sociological science

Supplement to:

Wodtke, T. Geoffrey, Kailey White, and Xiang Zhou. 2026. “Poor Neighborhoods, Bad Schools? A High-Dimensional Model of Place-Based Disparities in Academic Achievement” Sociological Science 13: 109-153.

Online Supplementary Materials for "Poor Neighborhoods, Bad Schools? A High-dimensional Model of Place-based Disparities in Academic Achievement"

Geoffrey T. Wodtke[1], Kailey White[2], and Xiang Zhou[3]

[1]Department of Sociology, University of Chicago

[2]Crime Lab and Education Lab, University of Chicago

[3]Department of Sociology, Harvard University

A Analyses with a Multidimensional Disadvantage Index

In this appendix, we replicate our descriptive decomposition and causal analyses using a multidimensional index of neighborhood disadvantage. This index is constructed through a principal components analysis of multiple tract-level characteristics, including the poverty rate, measures of educational attainment and family structure, the unemployment rate, and racial composition. We estimate observed, adjusted, and counterfactual disparities that compare students from neighborhoods in the top quintile of this index to those from the bottom four quintiles. Results from the descriptive decomposition are shown in Figure A.1, while estimates of the disparity eliminated appear in Table A.1. Overall, these results are nearly identical to those presented in the main text, which are based on the poverty rate alone.

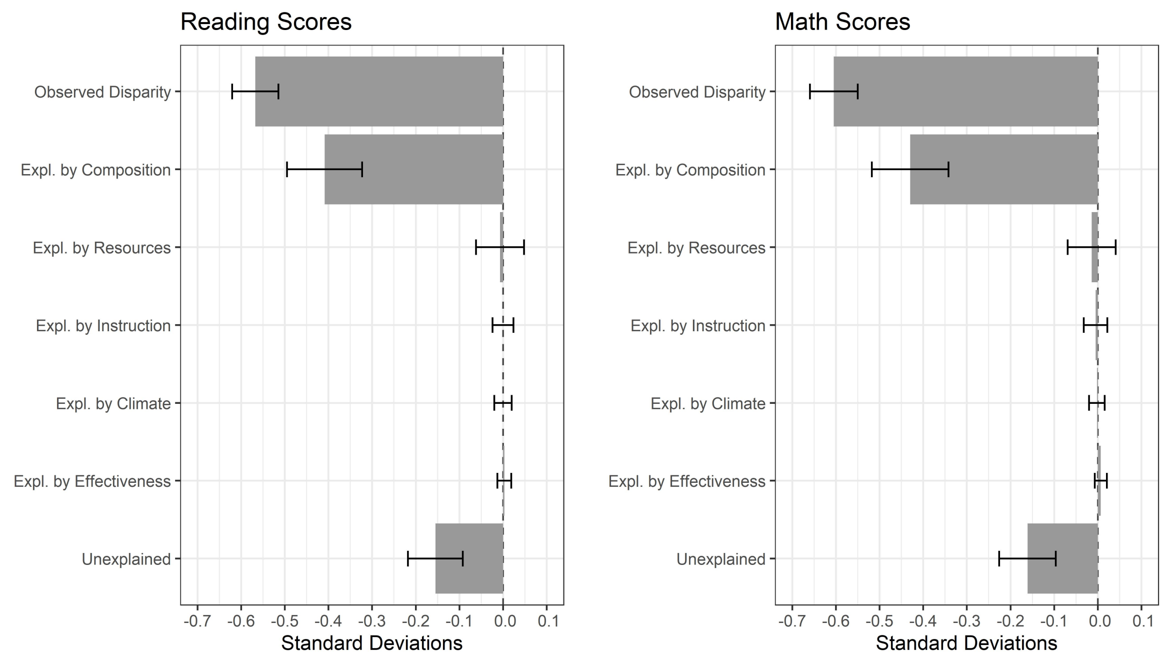

Figure A.1: Descriptive Decomposition of the Observed Disparity in Achievement Test Scores between Students from High- and Low-poverty Neighborhoods Defined in terms of a Multidimensional Index of Socioeconomic Disadvantage.

Note: This plot contains de-biased machine learning estimates of the “disparity explained” by different dimensions of school quality. All estimates are reported in standard deviation units. The disparities contrast the top with the bottom four quintiles of a multidimensional disadvantage index, which comes from a principal components analysis of tract poverty, educational composition, family structure, unemployment, and racial composition. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), “School District Finance Survey (F-33),” fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), “Local Education Agency Universe Survey,” “Public Elementary/Secondary School Survey,” 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

Table A.1: De-biased Machine Learning Estimates of the Disparity Eliminated between Students from High- and Low-poverty Neighborhoods Defined in terms of a Multidimensional Index of Socioeconomic Disadvantage.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.568	[-0.621, -0.515]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.012	[-0.067, 0.043]	0.02
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.012	[-0.073, 0.049]	0.02
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.013	[-0.074, 0.048]	0.02
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.012	[-0.073, 0.049]	0.02
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.012	[-0.071, 0.047]	0.02
Math Test Scores				
Observed disparity	Δ	-0.605	[-0.660, -0.550]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	0.001	[-0.068, 0.070]	0.00
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.008	[-0.082, 0.066]	0.01
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.009	[-0.083, 0.065]	0.01
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.008	[-0.088, 0.072]	0.01
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.007	[-0.081, 0.067]	0.01

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. The disparities contrast the top with the bottom four quintiles of a multidimensional disadvantage index, which comes from a principal components analysis of tract poverty, educational composition, family structure, unemployment, and racial composition. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

B School Characteristics in the ECLS-K

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality.

Label	Source Varname(s)	Item Text	Values	Mean	SD
School Composition					
Student Demographics					
% Free lunch	x4fmeal_i; x4rmeal_i	Percentage of students eligible for free meals; percentage of students eligible for reduced-priced meals	Proportions	0.50	0.31
% Black	s4blacpt	Approximately what percentage of the children in your school belongs to each of the following racial/ethnic groups? Black or African American, not Hispanic or Latino	Proportions	0.15	0.23
% White	s4whitpt	Approximately what percentage of the children in your school belongs to each of the following racial/ethnic groups? White, not Hispanic or Latino	Proportions	0.51	0.34
% Hispanic	s4hisppt	Approximately what percentage of the children in your school belongs to each of the following racial/ethnic groups? A9A2. Hispanic/Latino of any race	Proportions	0.22	0.27
% ELL	s4totell	What percentage of children in this school and in first grade are english language learners (ELL)? Percentage of ELL among all students in school	Proportions	0.14	0.19
% Gifted	s4gifpct	Approximately what percentage of your first-graders are in each of the following instructional programs? A gifted and talented program	Proportions	0.01	0.04
% Special ed.	s4spdpct	Approximately what percentage of your first-graders are in each of the following instructional programs? A gifted and talented program	Proportions	0.06	0.06
Staff Demographics					
Tchr male	a4tgend	What is your [teacher] gender?	0 (Female), 1 (Male)	0.03	0.17
Tchr white	a4white	Which best describes your [teacher] race?	0 (Not White), 1 (White)	0.80	0.40
Prcpl male	s4gender	What is your [school administrator] gender?	0 (Female), 1 (Male)	0.31	0.46

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Prcpl white	s4white	Which best describes your [school administrator] race? White	0 (Not White), 1 (White)	0.80	0.40
School Resources					
Type					
Pub schl	x4pubpri	Public or private school	0 (Private), 1 (Public)	0.91	0.28
Yr-round Schl	x4yrrnd	Year round school	0 (Not a Year-round school), 1 (Year-round school)	0.02	0.12
K lowest grd	x4lowgrd	Lowest grade at the school	0 (Pre-K), 1 (K or higher)	0.53	0.50
6th highest grd	x4higgrd	Highest grade at the school	0 (5th or lower), 1 (6th or higher)	0.37	0.48
Funding					
Schl funds decl	s4funddc	During the past year, to what extent did any of the following changes occur at your school? Funding levels decreased	0 (Not at all); 1 (At least to some extent)	0.87	0.34
Schl salary decl	s4csalde	During the past year, to what extent did any of the following changes occur at your school? Salaries decreased	0 (Not at all), 1 (At least to some extent)	0.19	0.39
Schl salary frz	s4cslyfz	During the past year, did any of the following changes occur at your school? Salaries were frozen	0 (Not at all), 1 (At least to some extent)	0.42	0.49
Schl salary inc	s4csalin	During the past year, to what extent did any of the following changes occur at your school? Salaries increased	0 (Not at all), 1 (At least to some extent)	0.30	0.46
Dist $ expend	ccd_dstr_exp4; ccd_dstr_enr4	CCD total district expenditures; CCD total number of students in district	Dollars ($1000s) per student	12.15	3.83
Staffing					
IT tchrs/stdnt	s4ctechf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time computer/technology teachers or support staff; total number of students	Staff per 100 students	0.08	0.13

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Elec tchrs/stdnt	s4artstf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time drama, music, or art teachers; total number of students	Staff per 100 students	0.18	0.20
Gym tchrs/stdnt	s4gymtf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time gym/PE or health teachers; total number of students	Staff per 100 students	0.16	0.17
ESL tchrs/stdnt	s4eslf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time esl/bilingual education/dual-language immersion teachers; total number of students	Staff per 100 students	0.24	0.53
Librar/stdnt	s4librf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time library media specialists/librarians; total number of students	Staff per 100 students	0.11	0.11
Paraprof/stdnt	s4paraf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time paraprofessionals; total number of students	Staff per 100 students	1.30	1.17
Psychs/stdnt	s4psycf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time school psychologists or social workers; total number of students	Staff per 100 students	0.06	0.13
Nurses/stdnt	s4nursf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time school nurses or health professionals; total number of students	Staff per 100 students	0.10	0.12
S/E tchrs/stdnt	s4spedf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time special education teachers and related service providers; total number of students	Staff per 100 students	0.66	0.48

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
G/T tchrs/stdnt	s4giftf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time teachers of gifted/talented students; total number of students	Staff per 100 students	0.10	0.25
Class tchrs/stdnt	s4rgtchf; sch_enr4	Approximately how many staff members does your school currently have in the following categories? Full-time regular classroom teachers; total number of students	Staff per 100 students	4.62	1.20
Schl translators	s4transl	Are any of the following services provided to families of children from households where a language other than english is spoken? Translators are made available to parents and/or meetings are conducted in the parents non-english language	0 (No), 1 (Yes)	0.84	0.37
Tchr turnover	s4teleft	Write in the approximate number of regular classroom teachers for each of the following. Number of regular classroom teachers who have left your school since [last year] and have not returned.	0 (No turnover), 1 (At least some turnover)	0.19	0.39
Qualifications					
Prcpl PhD/EdD	s4edlvl	What is the highest level of education you [administrator] have completed?	0 (No EdD/PhD), 1 (EdD/PhD)	0.41	0.49
Prcpl tenure	s4ystch	How many years experience do you [administrator] have in each of the following positions? Years as a teacher before becoming a principal	Years	12.37	6.35
Prcpl yrs exp	s4totpri	How many years experience do you [administrator] have in each of the following positions? Total number of years as a principal	Years	8.87	6.59
Tchr MA	a4hghstd	What is the highest level of education you [teacher] have completed?	0 (Less than graduate degree), 1 (Graduate degree)	0.51	0.50
Tchr yrs exp	a4yrstch	Counting this school year, how many years have you [teacher] been a school teacher?	Years	14.99	9.81

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Tchr state cert	a4statct	Which of the following describes the teaching certificate you [teacher] currently hold in this state?	0 (No certification), 1 (State certification)	0.93	0.25
Tchr passed board	a4natexm	Have you [teacher] taken the exam for national board for professional teaching standards certification?	0 (No), 1 (Yes)	0.21	0.41
Tchr tenure	a4yrsch	Counting this school year, how many years have you [teacher] been teaching at this school?	Years	9.75	7.79
Facilities					
Qlty classrms	s4clssok	In general, how adequate are each of the following school facilities for meeting the needs of the children in your school? Classrooms	0 (Not adequate/do not have), 1 (Adequate)	0.81	0.39
Qlty auditorium	s4audtok	In general, how adequate are each of the following school facilities for meeting the needs of the children in your school? Auditorium	0 (Not adequate/do not have), 1 (Adequate)	0.23	0.42
Qlty library	s4lbryok	In general, how adequate are each of the following school facilities for meeting the needs of the children in your school? Library/Media Center	0 (Not adequate/do not have), 1 (Adequate)	0.80	0.40
Qlty gym	s4gymok	In general, how adequate are each of the following school facilities for meeting the needs of the children in your school? Gymnasium	0 (Not adequate/do not have), 1 (Adequate)	0.61	0.49
Qlty playground	s4playok	In general, how adequate are each of the following school facilities for meeting the needs of the children in your school? Playground	0 (Not adequate/do not have), 1 (Adequate)	0.75	0.43
Qlty music rm	s4muscok	In general, how adequate are each of the following school facilities for meeting the needs of the children in your school? Music Room	0 (Not adequate/do not have), 1 (Adequate)	0.65	0.48
Qlty art rm	s4artok	In general, how adequate are each of the following school facilities for meeting the needs of the children in your school? Art Room	0 (Not adequate/do not have), 1 (Adequate)	0.56	0.50
Qlty IT lab	s4compok	In general, how adequate are each of the following school facilities for meeting the needs of the children in your school? Computer Lab	0 (Not adequate/do not have), 1 (Adequate)	0.65	0.48

