sociological science

Poor Neighborhoods, Bad Schools? A High-Dimensional Model of Place-Based Disparities in Academic Achievement

Geoffrey T. Wodtke,[a] Kailey White[a], Xiang Zhou[b]

a) University of Chicago; b) Harvard University

Abstract: Persistent disparities in academic achievement between students from high- and low-poverty neighborhoods are widely attributed to differences in school quality. Using nationally representative data from more than 18,000 students and nearly 1,000 elementary schools, we examine how the schools serving students from different neighborhoods vary across more than 160 characteristics, including detailed measures of their composition, resources, instruction, climate, and effectiveness. Our findings document significant differences in demographic composition between schools serving high- and low-poverty neighborhoods but comparatively little variation in other dimensions of the school environment. With novel machine learning methods tailored for high-dimensional data, we estimate that equalizing all these different factors would reduce the achievement gap by less than 10 percent, primarily through changes in school composition. These results suggest that the main drivers of place-based disparities in achievement lie outside of elementary schools, underscoring the need to address broader structural inequalities as part of any effort to reduce achievement gaps.

Keywords: poverty; neighborhoods; schools; achievement; disparities; gaps

Reproducibility Package: Code and instructions for accessing the data necessary to reproduce the results presented in this article are available at https://doi.org/10.5281/zenodo.17634676.

Citation: 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.

Received: September 8, 2025

Accepted: December 12, 2025

Published: February 6, 2026

Editor(s): Arnout van de Rijt, Jeremy Freese

DOI: 10.15195/v13.a6

By the end of third grade, students from high-poverty neighborhoods exhibit substantial disadvantages in academic achievement compared to their peers from low-poverty neighborhoods. For example, students from neighborhoods with a poverty rate exceeding 20 percent score, on average, about half a standard deviation (SD) lower on standardized tests of reading and math skills than those from more affluent neighborhoods.[1] Although standardized tests have well-documented limitations (Jackson 2018; Jennings et al. 2015), the disparities they reveal are concerning, especially given that test scores reliably predict many other outcomes, such as educational attainment, earnings, and health later in adulthood (Auld and Sidhu 2005; Heckman, Stixrud, and Urzua 2006).

Although many factors contribute to the achievement gap between students from high- and low-poverty neighborhoods–including inequalities in family background (Duncan and Murnane 2011), exposure to violent crime (Sharkey et al. 2014), and environmental health hazards (Wodtke, Ramaj, and Schachner 2022), among a variety of other possibilities–differences in school quality are widely regarded as an important driver of this disparity (Ainsworth 2002; Arum 2000; Sanbonmatsu et al. 2006). According to institutional resource theory (Jencks and Mayer 1990; Johnson

2012), along with media accounts and public narratives (e.g., Kozol 2012; Semuels 2016), students from high-poverty neighborhoods often attend schools with fewer resources, less effective teaching, and less supportive learning environments. In other words, children from "poor neighborhoods" are widely believed to attend "bad schools," which harm student learning and contribute to the persistent gaps in achievement observed across neighborhoods with varying levels of poverty.

Despite widespread assertions that students from high-poverty neighborhoods attend lower-quality schools, rigorous evidence evaluating this claim is surprisingly scarce (Owens and Candipan 2019). Most prior studies have focused on a narrow set of easily quantifiable school characteristics that are recorded in national administrative data, such as free lunch participation rates, student-teacher ratios, and per-pupil expenditures. Moreover, findings from this body of research are mixed: schools serving poor neighborhoods appear disadvantaged on some measures but do not consistently under-perform across all, or even most, of these metrics (Owens and Candipan 2019; Wodtke and Parbst 2017). Similarly, another set of recent studies drawing on indicators of school value added have also failed to reveal large differences in the academic skills imparted by schools serving high- versus low-poverty communities (Downey, Quinn, and Alcaraz 2019; Hanselman and Fiel 2017; Wodtke et al. 2023). Taken together, these findings suggest that schools may have a more limited influence on socioeconomic gaps in academic achievement, rather than acting as a primary source of these disparities (Condron, Downey, and Kuhfeld 2021; Downey 2020).

Nevertheless, prior studies are constrained by how they conceptualize and measure school quality. School quality is a broad and multifaceted concept that is exceedingly difficult to define and operationalize (Ladd and Loeb 2013; Raudenbush and Eschmann 2015). Studies that rely on convenient indicators from administrative data (e.g., Wodtke and Parbst 2017) or noisy estimates of school value added (e.g., Wodtke et al. 2023) capture only fragmented aspects of what might make a school "good" or "bad."

In reality, school quality encompasses a myriad of different factors–everything from physical infrastructure to classroom management strategies. These diverse elements of the school environment may vary across neighborhoods and affect students in complex, intertwined ways. Furthermore, what constitutes a quality school may depend on the characteristics of the students it serves, as instructional approaches that benefit one student may not work for another. Even though school quality is multidimensional and contingent on student needs, it is typically reduced to a simplistic, low-dimensional set of indicators that are naively assumed to affect all students uniformly. These limitations risk oversimplifying–and potentially misrepresenting–the role schools play in shaping academic disparities.

In this study, we examine how the elementary schools serving students from high- and low-poverty neighborhoods differ, using a multidimensional approach to conceptualizing and measuring the school environment. We then investigate the extent to which these differences are associated with achievement gaps linked to concentrated poverty. Finally, we estimate the impact of hypothetical interventions designed to equalize school environments across neighborhoods, evaluating their potential to reduce achievement gaps.

Analyzing the link between schools, neighborhood poverty, and academic achievement presents a number of methodological challenges. The first challenge involves the difficulty of measuring school quality, a multidimensional and contingent concept that cannot be easily captured by a few basic indicators. The second challenge lies in accurately modeling the association between differences in school quality and gaps in achievement. Given the many dimensions of the school environment, this relationship is likely to be complex, involving nonlinearity, interactions among different school characteristics, and intersectionality in how these factors influence students. Identifying the effects of hypothetical interventions presents a third major challenge. It requires strong assumptions about the absence of unobserved confounding, which are difficult to satisfy by design and cannot be empirically verified.

We address these challenges by leveraging data from the Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K; Tourangeau et al. 2015), along with novel methods tailored for high-dimensional data. Using the ECLS-K, we measure more than 160 distinct characteristics of the school environment across five broad dimensions: composition, resources, instruction, climate, and effectiveness. School *composition* refers to the demographic and socioeconomic makeup of the student body and staff, whereas school *resources* encompass the funding, staffing, and facilities that support student learning. A school's *instruction* includes the teaching methods, curriculum content, and assessment practices employed by teachers. School *climate* captures the social, emotional, and cultural atmosphere of the school, measured through factors such as classroom disorder, administrative leadership, and teacher morale. And finally, school *effectiveness* gauges the impact of the school on its students' learning, assessed using conventional estimates of value added.

To analyze how differences in these dimensions of the school environment are linked to achievement gaps, we use a novel set of multiply robust estimating equations together with supervised machine learning. This approach employs flexible, data-adaptive algorithms that inductively learn to approximate complex relationships in high-dimensional data. With these methods, we perform both a descriptive and causal decomposition of achievement disparities between students from high- and low-poverty neighborhoods. The descriptive decomposition identifies the specific dimensions of the school environment that vary significantly across neighborhoods and that are most strongly associated with student achievement, whereas the causal decomposition focuses on the impact of hypothetical interventions that equalize characteristics of the schools serving different communities. To address the challenge of causal identification in this analysis, we adjust for a comprehensive set of potential confounders, including baseline measures of student achievement, and we conduct a sensitivity analysis to evaluate the robustness of our inferences against bias from unobserved selection.

Our analysis reveals large differences in the demographic composition of schools attended by students from high- and low-poverty neighborhoods. However, we find relatively few, and mostly small, differences in nearly all other dimensions of the school environment. Furthermore, our estimates suggest that equalizing all these dimensions of schools would, collectively, reduce the achievement gap by less than 10 percent, with most of this reduction driven by changes in school composition. These results are highly robust to bias from unobserved confounding, which suggests that differences in school quality are not a primary driver of disparities in achievement between students from high- and low-poverty neighborhoods.

This study makes three contributions to research on neighborhood, school, and educational inequalities. First, methodologically, it introduces a new approach to operationalizing school quality and decomposing achievement disparities using high-dimensional data. Second, it empirically challenges the widely held view that disparities in achievement between students from high- and low-poverty neighborhoods are driven by differences in access to quality schools. Third, theoretically, these findings prompt a reconsideration of the institutional resource perspective. Contrary to this theoretical model, our results suggest that elementary schools likely play a more benign role in the etiology of academic disparities across more versus less disadvantaged neighborhoods.

What Is School Quality?

School quality is a complex construct. It is shaped by many different characteristics of the school environment, each of which may influence individual students in diverse ways. However, studies of achievement disparities often simplify this complexity by focusing on a few isolated features of schools and assuming these factors contribute uniformly to student learning. For example, although Wodtke et al. (2023) recognize that school quality is multifaceted and adopt a more comprehensive measurement strategy than many earlier studies, they still rely on a limited set of basic indicators and constrain their influence on achievement to be invariant across different types of students. This persistent tension between theoretical conceptions of school quality as multidimensional and contingent, and empirical approaches that reduce this complexity to facilitate analysis, may distort our understanding of how access to quality schools varies across neighborhoods and contributes to disparities in achievement.

We define a school's quality as encompassing both the investment and consumption value of the education it provides to students, each broadly construed (Ladd and Loeb 2013). The investment value of schooling refers to its tangible benefits, such as improved skills, greater knowledge, better health, higher earnings, and stronger social networks. In contrast, the consumption value of a school reflects the immediate, intrinsic gratification that students gain from learning and from their social interactions within the school environment. However, because investment benefits may take years to materialize and the consumption value of education is inherently difficult to quantify, directly measuring school quality is essentially impossible. As a result, researchers must always rely on measurable proxies for this construct (Wodtke et al. 2023).

Proxies for school quality can be broadly categorized into school inputs, school outputs, and the internal structure, culture, and practices that transform inputs into outputs. School inputs encompass both the characteristics of the student body and the resources available to support student learning. Common measures of school inputs include the racial and socioeconomic composition of students (Lauen and

Gaddis 2013; Willms, 2010), as well as different types of educational resources, such as financial expenditures, student-teacher ratios, and staff qualifications (Clotfelter, Ladd, and Vigdor 2010; Jackson, Johnson, and Persico 2016). The demographic composition of a school, along with the availability of resources, is widely thought to influence the quality of the education it provides to students (Arum 2000; Wodtke and Parbst 2017).

Another approach to measuring school quality focuses on outputs–typically, a limited set of short-term outputs that are relatively easy to quantify. For example, school quality is often assessed using value-added scores, which capture how effectively a school improves the performance of their students on standardized tests (Downey et al. 2019; Wodtke et al. 2023). Value-added scores are intended to isolate a school's unique contribution to student learning by adjusting for other confounding factors that operate outside the school environment. Although these scores capture certain aspects of student growth and predict some long-term benefits of education (Chetty, Friedman, and Rockoff 2014a,b), they are not without limitations. In particular, they do not accurately account for the full range of school outputs that students, families, or society may value. In addition, they often exhibit low reliability and can fluctuate from one school year to the next (Emslander et al. 2022; von Hippel 2009).

School quality can also be evaluated by examining a school's "technology." Within the framework of an educational production function, the technology of a school refers to its internal structure, culture, and practices that transform inputs into outputs (Ladd and Loeb 2013). This technology includes many different components, such as curricular content, instructional techniques, and the academic climate, which collectively shape the social and learning environment of a school, influencing its value to students as both an investment vehicle and a source of immediate enrichment (Crosnoe et al. 2010; Raudenbush 2008; Wang and Degol 2016). Despite extensive research on the many different components of a school's technology, consensus on which specific features are optimal remains somewhat elusive (Ladd and Loeb 2013). As but one example, the debate over the effectiveness of phonics, whole language, or balanced approaches to reading instruction continues on, even after decades of research in this area (National Reading Panel 2000; Pressley, Allington, and Pressley 2023).

Although school quality can be broadly linked to inputs, technology, and outputs, each of these categories is itself complex and multidimensional, with components that can be disaggregated even further. Consider, for example, the category of school inputs, which encompasses the characteristics of the students enrolled and the resources available to educate them. Each of these subcategories subsumes many more specific features of the school environment. The resources available to a school include its funding, staffing levels, teacher qualifications and skills, physical infrastructure, and instructional equipment. And these particular resources can be unpacked further still–into funding from different sources, staffing levels across different subject areas, and the suitability of specific equipment on campus. Similarly, the technology of a school includes its curriculum design and instructional practices, which cover everything from the amount of class time allocated to different topics to the teaching methods employed in the classroom. A school's technology also

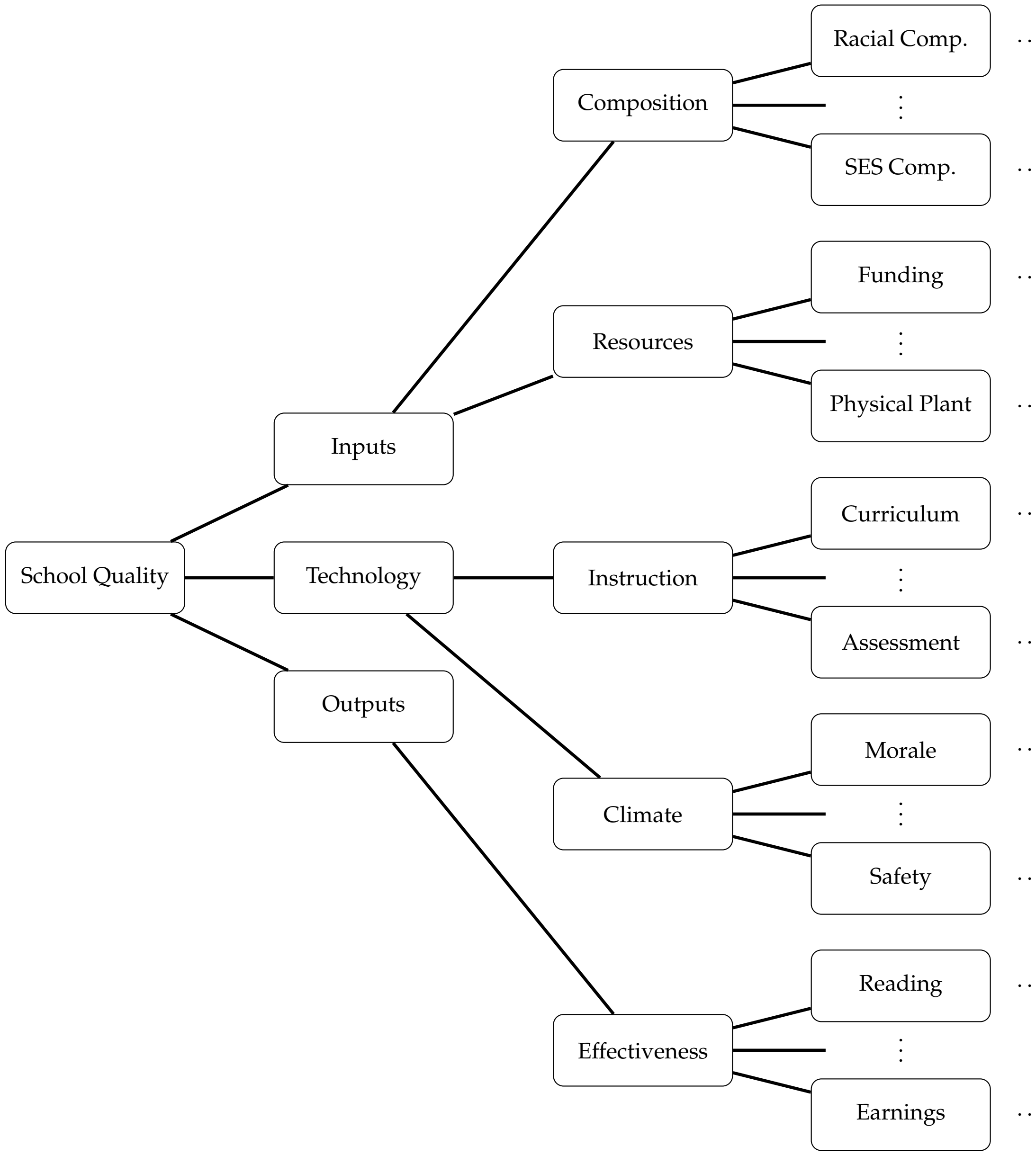

Figure 1: School quality as a multidimensional taxonomy.

encompasses its social and academic climate, which is shaped by teacher morale, classroom disorder, and parental engagement, among a variety of other factors.

