

The Geography of Mathematical (Dis)Advantage: An Application of Multilevel Simultaneous Autoregressive (MSAR) Models to Public Data in Education Research

Manuel S. González Canché 

The University of Pennsylvania

Research has shown that mathematical proficiency gaps are related to students' and schools' indicators of poverty, with fewer studies on neighborhood effects on achievement gaps. Although this literature has accounted for students' nesting within schools, so far, methodological constraints have not allowed researchers to formally account for multilevel and spatial effects. I contribute to this discussion by simultaneously considering test-takers' own socioeconomic standing and the impact of their nesting schools and neighborhood structures. Multilevel simultaneous autoregressive (MSAR) models and population-level data of 2.09 million test-takers, whose standardized performances were measured at Grades 3–8 in New York State, revealed the presence of geography of mathematical (dis)advantage. Because mathematical performance is spatially dependent across schools and neighborhoods, moving forward, applied researchers should rely on MSAR to account for sources of spatially driven bias that cannot be handled with multilevel models alone. Full replication code and data are provided at <https://cutt.ly/N4zRstL>.

Keywords: *achievement gap, data reproducibility, data science applications, econometric analysis, elementary schools, geographical network modeling, geography of mathematical (dis)advantage, hierarchical linear modeling, interactive data visualization, machine learning applications, multicollinearity in spatial or place-based indicators, multilevel simultaneous autoregressive (MSAR), observational research, performance assessment, regression analyses, research methodology, secondary data analysis, social class, social context, social stratification, spatial dependence, spatial socioeconometric modeling*

Introduction

Mathematical proficiency is a gateway to academic and professional success. Mathematical proficiency predicts reading achievement more strongly than early reading performance (Duncan et al., 2007; Farran et al., 2006; Lerkkanen et al., 2005; Pagani et al., 2010) and is a better predictor of later school success than many other academic indicators (Duncan et al., 2007; Wright et al., 2020). Relatedly, given the increasing relevance of science, technology, engineering, and mathematics (STEM) in the globalized economy (González Canché, 2017; Pearman, 2019; Rothwell, 2013), efforts to strengthen students' mathematical proficiency levels as early as possible would also affect students' prospects for economic and professional success (Aber et al., 2013; Pearman, 2019).

Mathematical proficiency levels, however, are not homogeneously distributed across student populations. Social-class differences have been consistently documented to influence students' prospects of mathematical performance as early as preschool and kindergarten (Case et al., 1999;

Denton & West, 2002; Starkey et al., 2004). Early competence in mathematics has also been shown to lead to later competence in mathematics, when “children who start ahead in mathematics generally stay ahead” (Siegler et al., 2012, p. 691). Notably, because the effects of mathematical proficiency also lead to success in other academic, economic, and professional areas, early success in mathematics leads to positive spillovers, where success in mathematics begets success in mathematics and in other academic content, thus increasing the possibility of subsequent success in life (Aber et al., 2013). Spillover effects based on mathematical proficiency (or the lack of thereof) can be negative as well, of course, and failure to reach mathematical proficiency in one grade may imply a negative vicious cycle in the learning process of subsequent mathematical topics (Duncan et al., 2007; Siegler et al., 2012; Stevenson & Newman, 1986).

The mechanisms behind the idea that students who start behind in mathematical proficiency tend to remain behind (Siegler et al., 2012) may be explained by the cumulative nature of mathematical knowledge—that is, knowledge



gaps in arithmetic foundations pose significant barriers to the learning of derivatives and integrals (Fogarty et al., 2018). In other words, considering this inherently cumulative nature of mathematical knowledge, the negative spill-overs of not reaching, or barely reaching, proficiency levels in one grade may impose greater challenges in mastering mathematical content in subsequent grades (Clotfelter et al., 2009). From this point of view, the fact that grade-level proficiency levels may decrease as students advance through grades (e.g., see National Assessment of Educational Progress [NAEP], 2020, 2022) may in part be explained by this cumulative nature but may also be influenced by students' socioeconomic standing (Case et al., 1999; Denton & West, 2002; Siegler et al., 2012; Starkey et al., 2004). In this respect, although summary indicators of mathematical proficiency are useful in depicting how mathematical proficiency may decrease across grades (Mervosh & Wu, 2022), it remains unclear how these gaps in proficiency levels are expected to vary when simultaneously considering the impact of poverty indicators of students, schools, and even the neighborhoods where those schools are nested or located—that is, it remains much less clear in the academic literature how students from economically disadvantaged backgrounds may fare when they attend schools located in high-poverty neighborhoods or how these same students may be affected when attending schools located in less economically challenged neighborhoods.

Building from research on academic achievement gaps (see Clotfelter et al., 2009; Hanushek & Rivkin, 2006; Phillips & Chin, 2004; Reardon & Galindo, 2009), it is clear that as early as kindergarten, children's knowledge and skills differ in relation to their families' socioeconomic indicators (Denton & West, 2002; Duncan et al., 2007; Siegler et al., 2012; Starkey et al., 2004). Although less prevalent in this literature, research has also shown an important association between achievement gaps and place-based socioeconomic disadvantages in the form of neighborhood effects (Considine & Zappalà, 2002; Pearman, 2019). Although this literature has accounted for students' nesting within schools by relying on multilevel models, so far, methodological constraints have not allowed researchers to formally account for multi-level and spatial or place-based effects on students' performance while also accounting for these students' own socioeconomic well-being. To address this gap in the literature, this study aims to offer more nuanced estimates of the magnitudes of expected proficiency changes in mathematics across grade levels by *simultaneously* considering students' socioeconomic standing as well as their school-level contexts and place-based poverty levels.

Considering the impact of socioeconomic contexts in mathematical proficiency is relevant because, although the achievement-gap literature has consistently found that as indicators of poverty in a child's life increase, their academic performance will tend to decline (Considine & Zappalà,

2002; Pearman, 2019), it remains unclear how low-income students attending schools located in wealthier zones perform in mathematical proficiency and how these gaps vary across grades. Similarly, this study assesses how economically disadvantaged students' performances compare to their non-economically disadvantaged peers attending low-income schools. In sum, this study contributes to research on mathematical competence by modeling changes in mathematical proficiency across grades while simultaneously considering students' own socioeconomic standing and the impact of their nested school and neighborhood structures.

To meet the goal of this study, I relied on standardized performance records of more than two million test-takers (see Table 3) whose individual mathematical proficiency¹ was reported at the grade level (i.e., not individual records but records aggregated per grade within 3,484 public schools) in Grades 3–8 in the state of New York in the academic years 2017–2018 (3,335 public schools) and 2018–2019 (3,372 public schools), for a total of 3,484 public schools. This administrative data set was complemented with American Community Survey (ACS) indicators and TIGER/Line Shapefiles from the U.S. Census Bureau to build a completely open-access data set (access to replication code and data for all models are available at <https://cutt.ly/N4zRstL> and in González Canché (2023a)). That accounts for students', schools', and neighborhoods' economic standing to address the following questions, which are captured in the conceptual map in Figure 1:

1. What is the expected change in mathematical proficiency levels observed among students in Grades 3–8? (addressed with Model “Aggregate” in Table 4)
2. How do these proficiency levels vary when considering students' own levels of economic disadvantage (i.e., receiving at least one form of public assistance²) while simultaneously considering their school-level sociodemographic and economic nesting structures? (addressed with Models 2–7 in Table 4)

This differentiation in mathematical performance by economically disadvantaged status was possible because the administrative records reported by the New York State Education Department (NYSED, 2020b) offered disaggregated proficiency levels in mathematics across schools and within grades by an indicator capturing whether students belonged to the “economically disadvantaged” class—that is, by whether students' families participated in any form of public assistance (see Table 1).

