to 48 percent while holding the average score at Payton fixed at 98.0 and reducing subsidized lunch representation just from 43 percent to 40 percent. Similarly, the steepness of the right portions of the curve means that a modification of a pure racial-preference plan that bases a portion of the preference on income rather than race can substantially increase low-income representation while only slightly reducing minority representation. For instance, the near verticality of the outermost curve for Payton near its right endpoint means that at average composite scores of 98.0 it is possible to increase low-income representation from 27 percent to 38 percent while only reducing minority representation from 53 percent to 50 percent. Together these two features suggest that policymakers who value both racial and income diversity would presumably have a preference for policies that explicitly considered both race and income.

Implementing a policy that achieves an interior point on a curve requires explicit use of both FRPL eligibility and minority status. Assuming this is possible, policies can be implemented using simple bonus schemes. Students have x bonus points added to their score if they qualify for FRPL and y bonus points added if they are a minority. Assignments are then made simply by accepting students in the order of their bonus-adjusted scores. The ratio x/y determines where we end up between the fewer minorities/more low-income or the more minorities/fewer low-income end of each curve. The magnitudes of the bonuses will determine whether we reach a curve with a higher or lower average composite score. For example, in the case of Payton, the rightmost (53 percent minority, 28 percent subsidized lunch) point on the average-score-98 curve is obtained by giving 6.4 points for minority status and no points for being low-income. The leftmost (37 percent minority, 43 percent subsidized lunch) point on the same curve is obtained by giving 6.9 points for low-income status and no points for being a minority. The allocation two points to the right on the same curve where minority representation is substantially higher (48 percent) and subsidized lunch representation is only a little lower (40 percent) are obtained by giving 2.6 points for minority status and 5 points for subsidized lunch eligibility, respectively.

B. *Within-School Heterogeneity in Race, Income, and Composite Scores*

A separate concern about affirmative action policies is that they may contribute to within-school achievement gaps that adversely affect minority students for reasons, e.g., stereotype formation, that we did not try to incorporate into our model.¹⁹ This might also be thought to be an advantage of race-neutral plans: when

¹⁹ See, e.g., Steele and Aronson (1995) and Austen-Smith and Fryer (2005) for further discussion.

preferences are given only to minority students, it is clear that the lowest scoring students in a school will be minorities. But, Ray and Sethi (2010) note that race-neutral plans do not lower achievement gaps when optimally designed. When the distribution of scores for minorities and non-minorities differ, an optimal plan admits students from a region of score distribution heavily populated by minorities, but below the region for the remaining students, which increases the mean achievement gap. In this section, we illustrate how the gap in composite scores is affected as more seats are reserved for minorities under a quota system and with CPS's shift to a race-neutral policy. We find that the Ray and Sethi (2010) concern is not purely academic. At any given level of minority representation the majority-minority achievement gap is wider under CPS's tier-based plan than it would be under a race-based plan.

CPS's race-neutral admissions policy has created substantially more heterogeneous student bodies, measured by race or income, at Payton and Northside than a pure score-based admissions process. The histograms in panel A of Figure 4 show what the composite score distribution would be at Payton and Northside under score-based admissions with no set-aside. Each histogram bar is divided into two portions: the lower dark gray and red portion reflects the number of underrepresented minority students with composite scores in the band and the upper light gray and red portion reflects the number of White or Asian American students. The upper histograms show that within-school racial achievement gaps are small under a purely score-based admissions policy. Without affirmative action, there would be little within-school variation at either school: about 60 percent of students at each school would have composite scores of at least 99 and all students would have scored at least 97. The average score gap between majority and minority students is 0.4 at Payton and 0.2 at Northside. The average FRPL score gap is 0.4 at both schools.

The second row of histograms in the figure illustrate the class composition at Payton and Northside under benchmark affirmative action policies that use race and FRPL data to make Payton and Northside as diverse as they are under the current CPS policy on both dimensions. Note that a bonus scheme in which students are assigned a school-independent number of bonus points for minority and FRPL status cannot simultaneously match the current demographic composition of both Payton and Northside. However, it seems unlikely that CPS would choose to implement a scheme with school-based bonuses. Hence, we have chosen throughout to report on two separate race and FRPL-based benchmarks. The Payton benchmark gives 2.9 points for minority status and 1.3 points for FRPL, while the Northside benchmark uses bonuses of 2.8 and 1.5 points, respectively.

There is relatively little increase in within-school heterogeneity under these benchmarks. The histograms in the second row show that achievement gaps would be only moderately larger if CPS had used these plans to bring minority and FRPL representation up to their current levels. For example, the left histogram in Panel B corresponds to a policy in which 36.8 percent of the students assigned to Payton are underrepresented minorities and 24.5 percent are eligible for FRPL. These benchmark plans replace many of the students with scores between 97 and 98.5 with minority and/or FRPL eligible students with scores between 95 and 96, but admit few students with scores under 95. The average majority-minority score gap is 1.8 at both schools, and the average FRPL gap is 1.5 at Payton and 1.4 at Northside.

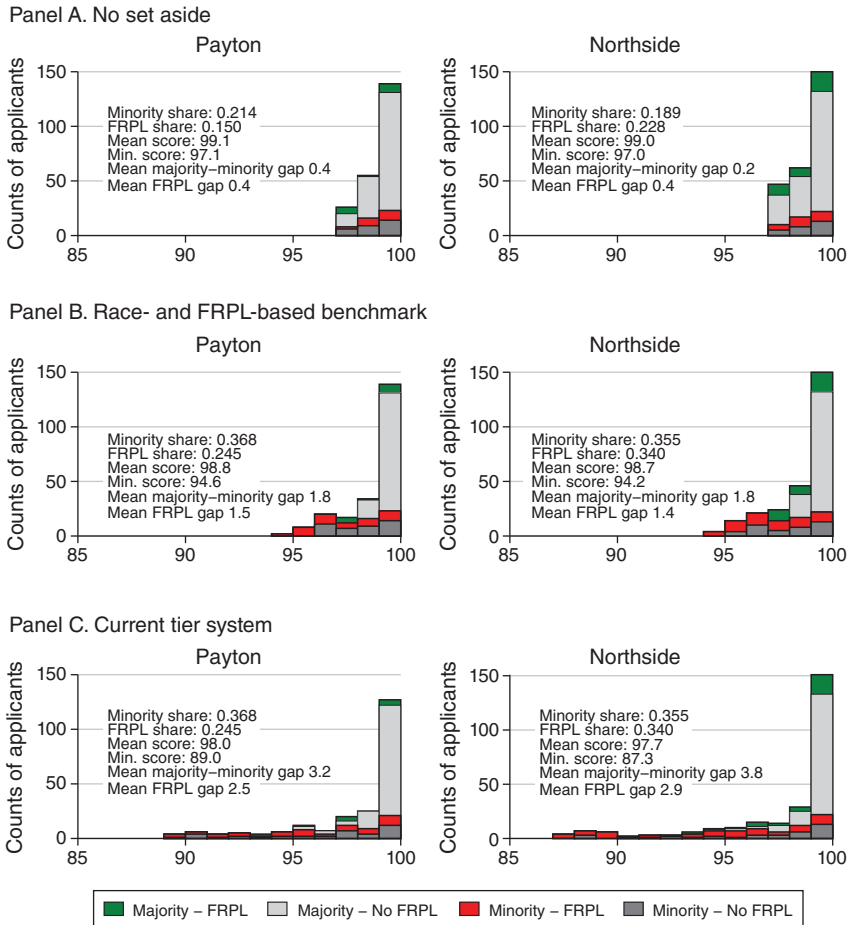


FIGURE 4. WITHIN-SCHOOL SCORE DISTRIBUTIONS UNDER ALTERNATE ADMISSIONS POLICIES

Notes: Each histogram bar is divided into four portions: green and light gray reflect the number of non-minority students and magenta and dark gray reflect the number of minority students. The subportions in green and magenta represent FRPL students in the respective subportion. Score refers to the composite score.