In this way, the school environment can be viewed as a nested taxonomy of inputs, outputs, and the technology that connects them, with each element further divided into its component parts. This taxonomy is illustrated in Figure 1, which presents a stylized diagram depicting how school quality can be broken down into different elements at varying levels of specificity.

Adding to this multidimensional complexity, the quality of a school is also contingent on the unique characteristics and dispositions of individual students.

Schools are not static environments experienced uniformly. Rather, the benefits that students derive from specific inputs, practices, or climatic traits are highly individualized and shaped by their particular interactions with different elements of the school environment. What benefits one student may be less effective–or even counterproductive–for another, depending on their learning styles, personality, prior experiences, or baseline abilities (National Research Council 2000; Tomlinson 2014). Whether a particular school is "good" or "bad," then, may reflect not only its resources, instruction, climate, and so on but also the extent to which these elements align with or can adapt to student idiosyncrasies. In other words, a school's quality depends on how its inputs and technology intersect with the individual needs of students.

Moreover, the myriad dimensions of a school often interact in complex ways to shape the quality of the education it provides (Bryk et al. 2010; Hassrick, Rosen, and Raudenbush 2017). For example, students with an aptitude for math may only reach their full potential if they attend a school that provides a well-equipped classroom, dedicates sufficient time to math instruction, keeps the size their math classes manageable, employs knowledgeable and enthusiastic math teachers, and cultivates a social and academic climate that supports and values interest in math. The absence of any one of these factors could undermine the quality of education for such students, as the value of particular resources, instructional practices, and the school's climate depends on how they interact with and reinforce one another. Different elements of the school environment work in concert, rather than in isolation, to foster student learning.

To summarize, school quality is a complex, multidimensional, and contingent construct that cannot be adequately represented by a few convenient proxies. A comprehensive, multifaceted approach to measurement is therefore essential, as is an analytic strategy that accommodates interaction and individual contingency.

Neighborhoods, Schools, and Achievement Disparities

How do the schools serving students from high- and low-poverty neighborhoods differ, and are these differences linked to disparities in achievement? *Institutional resource theory* suggests that children from poor neighborhoods perform worse academically because they attend lower-quality schools (Arum 2000; Jencks and Mayer 1990; Johnson 2012). In many U.S. districts, public school enrollment is based on geographic assignment rules that establish a direct link between neighborhood demographics and school composition. Although charter schools, magnet schools, and intra-district open enrollment policies provide some flexibility, travel constraints generally limit families to schools within reasonable commuting distance of their homes. As a result, children from high-poverty neighborhoods typically attend schools with a higher concentration of disadvantaged peers (Saporito and Sohoni 2007), even amid the expansion of school choice programs (Rich, Candipan, and Owens 2021).

Schools with a high concentration of disadvantaged students are thought to face many educational challenges, including limited access to resources, difficulties in delivering rigorous instruction, and obstacles to fostering a supportive climate (Boyd et al. 2005; Clotfelter et al. 2010; Kahlenberg 2001). Together, these challenges may constrain a school's capacity to impart academic skills and promote student development.

According to this perspective, schools serving low-income communities often suffer from significant resource constraints (Jencks and Mayer 1990; Johnson 2012). These schools may receive less funding due to the reliance of public school financing on local property taxes, which generate less revenue in economically disadvantaged areas. However, compensatory funds from state and federal programs now largely mitigate the financial disparities created by differences in local government support (Owens and Candipan 2019). Nevertheless, resource gaps may still persist within districts because families in high-poverty neighborhoods have less capacity to supplement government funding for schools through private contributions or in-kind support (Kahlenberg 2001). As a result, schools serving disadvantaged communities may struggle to attract and retain qualified, experienced teachers, and they may operate with larger class sizes, substandard facilities, or outdated instructional materials (Borman and Dowling 2008; Boyd et al. 2005; Owens and Candipan 2019).

Schools serving large numbers of low-income students may also encounter challenges that constrain the rigor of their curriculum and instructional practices. These schools often enroll students who are not as well prepared academically and who exhibit higher rates of behavioral problems. In response, instructional practices may be adapted to accommodate these challenges, leading to a slower pace of instruction, a less demanding curriculum, and a greater focus on disciplinary enforcement (Kahlenberg 2001; Lauen and Gaddis 2013; Willms 2010). Difficulties in attracting and retaining experienced staff, along with larger class sizes, might also limit the use of appropriate teaching methods or individualized instruction. These disparities in instruction reflect upstream differences in resources and school composition, and they may further compromise the quality of the education provided to students from high-poverty neighborhoods.

A high concentration of disadvantaged students can also influence a school's climate. Because schools in high-poverty neighborhoods may contend with higher rates of absenteeism, more frequent classroom disruptions, and more limited parental engagement, they might struggle to cultivate a social milieu that promotes academic excellence (Kahlenberg 2001; Willms 2010). In schools beset by the harms of poverty, academic expectations can decline, teacher morale may wane, and the school's focus may shift primarily to maintaining order. In addition, exposure to crime and violence in the surrounding community–or even within the school itself–can erode interpersonal trust and entrench an authoritarian disciplinary culture (Devine 1996; Nolan 2011). These challenges may engender a school climate that, rather than supporting student learning, can actually impede academic progress.

In summary, schools serving disadvantaged communities are thought to face significant constraints on their resources, instruction, and climate, all of which may limit their effectiveness at imparting skills and knowledge. These limitations erode a school's capacity to provide a high-quality education and hinder their students' achievement. Thus, institutional resource theory asserts that neighborhood poverty directly shapes the quality of schools that children attend, which, in turn, contributes to disparities in their achievement.

Conversely, an alternative perspective, which we refer to as *compensatory schooling theory*, suggests that schools may play a neutral–or even equalizing–role in shaping academic disparities (Condron et al. 2021; Downey 2020). According to this view, schools can mitigate some of the disadvantages associated with concentrated poverty by acting as buffers, rather than amplifiers, of inequality. Although neighborhood and school composition are closely intertwined, the resources, instruction, climate, and effectiveness of elementary schools serving high- and low-poverty neighborhoods may not differ all that much due to legal regulations and other organizational constraints that limit their variability. In addition, the complexities of parental decision making about where to send their children to school may weaken the relationship between demographic composition and other dimensions of the school environment.

Supplementary funding from state and federal sources, coupled with bureaucratic standardization of staffing and facility requirements, systematically reduce variation in per-pupil spending, class size, teacher qualifications, and physical infrastructure between schools in high- and low-poverty neighborhoods. Elementary school curricula are also fairly uniform, often focusing on foundational skills and instructional methods that align well with the learning needs of disadvantaged students (Downey 2020). As a result, schools serving low-income neighborhoods may not be any less effective at fostering student learning (Downey et al. 2019; Wodtke et al. 2023). Instead, schooling may compress or merely channel, rather than exacerbate, socioeconomic disparities in achievement–gaps that are already pronounced before children enter kindergarten and that tend to remain stable thereafter (von Hippel, Workman, and Downey 2018).

Elementary education is largely delivered by individual teachers in specific classrooms. Research indicates that variation in teacher effectiveness–measured by value added to student test scores–is much greater within schools than between them (Hanushek and Rivkin 2012). This suggests that the quality of instruction received by students is shaped mainly by which teacher they are assigned, not which school they attend. Moreover, differences in teacher value added between high- and low-income students appear minimal, sometimes amounting to less than one-hundredth of a SD (Hanselman 2018; Isenberg et al. 2022). The pronounced variation in teacher effectiveness within schools, rather than between them, coupled with the negligible differences across socioeconomic groups, aligns with the view that place-based inequality in students' school experiences may not contribute to disparities in achievement.

Variation in school quality across high- and low-poverty neighborhoods may also be reduced by the imperfect information available to parents. Although many parents leverage their financial resources in an effort to secure better educational opportunities for their children, their choices are often based on incomplete or inaccurate information about the schooling options available to them (Abdulkadiroglu et al. 2020; DeLuca and Rosenblatt 2010). Even high-income parents, despite their advantages, may struggle to identify schools that best match their children's specific needs and learning styles. Instead, they often use convenient but unreliable heuristics, such as informal reputations within parental networks or the demographic composition of the student body (Billingham and Hunt 2016; Thompson

2024), when making these choices. As a result, the link between a family's socioeconomic status, their neighborhood composition, and the quality of the schools their children attend may be weaker than is commonly assumed.

Even when accurate information about school quality is available and actionable, parents also make enrollment decisions partly based on non-academic factors, such as athletic programs, social relationships, or transportation convenience, which reflect their idiosyncratic preferences or broader developmental goals for their children. These diverse priorities introduce additional variability into how families sort themselves across social contexts in pursuit of educational resources and may further attenuate the link between neighborhood composition and school quality (Wodtke et al. 2023).

Given the complexities of parental decision making and the structural constraints on formal education, the compensatory schooling perspective suggests that the schools serving disadvantaged neighborhoods may perform better than anticipated by institutional resource theory. By providing critical support for low-income students, these schools may potentially mitigate some of the challenges they face outside the school environment. At minimum, differences between the schools attended by students from high- and low-poverty neighborhoods are unlikely to contribute to or deepen disparities in academic achievement, according to this perspective.

Despite these conflicting theoretical arguments, existing research on the relationship between schools and disparities in student achievement is limited, and its results are mixed. Some studies find that schools with a higher concentration of disadvantaged students negatively impact their academic performance, thereby contributing to achievement gaps associated with poverty and other indicators of socioeconomic status (Owens 2010; Rumberger and Palardy 2005; Willms 2010). In contrast, other research suggests that schools play only a modest role in shaping achievement disparities, particularly across high- and low-poverty neighborhoods, with factors outside the school seemingly driving most of the observed differences (Lauen and Gaddis 2013; Wodtke and Parbst 2017; Wodtke et al. 2023). Similarly, evaluations of a wide range of school reforms indicate that many have left achievement disparities between different groups of students essentially unchanged (Angrist and Lang 2004; Bifulco and Ladd 2006; Cullen, Jacob, and Levitt 2005), whereas other studies have identified interventions that appear to substantially narrow these disparities in certain situations (Dobbie and Fryer 2011; Hassrick et al. 2017; Jackson et al. 2016).

These mixed findings leave the question of how schools contribute to achievement gaps–and the extent to which school-based interventions can help to close them–largely unresolved. A key source of inconsistency in prior research may stem from limitations in how school quality has been operationalized. In this study, we examine differences in school quality between high- and low-poverty neighborhoods and assess whether these disparities contribute to gaps in achievement, while measuring the school environment and modeling its relationship with student outcomes as flexibly as current data and methods allow.

Methods

Data

To this end, we use data from the Early Childhood Longitudinal Study, Kindergarten Cohort of 2010–2011 (ECLS-K; Tourangeau et al. 2015). The ECLS-K is a nationally representative, longitudinal survey of American schools and their students that began in fall of 2010 when participants first entered kindergarten. All schools offering a kindergarten program, both public and private, were eligible for inclusion in the study, and within selected schools, kindergarten students were sampled with approximately equal probability. The study collected data at both the fall and spring of kindergarten, first grade, and second grade, as well as in the spring of third, fourth, and fifth grades. The analytic sample for this study consists of all $n \cong 18,170$ children attending $h \cong 950$ schools who were enrolled in the study at the start of kindergarten.[2]

The ECLS-K collected a wide array of data, covering students, their families, and the school environments they experienced. These data are particularly well-suited for a multidimensional analysis of school quality, as they include hundreds of measures related to the schools children attended, all derived from surveys completed by teachers and school administrators. Standardized assessments of reading and math achievement were also administered at each wave of the study, enabling detailed analyses of student learning over time. In addition, the ECLS-K provides restricted access to geographic and school identifiers, which we use to link the survey with external data on children's neighborhoods and schools.[3]

Specifically, we use census tract identifiers for each sample member's home address to link them with data on the demographic composition of their residential neighborhoods. These demographic indicators were drawn from the Neighborhood Change Database (NCDB), which integrates tract-level data from the 2010 U.S. Census and the 2006–2010 American Community Surveys (GeoLytics 2012). Using school identifiers from the ECLS-K, we also linked sample members to additional information on their schools from the U.S. National Center for Education Statistics. Public school students were matched to information from the Common Core of Data (CCD), whereas private school students were matched to data from the Private School Universe Survey (PSS). Although the ECLS-K itself already contains an abundance of information about sampled schools, the CCD and PSS provide additional data on district finances and school staffing.

Measures

Academic Achievement

Academic achievement is measured using item-response theory (IRT) scores from the ECLS-K assessments of reading and math abilities. These scores provide vertically scaled, equal-interval measures of achievement, enabling meaningful comparisons across time. The reading assessment evaluates a range of skills, including vocabulary knowledge and reading comprehension, whereas the math assessment covers number properties, arithmetic, and measurement as well as basic algebra and geometry. The IRT scores for both reading and math exhibit high reliability, strong construct validity, and minimal differential item functioning (Najarian et al. 2019).

Neighborhood Poverty

Neighborhood poverty is measured using linked data from the NCDB. Specifically, we measure neighborhood poverty as the proportion of families in a sample member's home census tract with incomes below the federal poverty threshold. Although poverty is a multidimensional construct that could be assessed using several different indicators of neighborhood composition, we focus on an income poverty rate because it is easy to interpret and strongly correlated with other dimensions of socioeconomic disadvantage. In part A of the online supplement, we include a conceptual replication of our key findings using a multidimensional index of neighborhood disadvantage, which incorporates not only the poverty rate but also other measures, such as the unemployment rate, the proportion of households headed by single parents, and racial composition.

School Quality

To measure school quality, we construct 168 indicators of different inputs, outputs, and features of a school's technology, using teacher and administrator surveys fielded as part of the ECLS-K, as well as matched data from the CCD and PSS. These measures span five key dimensions of the school environment: composition, resources, instruction, climate, and effectiveness.

Specifically, we include 11 measures of school composition, reflecting the racial and socioeconomic makeup of the student body and the proportion of students enrolled in gifted and talented, special education, and English as a second language (ESL) programs. We also include 40 measures of school resources, which capture funding and staffing levels, class sizes, staff qualifications, and the quality of a school's facilities. In addition, we incorporate 91 distinct features of a school's instructional practices, including measures of the time allocated to different subjects as well as the use of different teaching methods, classroom management strategies, and curriculum content, among a variety of other factors. School climate is assessed with 24 measures that variously capture classroom misbehavior, student absenteeism, communication with parents, staff morale, parental and community support, and administrative leadership. Finally, we include two indicators of school effectiveness, which are based on value-added scores for reading and math achievement in first grade.

Value-added scores estimate the average learning gains that a school provides to its students, above and beyond what would be expected based on their prior achievement and demographic characteristics. These estimates are calculated using linear mixed models with school-level random effects fit to data from the ECLS-K. The models predict students' reading and math test scores at the end of first grade while controlling for prior test scores at both the beginning and end of kindergarten, as well as for gender, race, and parental education. To account for variations in the timing of assessments across schools, the models also adjust for the number of months that elapsed between test administrations. With these models, a school's value-added score is given by the empirical Bayes estimate of its random effect on achievement at the end of first grade. By incorporating value-added scores as measures of school effectiveness, we aim to capture the influence of unobserved or imperfectly measured inputs, instructional practices, and climatic characteristics on several key outputs–specifically, reading and math achievement during the early elementary years.

A complete list of all the school characteristics used in this analysis is provided in part B of the online supplement. It includes a description of each measure, a key with the abbreviated labels used to identify them throughout the results section, and basic summary statistics describing their distribution in our sample.

Before including all 168 measures of the school environment as separate variables, we explored whether these data could be reduced to a smaller set of underlying factors or discrete classes using a range of dimension reduction techniques. However, our analyses demonstrated that these measures cannot be effectively condensed, indicating that each variable captures a distinct aspect of the school environment. Results from these analyses are provided in part C of the online supplement.

Baseline Controls

In our causal analyses, we address the potential for confounding by adjusting for a wide range of child and family characteristics, all measured at the fall or spring of kindergarten. Specifically, we control for a sample member's gender, race, and birth weight. In addition, we adjust for a sample member's test scores in reading and math at baseline, as well as several other developmental indicators measured during the fall of kindergarten. These include teacher-reported measures of a child's propensity for engaging in internalizing and externalizing behaviors, interviewer-reported measures of the child's attention, motivation, and cooperation, and a parent-reported measure of the child's overall health.

The family controls included in this analysis are parental age, marital status, employment status, education, and occupation, along with family income, household size, the primary language spoken at home, and indicators denoting receipt of several forms of government assistance. We also account for parental engagement, as measured by the frequency with which parents read books or practice numbers with their child, as well as parental expectations regarding their child's educational attainment.

Finally, we include controls for several geographic characteristics of a sample member's residence at baseline. Specifically, we adjust for urban versus rural residence, categorized as living in a large, medium, or small city, a suburb, or a rural area. We also control for geographic region using an indicator for whether a sample member lived in the "northeast," "midwest," "south," or "west." Additional details on how all control variables were measured, along with a full set of summary statistics, are provided in part D of the online supplement.