3. How do performance gaps in mathematical proficiency vary for economically disadvantaged students attending schools located in high-poverty neighborhoods, and how do these gaps compare to those realized by their non-economically disadvantaged peers

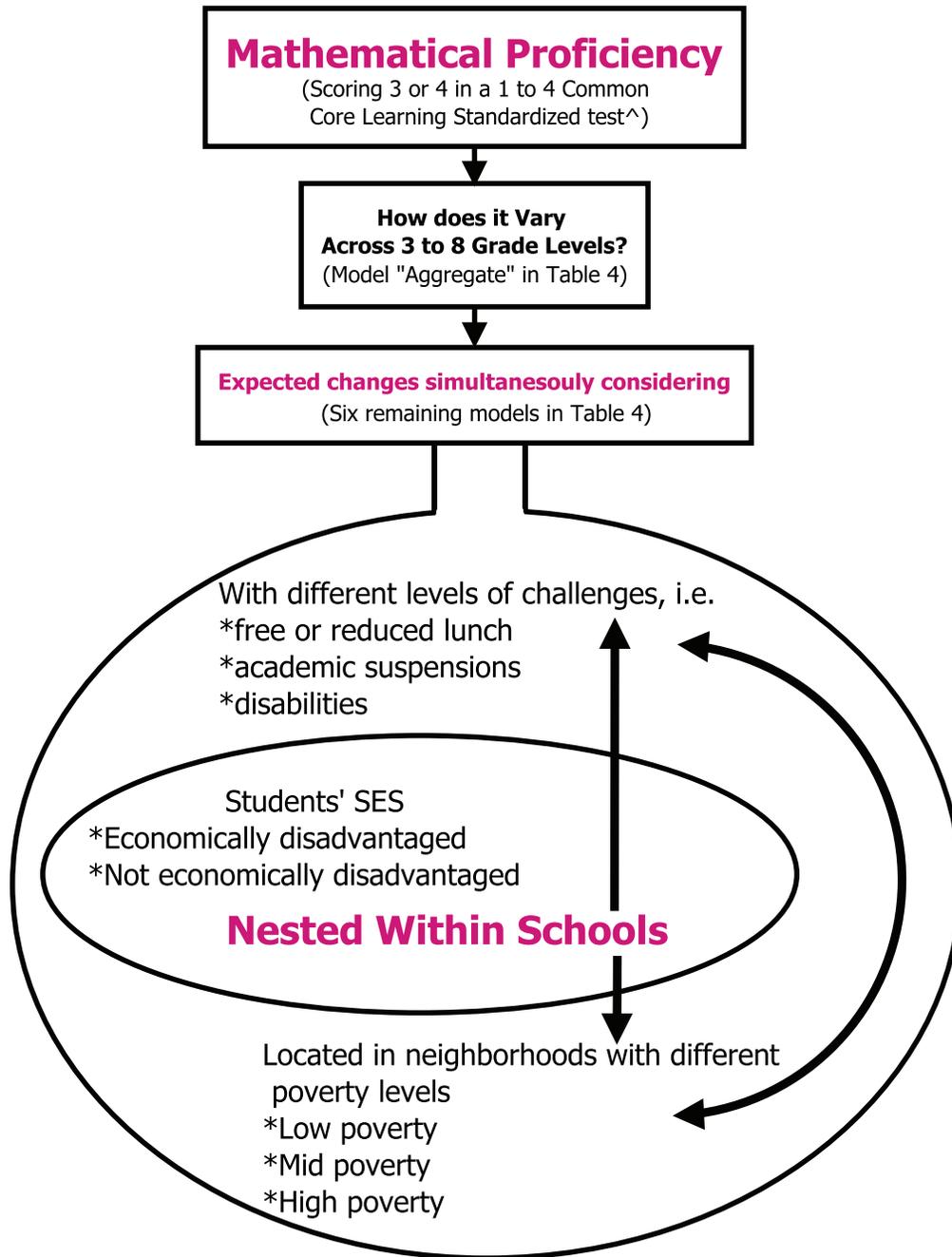


FIGURE 1. *Conceptual representation of modeling framework.*

^Tests administered at public elementary and middle schools in the state of New York (NYSED, 2020a, 2020b).

- attending schools located in the least economically challenged neighborhoods? (addressed with Table 3 and Models 2–7 in Table 4)
4. How do other place-based sociodemographic and economic indicators influence the variation of mathematical proficiency? (addressed with Models 1–7 in Table 4)

To address these questions, this study relied on methodological advancements that are yet to be widely implemented in education research (see Dong & Harris, 2015; Dong et al., 2015). The main method of analysis was MSAR models that accounted for the nested structure of the data (i.e., grades nested within schools, schools nested within regions) while also testing for whether spatial dependence was present at

TABLE 1
Example of Data Structure

ENTITY_CD	ASSESSMENT_NAME	SUBGROUP_NAME*	**PER_PROF
10100010014	MATH3	Economically_Disadvantaged	25
10100010014	MATH4	Economically_Disadvantaged	44
10100010014	MATH5	Economically_Disadvantaged	23
10100010014	MATH5	Not_Economically_Disadvantaged	83
10100010014	MATH3	Not_Economically_Disadvantaged	67
10100010016	MATH4	Not_Economically_Disadvantaged	72
10100010016	MATH5	Economically_Disadvantaged	15
10100010016	MATH3	Economically_Disadvantaged	38
10100010016	MATH4	Economically_Disadvantaged	9
10100010016	MATH3	Not_Economically_Disadvantaged	47
10100010016	MATH4	Not_Economically_Disadvantaged	54
10100010018	MATH5	Not_Economically_Disadvantaged	39
10100010018	MATH5	Economically_Disadvantaged	20
10100010018	MATH3	Economically_Disadvantaged	13
10100010018	MATH4	Economically_Disadvantaged	19
10100010018	MATH3	Not_Economically_Disadvantaged	25
10100010018	MATH4	Not_Economically_Disadvantaged	27
10100010019	MATH5	Not_Economically_Disadvantaged	25
10100010019	MATH5	Economically_Disadvantaged	3
10100010014	MATH3	Economically_Disadvantaged	21

*As reported by NYSED (2020b).

**Outcome of interest. Data retrieved from NYSED (2020b).

the school and region levels and, if so, controlling for those sources of variation. The results indicate that mathematical performance is spatially dependent across schools and neighborhoods. Accordingly, moving forward, applied researchers should consider relying on MSAR to account for sources of spatially driven bias that cannot be handled with multilevel models alone. From this perspective, the distribution of the replication code and data set with this submission is of particular importance for researchers who may want to have access to a fully reproducible example.

The remainder of the paper is organized as follows. It starts with a review of the literature, including the most prevalent strategies that researchers have employed to estimate the variation of mathematical proficiency, considering individual-level attributes, such as ethnicity and socioeconomic status, as well as school and neighborhood effects. It then discusses the conceptual lenses guiding the variable selection as well as indicating hypothetical relationships expected to be observed when considering the literature reviewed and the conceptual framework. Subsequently, the data and methods sections formally discuss MSAR and the relevance of relying on machine learning methods to select predictor and control indicators that remain relevant in the explanation of the main outcome of interest (i.e., mathematical proficiency). This feature or variable selection is particularly important when dealing with place-based indicators

that, as described in the conceptual framework section, may be highly correlated. Then, findings are presented, and the paper closes with implications, limitations, and further areas of inquiry that, although important, go beyond the scope of this study (for example, measuring the impact of the COVID-19 pandemic on the academic performance of economically disadvantaged students).