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Qlty multi-use rm	s4multok	In general, how adequate are each of the following school facilities for meeting the needs of the children in your school? Multi-Purpose Room	0 (Not adequate/do not have), 1 (Adequate)	0.34	0.47
Qlty cafeteria	s4cafeok	In general, how adequate are each of the following school facilities for meeting the needs of the children in your school? Cafeteria	0 (Not adequate/do not have), 1 (Adequate)	0.79	0.40
Instruction					
Instructional Time					
Rd-days/wk	a4oftrdl	How often does the typical child in your class usually work on lessons or projects in the following general subject areas, whether as a whole class, in small groups, or in individualized arrangements? Reading and language arts	1 (Less than 3 days a week), 2 (3-4 days a week), 3 (5 days a week)	2.94	0.30
Rd-hrs/day	a4txrdla	On the days children work in these areas, how much time does the typical child in your class usually work on lessons or projects in the following general subject areas? B4A. Reading and language arts	1 (Less than 1 hour) to 6 (3 or more hours)	3.94	1.36
Mth-days/wk	a4oftmth	How often does the typical child in your class usually work on lessons or projects in the following general subject areas, whether as a whole class, in small groups, or in individualized arrangements? Mathematics	1 (Less than 3 days a week), 2 (3-4 days a week), 3 (5 days a week)	2.92	0.34
Mth-hrs/day	a4txmth	On the days children work in these areas, how much time does the typical child in your class usually work on lessons or projects in the following general subject areas? Mathematics	1 (Less than 1 hour) to 6 (3 or more hours)	2.26	0.90
Teaching Methods					
Sml grp hrs/day	a4wksgrp	In a typical day, how much time does a child in your class spend in the following types of activities? Working in small groups with teacher	1 (Half hour or less) to 4 (3 hours or more)	1.79	0.74

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Lrg grp hrs/day	a4wklgrp	In a typical day, how much time does a child in your class spend in the following types of activities? Teacher lecture with large group and/or large group discussion led by teacher	1 (Half hour or less) to 4 (3 hours or more)	2.34	0.88
Peers hrs/day	a4wkpeer	In a typical day, how much time does a child in your class spend in the following types of activities? Working with peers under teacher direction	1 (Half hour or less) to 4 (3 hours or more)	2.01	0.81
Ind act hrs/day	a4wkindv	In a typical day, how much time does a child in your class spend in the following types of activities? Working on individual tasks under teacher direction	1 (Half hour or less) to 4 (3 hours or more)	2.41	0.82
Rd grp days/wk	a4divrd	In an average week, how often do you divide your class into instructional groups based on achievement or ability levels for reading and math activities or lessons? Reading	1 (Never) to 7 (Everyday)	5.94	1.72
Mth grp days/wk	a4divmth	In an average week, how often do you divide your class into instructional groups based on achievement or ability levels for reading and math activities or lessons? Math	1 (Never) to 7 (Everyday)	3.58	2.26
Rd decodables	a4usedecb	How often do you use the following resources to teach reading in your class? Decodable books, sound/symbols book	1 (Never or hardly ever) to 4 (Almost every day)	3.28	0.94
Rd kits	a4usekit	How often do you use the following resources to teach reading in your class? Reading kits	1 (Never or hardly ever) to 4 (Almost every day)	2.27	1.22
Rd basal series	a4usebsl	How often do you use the following resources to teach reading in your class? Basal reading series	1 (Never or hardly ever) to 4 (Almost every day)	3.20	1.20
Rd big books	a4usebgbk	How often do you use the following resources to teach reading in your class? Big books	1 (Never or hardly ever) to 4 (Almost every day)	2.83	0.93
Rd computer	a4usecmp	How often do you use the following resources to teach reading in your class? Computer software for reading instruction	1 (Never or hardly ever) to 4 (Almost every day)	2.56	1.21
Rd manipulatvs	a4useman	How often do you use the following resources to teach reading in your class? Manipulatives (plastic letters, picture cards, tiles, etc.)	1 (Never or hardly ever) to 4 (Almost every day)	3.49	0.76

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Rd anthology	a4useanth	How often do you use the following resources to teach reading in your class? Anthology (collections of literary works)	1 (Never or hardly ever) to 4 (Almost every day)	2.74	1.09
Rd leveled	a4uselev	How often do you use the following resources to teach reading in your class? Leveled or guided reading books	1 (Never or hardly ever) to 4 (Almost every day)	3.79	0.56
Rd audiobooks	a4useaubk	How often do you use the following resources to teach reading in your class? Read-along books paired with audiobooks	1 (Never or hardly ever) to 4 (Almost every day)	2.75	1.12
Rd news/mags	a4usenew	How often do you use the following resources to teach reading in your class? Newspapers and/or magazines	1 (Never or hardly ever) to 4 (Almost every day)	1.74	0.81
Rd trade books	a4usetrd	How often do you use the following resources to teach reading in your class? A variety of trade books	1 (Never or hardly ever) to 4 (Almost every day)	3.38	0.90
Rd other	a4useoth	How often do you use the following resources to teach reading in your class? Materials from other subjects	1 (Never or hardly ever) to 4 (Almost every day)	3.16	0.83
Curriculum Content					
Rd glossaries	a4useglos	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Using text features such as glossaries and other references to learn word meanings	1 (Not taught) to 6 (On more than 80 days)	4.45	1.40
Rd-evidence	a4reassup	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Identifying the reasons an author gives to support points in an opinion piece	1 (Not taught) to 6 (On more than 80 days)	4.23	1.64
Rd-char/plot	a4charplot	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Identifying character, setting, and plot	1 (Not taught) to 6 (On more than 80 days)	5.54	0.86

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Rd-sim/diff	a4simdiff	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Describing similarities and differences between two reading selections	1 (Not taught) to 6 (On more than 80 days)	4.66	1.33
Rd-narrator	a4whotell	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Identifying who is telling a story at different points in a text	1 (Not taught) to 6 (On more than 80 days)	4.51	1.46
Rd-char Qs	a4gencsp	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Generating questions about character, setting and plot	1 (Not taught) to 6 (On more than 80 days)	5.37	1.03
Rd-main id inf	a4maintext	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Identifying main ideas and details in informational text	1 (Not taught) to 6 (On more than 80 days)	5.10	1.14
Rd-feelings	a4senses	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Identifying words and phrases that suggest feelings or appeal to the senses	1 (Not taught) to 6 (On more than 80 days)	4.77	1.25
Rd-main id stry	a4mainid	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Identifying the main idea in a story	1 (Not taught) to 6 (On more than 80 days)	5.27	1.09

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Rd-fic/nonfic	a4ficnonf	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Recognizing the differences between fiction and nonfiction	1 (Not taught) to 6 (On more than 80 days)	5.29	1.10
Rd-sen context	a4senctxt	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Using sentence-level context to gain meaning of a word or phrase	1 (Not taught) to 6 (On more than 80 days)	5.17	1.11
Rd-accuracy	a4rdaccr	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Reading accurately and fluently to support comprehension	1 (Not taught) to 6 (On more than 80 days)	5.77	0.65
Rd-retelling	a4retell	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Retelling stories, including main ideas and details	1 (Not taught) to 6 (On more than 80 days)	5.47	0.92
Rd-pace	a4paceint	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Reading and re-reading passages orally with guidance on pacing, intonation, and expression	1 (Not taught) to 6 (On more than 80 days)	5.72	0.72
Rd-predict	a4predict	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Predicting what might occur next in the text	1 (Not taught) to 6 (On more than 80 days)	5.61	0.82
Rd-blend snds	a4sndwrd	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Blending sounds to form words	1 (Not taught) to 6 (On more than 80 days)	5.82	0.63

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Rd-characters	a4deschar	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Describing characters, settings, and major events in a story.	1 (Not taught) to 6 (On more than 80 days)	5.53	0.86
Rd-info text	a4cmpxinf	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Reading informational selections of appropriate complexity for this grade	1 (Not taught) to 6 (On more than 80 days)	5.19	1.07
Rd-word segs	a4segword	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Segmenting words into phonemes	1 (Not taught) to 6 (On more than 80 days)	5.70	0.80
Rd-form words	a4manpho	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Manipulating phonemes to form new words.	1 (Not taught) to 6 (On more than 80 days)	5.68	0.81
Rd-poetry	a4cmpxpro	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Reading prose and poetry of appropriate complexity for this grade	1 (Not taught) to 6 (On more than 80 days)	4.49	1.39
Rd-irreg spl	a4irregwd	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Reading irregularly spelled words.	1 (Not taught) to 6 (On more than 80 days)	5.61	0.84
Wrt-info	a4infpiec	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Writing an informational piece that includes some facts on the topic	1 (Not taught) to 6 (On more than 80 days)	4.19	1.38

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Wrt-opinion	a4opinion	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Writing an opinion piece, giving reasons for the opinion	1 (Not taught) to 6 (On more than 80 days)	3.64	1.64
Wrt-narrative	a4narrtv	From the first day of school until today, please indicate how many days each of the following reading skills and concepts has been covered in your class. Writing a narrative with two or more appropriately sequenced events	1 (Not taught) to 6 (On more than 80 days)	4.70	1.36
Mth-length	a4arr3obj	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Arranging three objects by length	1 (Not taught) to 6 (On more than 80 days)	3.22	1.34
Mth-comp lngth	a4lng2by3	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Comparing the length of two objects indirectly by using a third object	1 (Not taught) to 6 (On more than 80 days)	3.12	1.33
Mt-msr lngth	a4lngmult	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Measuring length of an object as a whole number of length units, by laying multiple copies of shorter object end to end	1 (Not taught) to 6 (On more than 80 days)	3.10	1.29
Mth-est lngth	a4estlng	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Estimating the length of object in standard units, such as inches, feet, centimeters, and/or meters	1 (Not taught) to 6 (On more than 80 days)	2.88	1.36

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Mth-msr tool	a4meatool	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Measuring length of an object in standrd units, using tools such as rulers, yardsticks, meter sticks, measuring tapes	1 (Not taught) to 6 (On more than 80 days)	3.02	1.36
Mth-rel quant	a4relqty	From the first day of school until today, indicate how many days each of the following math skills and concepts has been covered in your class. Labeling relative quantity using terms "greater than," "less than," "equal to," "fewest," "most"	1 (Not taught) to 6 (On more than 80 days)	5.00	1.13
Mth-symbols	a4relsym	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Relative quantity when comparing two-digit numbers, using the symbols ¿, = and ¡	1 (Not taught) to 6 (On more than 80 days)	4.16	1.38
Mth-shp togthr	a4dimcomp	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Putting two-dimensional or three-dimensional shapes together to create a composite shape	1 (Not taught) to 6 (On more than 80 days)	3.05	1.36
Mth-count 120	a4cnt120	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Counting to 120	1 (Not taught) to 6 (On more than 80 days)	4.64	1.56
Mth-add by 3s	a4slvadd3	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Solving word problems by adding three numbers whose sum is 20 or less	1 (Not taught) to 6 (On more than 80 days)	4.13	1.58

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Mth-word probs	a4slvadsb	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Solving word problems by adding or subtracting numbers equal to 20 or less	1 (Not taught) to 6 (On more than 80 days)	5.26	0.99
Mth-shp attrib	a4attrshp	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Identifying the difference between defining attributes of shapes vs. non-defining attributes	1 (Not taught) to 6 (On more than 80 days)	3.57	1.27
Mth-solve #	a4slvuknm	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Solving for an unknown whole number in an addition or subtraction equation	1 (Not taught) to 6 (On more than 80 days)	4.47	1.37
Mth-draw graph	a4drwgrph	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Drawing a picture graph and/or a bar graph to represent a data set with up to four categories	1 (Not taught) to 6 (On more than 80 days)	3.81	1.30
Mth-use graph	a4ansgrph	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Answering questions about the data in a picture graph and/or bar graph	1 (Not taught) to 6 (On more than 80 days)	4.03	1.27
Mth-equal sign	a4eqlsign	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. The meaning of the equal sign	1 (Not taught) to 6 (On more than 80 days)	5.11	1.33

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Mth-equation	a4sidequa	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Determining if both sides of and equation are equal or not equal using subtraction or addition	1 (Not taught) to 6 (On more than 80 days)	4.34	1.51
Mth-# vs quant	a4numqty	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Identifying the correspondence between number and quantity of quantities larger than 10	1 (Not taught) to 6 (On more than 80 days)	5.05	1.22
Mth-$ probs	a4slvcoin	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Solving word problems involving quarters, dimes, nickels, and pennies	1 (Not taught) to 6 (On more than 80 days)	3.59	1.62
Mth-write time	a4wrttime	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Writing time in hours and half hours	1 (Not taught) to 6 (On more than 80 days)	3.88	1.47
Mth-tell time	a4telltime	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Telling time in hours and half hours	1 (Not taught) to 6 (On more than 80 days)	4.00	1.46
Mth-skip count	a4skipcnt	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Skip counting by 5's, 10's, and/or 100's	1 (Not taught) to 6 (On more than 80 days)	5.27	1.09
Mth-10s/1s plc	a4tenones	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Identifying the numbers that represent the tens and ones places in a two-digit number	1 (Not taught) to 6 (On more than 80 days)	5.16	1.08