Switching to CPS’s tier-based affirmative action affects within-school achievement gaps in two opposing ways. One effect that works to reduce racial gaps is that the policy admits some relatively low-scoring White and Asian American students who live in low SES neighborhoods. Working in the opposite direction, however, is the fact that a number of high-scoring minority students living in medium to high SES neighborhoods are now being denied admission and that almost all of the lowest-scoring admits (all of whom come from tier 1) are minorities. Given the opposing effects, it is not a priori clear how the shift to a race-neutral policy will affect within-school gaps.

Panel C shows that the tier-based plan increases the within-school racial gaps. Under the race- and FRPL-based benchmark, the majority-minority score gap in average scores is 1.8 points at both Payton and Northside. Under the CPS plan, it is 3.2 points at Payton and 3.8 points at Northside. (As before, the larger gap

at Northside reflects that the CPS plan is less efficient at Northside.) A comparison of the histograms in the second and third rows of Figure 4 brings this out in more detail. At both schools, the shift to the race-neutral plan reduces the number of majority students with scores in the 98 to 99 range. The CPS plan does add a large number of majority students with scores between 95 and 97, but almost all of the added students with scores below 95 are underrepresented minorities.

The most striking difference is that the left tail of low-scoring students now extends down to 89 at Payton and 87 at Northside. This means that teaching and curricula in the two schools must accommodate a substantially more heterogeneous student population relative to an affirmative action policy which uses race and FRPL status. The histogram also shows that the CPS policy has displaced most students with composite scores of between 98 and 99. As a result, the classes at Payton and Northside have an unusual composition: the majority of students have scores of at least 99, but the rest are spread out over a wide range. Since the policies in panel B and C result in the same level of diversity, we conclude that most of the increase in within-school heterogeneity could have been avoided if there were no restrictions on the type of affirmative action which could be practiced.

Within-school majority-minority achievement gaps are larger under the tier-based policy for any fixed level of minority representation. This fact can be seen in Figure 5, which plots the difference between the average composite score of admitted majority and minority students under CPS-like and pure racial-preference policies for each level of minority representation. The solid lines illustrate racial quota policies varying the numbers of reserved slots, with blue circles for Payton and red triangles for Northside. At the left endpoint, the curves have a value of 0.2, which reflects the small difference in average scores of admitted majority and minority students under a purely score-based admissions process. Achievement gaps grow by about 1 point for every 10 points of minority share: they are slightly above 1 point when the school is 30 percent minority, around 2 points when the schools are 40 percent minority, and around 3 points at 50 percent minority. The dashed lines give corresponding numbers for policies based on the CPS tiers: for each desired level of minority representation, we find the CPS tier reservation percentage that produces that level of minority representation and report the majority-minority score gap under that policy. At Payton, the majority-minority gaps are small at first, but grow once enough seats are allocated by tier to increase minority representation beyond 30 percent. The gaps grow even more quickly at Northside, again illustrating that the lower efficiency of the CPS plan in the Northside environment is associated with a larger majority-minority score gap. Compared to a racial quota, the CPS policy roughly doubles the magnitude of the majority-minority score gap at Northside.²⁰ Figure 5 also reports the majority-minority gap in scores for the twenty-fifth percentile of the distribution. The gap at this quantile are wider than the gap at the mean at both Payton and Northside.

²⁰While we focus on the effect of CPS's race-neutral affirmative action system on Payton and Northside, CPS's tier-based plan produces a higher SD at every school relative to the benchmark. The effects are largest at Jones and Young, which are the next two schools in the selectivity hierarchy following Payton and Northside. Online Appendix Figure A.1 presents histograms similar to those in Figure 4 illustrating the impact of the CPS plan on the other schools.

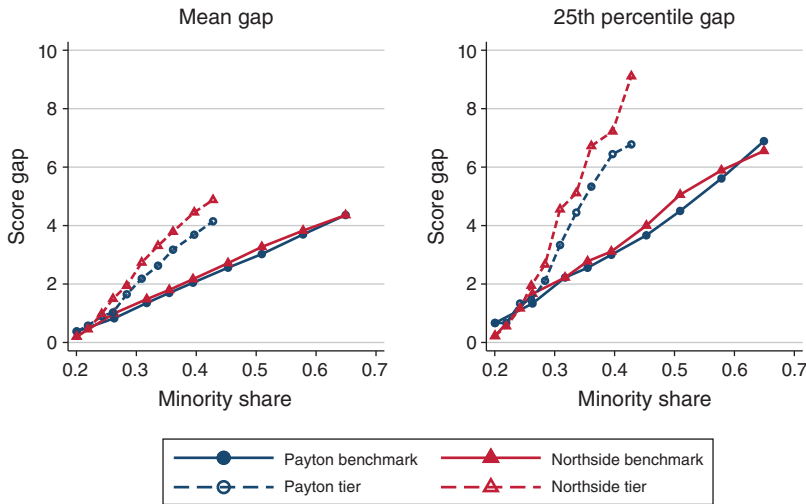


FIGURE 5. WITHIN SCHOOL GAPS IN AVERAGE COMPOSITE SCORES: QUOTAS VERSUS CPS TIER-BASED PLAN

Notes: Each point plots the composite score gap between majority and minority students for a given admissions policy. In the left panel, the score gap is measured as the difference in mean majority and mean minority scores. The right panel considers the gap at the twenty-fifth percentile of each distribution of scores. The solid line represents benchmark policies, varying the magnitude of the weight for minority status. The dashed line represents CPS tier policies, increasing the percentage of reserved seats. Blue circles indicate Payton and red triangles indicate Northside.

C. Targeting Disadvantaged Applicants

Compared to the benchmark, the CPS plan must be rejecting some students with high composite scores and instead accepting lower-scoring students with the same minority-FRPL status. In this section, we investigate further to what extent admissions probability depends on student characteristics and examine the attributes of students who are displaced.

Figure 6 reports the probability of admissions for students stratified by FRPL and minority status at Payton and Northside under the CPS policy. The vertical line in each panel corresponds to the lowest composite score of an admitted student under the race- and FRPL-based benchmark plan. That is, at Payton, any minority and FRPL student with a composite score of at least 94.5 would have been guaranteed admission under this alternative. In contrast, under the CPS policy, students with scores of 89 have a positive probability of being admitted. Their positive probability comes at the expense of higher scoring minority/FRPL students who do not obtain admissions for certain. At Northside, minority/FRPL students with scores of 87 have a positive admissions probability, when the lowest score to gain admittance under the benchmark would be 94.2, and the admissions probability for students with scores 90–95 is less than 0.5. The positive probability to the left of the vertical line, and the fact that points to the right of the vertical are not equal to 1 provide another visualization of the inefficiency of the CPS plan.