Literature Review

Studies on academic achievement gaps in elementary and middle school education have followed two major trends. The first has focused on studying ethnic and racial gaps to see whether performance gaps have narrowed or expanded among students belonging to different racial/ethnic groups *within grades*, typically using White students' academic performance as the *reference or comparison group* to measure achievement gaps (Clotfelter et al., 2009; Hanushek & Rivkin, 2006; Phillips & Chin, 2004; Reardon & Galindo, 2009). For example, Hanushek and Rivkin (2006) showed that the achievement gap in Grade 1 between Black and White students was 12.3 points, favoring the latter. They then indicated that by Grade 3, this performance gap between Black and White students had increased to 18 points (Hanushek and Rivkin, 2006, p. 10). Similarly, using departures from White students' performance, Clotfelter et al. (2009) showed that Black-White disparities

were the largest among all racial group comparisons in their analytic sample (i.e., comparing White students' performance with Hispanic, Asian, and American Indian students' performance). They also showed that in Grade 3, "more than a quarter of [B]lack [students] tested below the 10th percentile of the White distribution" (Clotfelter et al., 2009, p. 408). This result was also observed when using eighth graders' scores. A consistent modeling strategy employed in these studies is that performance gaps have been measured within grades and between groups.

The second major trend in the achievement gap literature consists of analyzing the association between achievement and socioeconomic disadvantage (Considine & Zappalà, 2002; Pearman, 2019). Indicators of economic distress have been measured at individual students' family levels as well as at school and neighborhood levels, with results consistently corroborating the negative association between economic disadvantage and academic achievement (Carlson & Cowen, 2015; Pearman, 2019). More to the point, studies have consistently shown that achievement gaps have been present as early as kindergarten (Case et al., 1999; Denton & West, 2002; Starkey et al., 2004), wherein children who start behind remain behind (Duncan et al., 2007; Siegler et al., 2012; Stevenson & Newman, 1986).

With respect to neighborhood and school effects, authors have analyzed these effects on students by standardizing students' performance across grades—like our data set analyzed herein. For example, Carlson and Cowen (2015) used data from all students attending the Milwaukee Public Schools between 2006–2007 and 2010–2011 who participated in the Wisconsin Knowledge and Concepts Examination. Then they estimated school and neighborhood contributions to 1-year gains in student test scores for elementary and middle school without considering performances within students' grades, as described above. More recently, Pearman (2019) estimated the effects of neighborhood poverty across children up to age 12. Specifically, Pearman tested the hypothesis that exposure to high-poverty neighborhoods influenced students' mathematical proficiency over and above their individual and school characteristics. Although Pearman did not formally rely on spatial econometric models, this author did present sensitivity tests that were robust to unobserved residual confounding. This remarkable contribution informed this study's goal of formally measuring and controlling for spatial dependence while also measuring how heightened poverty indicators may interact with students' and schools' sociodemographic and socioeconomic indicators. Having said this, note that in this neighborhood and school effects line of research, the goal has been to measure the overarching effect of poverty on academic performance rather than to measure the effect of poverty on mathematical proficiency changes across groups and grade levels, as conducted in this study.

Together, the literature review indicates that studies have yet to estimate the effect of poverty on mathematical performance

across grades by moving beyond purely descriptive estimates, as provided by the NAEP (2020, 2022), where, as of 2018–2019, the gaps between Grades 4–8 reached 7 percentage points nationwide. In the case of New York State, this observed gap was 3 percentage points (NAEP, 2020). Moreover, studies have not yet formally tested for spatial dependence or an autocorrelation effect that may affect the outcome of interest while simultaneously modeling the socioeconomic circumstances of students who attend schools nested within neighborhoods with differing poverty levels—a modeling approach implemented in this study.

School Context and Place-Based Lenses

The literature on school and neighborhood effects clearly states that school context influences students' academic performance (Carlson & Cowen, 2015; Goldstein, 1997; González Canché, 2019, 2022, 2023b; Konstantopoulos & Borman, 2011) and that living in high-poverty neighborhoods is negatively associated with mathematical achievement (Anderson et al., 2014; Pearman, 2019), both of which pose significant barriers for children's upward social mobility (Chetty et al., 2016, 2020). Carlson and Cowen further noted that "analyses of neighborhood contributions to academic achievement that do not account for a student's schooling context may result in biased estimates of neighborhood contributions" (2015, p. 49). Heeding their cautionary note, and as indicated in the conceptual map shown in Figure 1, this study accounts for school-context indicators (e.g., percentages of English language learners, students with disabilities, and students receiving free and reduced-price lunch³), schools' neighborhood socioeconomic and demographic characteristics (e.g., poverty and crime, academic attainment levels) employed in previous studies (see Carlson & Cowen, 2015; González Canché, 2022, 2023b; Jargowsky & El Komi, 2011; Pearman, 2019), and measurement of mathematics proficiency levels conditional on test-takers' economic disadvantage standing.

More specifically, the analytic framework implemented in this study pays close attention to measuring how schools' neighborhood poverty levels may affect students' mathematical proficiency rates across grades, after considering students' own socioeconomic standing and the nesting structures at the school and neighborhood levels (see Figure 1). This emphasis on neighborhood poverty is important because, as Pearman (2019) stated, regardless of whether children themselves live in low-income families, when these students live in high-poverty neighbors or attend schools located in high-poverty zones, they may increasingly find themselves surrounded by poverty. That is, students walking in those streets may observe homelessness, crime, and unemployment firsthand, even if in their homes poverty is not necessarily present (González Canché, 2022, 2023b). Pearman validated this statement by showing that 30% of all children, regardless of

their own poverty status, live in neighborhoods with poverty rates greater than 20%. In the state of New York, the panorama depicted by Pearman is grimmer: Internal Revenue Service (IRS, 2020) data show that 77% of all students included in this study attended schools where 20% of the households lived at least \$7,500 below the state's poverty line.⁴ Students attending schools located in high-poverty areas, however, are not randomly distributed; instead, their attendance is explained by their families' socioeconomic standing (González Canché, 2023b). That is, whereas 89% of students classified as economically disadvantaged⁵ attend schools within neighborhoods that have poverty levels of at least 20%, their non-economically disadvantaged peers are about 30 percentage points less likely to attend these types of schools (61.6%). Considering these statistics, I was compelled to rely on spatially driven conceptual frameworks to inform this study's variable selection and methods.

Conceptual Lenses: Neighborhood Effects and the Geography of Disadvantage

Given the relevance of place and space for this study, let us formally describe the notions of concentrated disadvantage (Elijah, 1990; González Canché, 2022, 2023b; Jargowsky & Tursi, 2015) and geography of opportunity (Pastor, 2001; Tate, 2008) or disadvantage (Pacione, 1997). Attending schools located in high-poverty areas may influence students' academic achievement not only because of the systematic heightened exposure to crime, housing instability, and vulnerable family composition typically associated with poverty levels (Carlson & Cowen, 2015; Jargowsky & El Komi, 2011; Pearman, 2019) but also because of academic achievement's relationship with school contextual indicators, such as teacher retention and overall school socioeconomic and demographic composition.

The notions of concentrated disadvantage and geography of disopportunity or disadvantage indicate that students' common exposure to spatially contextualized situations shapes their opportunities for upward social mobility (Chetty et al., 2016, 2020; González Canché, 2022, 2023b) by comprehensively affecting their cultural, racial, and socioeconomic identities (Rosen, 1985). Growing up in lower income neighborhoods where the vast majority of individuals do not finish high school or enter college translates into reduced opportunities to learn about careers that require college education, which may not only shape students' aspirations but also affect their employment prospects, salary levels, and exposure to crime and incarceration rates (Chetty et al., 2020; González Canché, 2022, 2023b; Iriti et al., 2018). Furthermore, even when students growing up in these types of neighborhoods observe a few individuals with some college or college degrees, they may form a belief that exposure to college does not help overcome difficulties in finding employment or increasing earnings (Rosen, 1985; Weicher, 1979), which may reinforce their

negative views about the long-term benefits associated with a college education (Iriti et al., 2018). On the other hand, growing up in more affluent neighborhoods, either since birth or after living in high-poverty housing at younger ages, has been found to causally affect individuals' prospects of upward income mobility (Chetty et al., 2016, 2020). For individuals experiencing life in more affluent neighborhoods, obtaining college degrees and securing employment are typically normalized, which translates into greater certainty about rates of return associated with investing in education and the expectation of success derived from college attendance (González Canché, 2022, 2023b). This certainty is obtained not only at home but also through community networks and resources (Iriti et al., 2018).