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Mth-add/sub	a4ctadsub	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Working with problems that demonstrate the relationship between counting, addition, and subtraction	1 (Not taught) to 6 (On more than 80 days)	5.31	1.01
Mth-count 20	a4cnt20qty	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Counting objects up to 20 to establish quantity	1 (Not taught) to 6 (On more than 80 days)	5.31	1.19
Mth-sum to 100	a4addto100	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Adding numbers that sum to 100 or less	1 (Not taught) to 6 (On more than 80 days)	4.08	1.67
Mth-add/sub 10	a4find10	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Finding 10 more or 10 less than a given two-digit number, without having to count	1 (Not taught) to 6 (On more than 80 days)	4.18	1.42
Mth-shp part	a4parteql	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Partitioning simple shapes into two and four equal shares	1 (Not taught) to 6 (On more than 80 days)	2.97	1.36
Mth-rd/wrt nums	a4nmrl120	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Reading and writing numerals up to 120	1 (Not taught) to 6 (On more than 80 days)	4.75	1.55
Mth-shp names	a4triquad	From the first day of school until today, please indicate how many days each of the following math skills and concepts has been covered in your class. Identifying triangles, quadrilaterals, pentagons, hexagons, and cubes	1 (Not taught) to 6 (On more than 80 days)	3.27	1.33

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Assessment					
Homework/wk	a4dpwhmwk	In an average week, how many days a week is homework assigned?	1 (0 days) to 6 (5 days)	5.04	0.91
Freq wrk samples	a4wrksam	How often do you use the following to assess your students? Work samples	1 (Never) to 6 (3 or more times/wk)	4.93	1.00
Freq projects	a4projct	How often do you use the following to assess your students? Individual or group projects	1 (Never) to 6 (3 or more times/wk)	3.49	1.30
Freq worksheets	a4wrksts	How often do you use the following to assess your students? Worksheets	1 (Never) to 6 (3 or more times/wk)	5.14	1.22
Freq quizzes	a4tstqz	How often do you use the following to assess your students? Classroom tests or quizzes	1 (Never) to 6 (3 or more times/wk)	4.88	0.70
Freq std tests	a4stntst	How often do you use the following to assess your students? State or local standardized tests	1 (Never) to 6 (3 or more times/wk)	2.49	1.08
Eval relatively	a4toclss	How important is each of the following in evaluating the children in your class for reporting to parents? Individual child's achievement relative to the rest of the class	1 (Not important) to 4 (Extremely important)	2.87	0.94
Eval standards	a4tostdr	How important is each of the following in evaluating the children in your class for reporting to parents? Individual child's achievement relative to local, state, or professional standards	1 (Not important) to 4 (Extremely important)	3.19	0.80
Eval effort	a4effrt	How important is each of the following in evaluating the children in your class for reporting to parents? Effort	1 (Not important) to 4 (Extremely important)	3.67	0.52
Eval improvement	a4impprg	How important is each of the following in evaluating the children in your class for reporting to parents? Individual improvement or progress over past performance	1 (Not important) to 4 (Extremely important)	3.74	0.47
Eval particip	a4clspar	How important is each of the following in evaluating the children in your class for reporting to parents? Class participation	1 (Not important) to 4 (Extremely important)	3.42	0.66

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Eval behavior	a4clsbhv	How important is each of the following in evaluating the children in your class for reporting to parents? Classroom behavior or conduct	1 (Not important) to 4 (Extremely important)	2.71	0.49
Eval co-op	a4cooprt	How important is each of the following in evaluating the children in your class for reporting to parents? Cooperativeness with other children	1 (Not important) to 4 (Extremely important)	3.48	0.64
Eval directions	a4flldir	How important is each of the following in evaluating the children in your class for reporting to parents? Ability to follow directions	1 (Not important) to 4 (Extremely important)	3.68	0.51
School Climate					
Attendance					
Attendance	s4ada	Approximately what is the average daily attendance for your school this year?	Proportion of students in attendance	0.96	0.02
Communication					
Freq parent-tchr confs	s4ptconf	Please indicate how often each of the following activities is provided by your school. Teacher-parent conferences	1 (Once a year or less) to 4 (7 or more times a year)	2.02	0.65
Freq rep cards	s4rptcrd	Please indicate how often each of the following activities is provided by your school. Reports of child's performance provided to parents	1 (At least 3 times a year) to 3 (7 or more times a year)	1.80	0.64
Freq test info	s4sttest	Please indicate how often each of the following activities is provided by your school. Information on the child's standardized assessment scores provided to parents	1 (Never) to 5 (7 or more times a year)	2.58	0.73
Freq admin-prnt mtgs	s4talkpt	Please estimate how many hours you spend on average per week in the following activities. Talking and meeting with parents	Hours	5.66	3.74
Support					

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Prnts supprtv	a4psupp	Please indicate the extent to which you agree or disagree with each of the following statements about your school. Parents are supportive of school staff	1 (Strongly disagree) to 5 (Strongly agree)	3.87	0.79
Commun support	s4spprt	Please indicate how much you agree or disagree with the following statements about the school's community and parents. The community served by this school is supportive of its goals and activities	1 (Strongly disagree) to 5 (Strongly agree)	4.28	0.71
Teacher Morale					
Tchr enjoys job	a4enjoy	To what extent do you agree or disagree with each of the following statements? I really enjoy my present teaching job.	1 (Strongly disagree) to 5 (Strongly agree)	4.42	0.74
Tchr make diff	a4mkdiff	To what extent do you agree or disagree with each of the following statements? I am making a difference in the lives of the children I teach.	1 (Strongly disagree) to 5 (Strongly agree)	4.54	0.57
Tchr teach again	a4teach	To what extent do you agree or disagree with each of the following statements? If I could start over, I would choose teaching again as my career.	1 (Strongly disagree) to 5 (Strongly agree)	4.28	0.94
Tchr accepted	a4accptd	Please indicate the extent to which you agree or disagree with each of the following statements about your school. I feel accepted and respected as a colleague by most staff members	1 (Strongly disagree) to 5 (Strongly agree)	4.46	0.70
Tchr low stds	a4stndlo	Please indicate the extent to which you agree or disagree with each of the following statements about your school. The academic standards at this school are too low	1 (Strongly disagree) to 5 (Strongly agree)	1.69	0.75
Admin. Leadership					
Admin priorities	a4setpri	Please indicate the extent to which you agree or disagree with each of the following statements about your school. The school administrator sets priorities, makes plans, and sees that they are carried out	1 (Strongly disagree) to 5 (Strongly agree)	4.01	0.87

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Admin encourage	a4encour	Please indicate the extent to which you agree or disagree with each of the following statements about your school. The school administration's behavior toward the staff is supportive and encouraging	1 (Strongly disagree) to 5 (Strongly agree)	4.24	0.76
Tchr new ideas	a4cntnlr	Please indicate the extent to which you agree or disagree with each of the following statements about your school. Teachers in this school are continually learning and seeking new ideas	1 (Strongly disagree) to 5 (Strongly agree)	4.06	0.90
Tchr recognition	a4recjob	Please indicate the extent to which you agree or disagree with each of the following statements about your school. In this school, staff members are recognized for a job well done	1 (Strongly disagree) to 5 (Strongly agree)	3.80	0.95
Tchr schl mission	a4missio	Please indicate the extent to which you agree or disagree with each of the following statements about your school. There is broad agreement among the entire school faculty about the central mission of the school	1 (Strongly disagree) to 5 (Strongly agree)	3.99	0.83
Tchr paperwork	a4paprwr	Please indicate the extent to which you agree or disagree with each of the following statements about your school. Routine administrative duties and paperwork interfere with my job of teaching	1 (Strongly disagree) to 5 (Strongly agree)	3.17	1.13
Consensus exp	a4cnsnss	Indicate how much you agree or disagree with the following statements about your school and staff. There is a consensus among administrators and teachers on goals and expectations	1 (Strongly disagree) to 5 (Strongly agree)	4.02	0.83
Behavior and Safety					
Class behaves	a4behvr	At this point in the school year, how would you rate the behavior of the children in your class?	1 (misbehave frequently) to 4 (behave exceptionally well)	2.43	0.81
Class disorder	s4disord	To the best of your knowledge how often do the following types of problems occur at your school? Widespread disorder in classrooms	1 (Never happens) to 5 (Daily)	1.29	0.52

Continued on next page

Table B.1: Source Variables, Labels, Item Text, and Summary Statistics for all Measures of School Quality

Label	Source Varname(s)	Item Text	Values	Mean	SD
Freq theft	s4theft	To the best of your knowledge how often do the following types of problems occur at your school? Theft	1 (Never happens) to 5 (Daily)	1.95	0.40
Freq bullying	s4bully	To the best of your knowledge how often do the following types of problems occur at your school? Student bullying	1 (Never happens) to 5 (Daily)	2.41	0.75
Freq fights	s4conflc	To the best of your knowledge how often do the following types of problems occur at your school? Physical conflicts among students	1 (Never happens) to 5 (Daily)	2.27	0.68
School Effectiveness					
Rd value-added	NA	Reading value-added in 1st grade estimated from linear mixed model	Test scores (standardized)	0.00	0.05
Mth value-added	NA	Reading value-added in 1st grade estimated from linear mixed model	Test scores (standardized)	0.00	0.05

Source: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), School District Finance Survey (F-33), fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), Local Education Agency Universe Survey, Public Elementary/Secondary School Survey, 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

C Dimension Reduction

In this appendix, we report results from several attempts to reduce the dimension of our 168 measures of the school environment. First, we applied a clustering algorithm known as partitioning around medoids (PAM) to group schools into discrete clusters that are internally similar but externally distinct. Figure C.1 displays the average silhouette width for a range of PAM solutions. This metric assesses how well each school fits within its assigned cluster relative to other clusters, where higher values indicate more clearly defined and homogeneous groupings. Values above 0.25 are generally considered to reflect at least some modest degree of clustering in the data, which would indicate that it might be reasonably summarized by a smaller number of discrete groups. In our case, however, the average silhouette width never exceeds 0.15 for any solution and remains below 0.05 for most, providing essentially no evidence of clustering in these data.

We also attempted to reduce the dimension of our school characteristics using principal components analysis (PCA) and factor analysis of mixed data (FAMD). These techniques aim to represent high-dimensional data using a smaller set of components or latent factors that retain most of its original variation. Figure C.2 presents a scree plot showing the proportion of variation explained by ten factors extracted via FAMD. No single factor accounts for more than 5% of the total variation in these data, and cumulatively, all ten explain less than 15%. Results from a PCA were nearly identical, even though this approach is not as suitable for data that include binary and ordinal variables.

In sum, these findings suggest that our high-dimensional measures of school quality are not easily reducible to a smaller set of underlying clusters or factors. Rather, they appear to capture distinct features of the school environment.

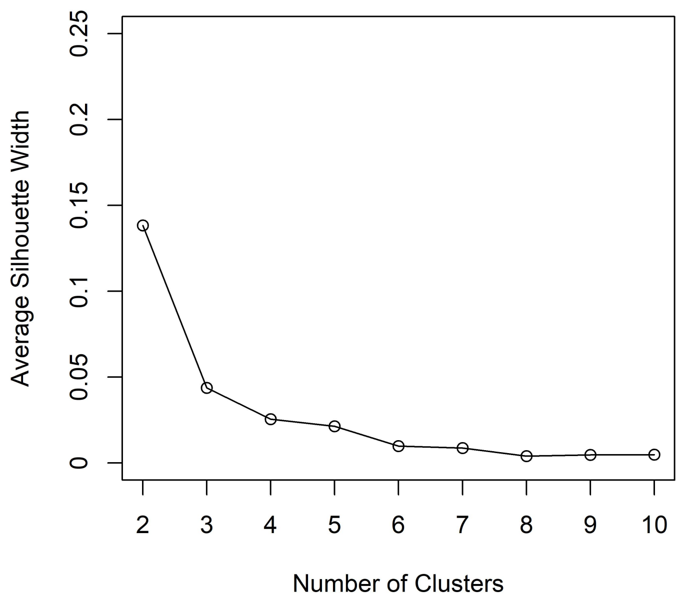

Figure C.1: Average Silhouette Width across Clustering Solutions based on Partitioning Around Medoids.

Note: This plot contains the average silhouette width for different clustering solutions obtained via partitioning around medoids. Source: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

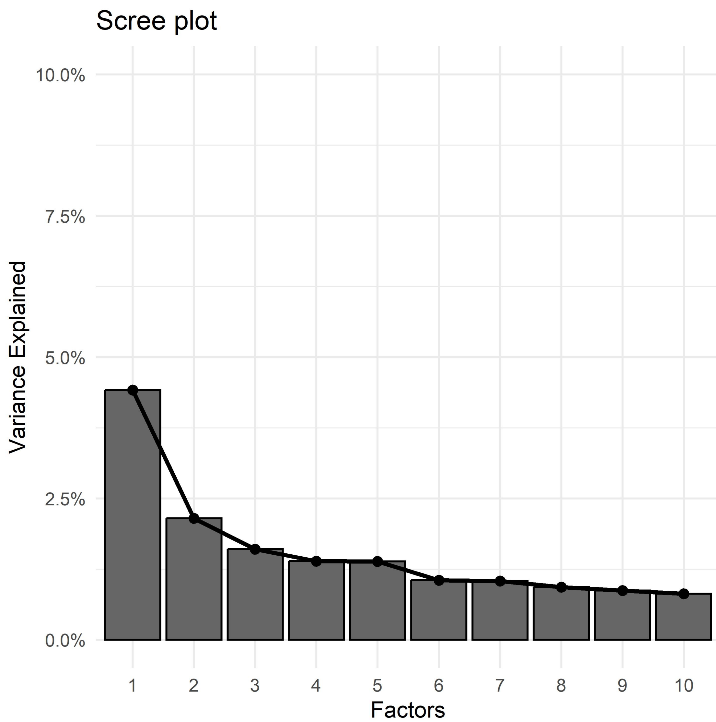

Figure C.2: Scree Plot based on a Factor Analysis of Mixed Data.