A Graphical Model

Figure 2 presents a graphical model illustrating our hypothesized relationships between neighborhood poverty, school quality, and academic achievement. In this

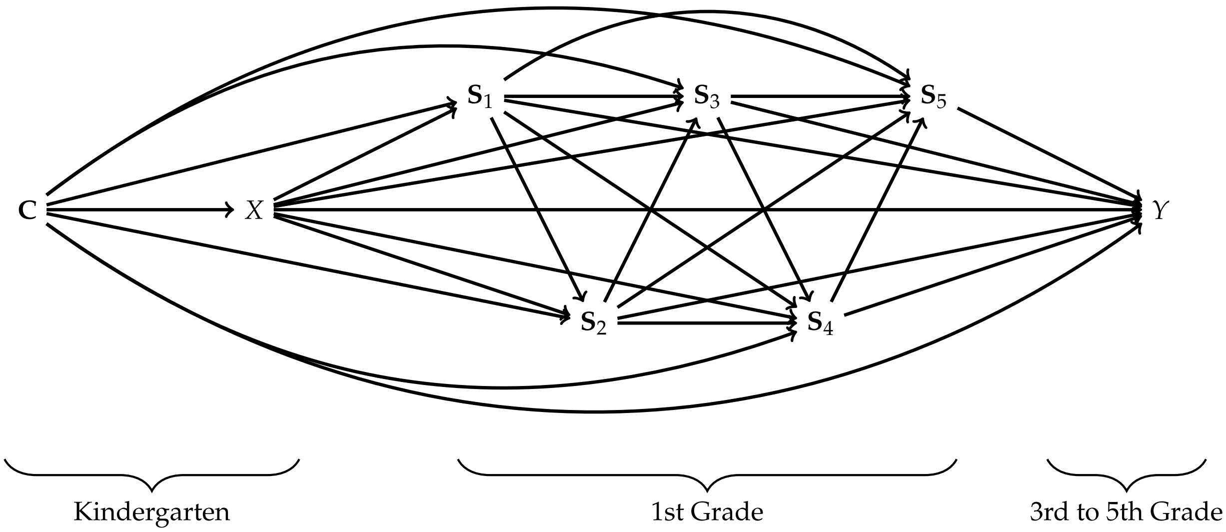

Figure 2: A graphical model of the relationships between neighborhood poverty, school quality, and academic achievement. *Note:* **C** denotes the vector of baseline confounders, X denotes neighborhood poverty, $\mathbf{S}_1$ denotes the set of variables measuring school composition, $\mathbf{S}_2$ denotes the set of variables measuring school resources, $\mathbf{S}_3$ denotes the set of indicators capturing a school's instruction, $\mathbf{S}_4$ denotes the set of measures reflecting a school's climate, $\mathbf{S}_5$ denotes the value-added scores measuring school effectiveness, and Y denotes a student's achievement test scores. The baseline confounders and neighborhood poverty are measured during the kindergarten waves of data collection. All school characteristics are based on the fall and spring waves of data collection during first grade. The outcome, student achievement, is measured later in the spring of third, fourth, or fifth grade.

model, X represents whether a student resides in a high-poverty neighborhood, defined as a census tract with a poverty rate greater than or equal to 20 percent. In addition, **C** denotes the set of baseline control variables, and Y represents a student's achievement test scores. The different dimensions of school quality are represented by the following sets of variables: $\mathbf{S}_1 = \{S_{1,1}, S_{1,2}, ..., S_{1,11}\}$, which denotes our 11 measures of school composition; $\mathbf{S}_2 = \{S_{2,1}, S_{2,2}, ..., S_{2,40}\}$, which denotes the 40 measures of school resources; $\mathbf{S}_3 = \{S_{3,1}, S_{3,2}, ..., S_{3,91}\}$, which captures the 91 indicators of instructional practices; $\mathbf{S}_4 = \{S_{4,1}, S_{4,2}, ..., S_{4,24}\}$, which denotes the 24 measures reflecting a school's climate; and $\mathbf{S}_5 = \{S_{5,1}, S_{5,2}\}$, which represents our two value-added measures of school effectiveness.

The temporal ordering of key variables is depicted by the brackets at the bottom of Figure 2. The baseline controls (**C**) and neighborhood poverty (X) are based on data from the kindergarten waves of the ECLS-K. All measures of school quality, represented by $\mathbf{S}_1$ through $\mathbf{S}_5$, are derived from the fall and spring waves of data collection during first grade. The outcome variable, student achievement (Y), is assessed later, during the spring of third, fourth, or fifth grade.

The model depicted in Figure 2 posits that neighborhood poverty influences every dimension of school quality, each of which is linked to student achievement in turn. It also accounts for inter-dependencies among the different dimensions of school quality. Inputs, such as school composition and resources, shape a school's technology, like its instructional practices and climate. These interconnected inputs and elements of a school's technology then influence its overall effectiveness. In addition, baseline characteristics of students and their families may influence all downstream factors, including their residential location, the quality of the school they attend, and their academic achievement.

In the results section, we focus on analyses of student achievement at the spring of third grade. Parallel analyses using test scores from the fourth and fifth grade assessments, presented in part E of the online supplement, yield similar results. Our analysis prioritizes the early years of elementary school because these grades offer the most reliable and comprehensive data on school quality. Furthermore, early elementary school is an important period for developing foundational skills that are critical for future academic success.

Analyses

Our analyses proceed in three stages. First, we describe differences in the characteristics of schools attended by students from high- and low-poverty neighborhoods. This analysis provides a broad overview of how school environments vary across these residential contexts.

Second, we conduct a descriptive decomposition of the achievement gap between students from high- and low-poverty neighborhoods. This decomposition examines whether and how disparities in achievement are associated with differences in the schools attended by children from different neighborhoods.

Finally, in the third stage, we perform a causal analysis of the achievement gap. This analysis estimates how disparities in test scores between students from high- and low-poverty neighborhoods would change under hypothetical interventions designed to equalize different dimensions of school quality.

School Differences by Neighborhood Poverty

In the first stage of our analysis, we begin by estimating standardized mean differences for each characteristic of the schools attended by students from high- and low-poverty neighborhoods. These differences can be formally expressed as follows:

$$D_{S_{j,k}} = \frac{\mathbb{E}\left[S_{j,k}|X=1\right] - \mathbb{E}\left[S_{j,k}|X=0\right]}{\sqrt{\mathrm{Var}\left[S_{j,k}\right]}}, \tag{1}$$

where $S_{j,k}$ denotes the kth measure in the jth vector of school characteristics, $X = 1$ indicates residence in a neighborhood with a poverty rate greater than or equal to 20 percent, and $X = 0$ indicates residence in a neighborhood with a poverty rate less than 20 percent. This metric quantifies the strength of the association between each characteristic of the school environment and neighborhood poverty, adjusting for differences in scale.

Because many of our school characteristics are binary or ordinal, the standardized mean difference may not always be the most appropriate measure of association. To address this, we also examined a number of other measures, including several tailored for categorical data, and all of them produced substantively similar results. Thus, for simplicity, we report only the standardized mean difference.

A Descriptive Decomposition of Place-Based Achievement Gaps

In the second stage of our analysis, we investigate the extent to which differences in the schools attended by students from high- and low-poverty neighborhoods are associated with the observed gap in achievement between these groups. Specifically, we conduct a descriptive decomposition of the achievement gap by partitioning it into a series of mean differences in test scores that are sequentially adjusted for different sets of school characteristics.

The observed gap in achievement is defined as follows:

$$\Delta = \underbrace{\mathbb{E}\left[Y|X=1\right]}_{\mu_1} - \underbrace{\mathbb{E}\left[Y|X=0\right]}_{\mu_0}, \tag{2}$$

where Y represents a student's observed test score and X denotes neighborhood poverty. In this expression, $\mu_1 = \mathbb{E}\left[Y|X=1\right]$ represents the mean test score for students in high-poverty neighborhoods, $\mu_0 = \mathbb{E}\left[Y|X=0\right]$ denotes the corresponding mean for those in low-poverty neighborhoods, and Δ captures the difference between them.

The achievement gap between students from high- and low-poverty neighborhoods can be partitioned into components that reflect its association with different sets of school characteristics (Opacic, Wei, and Zhou 2025). Formally, we decompose the gap as follows:

$$\begin{aligned} \Delta &= \mathbb{E}\left[Y|X=1\right] - \mathbb{E}\left[Y|X=0\right], \\ &= \sum_{j=1}^{J} \underbrace{\left(\mu_{1,\underline{\mathbf{S}}_{j-1}} - \mu_{1,\underline{\mathbf{S}}_j}\right)}_{\Delta_{\mathbf{S}_j}} + \underbrace{\left(\mu_{1,\underline{\mathbf{S}}_J} - \mu_0\right)}_{\Delta_\perp}, \end{aligned} \tag{3}$$

where $\mu_{1,\underline{\mathbf{S}}_j} = \mathbb{E}\left[\mathbb{E}\left[Y|X=1,\underline{\mathbf{S}}_j\right]|X=0\right]$, $\underline{\mathbf{S}}_j = \{\mathbf{S}_1, ..., \mathbf{S}_j\}$ represents all school characteristics up through dimension j, and $\underline{\mathbf{S}}_0 = \emptyset$ is an empty set, such that $\mu_{1,\underline{\mathbf{S}}_0} = \mu_1$. In substantive terms, $\mu_{1,\underline{\mathbf{S}}_j}$ captures the mean test score among students in high-poverty neighborhoods, adjusted for differences in the school characteristics included in $\underline{\mathbf{S}}_j$. For example, $\mu_{1,\underline{\mathbf{S}}_2}$ represents the mean test score among students from high-poverty neighborhoods after accounting for differences in the composition ($\mathbf{S}_1$) and resources ($\mathbf{S}_2$) of the schools they attend, compared with the schools attended by their counterparts in low-poverty neighborhoods.

Equation 3 represents a nonparametric analog of the Kitagawa-Oaxaca-Blinder (KOB) decomposition. In this formulation, $\Delta_{\mathbf{S}_j}$ quantifies the portion of the achievement gap that can be "statistically explained" by differences in the school characteristics included in $\mathbf{S}_j$, net of differences in all preceding dimensions of the school environment. For example, $\Delta_{\mathbf{S}_1}$ reflects the portion of the gap explained by differences in school composition across high- and low-poverty neighborhoods, whereas $\Delta_{\mathbf{S}_2}$ represents the contribution of differences in school resources, net of composition. The second term in the decomposition, $\Delta_{\perp}$, captures the residual portion of the achievement gap that cannot be explained by any of the school characteristics we measure. Conversely, the sum of $\Delta_{\mathbf{S}_j}$ across each dimension of school quality, from $j = 1$ to 5, represents the portion of the gap explained by the combined influence of all school characteristics. This version of the KOB decomposition is nonparametric in that it does not impose any restrictions, such as linearity or additivity, on the relationships between school characteristics, neighborhood poverty, and student test scores.

To implement this decomposition, we need to estimate μ_1, μ_0, and $\mu_{1,\underline{\mathbf{S}}_j}$ for each set of school characteristics j. The observed means, μ_1 and μ_0, are estimated using their sample analogs. To estimate the adjusted means, $\mu_{1,\underline{\mathbf{S}}_j}$, we use a debiased machine learning (DML) approach. DML combines flexible, data-adaptive machine learning methods with targeted estimators derived from nonparametric efficiency theory (Hines et al. 2022). This approach ensures precise estimation and valid inference for the parameters of interest, even when they involve complex, high-dimensional data. By addressing potential biases due to over-fitting and regularization that can arise with certain machine learning methods, DML produces estimates with desirable asymptotic properties similar to those of traditional parametric estimators. Crucially, it achieves this without imposing any restrictions on the distribution of the data or relationships among variables. This makes DML particularly well-suited for the decomposition in Equation 3, which involves a high-dimensional set of school characteristics whose relationships with test scores and neighborhood poverty may involve complex forms of nonlinearity, interaction, and heterogeneity.

Specifically, we implement DML using the following estimator for $\mu_{1,\underline{\mathbf{S}}_j}$:

$$\begin{aligned}\hat{\mu}_{1,\underline{\mathbf{S}}_j} = \frac{1}{n}\sum \frac{X}{\hat{P}(X=0)} \frac{\hat{P}\left(X=0|\underline{\mathbf{S}}_j\right)}{\hat{P}\left(X=1|\underline{\mathbf{S}}_j\right)}\left(Y - \hat{\mathbb{E}}\left[Y|X=1,\underline{\mathbf{S}}_j\right]\right) \\ + \frac{1-X}{\hat{P}(X=0)}\hat{\mathbb{E}}\left[Y|X=1,\underline{\mathbf{S}}_j\right], \qquad (4)\end{aligned}$$

where "hats" denote estimated values (Opacic et al. 2025). In this expression, $\hat{P}(X=x)$ denotes the estimated marginal probability of residing in a high-poverty $(X=1)$ or a low-poverty $(X=0)$ neighborhood. Similarly, $\hat{P}\left(X=x|\underline{\mathbf{S}}_j\right)$ represents the estimated probability of living in a neighborhood with a poverty level given by x, conditional on the set of school characteristics $\underline{\mathbf{S}}_j$. The term $\hat{\mathbb{E}}\left[Y|X=1,\underline{\mathbf{S}}_j\right]$ denotes the estimated mean test score for students attending schools with characteristics $\underline{\mathbf{S}}_j$, after setting X to indicate residence in a high-poverty neighborhood.

To compute $\hat{\mu}_{1,\underline{\mathbf{S}}_j}$, we first estimate each of its component terms. Estimates of the marginal probabilities, $\hat{P}(X=1)$ and $\hat{P}(X=0)$, are obtained from their corresponding sample proportions. Then, for the conditional means and probabilities, we train a series of "super learners" to model $\mathbb{E}\left[Y|X,\underline{\mathbf{S}}_j\right]$ and $P\left(X|\underline{\mathbf{S}}_j\right)$, from which we obtain predictions for all the remaining terms in the estimator.

A super learner is a stacking algorithm that combines multiple models to create an ensemble predictor that is guaranteed to perform at least as well as the most accurate individual model (Van der Laan, Polley, and Hubbard 2007). When the models included in this ensemble are flexible, data-adaptive machine learning algorithms, a super learner can approximate complex functions in high-dimensional data with exceptional accuracy. In our analysis, we specify the super learners to include a random forest (RF; Breiman 2001) and a gradient boosted tree (GBT; Friedman 2001), along with a conventional generalized linear model (GLM) for reference.

RFs and GBTs use recursive partitioning to create regression or classification trees by iteratively dividing the sample into subgroups based on binary partitions of the predictors. Each split is chosen to minimize a loss function that measures prediction error. RFs are composed of many complex trees, each constructed using a random sample of observations drawn with replacement from the observed data and a random subset of predictors selected as candidates for partitioning at each step. Predictions from a RF are obtained by averaging the predictions from all the individual trees that compose it. In contrast, GBTs build a sequence of simple trees with fewer partitions, but where each tree is trained to correct the prediction errors of its predecessors. Predictions from a GBT are given by a weighted sum of the predictions from all the trees in the sequence. Tree-based learners such as RFs and GBTs excel at accurately modeling complex relationships in high-dimensional data because they naturally handle interactions and non-linearities by partitioning the predictor space into smaller, more homogeneous regions.

When using machine learning models to compute an estimator like $\hat{\mu}_{1,\underline{\mathbf{S}}_j}$, it is important to implement them in conjunction with a sample-splitting procedure (Chernozhukov et al. 2018). This involves dividing the data into separate subsamples, with one portion used to train the models and another used to evaluate the estimator. Sample splitting eliminates the risk of bias that can arise if machine learning models are trained on the same data used for estimation. Thus, we employ a repeated cross-fitting procedure, which initiates by randomly dividing the data into five separate subsamples. Next, at the first iteration of the procedure, four subsamples are used to train the machine learning models, whereas the fifth is used to estimate the target parameter of interest. The procedure then rotates through all the subsamples, alternating whether they are used for training or estimation, and it produces a final estimate by averaging the results across iterations.

DML estimation is optimal in that $\hat{\mu}_{1,\underline{\mathbf{S}}_j}$ will be $\sqrt{n}$-consistent, asymptotically normal, and nonparametrically efficient as long as the super learners converge to $\mathbb{E}\left[Y|X,\underline{\mathbf{S}}_j\right]$ and $P\left(X|\underline{\mathbf{S}}_j\right)$ at a rate faster than $n^{1/4}$. This rate can be achieved by all the models we submit as candidates to our super learner, including RFs (Wager and Walther 2015) and GBTs (Zhang and Yu 2005). Because super learners inherit the fastest convergence rate of their candidate models, our approach ensures accurate estimation and valid inference for a decomposition that incorporates a high-dimensional representation of school quality.