These social and cultural dynamics help explain why students in wealthier neighborhoods tend to realize higher proficiency rates (Pearman, 2019). These dynamics also highlight the need to control for these differences in observed and unobserved factors that may influence mathematical performance over and above student and school-based indicators. From this perspective, the nesting of schools within neighborhoods and grades within schools, along with the inclusion of indicators measured at all these levels, is an important requirement to offer estimates that, as comprehensively as possible, may capture the influence of multilevel and spatially contextualized factors affecting changes in mathematical proficiency across grades. Moreover, and to obtain a more comprehensive depiction of outcome variation, the models used in this study purposefully highlighted students' economic disadvantaged status in addition to their schools' contextual indicators and place-based poverty levels.

Hypothesized Relationships

As seen in the research questions and the conceptual model in Figure 1, neither offers a clear directionality of the signs of the relationships to be estimated. Nonetheless, based on the literature review (see Clotfelter et al., 2009; Hanushek & Rivkin, 2006; NAEP, 2020; Phillips & Chin, 2004; Reardon & Galindo, 2009) and conceptual frameworks on the association between achievement and socioeconomic disadvantage, including neighborhood and school effects (Considine & Zappalà, 2002; Pearman, 2019), I am better positioned to hypothesize expected variations of the estimates of interest.

First, considering that mathematical proficiency is highly dependent on previous knowledge and skills, I hypothesize that as students advance throughout their elementary school years, these proficiency rates may decrease. Although this direction has been shown in aggregated reports, such as those offered by the NAEP (2020), such reports have not accounted for students' or schools' and neighborhoods' exposures to differing poverty levels. Similarly, considering the persistent gaps in proficiency by socioeconomic status, students experiencing

economic disadvantages may be less proficient in mathematics at earlier grades than their non-economically disadvantaged peers, and these proficiency gaps across advantaged and disadvantaged groups may grow over time at disparate levels, with economically disadvantaged students likely experiencing greater losses in mathematical proficiency levels compared to their non-economically disadvantaged peers. Finally, based on nesting effects from multilevel and spatial frameworks, school-level sociodemographic and economic indicators as well as the poverty level of the location where a school resides may further influence the observed mathematics proficiency rates. Highest poverty areas may experience the greatest proficiency drops, but non-economically disadvantaged students may be somehow protected from experiencing these drops at similar rates as their economically disadvantaged peers. The multi-level and spatial nature of the data sources to be modeled required the implementation of an analytic approach designed to account for these nesting structures, as discussed next.

Data and Methods⁶

The literature on students, schools, and neighborhood effects (Chetty et al., 2014, 2016, 2020; González Canché, 2022, 2023b) and the application of the concentrated disadvantage and geography of disadvantage frameworks (Elijah, 1990; Jargowsky & Tursi, 2015; Pastor, 2001; Tate, 2008; Weicher, 1979) informed the data selection process and the methodological approach employed in this study. All indicators shown in Table 2 were retrieved from data officially released by the NYSED, the IRS, the ACS, the Federal Bureau of Investigation (FBI), and the U.S. Department of Agriculture Economic Research Service (USDA ERS), as described next.

As shown in Figure 1, a contribution of this study consists of simultaneously accounting for the impact of schools and schools' locations on the variation of mathematical proficiency at the grade level, after accounting for schools' and neighborhoods' nesting and spatial interactions at both levels while considering grade performances being nested within schools—that is, the analyses controlled for schools' outcome dependence with other nearby schools (see Moran's *I* School estimate in Table 3) and neighborhoods' outcome dependence with other neighborhoods (see Moran's *I* Region estimate in Table 4) in addition to measuring grade performance dependence given grades' within schools nesting. These different levels required that the feature or variable selection captured sociodemographic and economic indicators measured at grade, school, and place levels—that is, resulting in a data set configured by indicators measured at different levels and nesting structures. With respect to the influence of location, to account for immediate as well as more contextual place-based indicators, the data set included zip code (more immediate) and county-level (more contextual) indicators. In both cases, these indicators were matched to schools' zip codes via crosswalking feature engineered

procedures (see González Canché, 2023b). This crosswalking was based on data made available by the U.S. Department of Housing and Urban Development (HUD)⁷ in its U.S. Postal Service ZIP Code Crosswalk Files.

For clarity, the next description presents these indicators as measured at these different grade, school, and place-based levels.

Outcome

All grades' ($N = 38,101$) and schools' ($N = 3,484$) data across 2 academic years (2017–2018, 2018–2019), representing 2.09 million test-takers,⁸ were retrieved from the NYSED (2020b). The main outcome, mathematical proficiency,⁹ was measured as the proportion of tested students scoring at levels 3 and 4 (on a 4-level scale) in the P–12 Common Core Learning Standards tests administered at public elementary and middle schools in the state of New York (NYSED, 2020b). Following the Definitions of Performance Levels (NYSED, 2020a), students in levels 1 and 2 were below proficient standards for their grade. Students in level 1 demonstrated limited knowledge, skills, and practices. Although their peers in level 2 were partially proficient in standards for their grade, this proficiency level was insufficient for the expectations at this grade (NYSED, 2020a). Students in level 3 were proficient in standards for their grade and met the expectations for this grade. Finally, students in level 4 excelled in standards, and their proficiency level was considered more than sufficient for the expectations for their grade level.

Table 1 shows an example of the data structure provided by the NYSED (2020b). This table shows 20 instances of proficiency rates reported by students' socioeconomic disadvantage statuses and across grade levels within schools. These data were added to school and place-level indicators, as described next.

School-Level Indicators

Similar to previous literature (i.e., Case et al., 1999; Denton & West, 2002; Pearman, 2019; Starkey et al., 2004), the models accounted for sociodemographic indicators that may have been associated with the variation of mathematical performance, such as the percentage of students at the school level participating in free and reduced-price lunch programs and annual attendance rates (the school's total actual attendance divided by the sum of the number of students in attendance on each day) for a school year. The models also controlled for the percentage of students classified as English language learners who required support to become proficient in English. In addition, the models included gender (female and male [reference or comparison group]), ethnicity of students (American Indian or Alaska native, Asian, Black, Hispanic, White [reference or comparison group], and multiracial), percentage of students with disabilities,¹⁰

TABLE 2
Indicator Sources, Levels (Grade, School, Area), and Definitions That Were Corroborated as Relevant, Based on Feature Selection