Note: This plot contains the proportion of total variation explained by each factor obtained from a factor analysis of mixed data (FAMD). Source: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

D Sample Characteristics in the ECLS-K

In this appendix, we report a full set of summary statistics and describe how key control variables were measured. Gender is coded as a binary variable, 1 for male and 0 otherwise. Race captures whether the sample member is White (non-Hispanic), Black (non-Hispanic), Hispanic, or another race. And birth weight is recorded in pounds. Parental age is measured in years, marital status is a binary indicator coded 1 for "married" and 0 otherwise, and family income is expressed in dollars. Employment status indicates whether a parent is "not in the labor force," "working fewer than 35 hours per week," or "working 35 hours or more per week." The highest level of education achieved by a child's parents is classified into categories ranging from "less than high school" to a "graduate degree," while parental occupation is measured using the highest occupational status score achieved by either parent. The primary language spoken at home is a binary variable coded 1 for "English" and 0 otherwise. Family size reflects the total number of people residing in the child's household. Parental involvement is measured by the frequency with which parents read to or practice numbers with the sample member, where responses range from "not at all" to "every day." Government assistance is captured with a series of binary indicators denoting whether the family received benefits from the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), the Supplemental Nutrition Assistance Program (SNAP), or the Temporary Assistance for Needy Families (TANF) program. Urban versus rural residence is categorized as living in a large, medium, or small city, a suburb, or a rural area. Geographic region is measured using an indicator for whether a sample member lived in the "Northeast," "Midwest," "South," or "West" census divisions. Summary statistics for all control variables, along with our measures of neighborhood poverty and test scores, are provided in Table D.1.

Table D.1: Sample Characteristics.

Variable	Percent	Mean	Std. Dev.
Outcomes			
Reading test scores, 3rd grade spring		2.94	0.37
Reading test scores, 4th grade spring		3.15	0.37
Reading test scores, 5th grade spring		3.36	0.43
Math test scores, 3rd grade spring		3.64	0.65
Math test scores, 4th grade spring		3.95	0.65
Math test scores, 5th grade spring		4.23	0.66
Neighborhood poverty			
Neighborhood poverty level			
High poverty ($\geq 20\%$)	23.57		
Low poverty ($< 20\%$)	76.43		
Neighborhood poverty rate		0.14	0.11
Baseline confounders			
Gender			
Male	48.78		
Female	51.22		
Race			
White (non-Hispanic)	47.62		
Black (non-Hispanic)	12.82		
Hispanic	25.52		
Asian	8.07		
Other	5.98		
Child age (months)		67.41	4.47
Birth weight (lbs)		7.24	1.33
Parents married at birth	66.30		
English is child's first language	82.19		

Continued on next page

Table D.1: Sample Characteristics

Variable	Percent	Mean	Std. Dev.
Household size		4.57	1.36
Parent 1 age		33.97	6.75
Parent 2 age		36.52	7.36
Parent 1 employment status			
Not in the labor force	38.86		
Less than 35 hours per week	18.41		
35 or more hours per week	42.73		
Parent 2 employment status			
Not in the labor force	8.58		
Less than 35 hours per week	4.53		
35 or more hours per week	86.89		
Parent education (highest level)			
Less than high school diploma	9.66		
High school diploma	20.58		
Vocational/technical degree	32.01		
Bachelor's degree	21.63		
Graduate degree	16.11		
Family income ($1000s)		63.41	58.06
Parent occupational status (highest level)		45.51	11.95
Family received WIC in past 6 months	50.94		
Family received food stamps in past year	27.12		
Family received TANF ever	4.80		
Parent is currently married	74.19		
Two biological parents in household	71.29		
Parent practices numbers with child			
Not at all	0.51		
Once or twice a week	5.92		
3-6 times a week	27.78		

Continued on next page

Table D.1: Sample Characteristics

Variable	Percent	Mean	Std. Dev.
Everyday	65.79		
Parent reads books to child			
Not at all	1.01		
Once or twice a week	12.31		
3-6 times a week	32.34		
Everyday	54.34		
Parental expectations for child			
No post-secondary education	4.35		
Some post-secondary education	12.19		
Bachelor's degree	47.07		
Graduate degree	36.39		
Locale			
Large city	15.67		
Medium city	9.71		
Small city	7.32		
Suburb	36.33		
Rural	30.97		
Region			
Northeast	15.44		
Midwest	22.08		
South	36.97		
West	25.50		
Child externalizing behaviors (teacher reported), fall K		1.58	0.62
Child internalizing behaviors (teacher reported), fall K		1.45	0.49
Child motivation level (interviewer reported), fall K		3.41	0.85
Child cooperation level (interviewer reported), fall K		3.93	0.79
Child attention level (interviewer reported), fall K		3.35	0.90
Child health (parent reported), fall K		1.55	0.79

Continued on next page

Table D.1: Sample Characteristics

Variable	Percent	Mean	Std. Dev.
Reading test scores, fall K		0.00	1.00
Math test scores, fall K		0.00	1.00

Note: All test scores are standardized using the mean and standard deviation from the fall of kindergarten assessment.

Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File.

E Analyses based on Fourth and Fifth Grade Test Scores

This appendix reports a set of parallel analyses using test scores from assessments administered by the ECLS-K during the spring of fourth and fifth grade. Figure E.1 presents results from the descriptive decomposition of disparities in fourth grade scores, while Figure E.2 shows the corresponding results for fifth grade. Similarly, Tables E.1 and E.2 report estimates of the disparity eliminated based on fourth and fifth grade test scores, respectively. Across both grades, the results are similar to those presented in the main text, which are based on assessments administered during the spring of third grade.

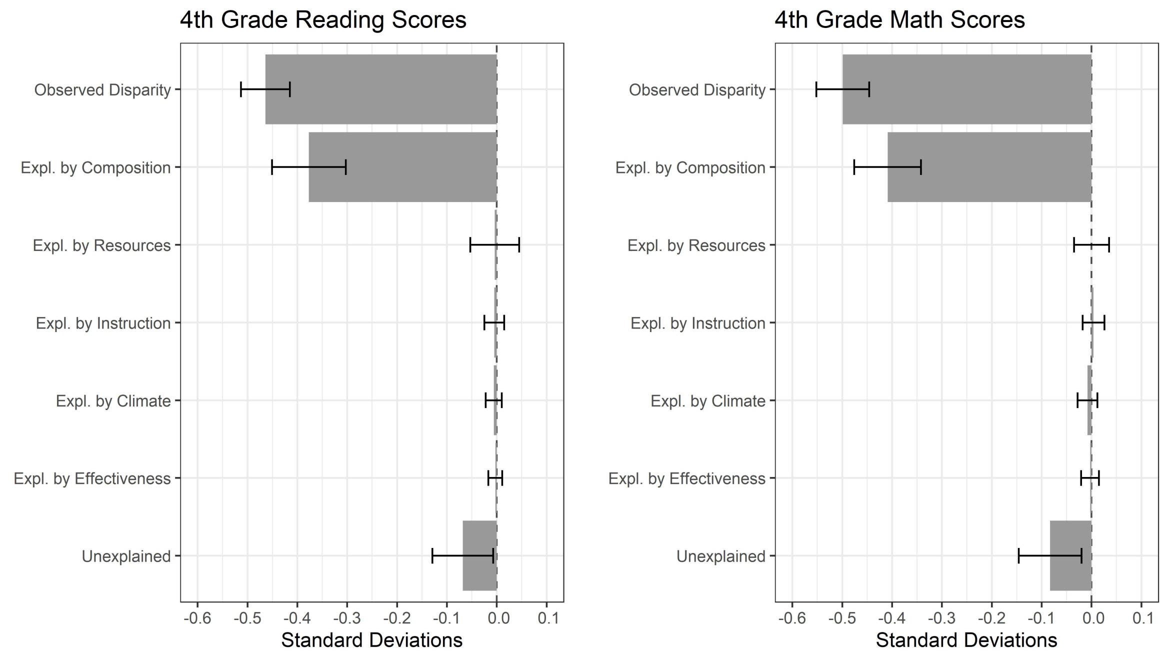

Figure E.1: Descriptive Decomposition of the Observed Disparity in 4th Grade Test Scores between Students from High- and Low-poverty Neighborhoods.

Note: This plot contains de-biased machine learning estimates of the "disparity explained" by different dimensions of school quality. All estimates are reported in standard deviation units. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

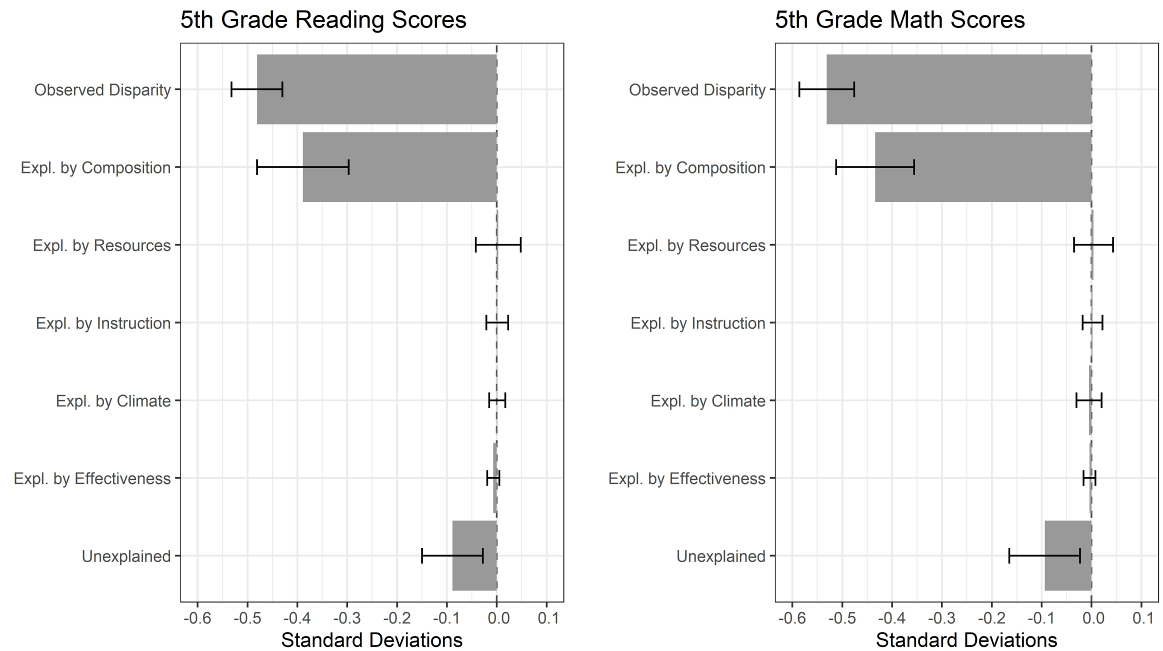

Figure E.2: Descriptive Decomposition of the Observed Disparity in 5th Grade Test Scores between Students from High- and Low-poverty Neighborhoods.

Note: This plot contains de-biased machine learning estimates of the "disparity explained" by different dimensions of school quality. All estimates are reported in standard deviation units. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

Table E.1: De-biased Machine Learning Estimates of the Disparity Eliminated between Students from High- and Low-poverty Neighborhoods based on Test Scores Measured during the Spring of 4th Grade.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.464	[-0.513, -0.415]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.021	[-0.068, 0.026]	0.05
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.022	[-0.067, 0.023]	0.05
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.021	[-0.070, 0.028]	0.05
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.024	[-0.071, 0.023]	0.05
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.022	[-0.069, 0.025]	0.05
Math Test Scores				
Observed disparity	Δ	-0.499	[-0.552, -0.446]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.011	[-0.060, 0.038]	0.02
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.017	[-0.066, 0.032]	0.03
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.012	[-0.061, 0.037]	0.02
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.014	[-0.063, 0.035]	0.03
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.016	[-0.065, 0.033]	0.03

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table E.2: De-biased Machine Learning Estimates of the Disparity Eliminated between Students from High- and Low-poverty Neighborhoods based on Test Scores Measured during the Spring of 5th Grade.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.481	[-0.532, -0.430]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.019	[-0.062, 0.024]	0.04
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.018	[-0.063, 0.027]	0.04
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.014	[-0.063, 0.035]	0.03
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.017	[-0.064, 0.030]	0.04
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.014	[-0.061, 0.033]	0.03
Math Test Scores				
Observed disparity	Δ	-0.531	[-0.586, -0.476]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.032	[-0.081, 0.017]	0.06
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.037	[-0.086, 0.012]	0.07
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.035	[-0.082, 0.012]	0.07
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.036	[-0.083, 0.011]	0.07
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.033	[-0.078, 0.012]	0.06

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

F DML Estimation of the Disparity Eliminated

In this section, we describe how the disparity eliminated can be nonparametrically identified and consistently estimated using machine learning (ML) methods. Nonparametric identification of the disparity eliminated depends on three assumptions: ignorability, positivity, and consistency.