Post-Intervention Gaps in Achievement

Although the descriptive decomposition illustrates how differences in school quality are associated with the achievement gap, it does not capture how equalizing school quality across neighborhoods would impact these disparities. In the third stage of our analysis, we examine the extent to which gaps in achievement between students from high- and low-poverty neighborhoods would change after hypothetical interventions aimed at equalizing different dimensions of the school environment (Lundberg 2024).

To define these post-intervention gaps, we use potential outcomes notation. Let $\underline{\mathbf{G}}_j$ represent a random vector of school characteristics drawn from $f(\underline{\mathbf{S}}_j|X = 0, \mathbf{C})$, the joint distribution of $\underline{\mathbf{S}}_j = \{\mathbf{S}_1, ..., \mathbf{S}_j\}$ observed among students from low-poverty neighborhoods ($X = 0$) with baseline covariates $\mathbf{C}$. In addition, let $Y\left(\underline{\mathbf{G}}_j\right)$ denote the test score a student would achieve if they attended a school with the characteristics given by $\underline{\mathbf{G}}_j$.

With this notation, we quantify the portion of the observed gap in achievement that would be eliminated by a stochastic intervention equalizing the distribution of school characteristics to match that observed among students from low-poverty neighborhoods. Formally, this estimand can be defined as follows:

$$\lambda_{\underline{\mathbf{S}}_j} = \underbrace{\mathbb{E}\left[Y|X=1\right]}_{\mu_1} - \underbrace{\mathbb{E}\left[Y(\underline{\mathbf{G}}_j)|X=1\right]}_{\eta_{1,\underline{\mathbf{S}}_j}}, \tag{5}$$

which represents the *disparity eliminated* by aligning the distribution of $\underline{\mathbf{S}}_j$ for students from different neighborhoods. In this expression, the first term, $\mu_1 = \mathbb{E}\left[Y|X=1\right]$, is the mean test score observed among students in high-poverty neighborhoods, as defined previously. The second term, $\eta_{1,\underline{\mathbf{S}}_j} = \mathbb{E}\left[Y\left(\underline{\mathbf{G}}_j\right)|X=1\right]$, represents the counterfactual mean score for these same students if they attended schools with the distribution of characteristics $\underline{\mathbf{S}}_j$ observed among their counterparts who have the same baseline covariates $\mathbf{C}$ but live in low-poverty neighborhoods. Thus, $\lambda_{\underline{\mathbf{S}}_j}$ captures how the observed gap in achievement would change under an intervention that aligns the distribution of school characteristics for students in high-poverty neighborhoods with that of their peers in low-poverty neighborhoods.

The disparity eliminated, $\lambda_{\underline{\mathbf{S}}_j}$, differs from some other estimands commonly examined in prior research on neighborhood and school inequality (e.g., Wodtke and Parbst 2017; Wodtke et al. 2023). Specifically, it does not represent a causal effect of living in one type of neighborhood versus another. Instead, it captures how the observed disparity in achievement between students from different neighborhoods would change under a hypothetical intervention that shifts access to schools of varying quality. In other words, rather than assess a causal effect of neighborhood poverty, it evaluates a causal effect of modifying access to different types of schools on the achievement gap between neighborhoods.

The disparity eliminated can be nonparametrically identified under an ignorability assumption, which requires that a student's potential outcomes are independent of their observed school characteristics, conditional on neighborhood poverty and the baseline controls. Formally, this condition can be expressed as $Y\left(\underline{\mathbf{s}}_j\right) \perp \underline{\mathbf{S}}_j \mid \{X = 1, \mathbf{C}\}$ for all j, where $Y\left(\underline{\mathbf{s}}_j\right)$ denotes a potential outcome under attendance at a school with the specific set of characteristics $\underline{\mathbf{s}}_j$. In substantive terms, this assumption requires that, after accounting for residence in a high-poverty neighborhood and the baseline controls, school assignment must be as good as random–that is, not confounded by unobserved factors that may influence selection into schools and student achievement.

This is a strong assumption that could be violated in our analysis. We attempt to mitigate this concern by adjusting for an extensive set of baseline covariates, including prior measures of achievement taken at kindergarten entry. In addition, we conduct a sensitivity analysis to assess the robustness of our results to hypothetical patterns of unobserved confounding.

Although ignorability of school assignment is a stringent assumption, it is weaker than the assumptions needed to identify the combined effects of neighborhood *and* school contexts together. Identifying these joint effects requires multiple, highly restrictive assumptions that are even harder to satisfy and more difficult to comprehensively evaluate in a sensitivity analysis. In contrast, the disparity eliminated depends solely on the causal effects of different school environments, making it comparatively easier to identify and consistently estimate than the joint effects of neighborhoods and schools together.

Under the assumption that school assignment is ignorable, the disparity eliminated can be consistently estimated using another DML approach. Specifically, for the counterfactual mean $\eta_{1,\underline{\mathbf{S}}_j}$, we implement this approach using the following estimator:

$$\begin{aligned}\hat{\eta}_{1,\underline{\mathbf{S}}_j} = \frac{1}{n}\sum & \frac{X}{\hat{P}(X=1)}\frac{\hat{P}(X=1|\mathbf{C})}{\hat{P}(X=0|\mathbf{C})}\frac{\hat{P}\left(X=0|\mathbf{C},\underline{\mathbf{S}}_j\right)}{\hat{P}\left(X=1|\mathbf{C},\underline{\mathbf{S}}_j\right)}\left(Y-\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]\right) \\ & + \frac{1-X}{\hat{P}(X=1)}\frac{\hat{P}(X=1|\mathbf{C})}{\hat{P}(X=0|\mathbf{C})}\left(\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]-\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]\right) \\ & + \frac{X}{\hat{P}(X=1)}\left(\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]-\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]|X=1\right]\right) \\ & + \hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X=0,\mathbf{C}\right]|X=1\right]. \end{aligned} \tag{6}$$

As with the descriptive decomposition, we predict the components of this estimator using repeated cross-fitting and a series of super learners that include a RF, GBT, and conventional GLM. However, for estimating $\eta_{1,\underline{\mathbf{S}}_j}$, these super learners are trained to model a different set of functions.

Specifically, we train the super learners to model $P(X|\mathbf{C})$, the probability of living in a high- versus low-poverty neighborhood conditional on the baseline covariates; $P(X|\mathbf{C},\underline{\mathbf{S}}_j)$, the same probability conditional on both baseline controls and the school characteristics in $\underline{\mathbf{S}}_j$; $\mathbb{E}\left[Y|X,\mathbf{C},\underline{\mathbf{S}}_j\right]$, the conditional mean of student test scores given neighborhood poverty, baseline controls, and the vector of school characteristics; and finally, $\mathbb{E}\left[\mathbb{E}\left[Y|X=1,\mathbf{C},\underline{\mathbf{S}}_j\right]|X,\mathbf{C}\right]$, an iterated expectation reflecting how this conditional mean varies with neighborhood poverty and the baseline controls alone. These models can be used to predict the key terms needed for estimating the counterfactual mean $\eta_{1,\underline{\mathbf{S}}_j}$, whereas the observed mean μ_1 in the disparity eliminated is simply estimated using its sample analog.

Provided that the super learners converge to these functions at sufficiently fast rates, the resulting estimator for the disparity eliminated is $\sqrt{n}$-consistent, asymptotically normal, and nonparametrically efficient. As before, this is achieved without imposing any functional form restrictions on the distribution of the data, allowing for complex relationships among neighborhood poverty, student achievement, and

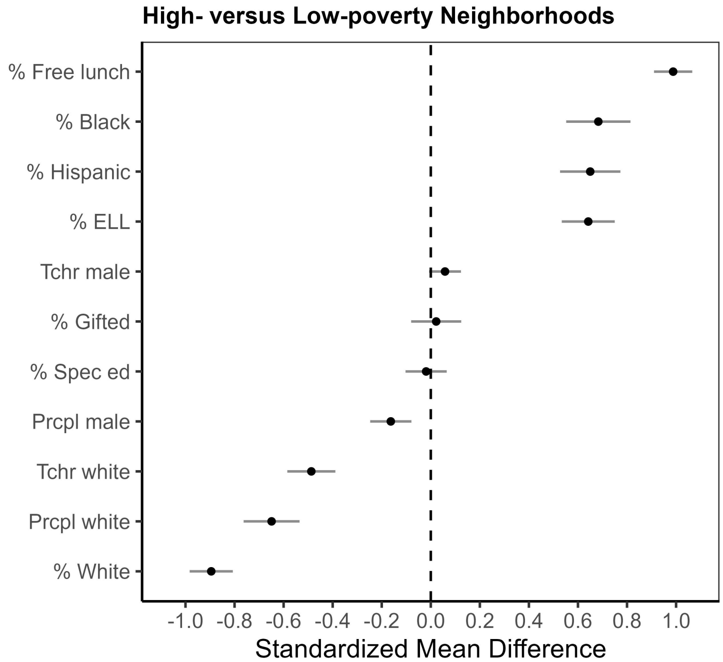

Figure 3: Standardized mean differences in measures of school composition between high- and low-poverty neighborhoods. *Note:* The standardized mean differences contrast residence in a neighborhood with a poverty rate $\geq$ 20 percent versus $<$ 20 percent. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File.

our high-dimensional representation of school quality. A detailed derivation of this estimator and its properties is provided in part F of the online supplement.

Inference

To account for the clustering of students within schools in the ECLS-K, we employ cluster-robust inferential statistics in all analyses. For our DML estimators, this involves modifying the cross-fitting algorithm to partition the sample by schools rather than students and then applying a cluster-robust estimator to compute the variance of both $\hat{\mu}_{1,\underline{\mathbf{S}}j}$ and $\hat{\eta}_{1,\underline{\mathbf{S}}j}$.

To address the challenge of missing data, we replicate our analysis across five multiply imputed data sets. For each of these data sets, missing values for all variables are imputed using chained RFs, a nonparametric approach that does not impose any restrictive assumptions on the distribution from which the imputed values are sampled (Hong et al. 2020). Results from across the imputed data sets are combined following Rubin (1987) to account for the additional uncertainty introduced by missing data. The proportion of missing information in our sample is 0.22, which is mainly due to a combination of item-specific non-response, survey non-response among teachers and administrators, and panel attrition from the ECLS-K.

Results

Differences in School Composition

Figure 3 displays standardized mean differences in school composition, comparing students from neighborhoods with a poverty rate of 20 percent or higher to those from neighborhoods with a poverty rate below 20 percent. Specifically, the figure presents a dot-and-whisker plot, where the horizontal axis shows the estimated differences in SD units, along with their 95 percent confidence intervals, and the vertical axis lists each measure of school composition, ordered by the magnitude of the differences.

School composition varies substantially across high- and low-poverty neighborhoods, reflecting the strong connection between residential location and school assignment. Students from high-poverty neighborhoods attend schools with much larger proportions of peers who are eligible for free lunch ($\hat{D}_{S_{1,k}} = 0.99$ SDs; % Free lunch), black ($\hat{D}_{S_{1,k}} = 0.68$ SDs; % black), Hispanic ($\hat{D}_{S_{1,k}} = 0.65$ SDs; % Hispanic), and English language learners ($\hat{D}_{S_{1,k}} = 0.64$ SDs; % ELL). Conversely, they have much lower exposure to white peers ($\hat{D}_{S_{1,k}} = -0.90$ SDs; % white), white principals ($\hat{D}_{S_{1,k}} = -0.65$ SDs; Prcpl white), and white teachers ($\hat{D}_{S_{1,k}} = -0.49$ SDs; Tchr white). There are no significant differences in the proportion of students enrolled in special education or gifted and talented programs. Thus, the racial and socioeconomic composition of students and staff differ sharply between high- and low-poverty neighborhoods, but these differences do not extend to participation in specialized programs.

Differences in School Resources

Figure 4 presents standardized mean differences in school resources, again contrasting students from high- versus low-poverty neighborhoods. Overall, differences in school resources between these residential contexts are substantially smaller than those observed for school composition.

There are few notable differences in measures of school funding, although students from high-poverty neighborhoods attend schools where staff are somewhat less likely to have received a salary increase in the past year ($\hat{D}_{S_{2,k}} = -0.18$ SDs; Schl salary inc). Staffing levels are also broadly similar across neighborhoods, except that students from high-poverty areas attend schools with more ESL teachers ($\hat{D}_{S_{2,k}} = 0.43$ SDs; ESL tchrs/stdnt) and translators ($\hat{D}_{S_{2,k}} = 0.21$ SDs; Schl translators), which reflects the larger number of English language learners in these settings. Differences in staff qualifications, including their educational attainment and certification, are generally small.

School facilities are also rated quite similarly across high- and low-poverty neighborhoods. Modest differences emerge for specific pieces of infrastructure, with multi-use rooms ($\hat{D}_{S_{2,k}} = -0.20$ SDs), art rooms ($\hat{D}_{S_{2,k}} = -0.16$ SDs), and music rooms ($\hat{D}_{S_{2,k}} = -0.13$ SDs) rated somewhat lower in quality at schools attended by students from high-poverty neighborhoods. However, no significant differences are observed in the quality ratings of regular classrooms, auditoriums, IT labs, libraries, or gyms.

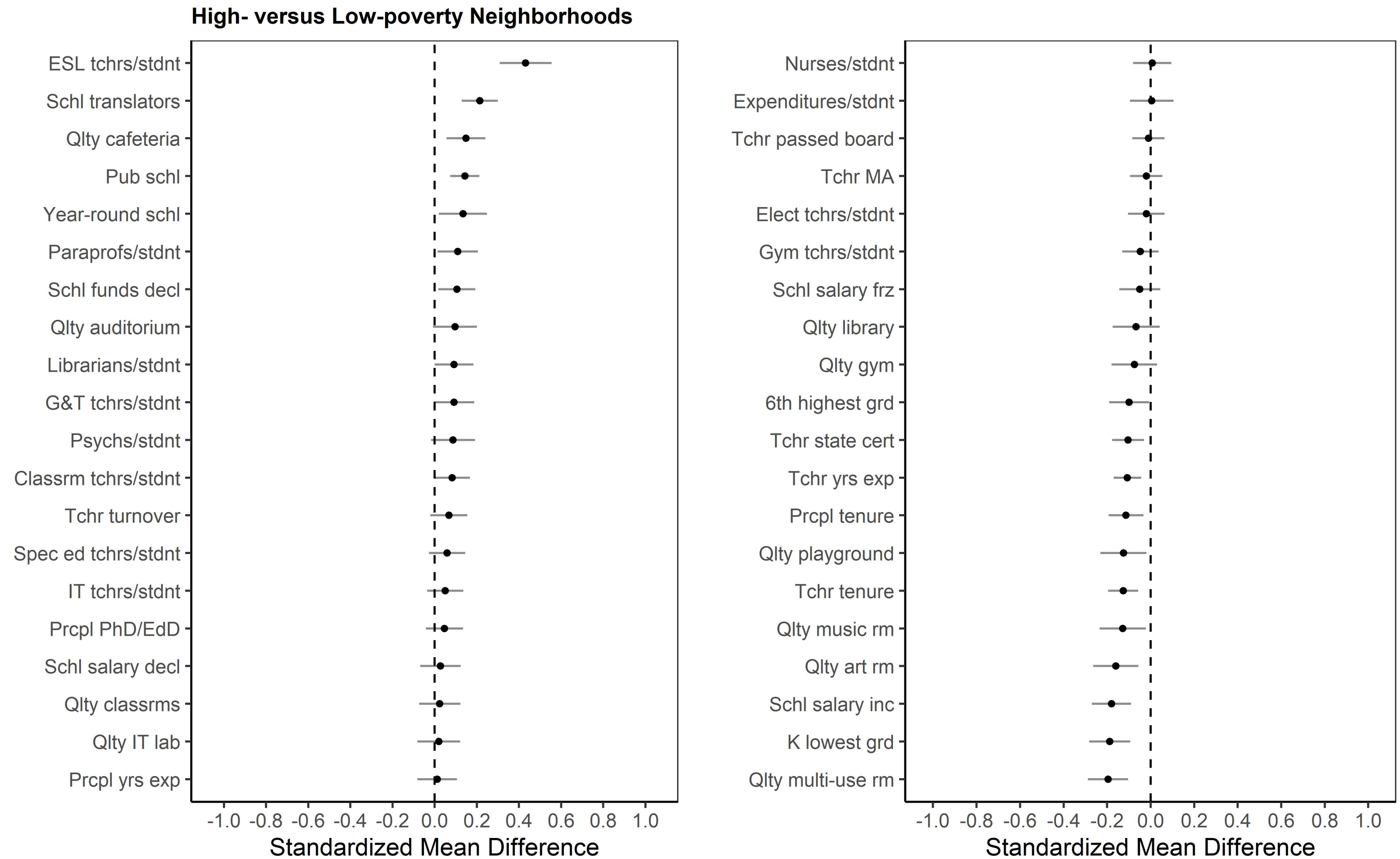

Figure 4: Standardized mean differences in measures of school resources between high- and low-poverty neighborhoods. *Note:* The standardized mean differences contrast residence in a neighborhood with a poverty rate $\geq$ 20 percent versus $<$ 20 percent. Source: Geolytics (2012); U.S. Department of 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, CCD, "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, CCD, "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–2012, Version Provisional 1a; and U.S. Department of Education, National Center for Education Statistics, PSS, 2011–2012.