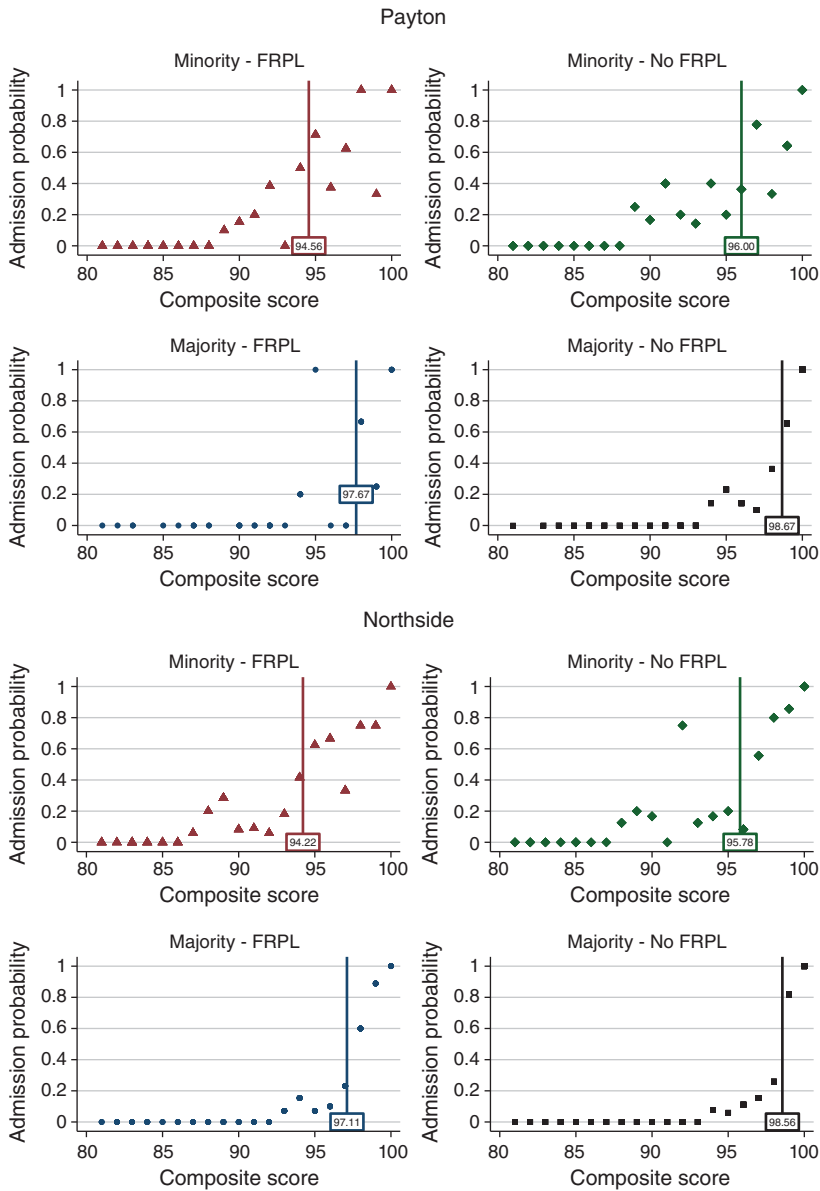


FIGURE 6. PROBABILITY OF ADMISSION AND COMPOSITE SCORE

Notes: This graph shows the probability of admission under the current CPS tier plan by composite score bucket, separately for four groups of minority and FRPL combinations. The vertical line represents the lowest composite score among students who would have been admitted under the race- and FRPL-based benchmark plan, subtracting the bonus points in the respective group.

The figure shows that the CPS’s plan does not simply admit the wrong minority-FRPL students. The other three panels illustrate that the CPS plans admits students with lower composite scores than the benchmark plan, and it also does not admit all pupils with higher scores. For example, even majority/non-FRPL students

with scores of 94 are admitted with positive probability at Payton and Northside, even though they do not increase diversity and have substantially lower composite scores than the composite score these types of applicants would have faced under the benchmark plan.²¹

III. Measuring the Efficiency of Affirmative Action Plans

A. Modeling Affirmative Action in Admissions

Several models show how race-based affirmative action can enhance welfare.²² In these models, restrictions on the use of race will always lead to welfare losses. Chan and Eyster (2003) first formulated this finding in a model in which an elite school values both diversity and the average quality of admitted students. They also noted that welfare losses can be severe and take on striking forms: constrained optimal plans can deny admission to some of the most-qualified students of every race; they can reject some minority students who are more qualified than some admitted majority students; and they can reduce average student quality as well as minority representation. Ray and Sethi (2010) note that absent the monotonicity restriction imposed by Chan and Eyster (2003) optimal policies will generically be non-monotone and accept some less-qualified students while simultaneously denying admission to some who are more qualified.

Consider a school system that serves a heterogeneous set of students. Each student has a type vector $(\theta, \mathbf{x}, \mathbf{z})$, where θ describes the best curriculum for the student, \mathbf{x} is a vector of student characteristics (such as race or socio-economic status), and \mathbf{z} is “proxy” that may be used in assigning students to schools. Suppose that the expected educational outcome of a student of type $(\theta, \mathbf{x}, \mathbf{z})$ when assigned to school s is

$$(1) \quad V_s(\theta, \mathbf{x}, \mathbf{z}) = h(\theta, \mathbf{x}, \mathbf{z}) - k(\theta - c_s)^2 - d \|\bar{\mathbf{x}}_s - \mathbf{x}^*\|,$$

where c_s is the curriculum at school s , $\bar{\mathbf{x}}_s$ is the mean of the vector of characteristics of students in school s , and \mathbf{x}^* is the composition of an optimally diverse school. The function h gives the component of a student’s expected outcome which does not depend on school s . The parameter k indexes the importance of providing students with a curriculum that is matched to their type θ . The parameter d indexes the importance of losses from schools having demographics that differ from \mathbf{x}^* . This loss term might reflect the value of discussions in diverse classrooms. An optimal affirmative action plan balances curriculum matching against concerns for diversity. An important assumption in this model is that the education benefits to diversity and optimal curriculum matching do not depend on whether the student is in the majority or minority group.

²¹ Figure A.2 in the online Appendix reports additional information on the tract characteristics of students displaced under the CPS plan. Almost all of the displaced students come from tracts that are above the median in the CPS SES index in its income and education components.

²² Coate and Loury (1993) develop a model based on statistical discrimination and skill investment, Athey, Avery, and Zemsky (2000) and Chung (2000) develop frameworks with mentoring and role models, and Fu (2006) considers investment incentives with exogenous asymmetries.

Suppose that the school system operates schools indexed by $s = 1, 2, \dots, S$. Assume that the school system chooses both a student assignment function $A : \Theta \times X \times Z \rightarrow S$ and the curricula at each school $\mathbf{c} = (c_1, c_2, \dots, c_s)$.²³ When affirmative action is unrestricted, we assume that the school system knows the distribution of student types and can choose any function $A(\theta, \mathbf{x}, \mathbf{z})$ and any curricula \mathbf{c} . Suppose the social welfare function aggregates student outcomes as follows:

$$(2) \quad W^{A,c} \equiv \sum_s \int_{\{\theta, \mathbf{x}, \mathbf{z} | A(\theta, \mathbf{x}, \mathbf{z}) = s\}} (h(\theta, \mathbf{x}, \mathbf{z}) - k(\theta - c_s)^2 - d \|\bar{\mathbf{x}}_s - \mathbf{x}^*\|) d\mu(\theta, \mathbf{x}, \mathbf{z}),$$

where μ is the distribution over types. The unrestricted optimal assignment policy thus maximizes:

$$\{A^*(\theta, \mathbf{x}, \mathbf{z}), \mathbf{c}^*\} = \arg \max_{A(\theta, \mathbf{x}, \mathbf{z}), \mathbf{c}} W^{A,c}.$$

Our utilitarian welfare function implicitly assumes that the motivation for maintaining diversity is that it affects educational outcomes of students in the system. But the same social welfare function could also capture situations where diversity benefits accrue to others, e.g., they could be realized when the current generation of students serves as role model for future students as in Chung (2000), or they could reflect the preferences of voters or politicians for diverse schools. The online Appendix develops some properties of the optimal assignment policy. The most important result is that when each school sets its curriculum optimally, school-specific welfare simplifies to

$$W_s^{A,c} = -k \text{Var}(\theta | A(\theta, \mathbf{x}, \mathbf{z}) = s) - d \|\bar{\mathbf{x}}_s - \mathbf{x}^*\|.$$

That is, welfare is determined by the combination of within-school variance in the curricula to which students are best matched and demographic diversity.

B. Measuring Relative Efficiency

The optimal admissions plan will often be infeasible for two reasons: (i) schools may be legally prohibited from basing admissions decisions on some dimension \mathbf{x} and (ii) schools may not observe some dimensions of \mathbf{x} . When this happens, school systems can only implement rules $A(\theta, \mathbf{z})$ that involve proxy variables \mathbf{z} imperfectly correlated with \mathbf{x} .

The difference in welfare from adopting a given race-neutral plan $\{\hat{A}(\theta, \mathbf{z}), \hat{\mathbf{c}}\}$ rather than $\{A^*(\theta, \mathbf{x}, \mathbf{z}), \mathbf{c}^*\}$ is $W^{A^*,c^*} - W^{\hat{A},\hat{c}}$. Several factors make estimating this quantity infeasible. Computing the optimal policy and evaluating the welfare function requires knowing what dimensions of diversity matter (to student achievement or school-board preferences) and estimates of the importance of curriculum mismatch and diversity for student achievement. Therefore, we take an alternative

²³Chan and Eyster (2003) and several subsequent papers consider a more general class of plans which may involve random assignment. The CPS plan was deterministic and we simplify the discussion that follows by only considering such plans.

approach that compares a given race-neutral policy to a well-defined race-based benchmark policy.