The ignorability assumption can be formally expressed as $Y\left(\underline{\mathbf{s}}_j\right) \perp \underline{\mathbf{S}}_j \mid X=1, \mathbf{C}$ for $j=1,\ldots,J$, where $Y\left(\underline{\mathbf{s}}_j\right)$ denotes a student's potential outcome under exposure to a school with characteristics given by $\underline{\mathbf{s}}_j$. Substantively, this assumption implies that, conditional on living in a high-poverty neighborhood ($X=1$) and the set of baseline covariates $\mathbf{C}$, school selection is as good as random.

The positivity assumption requires that $P\left(\underline{\mathbf{S}}_j = \underline{\mathbf{s}}_j \mid X, \mathbf{C}\right) > 0$ for any $\underline{\mathbf{s}}_j$. In words, this condition requires that there must be a nonzero probability of attending schools with any combination of characteristics in $\underline{\mathbf{S}}_j$ among individuals with different values on the baseline covariates $\mathbf{C}$ and who live in neighborhoods with different levels of poverty.

Finally, the consistency assumption requires that a student's observed outcome must correspond to their potential outcome under the set of school characteristics they did in fact experience. Formally, this condition can be expressed as $Y = Y\left(\underline{\mathbf{S}}_j\right)$.

To identify the disparity eliminated, it suffices to identify $\eta_{1,\underline{\mathbf{S}}_j} = \mathbb{E}\left[Y\left(\underline{\mathbf{G}}_j\right)|X=1\right]$, where $\underline{\mathbf{G}}_j$ represents a random vector of school characteristics drawn from the joint distribution observed among students from low-poverty neighborhoods ($X=0$) with baseline covariates $\mathbf{C}$. Under the assumptions of ignorability, positivity, and consistency, $\eta_{1,\underline{\mathbf{S}}_j}$ is nonparametrically identified by the following function of observable data:

$$
\begin{aligned}
\eta_{1,\underline{\mathbf{S}}_j} &= \mathbb{E}\left[Y\left(\underline{\mathbf{G}}_j\right)|X=1\right] \\
&= \sum_{\mathbf{c},\underline{\mathbf{g}}_j} \mathbb{E}\left[Y\left(\underline{\mathbf{g}}_j\right)|X=1,\mathbf{C}\right] P\left(\underline{\mathbf{G}}_j = \underline{\mathbf{g}}_j|\mathbf{C}\right) P\left(\mathbf{C}|X=1\right) && \text{by iterated expectations} \\
&= \sum_{\mathbf{c},\underline{\mathbf{s}}_j} \mathbb{E}\left[Y\left(\underline{\mathbf{s}}_j\right)|X=1,\mathbf{C}\right] P\left(\underline{\mathbf{S}}_j = \underline{\mathbf{s}}_j|X=0,\mathbf{C}\right) P\left(\mathbf{C}|X=1\right) && \text{by definition of } \underline{\mathbf{G}}_j \\
&= \sum_{\mathbf{c},\underline{\mathbf{s}}_j} \mathbb{E}\left[Y\left(\underline{\mathbf{s}}_j\right)|X=1,\mathbf{C},\underline{\mathbf{S}}_j = \underline{\mathbf{s}}_j\right] P\left(\underline{\mathbf{S}}_j = \underline{\mathbf{s}}_j|X=0,\mathbf{C}\right) P\left(\mathbf{C}|X=1\right) && \text{by ignorability/positivity} \\
&= \sum_{\mathbf{c},\underline{\mathbf{s}}_j} \mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j = \underline{\mathbf{s}}_j\right] P\left(\underline{\mathbf{S}}_j = \underline{\mathbf{s}}_j|X=0,\mathbf{C}\right) P\left(\mathbf{C}|X=1\right) && \text{by consistency} \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]|X=1\right]. && \text{by definition}
\end{aligned}
$$

The final equality provides the nonparametric identification formula for $\eta_{1,\underline{\mathbf{S}}_j}$. The innermost expectation in this expression captures the average value of the outcome Y for individuals in high-poverty neighborhoods

$(X=1)$, conditional on both their school characteristics $\underline{\mathbf{S}}_j$ and baseline covariates $\mathbf{C}$. This expected value is then averaged over the distribution of school characteristics observed in low-poverty neighborhoods $(X=0)$, conditional on $\mathbf{C}$. Finally, the outermost expectation averages this result again, now over the distribution of the baseline covariates $\mathbf{C}$ observed among those in high-poverty neighborhoods $(X=1)$. This identification formula is a special case of the generalized mediation functional introduced by Zhou (2022).

Under the same set of assumptions, the identification formula for $\eta_{1,\underline{\mathbf{S}}_j}$ can also be expressed as follows:

$$\begin{aligned}
\eta_{1,\underline{\mathbf{S}}_j} = \mathbb{E}\Big[& \underbrace{\frac{X}{P(X=1)}\frac{P(X=1|\mathbf{C})}{P(X=0|\mathbf{C})}\frac{P\left(X=0|\mathbf{C},\underline{\mathbf{S}}_j\right)}{P\left(X=1|\mathbf{C},\underline{\mathbf{S}}_j\right)}\left(Y-\mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]\right)}_{\text{term 1}} \\
& + \underbrace{\frac{1-X}{P(X=1)}\frac{P(X=1|\mathbf{C})}{P(X=0|\mathbf{C})}\left(\mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]-\mathbb{E}\left[\mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]\right)}_{\text{term 2}} \\
& + \underbrace{\frac{X}{P(X=1)}\left(\mathbb{E}\left[\mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]-\mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]|X=1\right]\right)}_{\text{term 3}} \\
& + \underbrace{\mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]|X=1\right]}_{\text{term 4}}\Big].
\end{aligned}$$

In this expression, term 4 corresponds exactly to the identification formula for $\eta_{1,\underline{\mathbf{S}}_j}$ derived previously, while the remaining terms—term 1, term 2, and term 3—have mean zero by construction. This alternative identification formula is useful for constructing a multiply robust estimator that is amenable for use with ML methods.

Specifically, it indicates that $\eta_{1,\underline{\mathbf{S}}_j}$ can be estimated using a combination of regression imputation and inverse probability weighting, as outlined by the following expression:

$$\begin{aligned}
\hat{\eta}_{1,\underline{\mathbf{S}}_j} = \frac{1}{n}\sum & \frac{X}{\hat{P}(X=1)}\frac{\hat{P}(X=1|\mathbf{C})}{\hat{P}(X=0|\mathbf{C})}\frac{\hat{P}\left(X=0|\mathbf{C},\underline{\mathbf{S}}_j\right)}{\hat{P}\left(X=1|\mathbf{C},\underline{\mathbf{S}}_j\right)}\left(Y-\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]\right) \\
& + \frac{1-X}{\hat{P}(X=1)}\frac{\hat{P}(X=1|\mathbf{C})}{\hat{P}(X=0|\mathbf{C})}\left(\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]-\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]\right) \\
& + \frac{X}{\hat{P}(X=1)}\left(\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]-\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]|X=1\right]\right) \\
& + \hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]|X=1\right].
\end{aligned}$$

In this estimator, $\hat{P}(X=x|\mathbf{C})$ denotes the estimated probability of living in a neighborhood with a poverty level given by x, conditional on the baseline covariates $\mathbf{C}$. Similarly, $\hat{P}(X=x|\mathbf{C},\underline{\mathbf{S}}_j)$ denotes an estimate for this same probability but now conditional on both the baseline covariates $\mathbf{C}$ and school characteristics $\underline{\mathbf{S}}_j$. The term $\hat{\mathbb{E}}[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j]$ denotes the estimated mean of student test scores given the baseline covariates

$\mathbf{C}$ and school characteristics $\underline{\mathbf{S}}_j$, setting X equal to one. The iterated expectation, $\hat{\mathbb{E}}[\hat{\mathbb{E}}[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j]|X=0,\mathbf{C}]$, estimates the average value of this mean, conditional on the baseline controls, now setting X equal to zero. Each of these components is referred to as a nuisance term—that is, a quantity not of immediate scientific interest, but still needed to construct an estimator for the target quantity $\eta_{1,\underline{\mathbf{S}}_j}$. The hats are used to distinguish estimates of the nuisance terms from their true but unknown values.

The estimator $\hat{\eta}_{1,\underline{\mathbf{S}}_j}$ has several desirable properties. First, it does not require modeling the distribution of $\underline{\mathbf{S}}_j$, which would be prohibitively difficult given its high dimensionality.

Second, $\hat{\eta}_{1,\underline{\mathbf{S}}_j}$ is multiply robust, meaning that it achieves consistency under multiple, distinct conditions. In particular, it is consistent if any one of the following three conditions holds, in addition to the identification assumptions outlined previously: (i) the outcome mean $\mathbb{E}\left[Y|X,\mathbf{C},\underline{\mathbf{S}}_j\right]$ and its iterated expectation $\mathbb{E}\left[\mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X,\mathbf{C}\right]$ are correctly modeled; (ii) the outcome mean $\mathbb{E}\left[Y|X,\mathbf{C},\underline{\mathbf{S}}_j\right]$ and the probability of group membership $P(X=x|\mathbf{C})$ are correctly modeled; or (iii) both probabilities of group membership, $P(X=x|\mathbf{C})$ and $P(X=x|\mathbf{C},\underline{\mathbf{S}}_j)$, are correctly modeled. When condition (i) holds, $\hat{\eta}_{1,\underline{\mathbf{S}}_j}$ converges to a consistent regression-imputation estimator. Under condition (ii), it converges to an imputation-based weighting estimator that is also consistent. And under condition (iii), it converges to a consistent inverse probability weighting estimator.

Although the multiple robustness property of $\hat{\eta}_{1,\underline{\mathbf{S}}_j}$ offers some protection against model misspecification, accurately estimating *any* of its nuisance terms can be very difficult in practice, particularly when the vector of school characteristics $\underline{\mathbf{S}}_j$ is high-dimensional. A third key advantage of this estimator, however, is its compatibility with data-adaptive ML algorithms for fitting the requisite nuisance terms. This advantage essentially obviates the need to correctly specify models for the outcome or for group membership.

The estimator's compatibility with ML methods stems from the form of its bias, which involves a sum of products of prediction errors across different pairs of nuisance terms. This multiplicative structure implies that $\hat{\eta}_{1,\underline{\mathbf{S}}_j}$ can attain $\sqrt{n}$-consistency even when the nuisance terms themselves are estimated with methods that converge at slower rates, such as those typically achieved by data-adaptive ML algorithms. Specifically, $\sqrt{n}$-consistency for $\hat{\eta}_{1,\underline{\mathbf{S}}_j}$ can be achieved as long as the product of the convergence rates for each relevant pair of nuisance terms exceeds $\sqrt{n}$. For example, if the methods used to estimate each nuisance term all converge at a rate faster than $n^{1/4}$, then $\hat{\eta}_{1,\underline{\mathbf{S}}_j}$ will be $\sqrt{n}$-consistent. This rate condition can be achieved by a wide range of ML algorithms, including those employed in the present study (Wodtke and Zhou Forthcoming).

G Alternative Contrasts between Neighborhoods

In this appendix, we replicate our descriptive decomposition and causal analyses using alternative definitions of high- and low-poverty neighborhoods. Specifically, we estimate observed, adjusted, and counterfactual disparities comparing students from neighborhoods with poverty rates of (i) $\geq 30\%$ versus $< 30\%$, (ii) $\geq 20\%$ versus $\leq 10\%$, and (iii) $\geq 30\%$ versus $\leq 5\%$. Results from the descriptive decomposition based on these alternative contrasts are presented in Figures G.1 to G.3, while estimates of the disparity eliminated are shown in Tables G.1 to G.3.

Across all of these contrasts, our results are substantively similar to those reported in the main text, which focused on neighborhoods with poverty rates of $\geq 20\%$ versus $< 20\%$. The main difference is that the observed disparities grow larger as the contrast between neighborhood poverty rates becomes more extreme. However, estimates for the disparity explained and the disparity eliminated remain broadly similar, both in absolute and relative terms. These findings suggest that our key conclusions are robust to variations in how high- and low-poverty neighborhoods are operationally defined.