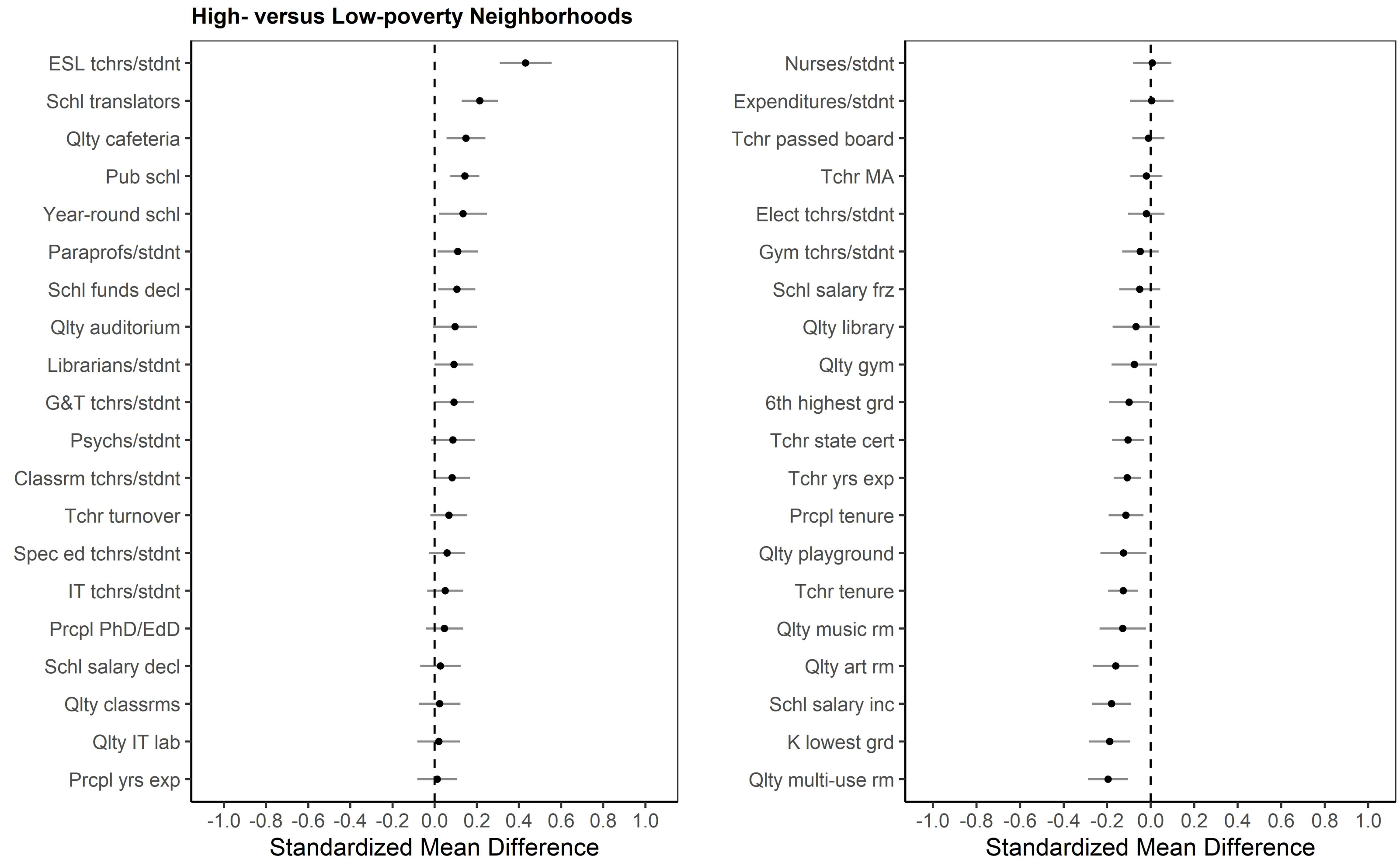

Figure 5: Standardized mean differences in measures of instructional practices and curriculum content between high- and low-poverty neighborhoods. *Note:* The standardized mean differences contrast residence in a neighborhood with a poverty rate ≥ 20 percent versus < 20 percent. Source: Geolytics (2012); 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.

Differences in Instruction

Figure 5 presents standardized mean differences in instructional practices and curricular content, comparing students from high- versus low-poverty neighborhoods. Overall, differences in instruction by neighborhood poverty are small and relatively infrequent, despite examining more than 90 distinct measures of this dimension of school quality. Moreover, the few and mostly minor differences that are observed do not clearly suggest superior instruction for either group. For example, students from high-poverty neighborhoods are more likely to receive reading instruction using decodables (Rd decodables), basal series (Rd basal series), and big books (Rd big books), where each measure is about one-fifth of a SD higher for these students compared to those from low-poverty neighborhoods. In general, though, the primary conclusion from Figure 5 is that the schools serving students from high- and low-poverty neighborhoods appear to offer strikingly similar instruction during first grade.

Differences in School Climate

Figure 6 displays standardized mean differences in school climate. Although these estimates are smaller than those observed for school composition, many are both statistically and practically significant, suggesting meaningful differences in the social and cultural environment of these schools.

For example, students from high-poverty neighborhoods attend schools where administrators perceive lower levels of parental ($\hat{D}_{S_{4,k}} = -0.43$ SD; parental support) and community support ($\hat{D}_{S_{4,k}} = -0.39$ SD; community support), which represent the largest differences observed in school climate. Students from high-poverty neighborhoods are also less likely to have teachers with high morale, as fewer of them report enjoying their jobs ($\hat{D}_{S_{4,k}} = -0.19$ SD; Tchr enjoys job) or express a desire to choose the same career again ($\hat{D}_{S_{4,k}} = -0.10$ SD; Tchr teach again). In addition, perceptions of administrative support are lower in schools serving students from high-poverty neighborhoods, with fewer teachers feeling accepted ($\hat{D}_{S_{4,k}} = -0.18$ SD; Tchr accepted), encouraged ($\hat{D}_{S_{4,k}} = -0.17$ SD; admin encouraging), or aligned with the school's mission ($\hat{D}_{S_{4,k}} = -0.15$ SD; aligned on mission). Moreover, class disorder is higher in these schools, as indicated by a greater frequency of physical altercations ($\hat{D}_{S_{4,k}} = 0.30$ SD; freq fights) and classroom misbehavior ($\hat{D}_{S_{4,k}} = 0.22$ SD; class disorder).

Thus, while still smaller than compositional differences, the disparities in school climate across high- and low-poverty neighborhoods are not trivial, particularly with respect to perceptions of parental, community and administrative support, teacher morale, and student behavior.

Differences in School Effectiveness

Figure 7 presents density plots of school effectiveness, as measured by value-added scores in reading and math during first grade. The distributions for students from high- versus low-poverty neighborhoods overlap substantially, suggesting that students from different residential contexts attend elementary schools with broadly

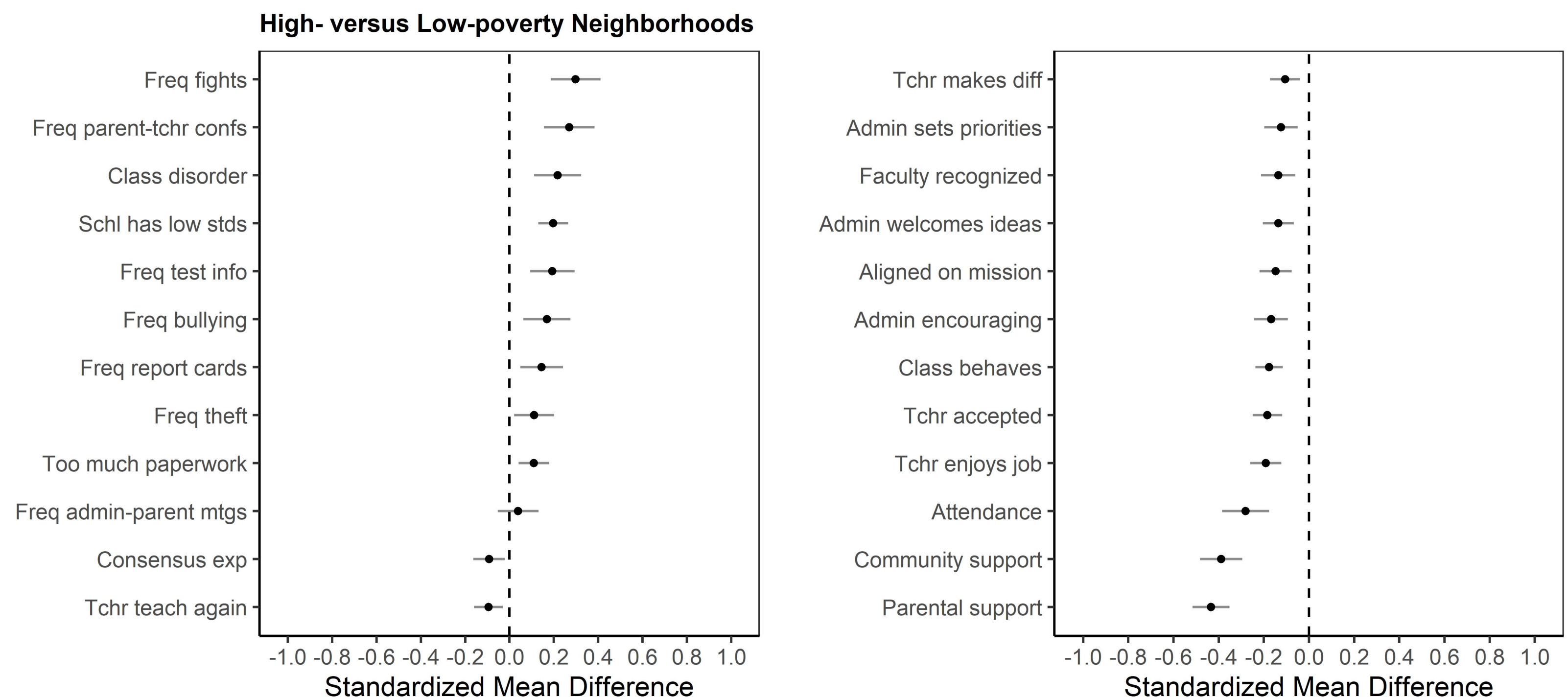

Figure 6: Standardized mean differences in measures of school climate between high- and low-poverty neighborhoods. *Note:* The standardized mean differences contrast residence in a neighborhood with a poverty rate $\geq$ 20 percent versus $<$ 20 percent. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File.

comparable effectiveness at imparting academic skills. On average, students from high-poverty neighborhoods attend schools whose value-added scores are about one-tenth of a SD lower than those from low-poverty neighborhoods, as shown by the vertical lines overlaying the estimated densities.

To summarize, the primary differences between schools attended by students from high- and low-poverty neighborhoods involve their composition. In contrast, differences in resources, instruction, and effectiveness are small, whereas disparities in school climate are evident and meaningful but still comparatively modest. Contrary to prevailing assumptions, then, we find rather limited evidence that students from high-poverty neighborhoods consistently attend low-quality schools. Although the schools serving students from disadvantaged neighborhoods differ in their composition and climate, these differences do not appear to undermine their quality of instruction or capacity to foster an effective learning environment. When school quality is comprehensively assessed along multiple dimensions, its distribution across neighborhoods is more complex, with substantial overlap between more and less disadvantaged communities in many cases.

The Association of School Differences with Disparities in Achievement

Figure 8 summarizes our descriptive decomposition of the achievement gap between students from high- and low-poverty neighborhoods. The top row displays the observed disparity in reading and math scores, measured during the spring of third grade. It indicates that students from high-poverty neighborhoods score approximately one-half of a SD below their peers from low-poverty neighborhoods.

The second row presents estimates for the portion of this disparity that can be "statistically explained" by differences in school composition. Specifically, it shows DML estimates of $\Delta_{\mathbf{S}_1}$, which captures the contribution of demographic differences across the schools serving high- versus low-poverty neighborhoods. These estimates indicate that differences in school composition account for roughly four-fifths of the disparity in achievement. This finding reflects the strong association between neighborhood poverty, the socioeconomic composition of schools, and student achievement.

The subsequent rows of Figure 8 display DML estimates for $\Delta_{\mathbf{S}_2}$ through $\Delta_{\mathbf{S}_5}$, which capture the extent to which differences in school resources, instruction, climate, and effectiveness might statistically explain the achievement gap. In contrast to school composition, none of these other factors explain an appreciable share of the disparity in test scores between students from high- and low-poverty neighborhoods. Once differences in composition are taken into account, adjusting for additional dimensions of school quality leaves the achievement disparity virtually unchanged.

Finally, the bottom row of Figure 8 shows estimates of $\Delta_{\perp}$, which represents the residual portion of the achievement gap that cannot be explained by any of the school characteristics included in our analysis. This unexplained component amounts to roughly one-fifth of the overall disparity in achievement.

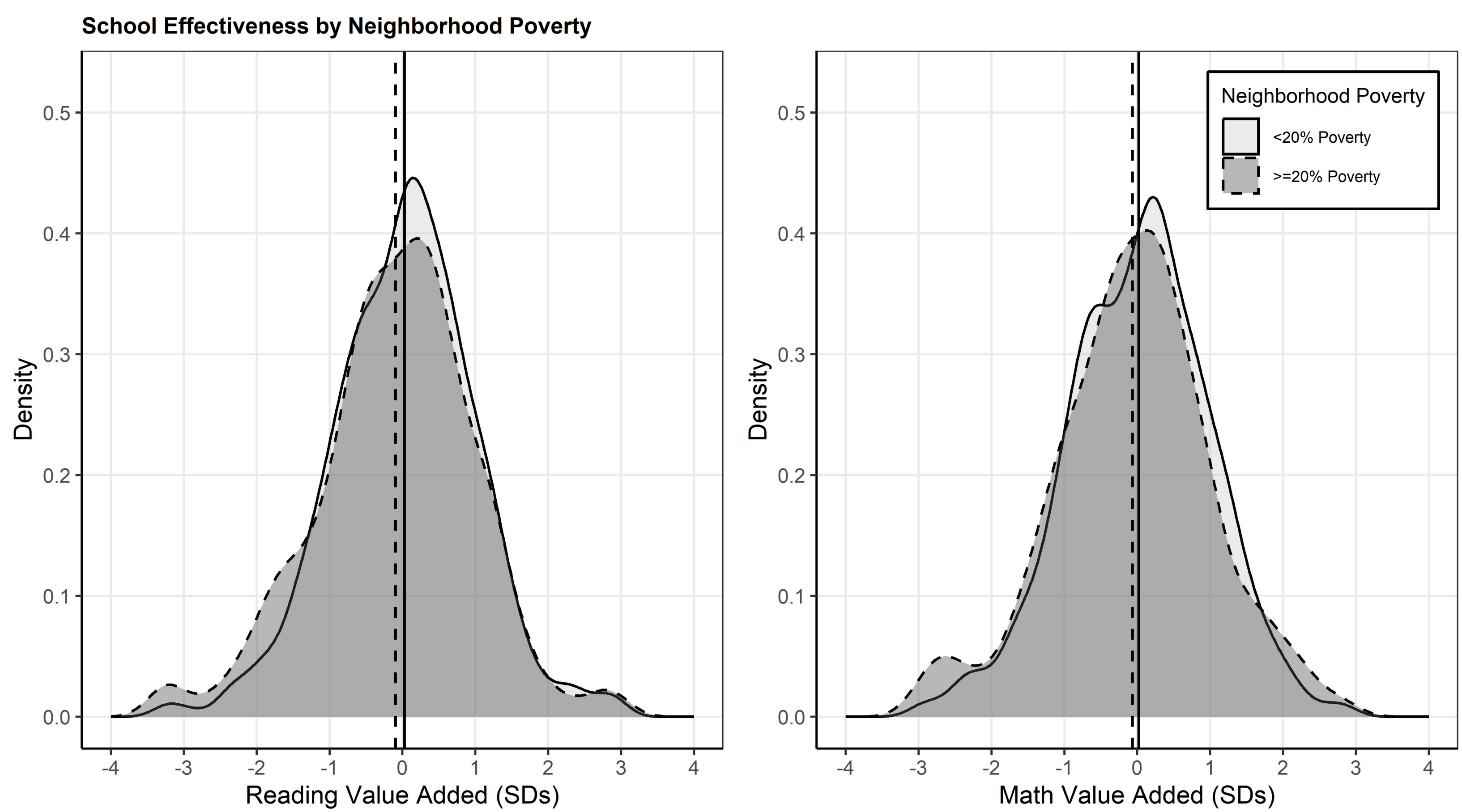

Figure 7: Differences in school effectiveness between high- and low-poverty neighborhoods. *Note:* This plot contains kernel density estimates for school effectiveness in neighborhoods with a poverty rate $\geq$ 20 percent versus $<$ 20 percent. Source: Geolytics (2012); U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011 (ECLS-K:2011), Restricted-Use Data File.