Indicator	Level	Source	Definition
math3 g	Grade	NYSEd	Mathematical proficiency Grade 3
math4 g	Grade	NYSEd	Mathematical proficiency Grade 4
math5 g	Grade	NYSEd	Mathematical proficiency Grade 5
math6 g	Grade	NYSEd	Mathematical proficiency Grade 6
math7 g	Grade	NYSEd	Mathematical proficiency Grade 7
math8 g	Grade	NYSEd	Mathematical proficiency Grade 8
not Economically Disadvantaged g	Grade	NYSEd	Students not classified as economically disadvantaged
mean number of students grade g	Grade	NYSEd	Mean number of students in given grade within school
pct students not tested g	Grade	NYSEd	Percentage of students in a grade who opted out of tests
attendance rate s	School	NYSEd	Attendance rate at the school level
pct teacher turnover s	School	NYSEd	Percentage of teacher turnover (last year only)
mean pct teacher turnover 5 years s	School	NYSEd	Mean percentage of teacher turnover in the past 5 years
pct ELL s	School	NYSEd	Percentage of ELLs at the school
pct Native American s	School	NYSEd	Percentage of student body identified as Native Americans
pct Black s	School	NYSEd	Percentage of student body identified as Black or African American
pct Hispanic s	School	NYSEd	Percentage of student body identified as Hispanic or Latina/o
pct Asian s	School	NYSEd	Percentage of student body identified as Asian
pct White s	School	NYSEd	Percentage of student body identified as White
pct_Multi Ethnicity s	School	NYSEd	Percentage of student body identified as more than one ethnicity
pct Women s	School	NYSEd	Percentage of student body identified as women
pct Economically Disadvantaged s	School	NYSEd	Percentage of student body identified as economically disadvantaged
pct with Disabilities s	School	NYSEd	Percentage of student body identified as having at least one disability
Large cities s	School	NYSEd	Indicator of school located in a large city, except for New York City
High Need Urban or Suburban s	School	NYSEd	Indicator of school located in a high-need area in urban or suburban areas, excluding New York City or large cities classification
NYC public s	School	NYSEd	Indicator of school located in New York City (excluding charter schools)
High Need Rural s	School	NYSEd	Indicator of school located in a high-need rural area
Average Need s	School	NYSEd	Indicator of school located in an average-need area (not rural and excluding other locations above)
Low Need s	School	NYSEd	Indicator of school located in a low-need area (not rural and excluding other locations above)
Charter School s	School	NYSEd	Indicator of school being a charter school, may be in New York City
pct free-reduced lunch s	School	NYSEd	Percentage of student body in a school receiving free or reduced-price lunch
pct suspensions s	School	NYSEd	Percentage of students being suspended
Gini z	ZCTA	ACS	Gini coefficient—5-year estimate (2015–2019)
pct EITC recipients z	ZCTA	IRS	Percentage of inhabitants in a ZCTA receiving EITC, fiscal years 2018 and 2019
pct mother only household z	ZCTA	ACS	Percentage of households in a ZCTA led by single mothers—5-year estimate (2015–2019)
pct crime z	ZCTA	FBI (crosswalked to ZCTA)	Crime rates of any type for the years 2018 and 2019
pct taxpayers below poverty line z	ZCTA	ACS	Percentage of households in a ZCTA earning at least \$7,500 below the poverty line for that area
pct children in poverty 5 to 17 c	County	ACS	Percentage of children ages 5–17 in a county classified below the poverty line based on their family size—5-year estimate (2015–2019)
death rate c	County	ACS	Total death rate in the county, including deaths of any type—5-year estimate (2015–2019)
net migration rate c	County	ACS	Net migration rate in the county—5-year estimate (2015–2019)
pct High school or less c	County	ACS	Percentage of adults (25 and older) in a county with up to a high school or equivalent degree—5-year estimate (2015–2019)
pct unemployment c	County	ACS	Percentage of unemployed and job-searching adults (25 and older) in a county—5-year estimate (2015–2019)
Median household income c	County	ACS	Median household income in a county—5-year estimate (2015–2019)

Note. ACS = American Community Survey; EITC = Earned Income Tax Credit; ELL = English language learner; FBI = Federal Bureau of Investigation; IRS = Internal Revenue Service; NYSED = New York State Education Department; ZCTA = zip code tabulated area.

TABLE 3
Summary Statistics Data Years 2017–2018 and 2018–2019

	Total (3,484 schools)	Lowest pov. tertile (1,200 schools)	Mid pov. tertile (1,070 schools)	Highest pov. tertile (1,214 schools)	<i>p</i> value
Outcomes: Mean (SD)					
math proficiency all grades	46.91 (24.61)	53.68 (23.51)	43.94 (23.08)	43.22 (25.72)	<0.001
math3 g Not Econ Dis_g	66.17 (19.92)	72.27 (15.95)	64.42 (18.30)	60.96 (23.56)	<0.001
math3 g Econ Dis_g	46.53 (21.05)	48.95 (20.57)	43.91 (19.77)	46.62 (22.47)	<0.001
Continuous or numeric features: Mean (SD)					
pct students not tested g	16.79 (20.33)	26.37 (23.66)	17.02 (18.75)	7.27 (12.31)	<0.001
pct Economically Disadvantaged s*	57.91 (24.77)	35.01 (19.98)	58.63 (16.44)	79.41 (13.54)	<0.001
attendance rate s	94.53 (2.27)	95.42 (1.54)	94.59 (1.78)	93.62 (2.87)	<0.001
pct teacher turnover s	2.59 (9.62)	0.69 (5.26)	1.66 (8.10)	5.34 (13.02)	<0.001
pct teacher turnover 5 years s	2.58 (9.84)	0.66 (5.00)	1.60 (8.00)	5.42 (13.65)	<0.001
pct ELL s	7.83 (10.07)	4.79 (5.85)	4.92 (8.05)	13.59 (12.32)	<0.001
pct Native American s	0.69 (3.03)	0.22 (1.43)	0.63 (2.07)	1.21 (4.52)	<0.001
pct Black s	17.03 (23.60)	8.42 (12.97)	14.62 (24.46)	27.71 (26.51)	<0.001
pct Hispanic s	24.01 (23.83)	18.53 (15.92)	14.98 (18.88)	38.10 (27.77)	<0.001
pct Asian s	7.83 (13.40)	8.52 (11.50)	4.97 (10.62)	9.95 (16.67)	<0.001
pct White s	47.53 (36.13)	60.83 (26.06)	61.61 (35.59)	20.97 (29.68)	<0.001
pct Multi Ethnicity s	2.84 (2.97)	3.42 (2.36)	3.13 (3.09)	2.00 (3.19)	<0.001
pct Women s	49.07 (4.90)	48.88 (3.37)	48.97 (3.53)	49.34 (6.88)	<0.001
pct with Disabilities s	17.50 (6.47)	16.08 (6.30)	16.87 (5.66)	19.50 (6.88)	<0.001
pct free-reduced lunch s	54.57 (24.91)	31.79 (20.07)	55.32 (16.60)	75.93 (14.38)	<0.001
pct suspensions s	2.91 (4.71)	1.58 (2.86)	3.31 (4.61)	3.82 (5.83)	<0.001
Gini z	0.45 (0.06)	0.43 (0.06)	0.43 (0.05)	0.47 (0.05)	<0.001
pct EITC recipients z	34.94 (12.59)	22.82 (6.99)	33.65 (6.91)	47.95 (7.68)	<0.001
pct mother only household z	28.25 (16.88)	15.99 (7.74)	26.33 (12.09)	42.01 (17.31)	<0.001
pct crime z	3.27 (1.50)	2.91 (1.34)	3.72 (1.13)	3.18 (1.82)	<0.001
pct children in poverty 5–17 c	17.84 (7.61)	13.85 (5.99)	17.27 (5.78)	22.25 (8.23)	<0.001
death rate c	8.50 (1.53)	8.41 (1.13)	9.11 (1.64)	7.99 (1.56)	<0.001
net migration rate c	-5.95 (5.85)	-3.49 (4.34)	-5.28 (5.87)	-8.98 (5.80)	<0.001
pct High school or less c	40.53 (8.02)	35.71 (5.84)	42.86 (6.56)	42.95 (8.99)	<0.001
pct unemployment c	4.27 (0.67)	3.92 (0.38)	4.43 (0.65)	4.46 (0.77)	<0.001
Median household income c	1.33 (29.70)	20.71 (30.00)	-7.91 (22.97)	-8.49 (25.78)	<0.001
pct taxpayers below poverty line z	0.36 (0.09)	0.26 (0.04)	0.35 (0.02)	0.46 (0.06)	<0.001
mean number of students grade g	54.87 (56.86)	62.16 (62.12)	47.49 (43.71)	54.96 (61.70)	<0.001
Categorical features: <i>n</i> (%)					
math3 g**	8,019 (21%)	2,825 (22%)	2,647 (21%)	2,547 (20%)	<0.001
math4 g**	7,864 (21%)	2,745 (22%)	2,618 (21%)	2,501 (19%)	<0.001
math5 g**	7,589 (20%)	2,599 (21%)	2,443 (19%)	2,547 (20%)	0.027
math6 g**	5,504 (14%)	1,730 (14%)	1,800 (14%)	1,974 (15%)	0.003
math7 g**	4,928 (13%)	1,483 (12%)	1,619 (13%)	1,826 (14%)	<0.001
math8 g**	4,197 (11%)	1,181 (9.4%)	1,454 (12%)	1,562 (12%)	<0.001
Not Econ Disadvantaged***	904,451 (43%)	521,139 (67%)	241,209 (40%)	142,103 (20%)	0.035
Large cities s	1,579 (4.1%)	303 (2.4%)	271 (2.2%)	1,005 (7.8%)	<0.001
High-Need Urban or Suburban s	2,866 (7.5%)	278 (2.2%)	1,155 (9.2%)	1,433 (11%)	<0.001
NYC public s	11,289 (30%)	2,330 (19%)	2,381 (19%)	6,578 (51%)	<0.001
High-Need Rural s	3,630 (9.5%)	89 (0.7%)	2,250 (18%)	1,291 (10.0%)	<0.001
Average Need s	11,832 (31%)	5,749 (46%)	5,600 (45%)	483 (3.7%)	<0.001