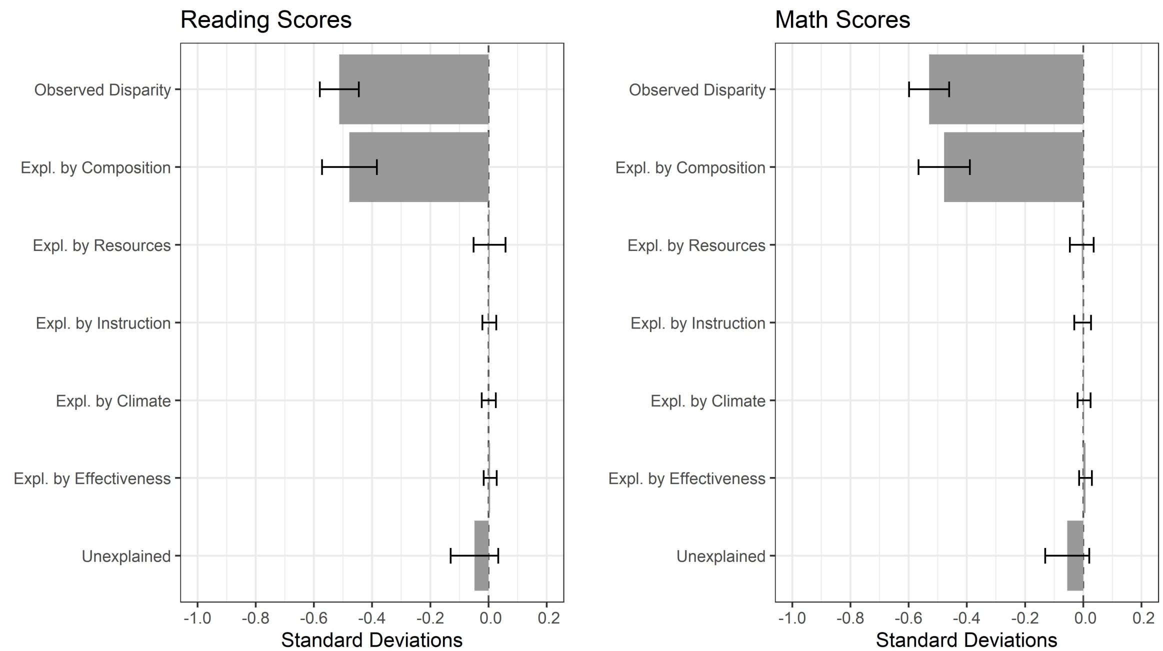

Figure G.1: Descriptive Decomposition of the Observed Disparity in Achievement Test Scores between Students from Neighborhoods with a Poverty Rate $\geq 30\%$ versus $< 30\%$.

Note: This plot contains de-biased machine learning estimates of the "disparity explained" by different dimensions of school quality. All estimates are reported in standard deviation units. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

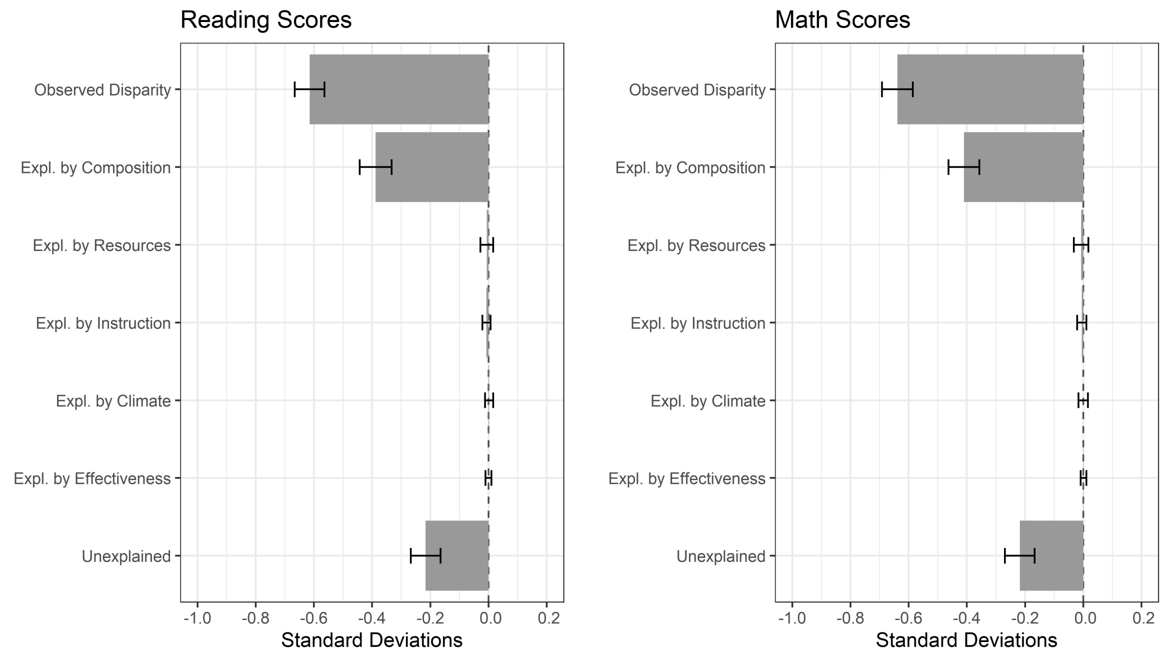

Figure G.2: Descriptive Decomposition of the Observed Disparity in Achievement Test Scores between Students from Neighborhoods with a Poverty Rate $\geq 20\%$ versus $\leq 10\%$.

Note: This plot contains de-biased machine learning estimates of the "disparity explained" by different dimensions of school quality. All estimates are reported in standard deviation units. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

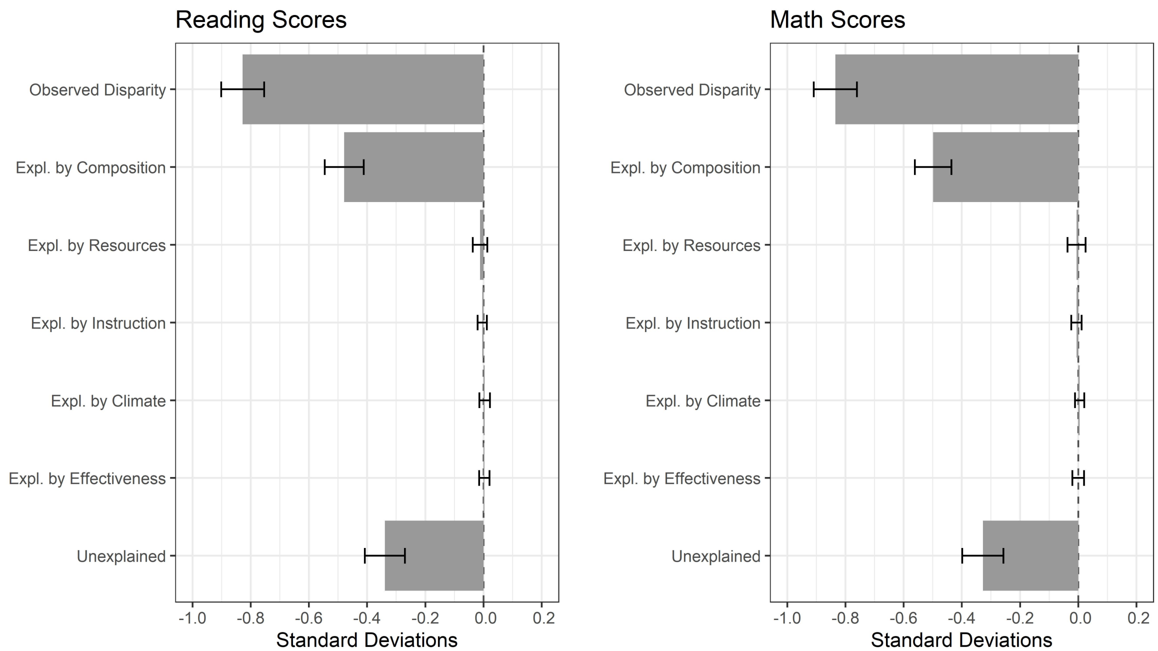

Figure G.3: Descriptive Decomposition of the Observed Disparity in Achievement Test Scores between Students from Neighborhoods with a Poverty Rate $\geq 30\%$ versus $\leq 5\%$.

Note: This plot contains de-biased machine learning estimates of the "disparity explained" by different dimensions of school quality. All estimates are reported in standard deviation units. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

Table G.1: De-biased Machine Learning Estimates of the Disparity Eliminated between Students from Neighborhoods with a Poverty Rate $\geq$ 30% versus $<$ 30%.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.513	[-0.580, -0.446]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.046	[-0.136, 0.044]	0.09
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.033	[-0.117, 0.051]	0.06
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.038	[-0.122, 0.046]	0.07
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.037	[-0.119, 0.045]	0.07
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.031	[-0.117, 0.055]	0.06
Math Test Scores				
Observed disparity	Δ	-0.530	[-0.599, -0.461]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.012	[-0.086, 0.062]	0.02
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.022	[-0.095, 0.051]	0.04
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.019	[-0.095, 0.057]	0.04
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.021	[-0.101, 0.059]	0.04
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.017	[-0.091, 0.057]	0.03

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. The disparities contrast the top with the bottom four quintiles of a multidimensional disadvantage index, which comes from a principal components analysis of tract poverty, educational composition, family structure, unemployment, and racial composition. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table G.2: De-biased Machine Learning Estimates of the Disparity Eliminated between Students from Neighborhoods with a Poverty Rate ≥ 20% versus ≤ 10%.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.615	[-0.666, -0.564]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.038	[-0.083, 0.007]	0.06
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.035	[-0.080, 0.010]	0.06
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.035	[-0.080, 0.010]	0.06
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.038	[-0.083, 0.007]	0.06
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.039	[-0.086, 0.008]	0.06
Math Test Scores				
Observed disparity	Δ	-0.639	[-0.692, -0.586]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.024	[-0.071, 0.023]	0.04
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.025	[-0.076, 0.026]	0.04
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.022	[-0.075, 0.031]	0.03
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.023	[-0.076, 0.030]	0.04
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.023	[-0.074, 0.028]	0.04

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table G.3: De-biased Machine Learning Estimates of the Disparity Eliminated between Students from Neighborhoods with a Poverty Rate $\geq 30\%$ versus $\leq 5\%$.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.828	[-0.902, -0.754]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.064	[-0.178, 0.050]	0.08
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.071	[-0.175, 0.033]	0.09
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.065	[-0.159, 0.029]	0.08
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.069	[-0.177, 0.039]	0.08
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.065	[-0.165, 0.035]	0.08
Math Test Scores				
Observed disparity	Δ	-0.835	[-0.909, -0.761]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.031	[-0.113, 0.051]	0.04
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.050	[-0.152, 0.052]	0.06
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.064	[-0.164, 0.036]	0.08
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.058	[-0.150, 0.034]	0.07
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.058	[-0.150, 0.034]	0.07

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

H The Disparity Eliminated by Marginal Equalization

In our main analysis, we assessed the impact of equalizing schools on disparities in achievement by focusing on post-intervention gaps that would result from equalizing the distribution of school characteristics between students from high- and low-poverty neighborhoods who otherwise share similar baseline covariates. Alternatively, we might consider an intervention that equalizes the distribution of school characteristics between students from high- and low-poverty neighborhoods regardless of their baseline attributes. The corresponding post-intervention gap can be defined as follows:

$$\gamma_{\underline{\mathbf{S}}_j} = \mathbb{E}\big[Y\big(\underline{\mathbf{H}}_j\big)|X=1\big] - \mathbb{E}\big[Y\big(\underline{\mathbf{H}}_j\big)|X=0\big], \tag{1}$$

where $\underline{\mathbf{H}}_j$ denotes a vector of school characteristics randomly drawn from $f(\underline{\mathbf{S}}_j|X=0)$—that is, from the joint distribution of school characteristics $\underline{\mathbf{S}}_j = \{\mathbf{S}_1, ..., \mathbf{S}_j\}$ observed among students in low-poverty neighborhoods, without conditioning on the baseline covariates $\mathbf{C}$. This alternative estimand captures the achievement gap that would persist if every student attended a school with characteristics randomly drawn from their joint distribution observed in low-poverty neighborhoods, regardless of the student's individual characteristics. Thus, unlike the intervention in our main analysis, this intervention constitutes a form of *marginal* as opposed to *conditional* equalization (Opacic et al., 2025). A corresponding version of the disparity eliminated can be defined as $\Delta - \gamma_{\underline{\mathbf{S}}_j}$, where $\Delta = \mathbb{E}[Y|X=1] - \mathbb{E}[Y|X=0]$.