One school assignment plan, available regardless of whether it is legal to condition on x , ignores both x and z and chooses the assignment plan that minimizes the average within-school variance in θ . Other plans involving affirmative action can improve diversity, but improvements come at the expense of increasing within-school variance in θ . The relative efficiency measure we define below focuses on how well plans make this trade-off, i.e., how much of a loss in curriculum matching is incurred for each unit of improvement in diversity.

Let \hat{s} be a school of interest. Let $\{\hat{A}(\theta, x, z), \hat{c}\}$ be some assignment plan of interest in which the curriculum is optimal with respect to the school assignments.²⁴ Suppose that \hat{A} assigns n_s students to each school s and write $\hat{x}_{\hat{s}}$ for the average demographics of school \hat{s} under this plan. To define our relative efficiency measure for plan \hat{A} as school \hat{s} , we first construct two benchmark assignment plans that serve as points of comparison. First, let A_{n_1, \dots, n_s}^{SB} be the “score-based” admission plan in which students are prioritized in order of θ and assigned in order to their most-preferred school provided it is not full given the capacities n_1, \dots, n_s . Under the score-based admissions plan, a student is assigned her most preferred school among those with available capacity when it is her turn to choose. Next, let $A_{\hat{s}, n_1, \dots, n_s, \hat{x}_{\hat{s}}}^{RB}$ be the “race-based” affirmative action plan which prioritizes students on the basis of θ plus an underrepresentation bonus and provides members of groups which would be less represented at school \hat{s} under plan A^{SB} with the smallest number of bonus points that are sufficient to make the demographics of school \hat{s} match the demographics, $\hat{x}_{\hat{s}}$, that it has under plan \hat{A} . We then define the *relative efficiency* of school assignment plan \hat{A} at school \hat{s} by

$$(3) \quad \text{Relative Efficiency}(\hat{A}, \hat{s}) = \frac{\text{SD}(\theta | A_{\hat{s}, n_1, \dots, n_s, \hat{x}_{\hat{s}}}^{RB}(\theta, x, z) = \hat{s}) - \text{SD}(\theta | A_{n_1, \dots, n_s}^{SB}(\theta, x, z) = \hat{s})}{\text{SD}(\theta | \hat{A}(\theta, x, z) = \hat{s}) - \text{SD}(\theta | A_{n_1, \dots, n_s}^{SB}(\theta, x, z) = \hat{s})}.$$

Relative efficiency can be thought of as measuring the fraction of the welfare losses from inferior curriculum-matching at school \hat{s} , which were necessary to achieve the level of diversity which plan \hat{A} achieves. The measure compares \hat{A} to another plan which uses bonuses based on x to achieve the same level of diversity. For example, if plan \hat{A} increases the standard deviation of θ within school \hat{s} by 4 units for every 0.01 increase in the representation of some group and the alternate plan A^{RB} increases the standard deviation of θ within school \hat{s} by only 1 unit for every 0.01 increase in the representation of that group, then we would say that plan \hat{A} is only 25 percent efficient at school \hat{s} . That is, 25 percent of the loss in curriculum-matching benefits that plan \hat{A} entails were necessary to achieve the diversity benefits that plan \hat{A} achieves while the remaining 75 percent were due to the restriction that prevents using x in the admissions process and/or to plan \hat{A} having been suboptimally designed relative to the constraints.

²⁴We will make this assumption about curricula throughout this section and henceforth describe assignment plans solely by their assignment functions.

School-specific welfare depends on within-school variance in curricula and demographic diversity. The score-based benchmark is a natural point of comparison because the equivalence of minimizing within-school standard deviation and maximizing across school standard deviation in mean scores means that a plan which considers only θ in admissions can be thought of as a plan designed to roughly maximize one dimension of welfare. The race-based benchmark, which focuses on maximizing average scores at school \hat{s} subject to a diversity constraint, is another natural point of comparison if we think of trying to reduce the system-wide standard deviation given a diversity constraint.²⁵

Cardinalization of the welfare losses in curriculum matching that have been incurred in the course of increasing diversity involves taking a position on the appropriate scale. While our model suggests using the within-school variance in composite scores as a measure of welfare losses, we use the within-school standard deviations, $SD(\theta | \hat{A}(\theta, x, z) = \hat{s})$, instead. We make this choice to be more conservative in labeling plans as inefficient, cognizant of the fact that the model-based argument for using variance instead would be relying heavily on the assumption that welfare losses are *quadratic* in the mismatch in composite scores. We are also trying to be conservative in omitting other potential welfare losses from our calculation. For example, we noted that the CPS plan has much larger majority-minority score gaps than our benchmark plans, and the impact that this may have on the formation of stereotypes is another welfare concern that might be regarded as making the CPS plan even less efficient than it appears in our measures.

We also report a measure of efficiency using the mean composite score in the online Appendix. For a given admissions plan, this measure takes the ratio of the difference in the average score under the optimal plan for corresponding level of diversity compared to purely score-based admissions to the difference in the average score under the plan minus the average under purely score-based admissions. That is, this metric simply replaces the standard deviation in equation (3) with the mean. In many cases, the standard deviation measure and the mean produce similar efficiency measures, so we opt to emphasize the standard deviation measure. The similarity is not surprising. For example, if all students have a score of 100 under the score-based plan, plan A has a fraction $1 - p$ of students having a score of 100 and fraction p having a score of $100 - y$, and plan A 's benchmark has a fraction $1 - p$ at 100 and fraction p at $100 - x$, then the mean and standard-deviation based definitions both measure the efficiency as x/y .

In the discussion above, we have defined the relative efficiency of a plan \hat{A} considering all demographic changes it entails. In practice, affirmative action plans change the demographics of schools in multiple dimensions: they affect the fraction of minority students, the fraction of low-income students, etc. Our primary measure, which we will refer to as overall efficiency, considers benchmark plans A^{RB} that achieve exactly what plan \hat{A} achieves on both of these dimensions. To better

²⁵The plan that minimizes within-school standard deviation in θ just at school \hat{s} , in contrast, would not be a sensible benchmark: one can reduce within-school standard deviation at any school almost to zero simply by denying admission to all of the highest-achieving students and instead accepting students from some point in the middle of the distribution where there are many students from every demographic group. Our benchmarks are designed to avoid the concern that one is achieving a favorable result at school \hat{s} at the expense of increasing within-school variance at other schools.

understand a plan's performance, we also separately report plans' relative efficiency as a means to improve single dimensions of diversity. For example, if plan \hat{A} increases the representation of minority students at \hat{s} from 15 percent to 30 percent and increases the representation of low-income students from 20 percent to 25 percent, we will measure the "minority-only efficiency" of the plan by a similar formula, but using in place of A^{RB} a benchmark that maximizes scores at school \hat{s} imposing only the constraint that minority enrollment must be at least 30 percent. We will measure the "income-only efficiency" by using in place of A^{RB} a benchmark that only imposes the constraint that low-income enrollment must be at least 25 percent. At Payton, the bonus points for race under the minority-only benchmark are 3.4 points and the bonus points for FRPL under the income-only benchmark are 2.9. At Northside, the bonus points for race under the minority-only benchmark are 2.8 and the bonus points for FRPL under the income-only benchmark are 2.1.

IV. How Efficient Are Chicago's Policies?

Our efficiency calculation evaluates an affirmative action policy in terms of the portion of the increase in within-school variation in θ that was necessary to achieve the level of diversity which is being achieved. For our efficiency calculations, we will take θ to be the composite score. In terms of histograms in Figure 4, the actual increase in standard deviation (SD) can be thought of as calculated by computing the difference in SD between panel C and panel A for each school. The necessary increase in SD would be the difference between the SD of rescaled versions of the panel B and panel A for each school. Here, this calculation says that the current CPS policy has an overall efficiency of 28 percent at Payton and 22 percent at Northside. Put differently, the CPS policy has increased within-school SD in composite scores by about four times as much as was necessary to achieve the diversity that is being achieved. We more directly study the CPS's plan as a tool for racial and socioeconomic integration.