(continued)

TABLE 3. (CONTINUED)

	Total (3,484 schools)	Lowest pov. tertile (1,200 schools)	Mid pov. tertile (1,070 schools)	Highest pov. tertile (1,214 schools)	<i>p</i> value
Low Need s	3,843 (10%)	3,544 (28%)	299 (2.4%)	0 (0%)	<0.001
Charter School s	3,054 (8.0%)	270 (2.1%)	625 (5.0%)	2,159 (17%)	<0.001
Academic year 2018–2019_s****	18,832 (49%)	6,231 (50%)	6,255 (50%)	6,346 (49%)	0.4
<i>N</i> of grades (<i>g</i>)	38,101	12,563 (33%)	12,581 (33%)	12,957 (34%)	
Total no. of students	2,090,513	780,953	597,501	712,059	

Note. EITC = Earned Income Tax Credit; ELL = English language learner; SD = standard deviation.

*In addition to this school-level indicator, each grade contained proficiency levels disaggregated by students' economically disadvantaged status.

**Number of grades that were classified in 3, 4, 5, 6, 7, or 8 grades in the analytic sample

***Number of analytic units that were classified as not being economically disadvantaged

****Number of analytic units represented in the academic year 2018–2019

and percentage of students suspended (suspensions of 1 full day or longer during the school year).

The models also included the average number of students per grade, as reported by the NYSED. These mean estimates were computed as “number of students registered in a given grade in math classes divided by the number of those classes” in each of the schools observed in the analytic samples. Additionally, an indicator of the percentage of students who opted out of testing was included, given its potential impact on outcome variation: The lowest poverty schools had the highest test opt-out rates (see Table 1). Each school was classified following the NYSED needs index, which separates schools in New York City (NYC) Public Schools (reference or comparison group), schools located in Large Cities, schools located in High-Need/Resource Urban-Suburban Districts, schools located in High-Need/Resource Rural Districts, schools belonging to Average-Need Districts, schools ascribed to Low-Need Districts, and Charter Schools. Finally, the models captured teacher turnover (i.e., the percentage of teachers in the prior school year as well as in the previous 5 years who did not return to a teaching position). In the tables presented, school-level indicators contained the subindexes *g* if measured at the grade level or *s* if measured at the school level.

Place-Based Indicators

Poverty rates and Earned Income Tax Credit (EITC) indicators were retrieved from the IRS (2020) data measured at the zip code tabulated area (ZCTA) level. EITC captures the prevalence of low-income taxpayers with children in a given ZCTA and has been found to increase college enrollment (Manoli & Turner, 2018). This indicator is conceptually relevant for parental education and social class and has been shown to influence mathematical proficiency (Case, et al., 1999; Denton & West, 2002; Starkey et al., 2004). Other ZCTA-level

indicators were retrieved from the ACS 5-year (2015–2019) data estimates. They included the Gini coefficient of income inequality and percentage of single-mother households in a zip code as a measure of family composition. Proportion of single-mother households has been considered an important indicator of geography of disopportunity or concentration of disadvantages (Elijah, 1990; González Canché, 2022, 2023b; Jargowsky & Tursi, 2015; Pastor, 2001; Tate, 2008). Crime rates were gathered from the FBI's Uniform Crime Reporting Program (FBI, 2019), and although they were reported at the county level, these measures were crosswalked to the ZCTA level to obtain a more localized impact of crimes on schools. All these indicators contained the subindex *z* to represent their ZCTA level of measurement.

To expand the place-based contextual scope, the analyses also included variables measured at the county level. These variables included the percentage of inhabitants ages 5–17 living in poverty, death rates, and net migration (including in- and outmigration). These migration-related indicators aimed to account for the potential compositional change of students based on their parents' decisions to move to other schools in the pursuit of better academic opportunities; these mobility decisions' potential impact on the study's estimates are discussed in the section titled “Limitations: How Would Student Composition and Parents' Wealth Affect the Findings?” within the limitations section. Guided by the tenets of geography of opportunity (Pastor, 2001; Tate, 2008) and neighborhood effects (Chetty et al., 2014, 2020), the models included unemployment rates, median household income, and percentage of adult inhabitants (25 years and older) with up to a high school or equivalent degree. These indicators contained the subindex *c* to ease their identification in the models. Finally, given that the data accounted for the academic years 2017–2018 and 2018–2019, a binary indicator was added to account for time fixed effects, similar to the approach used by Dong et al. (2015).

TABLE 4

Multilevel SAR Models for Posterior Means of Mathematical Proficiency Rates—Academic Years 2017–2018 and 2018–2019