The post-intervention gap $\gamma_{\underline{\mathbf{S}}_j}$ can be nonparametrically identified under the same assumptions of ignorability, positivity, and consistency discussed in Appendix F. To identify the post-intervention gap, it suffices to identify $\zeta_{1,\underline{\mathbf{S}}_j} = \mathbb{E}[Y(\underline{\mathbf{H}}_j)|X=1]$ and $\zeta_{0,\underline{\mathbf{S}}_j} = \mathbb{E}[Y(\underline{\mathbf{H}}_j)|X=0]$—that is, the post-intervention mean scores for students from high-poverty and low-poverty neighborhoods, respectively. Without loss of generality, let us consider identifying $\zeta_{1,\underline{\mathbf{S}}_j}$. Under the assumptions of ignorability, positivity, and consistency, $\zeta_{1,\underline{\mathbf{S}}_j}$ is nonparametrically identified by the following function of observable data:

$$\begin{aligned}
\zeta_{1,\underline{\mathbf{S}}_j} &= \mathbb{E}\big[Y\big(\underline{\mathbf{H}}_j\big)|X=1\big] \\
&= \sum_{\mathbf{c},\underline{\mathbf{h}}_j} \mathbb{E}\big[Y\big(\underline{\mathbf{h}}_j\big)|X=1,\mathbf{C}\big] P\left(\underline{\mathbf{H}}_j = \underline{\mathbf{h}}_j|\mathbf{C}\right) P\left(\mathbf{C}|X=1\right) && \text{by iterated expectations} \\
&= \sum_{\mathbf{c},\underline{\mathbf{s}}_j} \mathbb{E}\big[Y\big(\underline{\mathbf{s}}_j\big)|X=1,\mathbf{C}\big] P\left(\underline{\mathbf{S}}_j = \underline{\mathbf{s}}_j|X=0\right) P\left(\mathbf{C}|X=1\right) && \text{by definition of } \underline{\mathbf{H}}_j \\
&= \sum_{\mathbf{c},\underline{\mathbf{s}}_j} \mathbb{E}\big[Y\big(\underline{\mathbf{s}}_j\big)|X=1,\mathbf{C},\underline{\mathbf{S}}_j=\underline{\mathbf{s}}_j\big] P\left(\underline{\mathbf{S}}_j = \underline{\mathbf{s}}_j|X=0\right) P\left(\mathbf{C}|X=1\right) && \text{by ignorability/positivity} \\
&= \sum_{\mathbf{c},\underline{\mathbf{s}}_j} \mathbb{E}\big[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j=\underline{\mathbf{s}}_j\big] P\left(\underline{\mathbf{S}}_j = \underline{\mathbf{s}}_j|X=0\right) P\left(\mathbf{C}|X=1\right) && \text{by consistency.}
\end{aligned}$$

The final equality provides the nonparametric identification formula for $\zeta_{1,\underline{\mathbf{S}}_j}$. The expectation $\mathbb{E}\big[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j=\underline{\mathbf{s}}_j\big]$ captures the average value of the outcome Y for students in high-poverty neighborhoods ($X=1$), conditional on both their school characteristics $\underline{\mathbf{S}}_j$ and baseline covariates $\mathbf{C}$. This expected value is then averaged over the distribution of school characteristics observed in low-poverty neighborhoods ($X=0$). Finally, this result is averaged again, now over the distribution of the baseline covariates $\mathbf{C}$ observed among those in high-poverty neighborhoods ($X=1$). Note that, unlike the identification formula for $\eta_{1,\underline{\mathbf{S}}_j}$, the expression above cannot be written as an iterated expectation.

Under the same set of assumptions, the identification formula for $\zeta_{1,\underline{\mathbf{S}}_j}$ can also be written as follows:

$$\begin{aligned}\zeta_{1,\underline{\mathbf{S}}_j} = \mathbb{E}\Big[&\underbrace{\frac{X}{P(X=1)}\frac{P\left(\underline{\mathbf{S}}_j|X=0\right)}{P\left(\underline{\mathbf{S}}_j|X=1,\mathbf{C}\right)}\left(Y-\mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]\right)}_{\text{term 1}} \\ &+ \underbrace{\frac{X}{P(X=1)}\sum_{\underline{\mathbf{s}}_j}\mathbb{E}\big[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j=\underline{\mathbf{s}}_j\big]P\left(\underline{\mathbf{S}}_j=\underline{\mathbf{s}}_j|X=0\right)}_{\text{term 2}}\Big].\end{aligned}$$

In this expression, the expected value of term 2 is equal to the identification formula for $\zeta_{1,\underline{\mathbf{S}}_j}$ derived previously, while term 1 has mean zero by construction. This alternative identification formula is useful for constructing a doubly robust estimator that is amenable for use with ML methods.

Specifically, it indicates that $\zeta_{1,\underline{\mathbf{S}}_j}$ can be estimated using a combination of inverse probability weighting and Monte Carlo simulation, as outlined by the following expression:

$$\begin{aligned}\hat{\zeta}_{1,\underline{\mathbf{S}}_j} = \frac{1}{n}\sum\Bigg(&\frac{X}{\hat{P}(X=1)}\underbrace{\hat{W}\left(Y-\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]\right)}_{\text{term 1}} \\ &+ \frac{X}{\hat{P}(X=1)}\underbrace{\sum_{\underline{\mathbf{s}}_j}\hat{\mathbb{E}}\big[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j=\underline{\mathbf{s}}_j\big]\hat{P}\left(\underline{\mathbf{S}}_j=\underline{\mathbf{s}}_j|X=0\right)}_{\text{term 2}}\Bigg),\end{aligned}$$

where $\hat{W}$ denotes an estimator for the weight $W=\frac{P\left(\underline{\mathbf{S}}_j|X=0\right)}{P\left(\underline{\mathbf{S}}_j|X=1,\mathbf{C}\right)}$. In this expression, $\hat{P}(X=1)$ represents the sample proportion of students residing in high-poverty neighborhoods, $\hat{P}\left(\underline{\mathbf{S}}_j=\underline{\mathbf{s}}_j|X=0\right)$ denotes an estimated probability of school characteristics within low-poverty neighborhoods, and $\hat{\mathbb{E}}[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j]$ denotes an estimated mean test score given the student's baseline covariates $\mathbf{C}$ and school characteristics $\underline{\mathbf{S}}_j$, after setting neighborhood poverty X to 1. Substantively, term 2 can be interpreted as the predicted test score for a student living in a high-poverty neighborhood with baseline characteristics $\mathbf{C}$ under an intervention that equalizes school characteristics marginally.

Because the school characteristics $\underline{\mathbf{S}}_j$ are high-dimensional, it is difficult to directly estimate their probability. Nonetheless, term 2 in the above expression can be evaluated through Monte Carlo simulation. Specifically, for each student, term 2 can be computed using an average of the estimated scores $\hat{\mathbb{E}}[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j]$ across a large number of different values for $\underline{\mathbf{S}}_j$ drawn randomly from their joint distribution observed in low-poverty neighborhoods.

Estimating term 1 is similarly complicated by the curse of dimensionality. This is because the weight W also involves probabilities of school characteristics within low- and high-poverty neighborhoods, which are hard to estimate directly. To circumvent this challenge, we leverage a novel technique called permutation weighting (Arbour et al., 2021).

In our implementation of this approach, we begin by using Bayes' rule to rewrite the weight W as follows:

$$\begin{aligned} W &= \frac{P\left(\underline{\mathbf{S}}_j|X=0\right)}{P\left(\underline{\mathbf{S}}_j|X=1,\mathbf{C}\right)} \\ &= \frac{P(\underline{\mathbf{S}}_j|X=0)}{P(\underline{\mathbf{S}}_j|X=0,\mathbf{C})}\frac{P(\underline{\mathbf{S}}_j|X=0,\mathbf{C})}{P(\underline{\mathbf{S}}_j|X=1,\mathbf{C})} \\ &= \underbrace{\frac{P(\underline{\mathbf{S}}_j|X=0)P(\mathbf{C}|X=0)}{P(\underline{\mathbf{S}}_j,\mathbf{C}|X=0)}}_{\text{term 1}}\underbrace{\frac{P(X=0|\underline{\mathbf{S}}_j,\mathbf{C})P(X=1|\mathbf{C})}{P(X=1|\underline{\mathbf{S}}_j,\mathbf{C})P(X=0|\mathbf{C})}}_{\text{term 2}} \end{aligned} \tag{2}$$

With this alternative representation of W, term 1 is a ratio that captures the dependency of the school characteristics $\underline{\mathbf{S}}_j$ and student attributes $\mathbf{C}$ within low-poverty neighborhoods ($X=0$). Specifically, the numerator is the product of their respective marginal probabilities among students residing in low-poverty neighborhoods, and the denominator is their joint probability among the same students.

As Arbour et al. (2021) show, this ratio can be estimated via permutation weighting. In our application, this consists of the following steps: (i) restrict the analytic sample to students residing in low-poverty neighborhoods ($X=0$); (ii) stack this subsample with a permuted version of itself, in which the school characteristics $\underline{\mathbf{S}}_j$ are reshuffled across rows and therefore independent of $\mathbf{C}$ in expectation; (iii) in the combined dataset, use a binary classification model to estimate the probability that a given observation belongs to the permuted sample ($Z=1$), conditional on $\mathbf{C}$ and $\underline{\mathbf{S}}_j$, which can be denoted as $\hat{P}(Z=1|\mathbf{C},\underline{\mathbf{S}}_j)$. If the model in step (iii) is correctly specified, then $\frac{\hat{P}(Z=1|\mathbf{C},\underline{\mathbf{S}}_j)}{1-\hat{P}(Z=1|\mathbf{C},\underline{\mathbf{S}}_j)}$ provides a consistent estimator for term 1 in Equation 2. To reduce chance dependence between $\mathbf{C}$ and $\underline{\mathbf{S}}_j$ in the permuted sample, we iterate steps (ii) and (iii) five times and use the average of $\frac{\hat{P}(Z=1|\mathbf{C},\underline{\mathbf{S}}_j)}{1-\hat{P}(Z=1|\mathbf{C},\underline{\mathbf{S}}_j)}$ over these iterations to estimate term 1.

The second term in Equation 2 is more straightforward to estimate. Specifically, term 2 in this expression can be estimated by first fitting two models predicting residence in a high- versus low-poverty neighborhood—one conditional on the baseline covariates $\mathbf{C}$ and one conditional on both the baseline covariates $\mathbf{C}$ and the school characteristics $\underline{\mathbf{S}}_j$. Next, we use these models to predict the corresponding probabilities—$\hat{P}(X =$

$0|\mathbf{C})$, $\hat{P}(X = 1|\mathbf{C})$, $\hat{P}(X = 0|\underline{\mathbf{S}}_j, \mathbf{C})$, and $\hat{P}(X = 1|\underline{\mathbf{S}}_j, \mathbf{C})$—and then substitute them for their population counterparts in Equation 2.

The resulting estimator $\hat{\zeta}_{1,\underline{\mathbf{S}}_j}$ has several desirable properties. First, it does not require modeling the distribution of $\underline{\mathbf{S}}_j$, which would be prohibitively difficult given its high dimensionality. Second, $\hat{\zeta}_{1,\underline{\mathbf{S}}_j}$ is doubly robust. In particular, under the identification assumptions outlined previously, it is consistent if either the outcome model for $\mathbb{E}\left[Y|X, \mathbf{C}, \underline{\mathbf{S}}_j\right]$ is consistently estimated or the weight W is consistently estimated, but not necessarily both. When the first condition holds, $\hat{\zeta}_{1,\underline{\mathbf{S}}_j}$ converges to a consistent regression-imputation estimator. When the second condition holds, it converges to a weighting estimator that is also consistent (Opacic et al., 2025).

Although the double robustness property of $\hat{\zeta}_{1,\underline{\mathbf{S}}_j}$ offers some protection against model misspecification, accurately estimating either the outcome model or the weight W can be very difficult in practice, particularly when the vector of school characteristics $\underline{\mathbf{S}}_j$ is high-dimensional. Fortunately, similar to our proposed estimator $\hat{\eta}_{1,\underline{\mathbf{S}}_j}$ for post-intervention disparities under conditional equalization, $\hat{\zeta}_{1,\underline{\mathbf{S}}_j}$ is also compatible with data-adaptive ML algorithms for fitting the requisite nuisance terms. This advantage essentially obviates the need to correctly specify models for the outcome or for group membership.

Figures H.1 and H.2 report our DML estimates of $\gamma_{\underline{\mathbf{S}}_j}$, the disparity in test scores that would remain after equalizing the school characteristics in $\underline{\mathbf{S}}_j$ across neighborhoods, regardless of students' baseline characteristics. Figure H.1 presents estimates for reading scores, while Figure H.2 displays the corresponding estimates for math. The results align closely with those presented in the main text, which focused on the disparity eliminated by equalizing school characteristics between students from high- and low-poverty neighborhoods who share similar covariates at baseline. This indicates that equalizing the distribution of school characteristics across neighborhoods—with or without attention to differences in students' prior achievement, family background, or other baseline factors—does not appear to appreciably reduce disparities in achievement test scores.