A. *The Policy as a Tool for Racial Integration*

The CPS policy departs from our race- and FRPL-based benchmark policy because it does not use student-level race data to increase minority representation, nor does it use student-level income data (measured by subsidized lunch eligibility) to increase low-income representation. Figure 7 examines the trade-off between increasing minority representation and reducing within-school SD in student ability/preparation at Payton and Northside. Specifically, the figure describes the classes that would be admitted to each school under various affirmative action plans on two dimensions. The x -axis is the percent of students who would be underrepresented minorities. The y -axis gives the within-school SD in composite scores. The y -axis is plotted on a flipped scale so that one can regard higher points on either axis as desirable. The left panel corresponds to Payton and the right panel to Northside.

The red circles describe the outcomes under a set of benchmark race-based policies which provide varying levels of advantage to underrepresented minority students. The upper left point on these curves correspond to the purely score-based

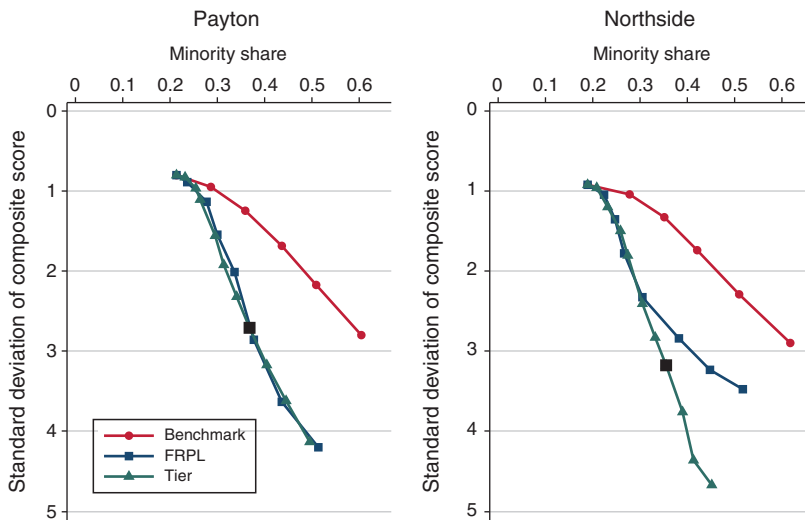


FIGURE 7. FEASIBLE COMPOSITE SCORE SPREAD AT VARIOUS LEVELS OF MINORITY REPRESENTATION

Notes: This figure compares the standard deviation of composite scores for several levels of minority representation. The upper left point of each curve corresponds to a purely score-based policy. Red circles correspond to policies that give bonus points for minority status, increasing by 1.5 point increments. Blue squares correspond to the analogous policy which instead rewards FRPL status. Green triangles represent simulations of admission schemes that reserve various fractions of seats to be allocated evenly across CPS tiers, increasing by 10 percent increments. The black square corresponds to the current CPS policy with 70 percent of seats reserved.

policy, which produces classes in which 21.4 percent of Payton students and 18.9 percent of Northside students are underrepresented minorities. The points further to the right on these curves correspond to increasing the minority preference: the first dot corresponds to a minority preference of one point, the second to a minority preference of two points, and so on. The flatness of the left portions of the curves is a further illustration that race-based affirmative action policies can substantially increase minority representation while only making small changes in the distribution of student preparation. The increasing steepness of the curves as we move to right shows that there is limited scope for making such low-cost changes. Larger increases in minority representation will require more substantial increases in within-school SD.

The green triangles describe the incoming classes when one implements CPS-like policies that reserve various fractions of the seats at Payton and Northside to be allocated evenly across the CPS tiers. The upper left point on these curves is the pure score-based admissions policy. Each subsequent point corresponds to allocating an additional 10 percent of the seats by tier. The large black square is the current CPS policy which involves a 70 percent reservation for tier seats and produces a class that is 37 percent minority. The blue squares describe the income classes under a set of FRPL-based benchmark policies, where bonus points are given for low-income status.

The vertical difference between the two curves provides another illustration of the inefficiency of the CPS policy. Our relative efficiency measure applied to this solely race-based comparison (i.e., minority-only efficiency) would say that the CPS policy is 26 percent efficient at Payton, i.e., the actual within-school SD relative to the score based policy $2.71 - 0.80 = 1.91$, is $1/0.258 = 3.9$ times as large

as the increase in SD ($1.29 - 0.80 = 0.49$) sufficient to achieve the same level of minority representation under the benchmark race-based policy.

The CPS policy is even less efficient as a means to increase minority representation at Northside. In the right panel of Figure 7, the green and red curves diverge even more quickly than in the left panel. The current CPS policy is only 19 percent minority-only efficient at Northside. The minority-only efficiency of the FRPL benchmark is close to that of the CPS policy at Payton, but at Northside the FRPL benchmark is midway between the race-based benchmark and the CPS policy at minority shares above 35 percent. The difference between the policies' performance at Payton and Northside is driven in part by the demand patterns of applicants to both schools. That is, place-based affirmative action plans may be less effective in environments similar to Northside than in environments similar to Payton, since Payton naturally draws students from nearby minority neighborhoods.

Note that the degree to which a social planner should care about relative efficiency would vary with the size of the change in school composition that is being contemplated. Near the left endpoint, low relative efficiency would not be a big concern because the increase in within-school SD (which is what enters into welfare) is still fairly small in absolute terms even for an inefficient policy. The form of the graphs, however, bolsters the case that relative efficiency is a useful concept in showing that it provides a measure which is somewhat robust to how intensively the policy level is used. For example, if one were to analyze the CPS-like policies which involve reserving 30 percent to 80 percent of the seats at Payton by CPS tier, one would find that they are all 26 percent to 33 percent efficient as means to increase minority representation.

B. *The Policy as a Tool for Socioeconomic Integration*

The nature of CPS policy's inefficiency as a means for admitting low-income students differs from its inefficiency as a means for admitting minority students since CPS faced a legal prohibition on using race in admissions, whereas it could have used FRPL eligibility if it had chosen to do so.

We nonetheless study the CPS policy's efficiency as a means for admitting low-income students for two reasons. First, many other cities and states have adopted place-based affirmative action policies, so it is of more general interest to know how efficiently one such policy is aiding low-income students.²⁶ Second, while the "low-income" inefficiency of the CPS policy could be eliminated by incorporating student-level data on FRPL eligibility, socioeconomic disadvantage involves much more than income. CPS does not have the student-level data to measure these other dimensions of disadvantage. For example, CPS does not know if a student's parent(s) are supportive of their children's education; how many books they have in their home; whether the student has been affected by domestic or neighborhood violence; whether the student's teacher encouraged them to pursue high achievement; etc.²⁷ Place-based

²⁶Texas's well-known top 10 percent policy for college admissions can be thought of as one example of a much coarser place-based policy which puts all students who attend the same public high school in an equivalence class.

²⁷The breadth of disadvantage motivates the multifaceted approach to addressing disadvantage in the Harlem Children's Zone studied in Dobbie and Fryer (2011). Fryer and Levitt (2004) show that the number of books in the

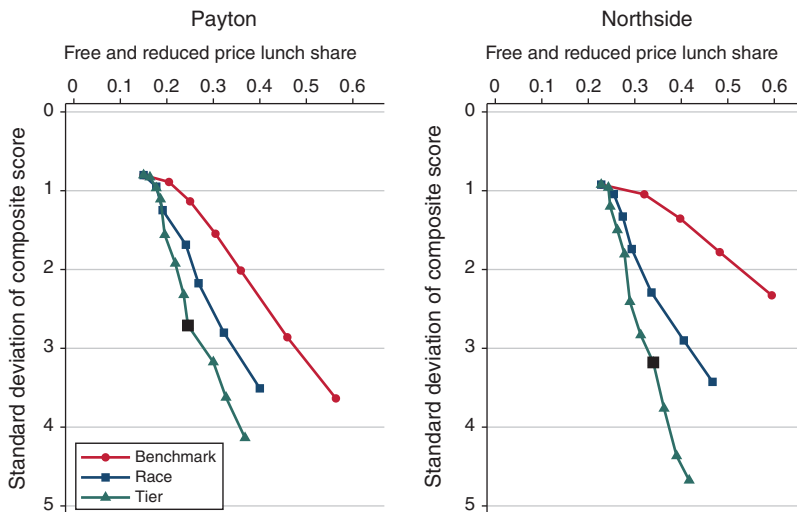


FIGURE 8. FEASIBLE COMPOSITE SCORE SPREAD AT VARIOUS LEVELS OF FREE OR REDUCED PRICE LUNCH ELIGIBILITY

Notes: This figure compares the standard deviation of composite scores for several levels of low income representations. The upper left point of each curve corresponds to a purely score-based policy. Red circles correspond to policies that give bonus points for FRPL status, increasing by 1.5 point increments. Blue squares correspond to the analogous policy which instead rewards minority status. Green triangles represent simulations of admission schemes that reserve various fractions of seats to be allocated evenly across CPS tiers, increasing by 10 percent increments. The black square corresponds to the current CPS policy with 70 percent of seats reserved.

policies are a possible way to address such disadvantages, and the lessons learned from examining the degree to which the CPS policy is benefiting low-income students may also apply to the effectiveness of place-based policies in aiding students who are disadvantaged on other dimensions on which schools (and we) lack data.