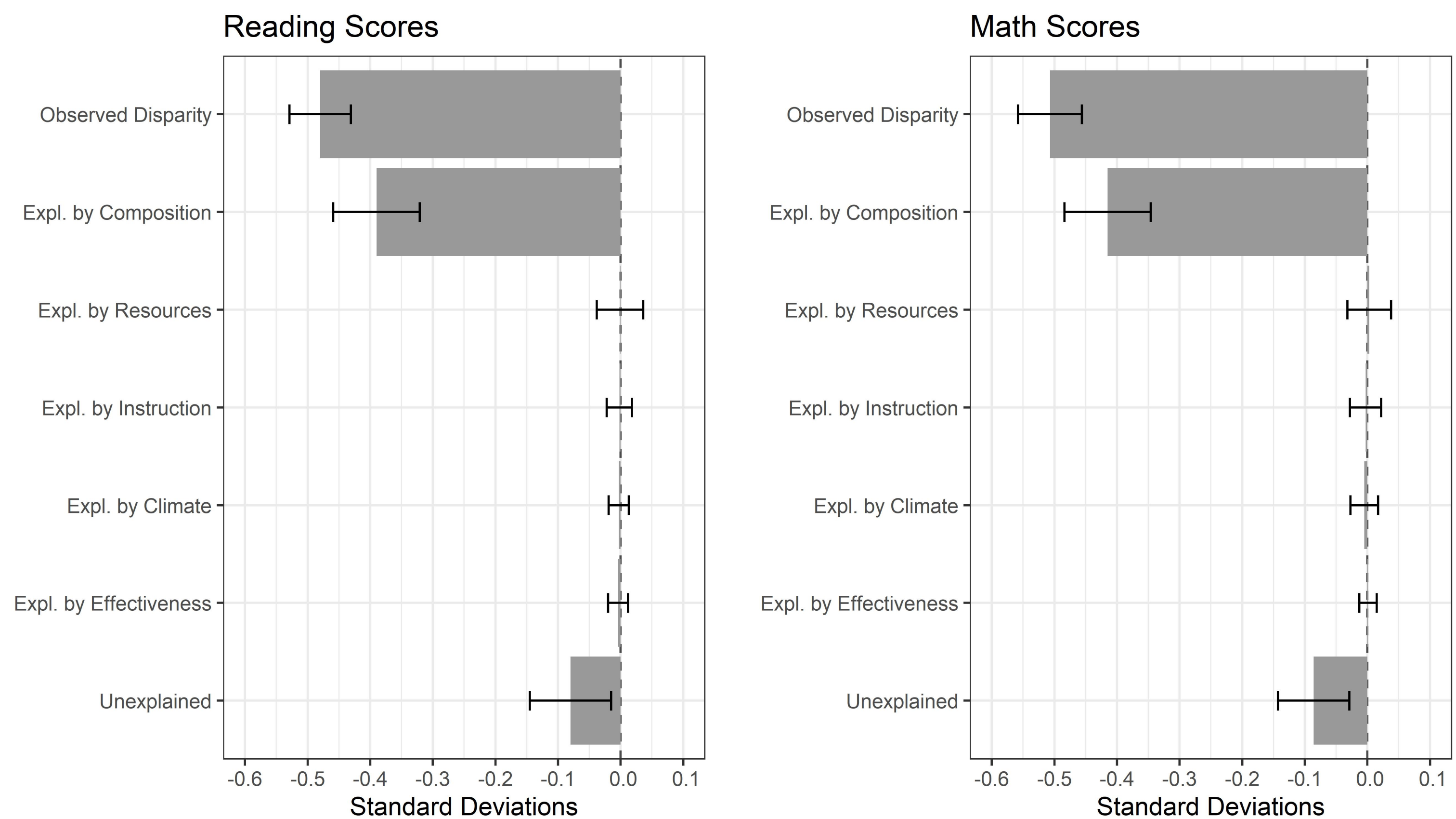

Figure 8: Descriptive decomposition of the observed disparity in achievement test scores between students from high- and low-poverty neighborhoods. *Note:* This plot contains DML estimates of the "disparity explained" by different dimensions of school quality. All estimates are reported in SD units. 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, CCD, "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, CCD, "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–2012, Version Provisional 1a; and U.S. Department of Education, National Center for Education Statistics, PSS, 2011–2012.

Taken together, these results suggest that differences in school composition are strongly associated with disparities in achievement between students from high- and low-poverty neighborhoods. This finding highlights the tight coupling of residential context, the demographic makeup of schools, and student achievement. In contrast, the other dimensions of school quality do not account for any meaningful portion of the achievement gap, above and beyond differences in composition, even when they are measured using rich, high-dimensional data.

The Impact of Equalizing Schools on Disparities in Achievement

The descriptive decomposition discussed previously does not have a causal interpretation–that is, it does not capture how changes in access to different types of schools might impact the achievement gap. To build on these descriptive results, Table 1 presents DML estimates of the disparity eliminated under hypothetical interventions aimed at equalizing various dimensions of school quality. These estimates assess how the achievement gap would change if students from high-poverty neighborhoods had access to schools with characteristics similar to those of their peers in low-poverty neighborhoods.

The results suggest that equalizing school characteristics between students from high- and low-poverty neighborhoods would lead to only modest reductions in the achievement gap. For example, aligning the composition of schools serving students from high- and low-poverty neighborhoods is estimated to reduce the disparity in reading scores by just 0.027 SDs, representing about 6 percent of the observed gap. A similar pattern holds for math scores, where the estimated reduction in the achievement disparity is somewhat smaller but substantively similar. Moreover, additionally aligning the resources, instruction, climate, or effectiveness of these schools also appears to have little impact on achievement disparities, above and beyond equalizing their composition.

Overall, equalizing *any* set of school characteristics is estimated to reduce the achievement disparity by no more than 0.03 SDs for either reading or math, and the proportion of the gap eliminated remains consistently below 10 percent. In fact, we estimate that aligning the distribution of *all* school characteristics together would reduce the disparity in reading scores by only 0.027 SDs and the gap in math scores by just 0.024 SDs. Although these point estimates are not entirely negligible, they are small in practical terms, and all of their confidence intervals span zero.

Our estimates of the disparity eliminated rely on the strong assumption that selection into different types of schools is not influenced by unobserved factors that also affect student test scores. This assumption cannot be empirically verified, and it may not hold in our analysis. Although we adjust for a large set of baseline covariates, it remains possible that unmeasured variables confound the relationship between school quality and student achievement. These variables might include unobserved characteristics of a student's family, like parental wealth, or unmeasured features of their home environment, such as the presence of environmental health hazards. To address potential distortions from unobserved confounding, we conduct a formal sensitivity analysis to evaluate the robustness of our causal inferences.

Table 1: De-biased machine learning estimates of the disparity eliminated between students from high- and low-poverty neighborhoods.

Label	Estimand	Point Est.	95% CI	Prop. Elim.
Reading test scores				
Observed disparity	Δ	-0.480	$[-0.529, -0.431]$	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.027	$[-0.072, 0.018]$	0.06
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.027	$[-0.074, 0.020]$	0.06
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.026	$[-0.073, 0.021]$	0.05
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.029	$[-0.076, 0.018]$	0.06
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.027	$[-0.074, 0.020]$	0.06
Math test scores				
Observed disparity	Δ	-0.507	$[-0.558, -0.456]$	
Disparity eliminated by equalizing:				
(1) School composition	$\lambda_{\underline{\mathbf{S}}_1}$	-0.018	$[-0.067, 0.031]$	0.04
(2) School resources + (1)	$\lambda_{\underline{\mathbf{S}}_2}$	-0.025	$[-0.072, 0.022]$	0.05
(3) Instructional practices + (2)	$\lambda_{\underline{\mathbf{S}}_3}$	-0.025	$[-0.076, 0.026]$	0.05
(4) School climate + (3)	$\lambda_{\underline{\mathbf{S}}_4}$	-0.024	$[-0.077, 0.029]$	0.05
(5) School effectiveness + (4)	$\lambda_{\underline{\mathbf{S}}_5}$	-0.024	$[-0.079, 0.031]$	0.05

Note: All estimates are reported in SD 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, CCD, "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center for Education Statistics, CCD, "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–12, Version Provisional 1a; and U.S. Department of Education, National Center for Education Statistics, PSS, 2011–12.

Specifically, we construct bias-adjusted estimates for the disparity eliminated that results from equalizing the full set of school characteristics (i.e., $\lambda_{\underline{\mathbf{S}}_5}$). These estimates are derived by subtracting multiples of a bias term, denoted by ψ, from our focal estimates reported in Table 1. We define ψ as the observed change in our point estimates for $\lambda_{\underline{\mathbf{S}}_5}$ after omitting all controls for family background from the super learners used to construct them. In other words, ψ captures the magnitude of the bias introduced to our estimates by excluding controls for parental income, education, occupation, employment status, and family structure. We then assess the robustness of our findings by subtracting multiples of ψ from our point and interval estimates, which allows us to quantify how strong any unobserved confounding would need to be–relative to that introduced by omitting controls for family background–to alter our main conclusions.

Figure 9 plots these bias-adjusted estimates of the disparity eliminated against multiples of ψ. Although our original estimates are substantively small, they do not appear particularly sensitive to unobserved confounding. The results in this figure suggest that even if we omitted a variable or set of variables with twice the confounding influence of family background, our conclusions about the limited impact of equalizing access to schools of different quality would remain largely intact.

Indeed, our estimates would *understate* the disparity eliminated only if there were unobserved factors whose confounding influence operated in the opposite direction to that of family background. For example, if students from high-poverty neighborhoods selected into lower-quality schools based on unobserved advantages, rather than disadvantages, this would attenuate our estimates of the degree to which equalizing school environments might reduce disparities in achievement. This pattern of unobserved confounding seems unlikely, although we cannot rule it out empirically. Overall, the results of our sensitivity analysis suggest that our inferences about the disparity eliminated are highly robust to unobserved confounding.

Why do hypothetical interventions aimed at equalizing school environments appear to have only modest effects on achievement disparities, even when allowing for potential bias due to unobserved confounding? To address this question, Figure 10 displays the estimated marginal effects of each school characteristic included in our analysis on student test scores. Specifically, the vertical axis of the plot shows the average marginal effect of each school characteristic on test scores, as computed from our super learners, whereas the horizontal axis represents the standardized mean difference in each of these characteristics, comparing students from high- versus low-poverty neighborhoods.

School characteristics located in the corners of the plot would represent those that (1) have large marginal effects on test scores and (2) differ substantially across high- and low-poverty neighborhoods. However, the roughly spherical distribution of estimates centered in the middle of Figure 10 suggests that few school characteristics strongly influence test scores, net of other factors, *and* exhibit pronounced differences by neighborhood poverty. In other words, the plot shows that the characteristics of schools that are most strongly linked to student performance are not those that vary significantly across neighborhoods with high- versus low-poverty rates. As a result, the corners of the plot are largely empty.

For example, while measures of school composition, such as the proportion of students who are eligible for free or reduced-price lunch, differ substantially between high- and low-poverty neighborhoods, they are not strongly associated with student test scores, controlling for other factors. Conversely, characteristics that are stronger predictors of student achievement, such as school value added, exhibit relatively little variation across neighborhood poverty levels. Moreover, most of the school characteristics measured in the ECLS-K are only weakly associated with both student achievement and neighborhood poverty. This is reflected in the dense cluster of unlabeled points near the origin of the plot at $(0, 0)$, which represent school characteristics that neither predict test scores nor differ much between high- versus low-poverty neighborhoods.

To summarize, our results provide little evidence that equalizing the distribution of school characteristics across high- and low-poverty neighborhoods would appreciably reduce disparities in reading or math achievement. This conclusion holds even under plausible patterns of unobserved confounding. It is driven primarily by the fact that the single dimension of school quality that differs markedly between

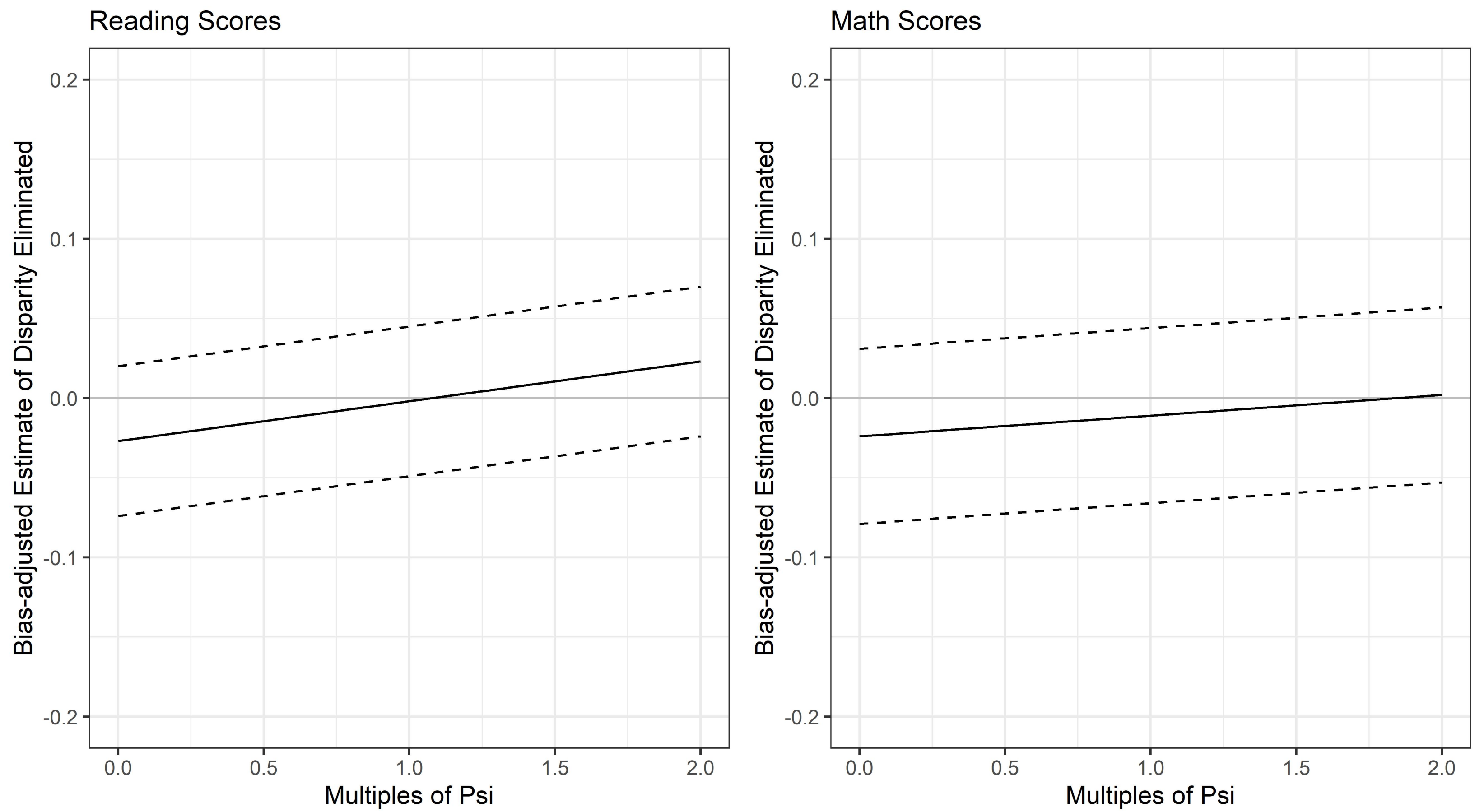

Figure 9: Bias-adjusted estimates of the disparity eliminated by equalizing all measures of school quality. *Note:* This plot contains bias-adjusted estimates of the "disparity eliminated" by equalizing the distribution of all measures of school quality. The parameter ψ represents the omitted variable bias that arises by virtue of excluding parental education, occupational status, and income from our vector of baseline controls. All estimates are reported in SD 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, CCD, "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, CCD, "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–2012, Version Provisional 1a; and U.S. Department of Education, National Center for Education Statistics, PSS, 2011–2012.