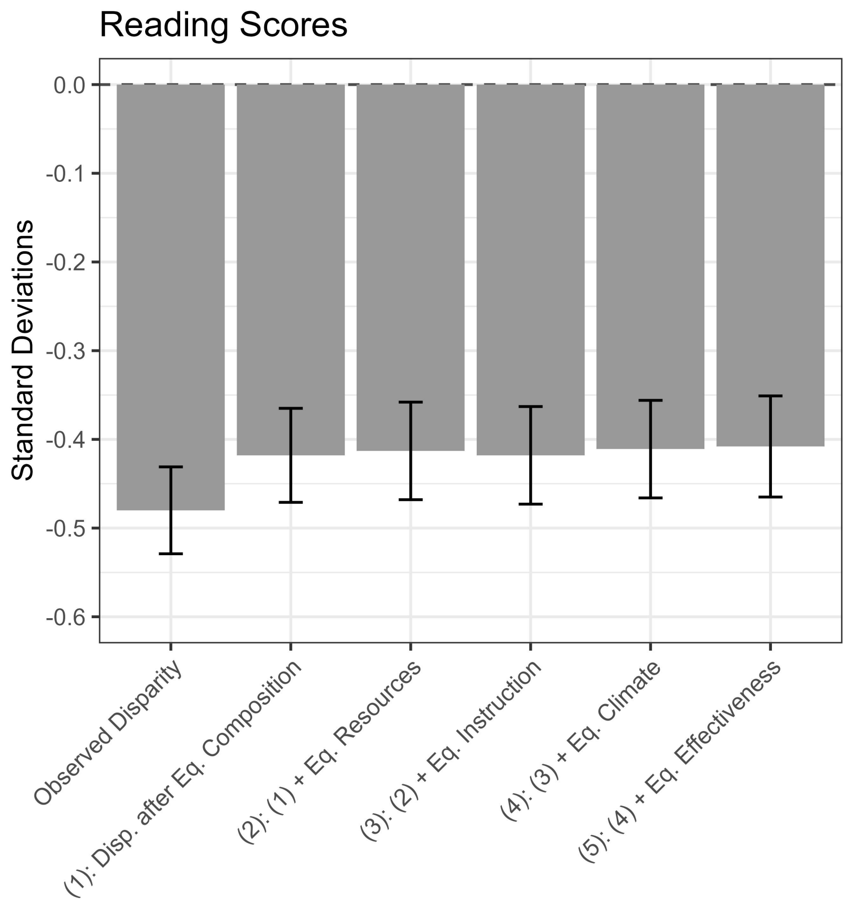

Figure H.1: Observed and Counterfactual Disparities in Reading Test Scores after a Hypothetical Intervention to Equalize the Distribution of School Characteristics Unconditionally.

Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

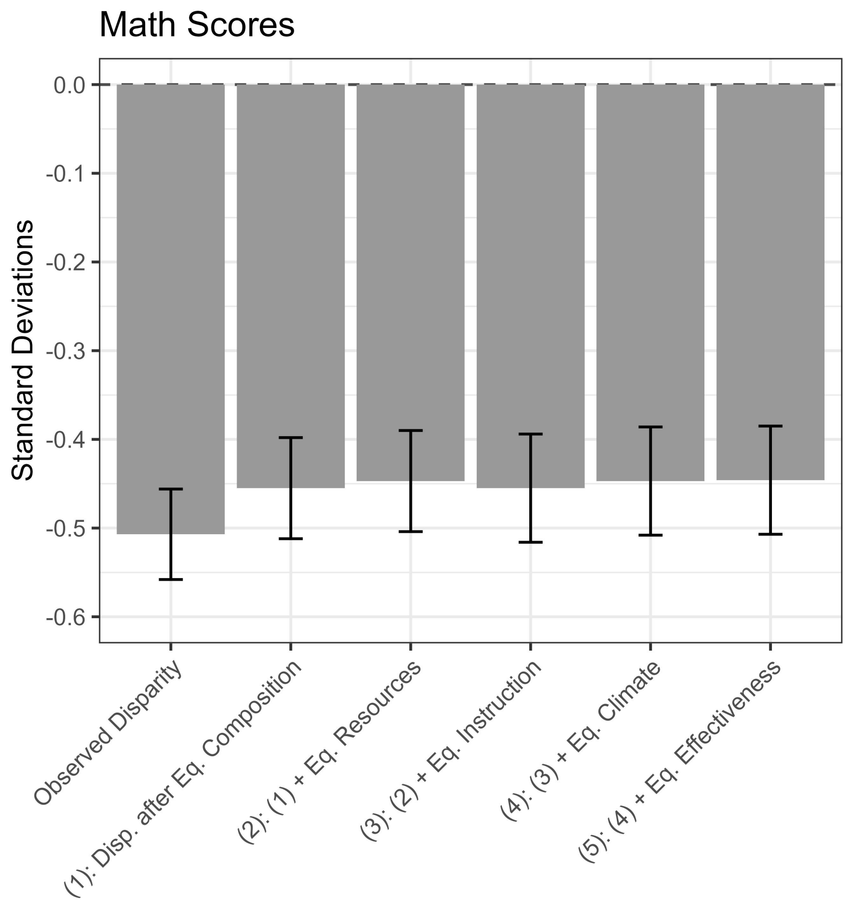

Figure H.2: Observed and Counterfactual Disparities in Math Test Scores after a Hypothetical Intervention to Equalize the Distribution of School Characteristics Unconditionally.

Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011-12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011-12.

I Subgroup Analyses

This appendix presents estimates of the disparity eliminated separately for specific subpopulations of students. Tables I.1 and I.2 report results for boys and girls, respectively. Tables I.3 through I.6 report estimates separately by race, focusing on those groups for which we have a large enough sample to produce reasonably precise estimates. Tables I.7 and I.8 present results for students with and without a college-educated parent. And finally, Tables I.9 to I.11 provide estimates for students living in different census regions.

Across all these subgroups, the results are broadly consistent with those presented in the main text for the total population of students. Point estimates of the disparity eliminated are somewhat larger for students living in the South (Table I.10) than those for other regions or subpopulations, but they are subject to considerable sampling variability and are not statistically significant at conventional thresholds. Overall, these findings suggest that equalizing access to schools of differing quality across high- and low-poverty neighborhoods would yield only modest reductions in the achievement gap, regardless of students' demographic background or geographic location.

Table I.1: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among Boys.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.485	[-0.548, -0.422]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.038	[-0.103, 0.027]	0.08
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.036	[-0.105, 0.033]	0.07
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.036	[-0.103, 0.031]	0.07
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.037	[-0.106, 0.032]	0.08
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.037	[-0.108, 0.034]	0.08
Math Test Scores				
Observed disparity	Δ	-0.540	[-0.607, -0.473]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.022	[-0.085, 0.041]	0.04
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.021	[-0.082, 0.040]	0.04
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.021	[-0.082, 0.040]	0.04
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.021	[-0.084, 0.042]	0.04
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.020	[-0.083, 0.043]	0.04

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table I.2: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among Girls.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.474	[-0.531, -0.417]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.015	[-0.070, 0.040]	0.03
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.014	[-0.081, 0.053]	0.03
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.011	[-0.072, 0.050]	0.02
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.011	[-0.072, 0.050]	0.02
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.012	[-0.073, 0.049]	0.03
Math Test Scores				
Observed disparity	Δ	-0.471	[-0.530, -0.412]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.012	[-0.069, 0.045]	0.03
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.014	[-0.075, 0.047]	0.03
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.012	[-0.071, 0.047]	0.03
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.010	[-0.073, 0.053]	0.02
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.013	[-0.076, 0.050]	0.03

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table I.3: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among White Students.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.293	[-0.381, -0.205]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.035	[-0.131, 0.061]	0.12
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.028	[-0.126, 0.070]	0.10
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.026	[-0.116, 0.064]	0.09
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.027	[-0.117, 0.063]	0.09
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.031	[-0.125, 0.063]	0.11
Math Test Scores				
Observed disparity	Δ	-0.284	[-0.370, -0.198]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.040	[-0.144, 0.064]	0.14
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.027	[-0.117, 0.063]	0.10
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.023	[-0.103, 0.057]	0.08
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.028	[-0.108, 0.052]	0.10
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.022	[-0.102, 0.058]	0.08

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table I.4: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among Non-White Students.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.410	[-0.465, -0.355]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.034	[-0.085, 0.017]	0.08
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.032	[-0.083, 0.019]	0.08
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.026	[-0.077, 0.025]	0.06
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.029	[-0.078, 0.020]	0.07
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.029	[-0.078, 0.020]	0.07
Math Test Scores				
Observed disparity	Δ	-0.433	[-0.492, -0.374]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.016	[-0.071, 0.039]	0.04
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.025	[-0.076, 0.026]	0.06
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.019	[-0.074, 0.036]	0.04
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.022	[-0.075, 0.031]	0.05
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.018	[-0.071, 0.035]	0.04

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table I.5: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among Black Students.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.282	[-0.376, -0.188]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.023	[-0.090, 0.044]	0.08
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.032	[-0.099, 0.035]	0.11
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.029	[-0.092, 0.034]	0.10
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.029	[-0.096, 0.038]	0.10
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.022	[-0.091, 0.047]	0.08
Math Test Scores				
Observed disparity	Δ	-0.334	[-0.428, -0.240]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.041	[-0.115, 0.033]	0.12
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.058	[-0.140, 0.024]	0.17
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.051	[-0.131, 0.029]	0.15
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.055	[-0.139, 0.029]	0.16
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.054	[-0.134, 0.026]	0.16

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table I.6: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among Hispanic Students.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.315	[-0.388, -0.242]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.038	[-0.105, 0.029]	0.12
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.032	[-0.108, 0.044]	0.10
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.027	[-0.103, 0.049]	0.09
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.030	[-0.106, 0.046]	0.10
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.025	[-0.103, 0.053]	0.08
Math Test Scores				
Observed disparity	Δ	-0.251	[-0.322, -0.180]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.005	[-0.079, 0.069]	0.02
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.015	[-0.084, 0.054]	0.06
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.009	[-0.078, 0.060]	0.04
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.010	[-0.081, 0.061]	0.04
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.006	[-0.080, 0.068]	0.02

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table I.7: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among Students without a College-educated Parent.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.245	[-0.308, -0.182]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.035	[-0.098, 0.028]	0.14
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.041	[-0.104, 0.022]	0.17
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.041	[-0.102, 0.020]	0.17
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.035	[-0.092, 0.022]	0.14
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.035	[-0.096, 0.026]	0.14
Math Test Scores				
Observed disparity	Δ	-0.281	[-0.344, -0.218]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.028	[-0.085, 0.029]	0.10
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.029	[-0.096, 0.038]	0.10
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.029	[-0.096, 0.038]	0.10
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.022	[-0.087, 0.043]	0.08
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.019	[-0.084, 0.046]	0.07

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table I.8: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among Students with a College-educated Parent.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.400	[-0.463, -0.337]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.024	[-0.093, 0.045]	0.06
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.025	[-0.092, 0.042]	0.06
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.018	[-0.087, 0.051]	0.04
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.022	[-0.087, 0.043]	0.05
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.014	[-0.085, 0.057]	0.03
Math Test Scores				
Observed disparity	Δ	-0.437	[-0.498, -0.376]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.015	[-0.078, 0.048]	0.03
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.015	[-0.082, 0.052]	0.03
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.011	[-0.076, 0.054]	0.03
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.017	[-0.084, 0.050]	0.04
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.015	[-0.084, 0.054]	0.03

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table I.9: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among Students in the Midwest and Northeast Census Regions.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.539	[-0.625, -0.453]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.021	[-0.103, 0.061]	0.04
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.005	[-0.085, 0.075]	0.01
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.006	[-0.084, 0.072]	0.01
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.009	[-0.089, 0.071]	0.02
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.016	[-0.094, 0.062]	0.03
Math Test Scores				
Observed disparity	Δ	-0.582	[-0.666, -0.498]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	0.003	[-0.103, 0.109]	-0.01
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.003	[-0.111, 0.105]	0.01
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	0.000	[-0.104, 0.104]	0.00
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.001	[-0.095, 0.093]	0.00
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.003	[-0.095, 0.089]	0.01

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table I.10: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among Students in the South Census Region.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.431	[-0.505, -0.357]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.052	[-0.125, 0.021]	0.12
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.066	[-0.140, 0.008]	0.15
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.057	[-0.128, 0.014]	0.13
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.057	[-0.128, 0.014]	0.13
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.058	[-0.134, 0.018]	0.13
Math Test Scores				
Observed disparity	Δ	-0.456	[-0.536, -0.376]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.052	[-0.130, 0.026]	0.11
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.063	[-0.136, 0.010]	0.14
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.051	[-0.124, 0.022]	0.11
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.055	[-0.126, 0.016]	0.12
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.056	[-0.129, 0.017]	0.12

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

Table I.11: De-biased Machine Learning Estimates of the Disparity Eliminated between High- and Low-poverty Neighborhoods among Students in the West Census Region.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading Test Scores				
Observed disparity	Δ	-0.451	[-0.553, -0.349]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.023	[-0.125, 0.079]	0.05
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.017	[-0.125, 0.091]	0.04
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.006	[-0.108, 0.096]	0.01
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.005	[-0.111, 0.101]	0.01
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.002	[-0.108, 0.104]	0.00
Math Test Scores				
Observed disparity	Δ	-0.450	[-0.546, -0.354]	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	0.014	[-0.078, 0.106]	-0.03
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	0.012	[-0.080, 0.104]	-0.03
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	0.014	[-0.076, 0.104]	-0.03
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	0.019	[-0.075, 0.113]	-0.04
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	0.020	[-0.072, 0.112]	-0.04

Note: All estimates are reported in standard deviation units. Est. = Estimate; CI = Confidence Interval; Prop. Elim. = Proportion Eliminated, which is calculated as the disparity eliminated divided by the observed disparity. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; U.S. Department of Education, National Center for Education Statistics, Private School Universe Survey (PSS), 2011–12.

References

Arbour, D., Dimmery, D., and Sondhi, A. (2021). Permutation weighting. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139, pages 331–341.

Opacic, A., Wei, L., and Zhou, X. (2025). Disparity analysis: A tale of two approaches. *Journal of the Royal Statistical Society Series A: Statistics in Society*. E-pub ahead of print.

Wodtke, G. T. and Zhou, X. (Forthcoming). *Causal Mediation Analysis*. Cambridge University Press.

Zhou, X. (2022). Semiparametric estimation for causal mediation analysis with multiple causally ordered mediators. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 84(3):794–821.