	Not economically disadvantaged			Economically disadvantaged			
	Aggregate	Low pov	Mid pov	High pov	Low pov	Mid pov	High pov
Intercept	-107.190*** (6.007)	-25.097*** (8.954)	-31.888*** (8.815)	-28.646*** (9.217)	-20.739** (9.263)	-51.300*** (8.706)	-34.542*** (8.558)
math4 g	-4.999*** (0.266)	-3.441*** (0.519)	-5.459*** (0.627)	-5.390*** (0.800)	-4.374*** (0.619)	-5.844*** (0.541)	-5.773*** (0.624)
math5 g	-9.483*** (0.271)	-6.576*** (0.527)	-10.911*** (0.639)	-11.411*** (0.790)	-7.597*** (0.636)	-10.599*** (0.560)	-10.665*** (0.621)
math6 g	-8.735*** (0.307)	-4.212*** (0.669)	-8.101*** (0.710)	-13.050*** (0.877)	-4.838*** (0.787)	-7.790*** (0.639)	-10.992*** (0.682)
math7 g	-12.795*** (0.328)	-7.922*** (0.742)	-13.096*** (0.779)	-17.785*** (0.905)	-7.655*** (0.864)	-12.716*** (0.692)	-13.745*** (0.718)
math8 g	-20.869*** (0.372)	-26.038*** (0.912)	-26.001*** (0.908)	-21.180*** (1.001)	-18.504*** (0.995)	-18.575*** (0.752)	-15.307*** (0.768)
not Economically Disadvantaged g	17.789*** (0.174)	NA (NA)	NA (NA)	NA (NA)	NA (NA)	NA (NA)	NA (NA)
Academic year 2018–2019_s	1.580*** (0.174)	1.219*** (0.353)	0.778 (0.404)	1.537*** (0.505)	2.156*** (0.426)	1.673*** (0.357)	2.066*** (0.396)
attendance rate s	1.502*** (0.057)	0.763*** (0.106)	0.837*** (0.099)	1.648*** (0.113)	0.572*** (0.114)	0.952*** (0.093)	1.392*** (0.091)
pct teacher turnover s	0.084*** (0.040)	-0.122 (0.156)	-0.076 (0.142)	0.219*** (0.084)	-0.155 (0.180)	0.323*** (0.122)	0.145*** (0.064)
pct teacher turnover 5 years s	0.065 (0.035)	0.345** (0.163)	0.17 (0.128)	-0.022 (0.068)	0.519*** (0.188)	-0.05 (0.112)	0.031 (0.052)
pct ELL s	-0.362*** (0.018)	-0.263*** (0.053)	-0.272*** (0.050)	-0.445*** (0.040)	-0.305*** (0.061)	-0.289*** (0.045)	-0.356*** (0.030)
pct Native American s^	-0.03 (0.038)	-0.011 (0.140)	0.091 (0.114)	-0.1 (0.070)	0.022 (0.164)	0.219** (0.105)	-0.038 (0.056)
pct Black s^	-0.137*** (0.011)	-0.145*** (0.029)	-0.145*** (0.022)	-0.236*** (0.035)	-0.131*** (0.020)	-0.082*** (0.020)	-0.093*** (0.027)
pct Hispanic s^	-0.042*** (0.012)	-0.006 (0.029)	-0.058* (0.029)	-0.083*** (0.035)	-0.037 (0.035)	-0.009 (0.025)	0.005 (0.028)
pct Asian s^	0.219*** (0.014)	0.162*** (0.025)	0.192*** (0.036)	0.153*** (0.042)	0.302*** (0.032)	0.278*** (0.034)	0.319*** (0.033)
pct Multi Ethnicity s^	-0.022 (0.048)	0.123 (0.103)	0.167 (0.090)	-0.097 (0.145)	-0.143 (0.127)	0.01 (0.082)	-0.298*** (0.117)
pct Women s	0.108*** (0.019)	0.009 (0.058)	0.176*** (0.064)	0.061 (0.038)	0.191*** (0.065)	0.154*** (0.053)	0.021 (0.031)
pct Economically Disadvantaged s	-0.244*** (0.022)	-0.138*** (0.053)	-0.158*** (0.054)	-0.427*** (0.058)	-0.083 (0.065)	-0.088 (0.048)	-0.434*** (0.048)
pct with Disabilities s	-0.376*** (0.020)	-0.331*** (0.042)	-0.425*** (0.049)	-0.389*** (0.055)	-0.303*** (0.049)	-0.450*** (0.040)	-0.501*** (0.043)
Large cities s^^	5.248*** (0.894)	-2.01 (2.096)	-1.279 (2.174)	11.061*** (1.896)	-5.055*** (2.499)	-0.529 (1.983)	10.366*** (1.577)
High-Need Urban/Suburban s^^	-1.026 (0.828)	-5.781*** (2.101)	-5.600*** (1.728)	4.910*** (1.832)	-4.095 (2.462)	-5.655*** (1.607)	4.328*** (1.577)
High-Need Rural s^^	-4.829*** (0.878)	-2.732 (3.062)	-7.376*** (1.734)	-3.66 (2.418)	-6.171 (3.500)	-7.924*** (1.586)	-2.891 (1.979)
Average Need s^^	-4.440*** (0.761)	-4.404*** (1.409)	-6.086*** (1.705)	-3.871 (2.460)	-10.057*** (1.686)	-7.425*** (1.558)	-4.608*** (2.007)
Low Need s^^	-6.101*** (0.857)	-6.120*** (1.399)	-5.160*** (2.254)	0.004 (9.933)	-11.138*** (1.660)	-7.338*** (2.073)	-0.096 (9.956)
Charter School s^^	9.106*** (0.902)	1.839 (3.048)	10.040*** (2.469)	8.737*** (1.916)	5.197 (3.360)	6.001*** (2.073)	10.548*** (1.556)
pct free-reduced lunch s	0.146*** (0.021)	-0.001 (0.053)	0.025 (0.053)	0.233*** (0.047)	-0.034*** (0.063)	0.026 (0.047)	0.221*** (0.039)
pct suspensions s	-0.136*** (0.026)	-0.394*** (0.118)	-0.07 (0.065)	-0.240*** (0.064)	0.031 (0.092)	-0.009 (0.051)	-0.189*** (0.050)
mean number of students grade g	0.028*** (0.001)	0.016*** (0.003)	0.019*** (0.007)	0.017 (0.014)	-0.012 (0.007)	-0.013*** (0.004)	-0.004 (0.003)
Gini z	-0.189 (3.264)	11.829*** (4.868)	7.107 (4.968)	-14.414 (7.985)	-1.619*** (5.230)	-6.648 (4.599)	-12.941 (7.469)

(continued)

TABLE 4. (CONTINUED)

	Aggregate	Not economically disadvantaged			Economically disadvantaged		
		Low pov	Mid pov	High pov	Low pov	Mid pov	High pov
pct EITC recipients z	0.013 (0.027)	-0.019 (0.058)	-0.045 (0.057)	-0.191** (0.088)	-0.095 (0.071)	-0.023 (0.051)	-0.195*** (0.078)
pct mother only household z	-0.011 (0.016)	0.009 (0.041)	-0.009 (0.026)	-0.009 (0.039)	-0.08 (0.052)	-0.045 (0.023)	-0.01 (0.035)
pct crime z	-0.086 (0.173)	0.187 (0.374)	0.294 (0.308)	-0.09 (0.335)	0.211 (0.429)	0.001 (0.287)	-0.185 (0.306)
pct children in poverty 5–17 c	0.407*** (0.051)	0.198** (0.096)	0.366*** (0.096)	0.742*** (0.146)	0.076 (0.109)	0.251*** (0.087)	0.552*** (0.137)
death rate c	-0.523*** (0.199)	0.363 (0.444)	0.012 (0.277)	-4.157*** (0.691)	-0.113 (0.521)	-0.168 (0.259)	-3.344*** (0.618)
net migration rate c	-0.033 (0.051)	-0.197 (0.108)	-0.046 (0.071)	0.397*** (0.166)	-0.044 (0.125)	-0.06 (0.067)	0.294 (0.151)
pct High school or less c	0.022 (0.040)	-0.102 (0.086)	-0.108 (0.063)	0.012 (0.118)	-0.097 (0.101)	0.013 (0.057)	0.015 (0.106)
pct unemployment c	0.053 (0.495)	1.671 (1.234)	1.415 (0.720)	-3.161*** (1.325)	3.303*** (1.437)	0.382 (0.657)	-0.627 (1.215)
Median household income c	0.114*** (0.014)	0.108*** (0.026)	0.168*** (0.029)	0.036 (0.052)	0.092*** (0.028)	0.089*** (0.025)	0.035 (0.048)
pct students not tested g	-0.142*** (0.007)	-0.165*** (0.013)	-0.157*** (0.015)	-0.209*** (0.025)	-0.069*** (0.015)	-0.051*** (0.013)	-0.121*** (0.021)
Pseudo R^2	0.535	0.504	0.435	0.459	0.388	0.453	0.568
Lambda M	16.492*** (1.288)	11.558*** (1.878)	7.887*** (1.981)	22.405*** (5.002)	14.997*** (2.820)	7.548*** (1.675)	22.353*** (3.723)
Rho W	274.441*** (2.390)	190.652*** (4.0411)	254.348*** (5.138)	395.242*** (9.114)	271.452*** (6.035)	198.199*** (4.125)	248.513*** (6.381)
Moran's I Regions ^{^^} Expectation	0.401*** (-0.0009)	0.313*** (-0.002)	0.092* (0.002)	0.412*** (-0.005)	0.297*** (-0.002)	0.239*** (-0.002)	0.571*** (-0.005)
Moran's I Schools ^{^^} Expectation	0.268*** (-0.00002)	0.3626*** (-0.0001)	0.329*** (-0.0001)	0.217*** (-0.0001)	0.371*** (-0.0001)	0.396*** (-0.0001)	0.295*** (-0.0001)
N of schools and grades	38,101	6,357	6,278	6,334	6,206	6,327	6,599
No. of students	2,090,513	521,139	241,587	141,725	259,814	356,828	569,420
Δ	[38101, 1028]	[6357, 396]	[6278, 421]	[6334, 215]	[6206, 386]	[6327, 420]	[6599, 215]
M	[1028, 1028]	[396, 396]	[421, 421]	[215, 215]	[386, 386]	[420, 420]	[215, 215]
W	[38101, 38101]	[6357, 6357]	[6278, 6278]	[6334, 6334]	[6206, 6206]	[6327, 6327]	[6599, 6599]