Figure 8 reports on the trade-off between increasing low-income representation and reducing within-school composite score SD at Payton and Northside. The figure describes the classes that would be admitted to each school under various affirmative action plans. Figure 8 differs from Figure 7 since the x -axis reports the fraction of students eligible for FRPL instead of the number of minority students.

As with racial diversity, bonus schemes can substantially increase low-income representation from its initial level with only small changes in the composite score distribution. The red circles in Figure 8 illustrate the trade-off that CPS faces when it gives bonus points to low-income students to increase their representation. The curves are nearly horizontal close to their left endpoints. This fact shows that it is possible to substantially increase low-income representation with only a slight change in the distribution of composite scores. But, the steeper slopes at higher levels of low-income representation show that more substantial increases in within-school variation are inevitable, even for levels of representation well below the CPS average FRPL eligibility of 86 percent. Comparing the two schools, Northside will have

home is a powerful predictor of student performance at the beginning of kindergarten and including this and a host of other controls can account for the racial achievement gap at that time.

substantially more low-income students than does Payton under any given level of low-income preference.

The green triangles describe the student bodies produced by variants of the CPS policy with different tier reserve sizes. Despite its focus on socioeconomic disadvantage, the CPS plan is not very efficient as a method for increasing the number of low-income students. At Payton, for example, the current CPS policy increases the proportion of students on FRPL from 15 percent to 25 percent at the cost of increasing the SD in composite scores from under 1 to almost 3. A simple FRPL quota could have achieved the same increase in subsidized lunch representation while only increasing the SD to about 1.1 SD. Our income-only efficiency metric says that the CPS policy is only 17 percent efficient as a method of increasing low-income representation. We take this as a strong cautionary message. Place-based affirmative action policies will provide a substantial benefit to students living in disadvantaged areas. But it is not at all clear that the benefit will go to students who are themselves disadvantaged.

The third series in Figure 8 provides another striking comparison: the CPS policy is worse than using racial quotas for low-income representation. The blue squares describe the classes that would have been admitted under a purely race-based plan that only explicitly favors students who belong to an underrepresented minority group. Strikingly, this graph lies *above* the CPS plan in the figure. That is, for every level of low-income representation, the CPS policy is actually worse than using racial quotas.

The fact that CPS's SES-based policy is less efficient than racial quotas as a means to increase low-income representation may at first seem paradoxical. CPS's SES index is based on several variables known to be highly correlated with poverty. How could it possibly be less efficient as a means to admit low-income students than a pure racial preference?

One explanation involves the fact that once we learn that a student from a low-SES census tract has an extremely high composite score, it means that the student is unusual for their census tract. This fact means that the relationship between tract and student characteristics is less reliable exactly for these kids. For example, most Black students in CPS are poor. An even higher percentage of students from very low-SES census tracts are poor. But once we learn that a Black student and a student from a low-SES census tract achieved a very high composite score, it is likely that neither one is poor. The reason racial quotas outperform, however, is that minority status retains more of its power as a predictor of poverty than the tract-level SES score.

The fact that the CPS plan is less efficient as a means to admit low-income students than racial quotas is a strong criticism of place-based plans like Chicago's. It has been recognized that plans that do not explicitly consider race will be less effective as a means to increase minority representation. It has been generally presumed that this disadvantage should be thought of as part of a trade-off, wherein race-neutral plans will benefit students who are themselves disadvantaged but not necessarily minority. Our analysis of the CPS plan shows that this potential offsetting benefit need not exist: the CPS place-based plan is both less efficient as a means to increase minority representation *and* less efficient as a means to benefit low-income students than a race-based plan.

V. How Efficient Can Race-Neutral Policies Be?

So far we have compared the CPS plan with benchmark race-based policies which are infeasible in the current legal/political environment. In this section, we explore the extent to which CPS might improve on its plan given current restrictions on the use of race. We do not construct the constrained-optimal plan: defining such a plan would require knowledge of how CPS values diversity and within-school variance and an understanding of political constraints. We instead pursue two more limited aims. Section VA discusses the extent to which small and potentially feasible changes to the CPS plan would improve minority-only efficiency. Section VB attempts to provide an upper bound on the efficiency CPS could have achieved by going beyond what might be legally or practically feasible, but still not directly using race.

A. Can Simple Changes Make the CPS Plan More Efficient in Increasing Minority Representation?

Several aspects of the CPS plan seem crude. The census tract SES index is an unweighted sum of six variables when some variables may be more important than others. The continuous tract-level CPS measure, which takes on 800 different values, is discretized into four bins before being used in the assignment process, suggesting it falls short of exploiting the full range of available information. And the plan makes no use of student-level data on FRPL eligibility. The decision not to make use of individual-level data potentially reflects the belief that SES tier is predictive of disadvantage, but it could alternately reflect a feasibility or political constraint.²⁸ We therefore explore whether two simple plans can make the CPS plan more efficient as a means to increase minority representation: one using only data CPS is now using and one also using individual-level FRPL data.²⁹

One class of potentially feasible modifications to the CPS plan constructs the SES index as a weighted average of any subset of the disadvantage indicators CPS is now using, and then gives preferences to students as a linear function of the continuous SES index. A rough intuition for what might make this type of plan work well is if the SES index is highly correlated with minority status among students with high composite scores. To find weights, we regressed an indicator for minority status on the six CPS variables, running the regression on the subsample of applicants with composite scores of at least 96. Some variables have the “wrong” sign in this regression, e.g., it suggests favoring students from higher-income census tracts, and it seems debatable whether CPS could use such an index, so we sequentially

²⁸ A reporter asked Richard Kahlenberg, the architect of the CPS plan, directly about this issue, “The new admissions policy uses the socioeconomic status of a student’s census tract, not an individual student. Is it possible that schools will draw, and admit, students who are the exceptions within their neighborhoods?” His response mentions both concerns about whether individual-level data would be reliable and confidence that the “rich” set of variables that go into the SES index will make it a reliable indicator of disadvantage. “There is that possibility. But there are problems with using the individualized data. FRPL data are presented by individual parents. There is usually very little verification that goes on. Using census data, we are able to get a rich set of factors. You could ask all these questions of people, but once it became clear that there was an advantage to being in one category, there might be an incentive for individuals to provide false information.” See Harris (2009).

²⁹ We focus only on minority efficiency because measured the income efficiency can be increased to 100 percent using the student-level data. As mentioned earlier, it is not clear whether this corresponds to 100 percent efficiency in addressing true disadvantage when disadvantage is multi-dimensional.