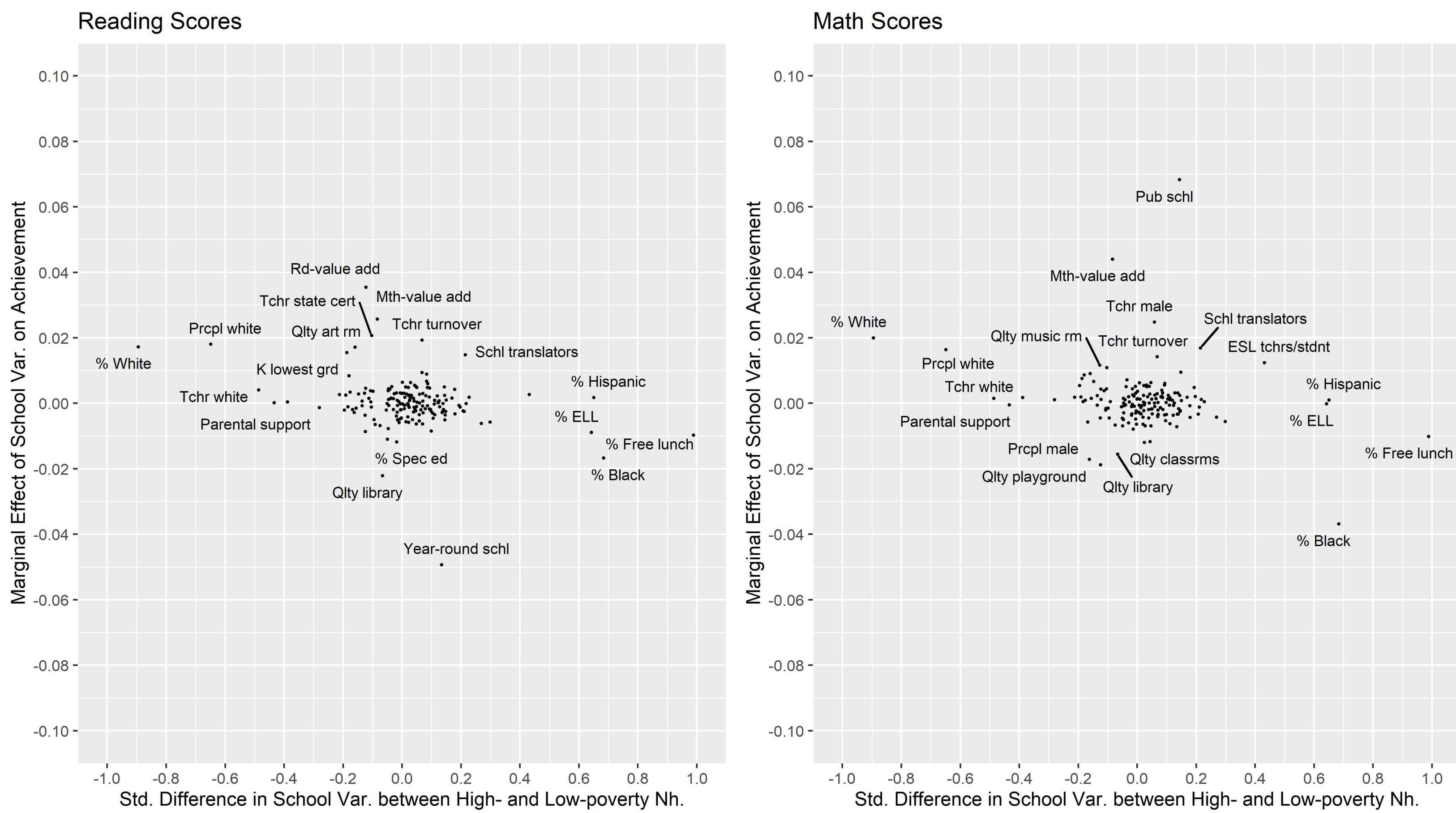

Figure 10: Average marginal effects of each school characteristic on test scores plotted against its standardized mean difference between high- and low-poverty neighborhoods. *Note:* This figure contains a plot of school characteristics according to (1) how strongly the characteristic predicts test scores, net of other factors, and (2) how strongly it differs by neighborhood poverty. Specifically, the vertical axis displays the average marginal effect for each school characteristic from a super learner used to model test scores, whereas the horizontal axis displays the standardized mean difference in a school characteristic across neighborhoods with a poverty rate $\geq$ 20 percent versus $<$ 20 percent. 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, CCD, "School District Finance Survey (F-33)," fiscal year 2012, Provisional Version 1a; U.S. Department of Education, National Center of Education Statistics, CCD, "Local Education Agency Universe Survey," "Public Elementary/Secondary School Survey," 2011–2012, Version Provisional 1a; and U.S. Department of Education, National Center for Education Statistics, PSS, 2011–2012.

high- and low-poverty neighborhoods–school composition–is only weakly linked to student achievement once baseline confounders have been controlled.

Ancillary Analyses

Beyond concerns about unobserved confounding, our inferences may also be influenced by several other analytic decisions that warrant special scrutiny. One such decision involves how we define high- and low-poverty neighborhoods. Throughout our analysis, we distinguished between neighborhoods with poverty rates of 20 percent or higher and those with rates below 20 percent. This threshold might obscure more pronounced disparities between students from extremely poor neighborhoods and those from areas with very low-poverty rates. To assess the sensitivity of our findings to these measurement choices, Part G of the online supplement replicates our analyses using alternative definitions of high- and low-poverty neighborhoods. Specifically, we compare students from neighborhoods with poverty rates of $\geq$ 30 percent versus $<$ 30 percent, $\geq$ 20 percent versus $\leq$ 10 percent, and $\geq$ 30 percent versus $\leq$ 5 percent. Across all specifications, our results are substantively similar to those presented previously.

We also focused on the disparity eliminated that results from equalizing the distribution of school characteristics between students from high- and low-poverty neighborhoods *who otherwise share similar covariates at baseline*. However, an alternative approach might consider the disparity that would remain after aligning school characteristics across neighborhoods *irrespective of student attributes*. Formally, this quantity can be expressed as follows:

$$\gamma_{\underline{\mathbf{S}}_j} = \mathbb{E}\big[Y\big(\underline{\mathbf{H}}_j\big)|X=1\big] - \mathbb{E}\big[Y\big(\underline{\mathbf{H}}_j\big)|X=0\big], \tag{7}$$

where $\underline{\mathbf{H}}_j$ denotes a random vector of school characteristics 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 disparity represents the counterfactual gap in achievement that would persist if all students attended schools with characteristics drawn from the distribution observed in low-poverty neighborhoods, regardless of their individual traits or family background.

In part H of the online supplement, we estimate this alternative disparity using a novel approach that combines permutation weighting with DML. These results are substantively similar to those presented previously, indicating that our conclusions are insensitive to whether school characteristics are equalized among students who are comparable at baseline or whether they are equalized with no attention to student covariates. In other words, our main findings are also robust to alternative conceptualizations of the hypothetical intervention aimed at equalizing access to quality schools.

Finally, our analyses focused on estimating the disparity eliminated across the total population of students. However, equalizing the distribution of school characteristics could potentially have greater impacts among specific subpopulations. To examine this possibility, part I of the online supplement presents ancillary analyses in which we estimate the disparity eliminated separately by gender, race, parental education, and geographic region. These subgroup estimates largely mirror those from our main analyses targeting the entire population of students as a whole. Thus, our key conclusion–that equalizing school environments would yield only modest reductions in the achievement gap between students from high- and low-poverty neighborhoods–appears consistent across a range of demographic and regional subpopulations.

Discussion

Students from high-poverty neighborhoods score significantly lower on standardized tests of academic achievement than their peers from more affluent communities. Although this disparity stems from multiple factors, unequal access to high-quality schools is widely viewed as a central cause. To investigate this possibility, the present study employed a high-dimensional approach to measure differences in school environments across high- and low-poverty neighborhoods, assessed whether these differences help explain the observed gap in achievement, and evaluated the potential for interventions aimed at equalizing school conditions to reduce disparities associated with concentrated poverty.

Using nationally representative data from more than 18,000 students in nearly 1,000 elementary schools, along with 168 distinct measures to characterize their learning environments, we document substantial differences in demographic composition between schools serving high- and low-poverty neighborhoods. Nevertheless, we find relatively limited variation in other dimensions of schools, including their resources, instruction, and effectiveness. Similarly, with novel machine learning methods tailored for high-dimensional data, we show that compositional differences are strongly associated with the achievement gap between students from high- and low-poverty neighborhoods, whereas other school characteristics account for little, if any, of the disparity in test scores. Moreover, our causal analyses suggest that equalizing the distribution of all school characteristics across neighborhoods would reduce the achievement gap by less than 10 percent, primarily through changes in student composition. This suggests that the primary drivers of disparities in achievement between high- and low-poverty neighborhoods originate outside the school environment.

These findings complicate widespread narratives that children from "poor" neighborhoods are routinely trapped in "bad" schools. They also challenge institutional resource theory, which contends that students from disadvantaged neighborhoods perform worse academically in part because they attend inferior schools that impede their learning and, by extension, widen achievement gaps. In contrast, our results suggest that the elementary schools serving students from high-poverty neighborhoods are, in general, "not so bad." Although these schools differ starkly in demographic composition–and, to a lesser extent, academic climate–their resources, instruction, and effectiveness are broadly similar to those serving less disadvantaged areas. Ill-equipped, dysfunctional, and ineffective schools certainly do exist in high-poverty communities, but our results suggest that they are not typical, nor are they prevalent enough to account for a sizable portion of the achievement disparities linked to neighborhood poverty.

Our findings align more closely with the compensatory schooling perspective (Condron et al. 2021; Downey 2020). Although we find no evidence that differences among schools attenuate achievement gaps between high- and low-poverty neighborhoods, our results suggest that elementary schools function less as engines of inequality and more like neutral institutions: they tend to reproduce, rather than amplify, disparities generated outside their walls.

Specifically, we find little evidence that schools in high- and low-poverty neighborhoods differ in ways that might produce large achievement gaps. This relative uniformity may reflect structural features of elementary education in the United States–such as curricular standardization, teacher certification requirements, building regulations, grade-level benchmarks, and accountability policies–that constrain variation across schools (Downey 2020). In addition, because instruction occurs mainly in individual classrooms, and because teacher effectiveness varies more within schools than between them (Hanushek and Rivkin 2012), students' learning opportunities may depend largely on classroom assignments rather than any broader differences in the schools serving their communities. Finally, while school sorting is often viewed as a central mechanism through which educational inequalities are generated and maintained, the modest differences we observe in most school characteristics across high- and low-poverty neighborhoods suggest that parental choices are shaped by diverse preferences, constrained options, and incomplete information, all of which may weaken the link between family resources and access to higher-quality schools (Wodtke et al. 2023).

Taken together, these institutional and behavioral constraints seem to engender a stratification process in which achievement disparities–largely established before school entry–persist with little disruption throughout the elementary years (von Hippel et al. 2018). Rather than actively producing inequality, schools appear to function more like conduits, merely channeling disparities generated by differences across families, neighborhoods, and other factors, with only modest alteration (Coleman et al. 1966).

Nevertheless, while our findings suggest that elementary schools currently function as relatively neutral institutions, this pattern is neither inevitable nor permanent. There is no universal law that prevents schools from becoming a disequalizing force. Our conclusions should therefore be considered historically and contextually bounded. Schools almost certainly produced or exacerbated inequality at certain points in the past, and they could do so again in the future. Indeed, the relative uniformity observed today at least partially reflects decades of policies aimed, however imperfectly, at promoting equality in elementary education. These include Title I of the Elementary and Secondary Education Act, the Individuals with Disabilities Education Act, school nutrition programs, English Language Learner (ELL) services, compensatory state funding policies, and the implementation of Common Core standards, among a variety of other efforts.

The current policy landscape is precarious, however. Proposals like "Project 2025," a conservative blueprint for government that recommends phasing out or converting many of these programs into unrestricted block grants (The Heritage Foundation 2023), suggest that institutional conditions could quickly change. Coupled with increasing income and wealth inequality, such changes could widen differences across schools serving high- and low-poverty communities and transform them from relatively passive conduits into more active generators of achievement disparities, deepening the chasmic gaps that already exist.

At the same time, there is also no reason that elementary schools could not be restructured to play a more compensatory role. A substantial body of research demonstrates that certain school interventions can improve outcomes for disadvantaged students and close achievement gaps (e.g., Dobbie and Fryer 2011; Hassrick et al. 2017). Our analyses, however, suggest that interventions focused solely on making the elementary schools in high-poverty neighborhoods more similar to those in low-poverty areas are unlikely to reduce achievement disparities–largely because the schools serving these communities already exhibit relatively little variation across most dimensions of quality. By the same logic, efforts that aim only to increase access to schools located outside one's residential neighborhood–for example, by expanding school choice or implementing intra-district open enrollment–are also unlikely to meaningfully narrow the achievement gaps linked to concentrated poverty.

Although our findings do not rule out the potential for schools to help close these gaps, they do suggest that effective reforms may have to transcend the goal of achieving equality of conditions or access. Instead, they may demand a more ambitious–and contentious–normative commitment to equity rather than equality, which would require tailoring the distribution of resources, supports, and practices to meet the needs of students from high-poverty neighborhoods.

The implications of this study for theory and policy should be interpreted cautiously in light of several limitations. First, although we incorporated an unusually wide range of school characteristics, it is still possible that we omitted some important indicators of quality. Schools are complex institutions, and social surveys, however sophisticated, cannot fully capture the depth and diversity of educational experiences within them.

Second, some of the measures we employed are imperfect. Teacher and administrator reports about their school's climate, infrastructure, and instructional practices may be subject to random measurement error or influenced by local reference points. For example, perceptions of what constitutes an "adequate" classroom may vary depending on location and the social context in which a school operates, potentially introducing bias into comparisons across high- and low-poverty communities. It is therefore possible that meaningful differences across communities go undetected simply because the measures used in ECLS-K lack the sensitivity to capture them. Similarly, our estimates of value-added scores are derived from relatively small samples of students within each school, which reduces their reliability and increases the influence of random noise. Nevertheless, these measures still appear to capture meaningful variation, as they rank among the strongest predictors of student achievement measured several years later, from third through fifth grades.

Third, although we used novel methods of causal inference designed to mitigate concerns about measurement and model misspecification, our identification strategy ultimately depends on the strong assumption that, conditional on a set of baseline covariates, school selection is effectively random. To bolster the plausibility of this assumption, we included a rich set of theoretically relevant controls, and we also conducted a sensitivity analysis to assess robustness to unobserved confounding. Nevertheless, the possibility of bias persists. Because some experimental and quasi-experimental studies have found larger effects of certain school factors than those estimated in observational studies like ours (e.g., Dobbie and Fryer 2011; Jackson et al. 2016), it remains possible that our results may understate the extent to which differences in schools contribute to achievement disparities between high- and low-poverty neighborhoods.

Fourth, our analysis focused exclusively on standardized test scores. These scores capture important academic skills, but they represent only a small subset of the broader competencies that schools aim to cultivate. Elementary education also attempts to foster social and emotional development, as well as knowledge of science, history, civics, and the arts. By concentrating solely on reading and math achievement, this study may obscure how schools contribute to place-based disparities in other aspects of student development.

Finally, our data on school quality are limited to the first grade, a period when school environments may be less variable due to shared curricular goals and more standardized instructional practices. However, as students progress through the education system, schools likely become more differentiated and unequal due to increased curricular tracking and more variable pedagogical priorities. These differences could exert a greater influence on educational outcomes during middle and high school, when both institutional variation and student sensitivity to their school environment may increase (Jennings et al. 2015). Future research should therefore examine how disparities in school quality evolve over time and how they contribute to achievement gaps between students from high- and low-poverty neighborhoods during later stages of schooling.

Despite these limitations, our study provides considerable evidence that the schools serving students from high- and low-poverty neighborhoods are more alike than different–and that whatever differences exist do not appreciably contribute to the large achievement disparities between these groups. Thus, schools in high-poverty neighborhoods should not be routinely targeted for blame, derision, or overhaul. Many of these schools resemble those serving more advantaged populations and appear to deliver instruction of comparable effectiveness, even as they work with student populations facing significant challenges and operate in an academic climate shaped by the strain of these circumstances. If schools are to play a more active role in reducing achievement disparities, rather than merely preventing these gaps from widening, the focus should be on equipping them with additional resources, supports, and practices that are responsive to their students' needs, not dismantling institutions that often keep their students moving forward in the face of strong headwinds.

Notes

1 Authors' calculations using the Early Childhood Longitudinal Study, Kindergarten Class of 2010–2011.

2 All sample sizes are rounded to the nearest 10 in accordance with U.S. Institute for Education Sciences (IES) disclosure risk guidelines.

3 The data used in this analysis are derived from restricted-access files provided by IES under contractual arrangements that preclude the authors from distributing them. Instructions for accessing these data are provided as part of the study's reproducibility package.

References

Abdulkadiroglu, Atila, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters. 2020. "Do Parents Value School Effectiveness?" *American Economic Review* 110:1502–39. https://doi.org/10.1257/aer.20172040

Ainsworth, James W. 2002. "Why Does It Take a Village? The Mediation of Neighborhood Effects on Educational Achievement." *Social Forces* 81(1):117–52. https://doi.org/10.1353/sof.2002.0038

Angrist, Joshua D. and Kevin Lang. 2004. Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review* 94(5):1613–34. https://doi.org/10.1257/0002828043052169

Arum, Richard. 2000. "Schools and Communities: Ecological and Institutional Dimensions." *Annual Review of Sociology* 26:395–418. https://doi.org/10.1146/annurev.soc.26.1.395

Auld, M. Christopher and Nirmal Sidhu. 2005. "Schooling, Cognitive Ability and Health." *Health Economics* 14:1019–34. https://doi.org/10.1002/hec.1050

Bifulco, Robert and Helen F. Ladd. 2006. "The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina." *Education Finance and Policy* 1(1):50–90. https://doi.org/10.1162/edfp.2006.1.1.50

Billingham, Chase M. and Matthew O. Hunt. 2016. "School Racial Composition and Parental Choice: New Evidence on the Preferences of White Parents in the United States." *Sociology of Education* 89(2):99–117. https://doi.org/10.1177/0038040716635718

Borman, Geoffrey D. and N. Maritza Dowling. 2008. "Teacher Attrition and Retention: A Meta-Analytic and Narrative Review of the Research." *Review of Educational Research* 78(3):367–409. https://doi.org/10.3102/0034654308321455

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2005. "Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students." *The American Economic Review* 95(2):166–71. https://doi.org/10.1257/000282805774669628

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45(1):5–32. https://doi.org/10.1023/A:1010933404324

Bryk, Anthony S., Penny B. Sebring, Elaine Allensworth, John Q. Easton, and Stuart Luppescu. 2010. *Organizing Schools for Improvement: Lessons from Chicago.* University of Chicago Press. https://doi.org/10.7208/chicago/9780226078014.001.0001

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. "Double/Debiased Machine Learning for Treatment and Structural Parameters." *The Econometrics Journal* 21(1):C1–68. https://doi.org/10.1111/ectj.12097

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014a. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104(9):2593–632. https://doi.org/10.1257/aer.104.9.2593

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014b. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104(9):2633–679. https://doi.org/10.1257/aer.104.9.2633

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2010. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." *Journal of Human Resources* 45(3):655–81. https://doi.org/10.1353/jhr.2010.0023

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic Weinfeld, and Robert L. York. 1966. "Equality of Educational Opportunity." Technical Report 1066-5684, U.S. Department of Health, Education, and Welfare, Office of Education.

Condron, Dennis J., Dougals B. Downey, and Megan Kuhfeld. 2021. "Schools as Refractors: Comparing Summertime and School-Year Skill Inequality Trajectories." *Sociology of Education* 94(4):316–40. https://doi.org/10.1177/00380407211041542

Crosnoe, R., F. Morrison, M. Burchinal, R. Pianta, D. Keating, S. L. Friedman, and K. A. Clarke-Stewart. 2010. "Instruction, Teacher-Student Relations, Math Achievement Trajectories in Elementary School." *Journal of Educational Psychology* 102(2):407–17. https://doi.org/10.1037/a0017762

Cullen, Julie B., Brian A. Jacob, and Steven D. Levitt. 2005. "The Impact of School Choice on Student Outcomes: An analysis of the Chicago Public Schools." *Journal of Public Economics* 89(5–6):729–60. https://doi.org/10.1016/j.jpubeco.2004.05.001

DeLuca, Stefanie and Peter Rosenblatt. 2010. "Does Moving to Better Neighborhoods Lead to Better Schooling Opportunities?" *Teachers College Record* 112:1143–491. https://doi.org/10.1177/016146811011200504

Devine, John. 1996. *Maximum Security: The Culture of Violence in Inner-City Schools*. Chicago, IL: University of Chicago Press.

Dobbie, Will and Roland G. J. Fryer. 2011. "Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone." *American Economic Journal: Applied Economics* 3(3):158–87. https://doi.org/10.1257/app.3.3.158

Downey, Douglas B. 2020. *How Schools Really Matter: Why Our Assumption about Schools and Inequality Is Mostly Wrong*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226733364.001.0001

Downey, Douglas B., David M. Quinn, and Melissa Alcaraz. 2019. "The Distribution of School Quality: Do Schools Serving Mostly White and High-SES Children Produce the Most Learning?" *Sociology of Education* 92(4):386–403. https://doi.org/10.1177/0038040719870683

Duncan, Greg J. and Richard J. Murnane. 2011. *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. Russell Sage Foundation, New York.

Emslander, Valentin, Jessica Levy, Ronny Scherer, and Antoine Fischbach. 2022. "Value-Added Scores Show Limited Stability Over Time in Primary School." *PLoS ONE* 17(12):e0279255. https://doi.org/10.1371/journal.pone.0279255

Friedman, Jerome H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." *Annals of Statistics* 29:1189–232. https://doi.org/10.1214/aos/1013203451

GeoLytics. 2012. CensusCD Neighborhood Change Database (NCDB): 1970-2010 tract data, version 2.1.

Hanselman, Paul. 2018. "Do School Learning Opportunities Compound or Compensate for Background Inequalities? Evidence from the Case of Assignment to Effective Teachers." *Sociology of Education* 91(2):132–58. https://doi.org/10.1177/0038040718761127

Hanselman, Paul and Jeremy E. Fiel. 2017. "School Opportunity Hoarding? Racial Segregation and Access to High Growth Schools." *Social Forces* 95(3):1077–104. https://doi.org/10.1093/sf/sow088

Hanushek, Eric A. and Steven G. Rivkin. 2012. "The Distribution of Teacher Quality and Implications for Policy." *Annual Review of Economics* 4(1):131–57. https://doi.org/10.1146/annurev-economics-080511-111001

Hassrick, Elizabeth M., Lisa Rosen, and Stephen Raudenbush. 2017. *The Ambitious Elementary School: Its Conception, Design, and Implications for Educational Equality*. Chicago: University of Chicago Press.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24(3):411–82. https://doi.org/10.3386/w12006

Hines, Oliver, Oliver Dukes, Karla Diaz-Ordaz, and Stijin Vansteelandt. 2022. "Demystifying Statistical Learning Based on Efficient Influence Functions." *The American Statistician* 76(3):292–304. https://doi.org/10.1080/00031305.2021.2021984

Hong, Shangzhi, Yuqi Sun, Hanying Li, and Henry S. Lynn. 2020. "Multiple Imputation Using Chained Random Forests." *arXiv preprint*. https://doi.org/10.48550/arXiv.2004.14823

Isenberg, Eric, Jeffrey Max, Ohilip Gleason, and Jonah Deutsch. 2022. "Do Low-Income Students Have Equal Access to Effective Teachers?" *Educational Evaluation and Policy Analysis* 44(2):234–56. https://doi.org/10.3102/01623737211040511

Jackson, C. Kirabo. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes." *Journal of Political Economy* 126(5):2072–107. https://doi.org/10.1086/699018

Jackson, C. Kirabo, Rucker C. Johnson, and Claudia Persico. 2016. "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms." *The Quarterly Journal of Economics* 131(1):157–218. https://doi.org/10.1093/qje/qjv036

Jencks, Christopher and Susan E. Mayer. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood." Pp. 111–86 in *Inner-City Poverty in the United States*, edited by Laurence Lynn and Michael G. H. McGeary, chapt. 4. Washington, DC: National Academy Press.

Jennings, Jennifer L., David Deming, Christopher Jencks, Maya Lopuch, and Beth E. Schueler. 2015. "Do Differences in School Quality Matter More Than We Thought? New Evidence on Educational Opportunity in the Twenty-First Century." *Sociology of Education* 88(1):56–82. https://doi.org/10.1177/0038040714562006

Johnson, Odis. 2012. "A Systematic Review of Neighborhood and Institutional Relationships Related to Education." *Education and Urban Society* 44(4):477–511. https://doi.org/10.1177/0013124510392779

Kahlenberg, Richard D. 2001. *All Together Now: Creating Middle Class Schools through Public School Choice*. Washington, DC: Brookings.

Kozol, Jonathan. 2012. *Savage Inequalities: Children in America's Schools*. Crown.

Ladd, Helen and Susanna Loeb. 2013. "The Challenges of Measuring School Quality: Implications for Educational Equity." Pp. 22–55 in *Education, Justice, and Democracy*, edited by Rob Reich and Danielle Allen. Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226012933.003.0002

Lauen, Douglas L. and S. Michael Gaddis. 2013. "Exposure to Classroom Poverty and Test Score Achievement: Contextual Effects or Selection?" *American Journal of Sociology* 118(4):943–79. https://doi.org/10.1086/668408

Lundberg, Ian. 2024. "The Gap-Closing Estimand: A Causal Approach to Study Interventions That Close Disparities across Social Categories." *Sociological Methods & Research* 53(2):507–70. https://doi.org/10.1177/00491241211055769

Najarian, Michelle, Karen Tourangeau, Christine Nord, Kathleen Wallner-Allen, and Nancy Vaden-Kierman. 2019. "Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011), Third-Grade, Fourth-Grade, and Fifth-Grade Psychometric Report." Technical Report NCES 2020-123, National Center for Education Statistics.

National Reading Panel. 2000. *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction*. National Institute of Child Health and Human Development.

National Research Council. 2000. *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: The National Academies Press.

Nolan, Kathleen. 2011. *Police in the Hallways: Discipline in an Urban High School*. Minneapolis, MN: University of Minnesota Press. https://doi.org/10.5749/minnesota/9780816675524.001.0001

Opacic, Aleksei, Lai Wei, and Xiang Zhou. 2025. "Disparity Analysis: A Tale of Two Approaches." *Journal of the Royal Statistical Society Series A: Statistics in Society*. E-pub ahead of print. https://doi.org/10.1093/jrsssa/qnaf008

Owens, Ann. 2010. Neighborhoods and Schools as Competing and Reinforcing Contexts for Educational Attainment." *Sociology of Education* 83(4):287–311. https://doi.org/10.1177/0038040710383519

Owens, Ann and Jennifer Candipan. 2019. "Social and Spatial Inequalities of Educational Opportunity: A Portrait of Schools Serving High- and Low-Income Neighbourhoods in US Metropolitan Areas." *Urban Studies* 56(15):3178–97. https://doi.org/10.1177/0042098018815049

Pressley, Tim, Richard Allington, and Michael Pressley. 2023. *Reading Instruction That Works: The Case for Balanced Teaching*. Guilford Publications.

Raudenbush, Stephen W. 2008. "Advancing Educational Policy by Advancing Research on Instruction." *American Educational Research Journal* 45(1):206–30. https://doi.org/10.3102/0002831207312905

Raudenbush, Stephen W. and Robert D. Eschmann. 2015. "Does Schooling Increase or Reduce Social Inequality?" *Annual Review of Sociology* 41:443–70. https://doi.org/10.1146/annurev-soc-071913-043406

Rich, Peter, Jennifer Candipan, and Ann Owens. 2021. "Segregated Neighborhoods, Segregated Schools: Do Charters Break a Stubborn Link?" *Demography* 58(2):471–98. https://doi.org/10.1215/00703370-9000820

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons. https://doi.org/10.1002/9780470316696

Rumberger, Russell W. and Gregory J. Palardy. 2005. "Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School." *Teachers College Record* 107(9):1999–2045. https://doi.org/10.1177/016146810510700905

Sanbonmatsu, Lisa, Jeffrey R. Kling, Greg J. Duncan, and Jeanne Brooks-Gunn. 2006. "Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment." *The Journal of Human Resources* 41(4):649–91. https://doi.org/10.3368/jhr.XLI.4.649

Saporito, Saporito and Deenesh Sohoni. 2007. "Mapping Educational Inequality: Concentrations of Poverty among Poor and Minority Students in Public Schools." *Social Forces* 85:1227–53. https://doi.org/10.1353/sof.2007.0055

Semuels, Alana. 2016. "Good School, Rich School; Bad School, Poor School: The Inequality at the Heart of America's Education System." *The Atlantic*.

Sharkey, Patrick, Amy E. Schwartz, Ingrid Ellen, and Johanna Lacoe. 2014. "High Stakes in the Classroom, High Stakes on the Street: The Effects of Community Violence on Student's Standardized Test Performance." *Sociological Science* 1:199–220. https://doi.org/10.15195/v1.a14

The Heritage Foundation. 2023. *Mandate for Leadership: The Conservative Promise*. Project 2025, Washington, D.C.

Thompson, Marissa E. 2024. "The Effect of Academic Outcomes, Equity, and Student Demographics on Parental Preferences for Schools: Evidence from a Survey Experiment." *Social Forces* 103(2):730–55. https://doi.org/10.1093/sf/soae101

Tomlinson, Carol A. 2014. *The Differentiated Classroom: Responding to the Needs of All Learners*. ASCD.

Tourangeau, Karen, Christine Nord, Thanh Le, Kathleen Wallner-Allen, Mary C. Hagedorn, John Leggitt, and Michelle Najarian. 2015. *Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011)*. U.S. Department of Education, National Center for Education Statistics, Washington, DC.

Van der Laan, Mark J., Eric C. Polley, and Alan E. Hubbard. 2007. "Super Learner." *Statistical Applications in Genetics and Molecular Biology* 6(1). https://doi.org/10.2202/1544-6115.1309

von Hippel, Paul T. 2009. "Achievement, Learning, and Seasonal Impact as Measures of School Effectiveness: It's Better to Be Valid Than Reliable." *School Effectiveness and School Improvement* 20:187–213. https://doi.org/10.1080/09243450902883888

Von Hippel, Paul T., Joseph Workman, and Douglas B. Downey. 2018. "Inequality in Reading and Math Skills Forms Mainly Before Kindergarten: A Replication, and Partial Correction, of 'Are Schools the Great Equalizer?'." *Sociology of Education* 91(4):323–57. https://doi.org/10.1177/0038040718801760

Wager, Stefan and Guenther Walther. 2015. "Adaptive Concentration of Regression Trees, with Application to Random Forests." arXiv preprint. arXiv:1503.06388. https://doi.org/10.48550/arXiv.1503.06388

Wang, Ming-Te and Jessica L. Degol. 2016. "School Climate: A Review of the Construct, Measurement, and Impact on Student Outcomes." *Educational Psychology Review* 28:315–52. https://doi.org/10.1007/s10648-015-9319-1

Willms, J. Douglas. 2010. "School Composition and Contextual Effects on Student Outcomes." *Teachers College Record* 112:1008–38. https://doi.org/10.1177/016146811011200408

Wodtke, Geoffrey T. and Matthew Parbst. 2017. "Neighborhoods, Schools, and Academic Achievement: A Formal Mediation Analysis of Contextual Effects on Reading and Mathematics Abilities." *Demography* 54(5):1653–76. https://doi.org/10.1007/s13524-017-0603-1

Wodtke, Geoffrey T., Sagi Ramaj, and Jared Schachner. 2022. "Toxic Neighborhoods: The Effects of Concentrated Poverty and Environmental Lead Contamination on Early

Childhood Development." *Demography* 59(4):1275–98. https://doi.org/10.1215/00703370-10047481

Wodtke, Geoffrey T., Ugur Yildirim, David J. Harding, and Felix Elwert. 2023. "Are Neighborhood Effects Explained by Differences in School Quality?" *American Journal of Sociology* 128(5):1472–528. https://doi.org/10.1086/724279

Zhang, Tong and Bin Yu. 2005. "Boosting with Early Stopping: Convergence and Consistency." *Annals of Statistics* 33(4):1538–579. https://doi.org/10.1214/009053605000000255

Acknowledgments: The authors thank Steve Raudenbush, Guanglei Hong, Ariel Kalil, Steven Durlauf, Eric Grodsky, and Lucienne Disch for helpful comments and discussions. This project was supported by a grant from the U.S. National Science Foundation (No. 2015613) and by the James M. and Cathleen D. Stone Foundation. We used ChatGPT, version 4o, for help with copyediting the manuscript and debugging R scripts. Responsibility for all content rests solely with the authors.

Geoffrey T. Wodtke: Department of Sociology, University of Chicago.
E-mail: wodtke@uchicago.edu.
Kailey White: Crime Lab and Education Lab, University of Chicago.
E-mail: kwhite10@uchicago.edu.
Xiang Zhou: Department of Sociology, Harvard University.
E-mail: xiang_zhou@fas.harvard.edu.