TABLE 3—EFFICIENCY OF ALTERNATE RACE-NEUTRAL AFFIRMATIVE ACTIONS PLANS

	Avg. score	Within school SD	Maj.- Min. Gap	Share minority	Share FRPL	Percent efficiency (SD)		
						Overall	Minority only	Income only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A. Walter Payton College Prep</i>								
Current CPS plan	98.0	2.71	3.17	0.368	0.245	28.3	25.8	17.1
Rewighted SES bonus	98.3	2.01	2.06	0.368	0.236	43.6	40.8	22.5
LASSO SES bonus	98.3	1.87	1.97	0.368	0.236	49.0	45.9	25.3
SES Bonus with FRPL	98.4	1.80	1.93	0.368	0.286	69.8	49.0	59.6
LASSO with FRPL	98.5	1.68	1.95	0.368	0.264	73.2	56.1	57.1
Race and FRPL-based benchmark	98.8	1.34	1.77	0.368	0.245	100.0	91.1	60.4
<i>Panel B. Northside College Prep</i>								
Current CPS Plan	97.7	3.18	3.79	0.355	0.340	22.1	18.5	9.7
Rewighted SES bonus	98.0	2.21	2.14	0.355	0.340	38.8	32.5	17.0
LASSO SES bonus	98.0	2.15	2.07	0.355	0.359	45.3	34.0	21.6
SES Bonus with FRPL	98.0	2.07	1.85	0.355	0.452	76.2	36.5	62.8
LASSO with FRPL	98.0	2.08	1.88	0.355	0.429	64.9	36.1	48.4
Race and FRPL-based benchmark	98.7	1.42	1.84	0.355	0.340	100.0	83.7	43.7

Notes: Reweighted SES bonus refers to a plan that uses a subset of the SES disadvantage indicators: CPS's Adult Education Index, Percentage of Single-Parent Families, and Local Elementary School ISAT Score variables. LASSO SES bonus adds the estimate for the probability of being a minority from a LASSO model as additional SES indicator. The SES bonus with FRPL and LASSO with FRPL refer to plans that additionally take the individual FRPL status into account. The race- and FRPL-based benchmark plan uses weights on individual minority and FRPL status to maximize the average composite score, while matching both the minority and FRPL share achieved in the current CPS plan.

drop wrong-sign variables from the regression until all weights are positive. We end up with a tract-level SES index that is a simple weighted average of just three variables: CPS's Adult Education Index, Percentage of Single-Parent Families, and Local Elementary School ISAT Score variables.³⁰

We do not know whether some political constraint or data quality concern led CPS to conclude that it should ignore student-level FRPL status and only use tract-level variables. Nevertheless, to demonstrate the additional predictive power of using such information, we also constructed a student-level SES index variable which turns out to be a weighted average of four variables: the individual-level FRPL indicator and the same three variables that are involved in our tract-level index.

We then considered alternative race-neutral assignment plans in which students' priorities at exam schools are the sum of their composite scores and a scale factor times of one of our two new SES indexes. We chose the scale factor so that the allocations produced would assign exactly as many minorities to Payton and Northside as CPS's actual plan. Table 3 provides statistics comparing the classes admitted under these plans with the classes admitted under the CPS plan. Panel A examines how the plans perform at Payton. Our simple modification raises minority-only efficiency from 26 percent to 41 percent. This increase corresponds to a reduction in within-school SD from 2.7 to 2.0. This reduction is quite far from eliminating the undesired side effect of the shift to race-neutral affirmative action, but it is a sizable enough shift to seem worthwhile. Another appealing feature of the modified plan is

³⁰Details on these regressions are provided in online Appendix Table A1.

that it reduces the majority-minority gap in average composite scores from 3.2 to 2.1 points. Adding individual FRPL data allows for a further improvement at Payton. Minority-only efficiency increases to 49 percent. The modified plan now also admits more FRPL-eligible students than the CPS plan.

Panel B reports comparable statistics for Northside. The version that uses only tract-level variables once again improves minority-only efficiency and simultaneously reduces the majority-minority gap in average composite scores. But, the fact that minority-only efficiency only reaches 33 percent provides further evidence for our initial conjecture that race-neutral affirmative action would be more difficult to carry out in an environment like Northside. Similar to Payton, the inclusion of individual-level FRPL status results in further improvement and raises minority-only efficiency to 37 percent.

It is possible that even these simple plans would run into legal trouble. Opponents of affirmative action could challenge the use of a weighted average, charging that the weights were derived from an attempt to influence racial outcomes of the school assignment plan rather than from their relative importance as components of diversity or disadvantage. We will not take a stance on this discussion and report in online Appendix Table A.6 that variants of these plans that use unweighted averages perform nearly as well as plans using weighted averages. The performance of several other plans, including neighborhood-based plans more directly analogous to the Texas top 10 percent plan, are also discussed in our working paper and in the online Appendix.

B. More Sophisticated Plans as an Upper Bound

We next ask whether substantially higher welfare gains could be reached with even more sophisticated race-neutral plans. We believe that the answer to this question is no. This view derives from an analysis of what could be done under a much more ambitious plan which already seems to go well beyond might be legally permissible.

To explore how far one might be able to increase efficiency, we began by manually browsing through the many thousands of tract-level variables available from Manson et al. (2010) and identified a subset of 145 that we thought might turn out to be useful predictors of minority status.³¹ The number of variables is sufficiently large relative to the number of high-scoring minority applicants to make overfitting an important concern, and it also seems implausible that CPS could choose an SES index with such a large number components. Accordingly, we use a LASSO regression to pick a parsimonious index that predicts minority status among students with high composite scores.³² Specifically, we estimate a LASSO regression on the subsample of applicants with composite scores of at least 96 with minority status as the dependent variable and both the CPS variables and our added 145 variables as potential explanatory variables.³³ The LASSO procedure led to a model that uses nine of the explanatory variables. The variables it chose magnify concerns that an

³¹ Not all variables are available for all tracts. We impute the mean value for missing values.

³² We also investigated the performance of random forest models. They did not fare quite as well in out-of-sample fit.

³³ The regularization parameter was chosen by cross validation. Details are in online Appendix Table A5.

approach of this form would not pass legal muster. Two variables it selects are the fractions of the foreign-born population in the census tract who come from Asia and Europe, which while not technically a use or race, run the risk of being seen by courts as having been chosen by the machine learning algorithm to penalize tracts with many Asians and Whites.

We then implemented race-neutral affirmative action plans as in the previous section. We treat the predicted probability of being a minority that comes out of the LASSO model as if it were an SES index and rank students on a weighted average of their composite scores and their predicted minority status. As above, we implemented two versions of each of the above plans. One uses a weight that makes the underrepresented minority share at Payton exactly match its value under the current CPS plan. The other exactly matches the current underrepresented minority share at Northside.

The LASSO model is able to improve on our simple modifications at both Payton and Northside. But the primary result that we would emphasize is the magnitudes: the results indicate that even such a fairly complex model cannot produce significantly larger gains than the plans presented previously. At Payton, the LASSO model achieves 46 percent minority efficiency when constructed without student-level FRPL data and 56 percent efficiency, when using such individual-level data. The improvement from LASSO without free lunch is only about 5 percentage points better than we were able to do achieve in our simple three variable model and remains far from 100 percent. At Northside, the LASSO model adds essentially nothing to what we were able to do earlier to improve minority efficiency.

We conclude that the moderate improvements that we were able to make to the CPS plan in the previous section are probably not far from the upper bound under the given constraints. Constraining affirmative action plans to be race-neutral and place-based will substantially increase the costs that are incurred if one tries to increase minority representation at Chicago's most elite public schools.

Table 3 also reports overall efficiency measures for the plans discussed above. They are as high as 76 percent for plans that use the individual-level FRPL data. When the overall efficiency evaluated using the mean, rather than the standard deviation, the overall efficiency is never greater than 62 percent.³⁴ Even these bounds should be regarded as an overly optimistic view of what can be accomplished in a race-neutral manner because it assumes that the FRPL variable that is now being used perfectly captures the socioeconomic disadvantage that CPS cares about, and gives our modified plan a great deal of credit for capturing such disadvantage accurately. We believe that the exercise of using an income proxy to examine how efficient a plan admits disadvantaged students is better motivated for plans that do not make use of an income proxy variable. It would be interesting to examine the overall efficiency of plans that use the income variable also considering how efficiently they admit students who are disadvantaged on other dimensions, e.g., coming from single-parent families or having parents who did not attend college, but we lack the data to do so.

It is also worth noting that the potential ways to increase efficiency we have examined only have a moderate impact on the number of Asian students. The proportion of Asian students is moderately lower under our benchmark plans that use

³⁴These and other additional results are shown in online Appendix Table A6.

race and FRPL data. At Payton, 12.3 percent of students at Payton are Asian under the CPS plan and 10.0 percent are under the race and FRPL-based benchmark. At Northside, 24.7 percent of students are Asian under CPS plan and 23.6 percent are under the race and FRPL-based benchmark. Some of the alternate race neutral plans we have considered slightly increase Asian representation at one or both schools. See online Appendix Table A6.

VI. Recap and Conclusions

This paper evaluates affirmative action plans when schools develop curricula tailored to students' ability or preparation levels and a diverse learning environment. Absent any restrictions on its form, we show that affirmative action in Chicago's exam school sector would be a powerful tool. There are sufficiently many minority and low-income students with composite scores within two or three points of the purely score-based cutoffs, which allows for a substantial increase in low-income and/or minority enrollment without a large impact on the distribution of admitted students' composite scores. For example, it is possible to roughly double minority and/or low-income enrollment by giving preferences on the order of two or three points.

Place-based affirmative action plans expand the definition of disadvantage to include neighborhood characteristics and Chicago's plan is one of the most ambitious and sophisticated of such plans. However, it performs worse than racial quotas in increasing minority or low-income access. This fact shows broadening the definition of disadvantage need not entail a trade-off between minority enrollment and other measures of disadvantage. The CPS plan incurs higher curriculum-matching losses and widens within-school racial achievement gaps.

We show that there is room to improve the efficiency of the CPS plan. By altering CPS's index of socioeconomic status, using continuous bonuses rather than discrete SES tiers and exploiting individual-level data on FRPL status, it is possible to achieve the same levels of minority and FRPL representation with significantly fewer distortions. However, a key finding is that there are still significant limits to the efficiency of place-based affirmative action policies. One main reason is that the high-achievers that Payton and Northside admits are by definition outliers for their neighborhood, which complicates the relationship between neighborhood characteristics and true measures of disadvantage.

While we have focused on Chicago, contentious debates about the lack of diversity at exam schools have taken place in other cities. The effectiveness of any affirmative action plan depends both on the design of the plan and the city in which it is implemented. Our ability to simulate the effects of different plans in Chicago have given us some scope to provide estimates of the range of efficiencies that are possible under different plans. The differing levels of efficiency at Payton and Northside demonstrate that a substantial range in efficiency exists both across plans and targets; while we achieved more than 50 percent minority-only efficiency at Payton, the most efficient plan reached just 36 percent at Northside. Our framework can be applied to other cities to help quantify the efficiency of affirmative action and evaluate the performance of alternatives.

Our framework to evaluate affirmative action plans is based on the premise that grouping together high ability students allows for schools to offer advanced

curricula that would not be available elsewhere. We have assumed that education production depends on curriculum matching and diversity and that the composite score, scaled appropriately, is a good proxy for curriculum matching. While the production function assumption provides a rationale for exam schools and our use of the composite score allows us to quantify efficiency loss, future work should try to provide more justification for the best proxies for diversity and curriculum matching and their importance in education production. Finally, we've also judged the performance of alternative schemes based on measures of exam school access. It would also be fruitful to study the effect of increased exam school access on subsequent educational outcomes.

REFERENCES

- ACT. 2020. "National Distributions of Cumulative Percents for ACT Test Scores ACT-Tested High School Graduates from 2015, 2016, and 2017." <https://www.act.org/content/dam/act/unsecured/documents/MultipleChoiceStemComposite2017-18.pdf>.
- Athey, Susan, Christopher Avery, and Peter Zemsky. 2000. "Mentoring and Diversity." *American Economic Review* 90 (4): 765–86.
- Austen-Smith, David, and Roland G. Fryer, Jr. 2005. "An Economic Analysis of 'Acting White.'" *Quarterly Journal of Economics* 120 (2): 551–83.
- Cestau, Dario, Dennis Epple, and Holger Sieg. 2017. "Admitted Students to Selective Education Programs: Merit, Profiling, and Affirmative Action." *Journal of Political Economy* 125 (3): 761–97.
- Chan, Jimmy, and Erik Eyster. 2003. "Does Banning Affirmative Action Lower College Student Quality?" *American Economic Review* 93 (3): 858–72.
- Chen, Yan, and Tayfun Sönmez. 2006. "School Choice: An Experimental Study." *Journal of Economic Theory* 127 (1): 202–31.
- Chicago Public Schools. 2014. "Chicago Public Schools: Student Administrative Data System." Chicago, Illinois.
- Chung, Kim-Sau. 2000. "Role Models and Arguments for Affirmative Action." *American Economic Review* 90 (3): 640–48.
- Coate, Stephen, and Glenn C. Loury. 1993. "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review* 83 (5): 1220–40.
- Corcoran, Sean Patrick, and E. Christine Baker-Smith. 2018. "Pathways to an Elite Education: Application, Admission, and Matriculation to New York City's Specialized High Schools." *Education Finance and Policy* 13 (2): 256–79.
- Cullen, Julie Berry, Mark C. Long, and Randall Reback. 2013. "Jockeying for Position: Strategic High School Choice under Texas' Top Ten Percent Plan." *Journal of Public Economics* 97: 32–48.
- Dobbie, Will, and Roland G. Fryer, Jr. 2011. "Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone." *American Economic Journal: Applied Economics* 3 (3): 158–87.
- Dur, Umut, Parag A. Pathak, and Tayfun Sönmez. 2020. "Explicit vs. Statistical Preferential Treatment in Affirmative Action: Theory and Evidence from Chicago's Exam Schools." *Journal of Economic Theory* 187: 104996.
- Eder, Derek, and Forest Gregg. 2014. "Chicago Public School Tiers." <http://cpstiers.opencityapps.org> (accessed September 19, 2020).
- Ellison, Glenn, and Parag A. Pathak. 2021. "Replication Data for: The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago's Exam Schools." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/120529V1>.
- Epple, Dennis, Richard Romano, and Holger Sieg. 2008. "Diversity and Affirmative Action in Higher Education." *Journal of Public Economic Theory* 10 (4): 475–501.
- Fryer, Roland G., Jr., and Steven D. Levitt. 2004. "Understanding the Black-White Test Score Gap in the First Two Years of School." *Review of Economics and Statistics* 86 (2): 447–64.
- Fryer, Roland G., Jr., and Glenn C. Loury. 2005. "Affirmative Action and Its Mythology." *Journal of Economic Perspectives* 19 (3): 147–62.
- Fryer, Roland G., Jr., Glenn C. Loury, and Tolga Yuret. 2008. "An Economic Analysis of Color-Blind Affirmative Action." *Journal of Law, Economics, and Organization* 24 (2): 319–55.

- Fu, Qiang.** 2006. "A Theory of Affirmative Action in College Admissions." *Economic Inquiry* 44 (3): 420–28.
- Haeringer, Guillaume, and Flip Klijn.** 2009. "Constrained School Choice." *Journal of Economic Theory* 144 (5): 1921–47.
- Harris, R.** 2009. "Q&A with Richard Kahlenberg, Consultant on Magnet School Admissions." *The Chicago Reporter*. <http://chicagoreporter.com/qa-richard-kahlenberg-consultant-magnet-school-admissions>.
- ISBE.** 2014. "Illinois State Board of Education Prairie State Achievement Examination."
- Li, Shengwu.** 2017. "Obviously Strategy-Proof Mechanisms." *American Economic Review* 107 (11): 3257–87.
- Manson, S., J. Schroeder, D. V. Riper, T. Kugler, and S. Ruggles.** 2010. "IPUMS National Historical Geographic Information System." Web. Accessed February 24, 2016.
- Pathak, Parag A., and Tayfun Sönmez.** 2013. "School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation." *American Economic Review* 103 (1): 80–106.
- Ray, Debraj, and Rajiv Sethi.** 2010. "A Remark on Color-Blind Affirmative Action." *Journal of Public Economic Theory* 12 (3): 399–406.
- Steele, C. M., and J. Aronson.** 1995. "Stereotype Threat and the Intellectual Test Performance of African Americans." *Journal of Personality and Social Psychology* 69 (5): 797–811.