Note. EITC = Earned Income Tax Credit; ELL = English language learner; MSAR = multilevel simultaneous autoregressive; NYC = New York City; SAR = simultaneous autoregressive.

^ Comparison group was White students.

^^ Comparison was NYC schools.

^^^ The parameter space for these variables was to be (-1,1), and these tests were conducted independently from MSAR, as can be seen in the replication code.

* $p < .05$. ** $p < .01$. *** $p < .0001$.

Multilevel Modeling With Spatial Interaction Effects

The data analyzed were multilevel and geographical in nature (González Canché, 2022, 2023b). They were multilevel because students' outcomes were nested within schools, and they were geographical because these schools were nested within zip codes and counties. In traditional multilevel approaches, nested units are assumed to have more similar (correlated) outcomes, based on their group dependence (Dong et al., 2015). However, these approaches ignore that schools themselves are also located in areas that have sociodemographic and economic indicators that may further influence the grade- and school-level outcomes—that is, in addition to grade- and school-level group dependence, schools' location and/or distance from other schools may lead to another form of dependence. This latter dependence

form is termed *place, contextual, or neighborhood effects* (Dong et al., 2015) and is typically modeled with geospatial or geostatistical techniques (Bivand et al., 2013). However, before addressing potential spatial dependence, consider a typical multilevel model specification, as expressed in equation (1):

$$y_{g \rightarrow s, z} = \beta_0 + x_{g \rightarrow s}^T \beta + x_z^T \gamma + u_z + \varepsilon_{g \rightarrow s, z} \quad (1),$$

where the index $g \rightarrow s$ is a composite identifier of grade IDs ascribed to schools' IDs, which in this study led to identifying 38,101 unique combinations of grades and schools (see Tables 2 and 3). Further, following Dong et al. (2015), $g \rightarrow s$ identifies lower level units, and z identifies higher level entities (or zip codes, totaling 1,028 in this study). $y_{g \rightarrow s, z}$ are grade-level outcomes as identified within schools,

which can also be conceptualized as being nested in z zip codes. From this perspective, the residuals u_z are captured at the zip code level, the residuals $\varepsilon_{g \rightarrow s, z}$ correspond to lower level residuals, and both are assumed to follow independent normal distributions, or $N(0, \sigma_u^2)$ and $N(0, \sigma_\varepsilon^2)$, respectively (Dong et al., 2015).

The covariates represented in equation (1) are measured at the zip code level¹¹ (x_z) and at the grade and school levels ($x_{g \rightarrow s}$), and each level has its corresponding coefficient estimates. However, this standard multilevel specification has not yet incorporated any spatial information, even though the outcome at a particular location may be influenced by its surrounding locations, and the intensity of this influence is expected to be higher the closer the units are to one another (Dong et al., 2015; Tobler, 1970). Whenever this dependence was observed, Dong et al. (2015) demonstrated that the multilevel models alone rendered estimates that were biased due to ignoring the spatial processes influencing the outcomes of interest. However, it is possible that spatial dependence was not present in the data analyzed, which would have nullified the need to rely on spatial analyses to obtain unbiased estimates. From this perspective, before relying on multilevel spatial models, the presence of spatial dependence at the school and the region nesting units should be tested. Accordingly, the following presents the steps followed to assess the presence of spatial dependence, which was corroborated at the $g \rightarrow s$ and z levels (see results in Moran's I School and Moran's I Region in Table 4).

Assessing Spatial Dependence. The assessment of spatial dependence required the identification of units that were located in close spatial proximity (neighboring units). The result of this identification process was stored in matrices of influence, which captured units based on geographic contiguity (i.e., among higher level nesting units that in this study were captured by ZCTAs) and proximity (i.e., identified among lower level units [schools] that met a given distance-based threshold, even if they were located in different ZCTAs), as shown in Figure 3.

Polygon-Based Neighbors. Specifically, matrices of influence or weight matrices are square matrices where the intersection between rows and columns captures the presence (typically expressed as 1 when the rows are unstandardized or $\frac{1}{n}$, where n is the number of neighbors in row-standardized matrices) or absence (expressed as 0) of vicinity or a source of influence. An absence indicates that row y and column z are not neighbors or do not meet a decision rule to be potentially influencing one another. When the units of interest are polygons (e.g., states, counties, ZCTAs), the decision rule to establish neighbors can be sharing a border, touching a point, or both (Bivand et al., 2013). Because the higher level units in this case were polygons, ZCTAs that

surrounded another ZCTA by sharing a border or touching a point were defined as neighbors in the matrix M depicted in equation (2), below. This decision rule is referred to as the queen's approach (Bivand et al., 2013), and although it requires more computing power, it also enables the identification and modeling of more information (Lloyd, 2010). As shown in Figure 3, all contiguous ZCTAs were neighbors whose influence could be modeled via Moran's I , as described below.

Radius-Based Neighbors. A second type of neighbor identification rule is based on a radius-based distance threshold. Here, units located within such a threshold, regardless of their nesting units of ascription (or ZCTA, in this case), are defined as neighbors in their corresponding matrix of influence. This matrix, referred to as W in equation (2), captured lower level vicinity, wherein schools nested in the same or in different ZCTAs were neighbors as long as they were located within the previously defined radius-based distance threshold. Relying on variograms, simulations, and related literature (see Dong & Harris, 2015), Dong et al. (2015) proposed different distance thresholds—1.5, 2, and 2.5 kilometers—consistently showing that 1.5 kilometers outperformed the latter two in all their findings. Following their recommendation, lower level units (schools) located within 1.5 kilometers from one another were considered neighbors in this study.¹²

Figure 3 shows the result of the identification strategies employed herein. In those figures, all surrounding ZCTAs were neighbors (matrix M), whereas schools that met the distance threshold within and across ZCTAs were identified as neighbors as well (see the white lines in the figure), representing matrix W .

The use of matrices M and W is further discussed below, in addition to a matrix Δ that combined M and W to capture place fixed effects. The way matrices M and W were used to test for spatial dependence is described as follows.

Moran's I . Mathematically, spatial dependence or autocorrelation measures the extent to which the outcomes or attributes of units located in close physical proximity are correlated (Bivand et al., 2013). Here, spatial dependence captured the correlation of an indicator with the average neighbor performance in this same indicator (Bivand et al., 2013). Taking the matrix of ZCTAs M as an example, the measure of spatial dependence can be identified as follows:

$$I = \frac{\frac{1}{n} \sum_{z=1}^n \sum_{z'=1}^n M_{zz'} (y_z - \bar{y})(y_{z'} - \bar{y})}{\sum_{z=1}^n (y_z - \bar{y})^2} \quad (2),$$

where y_z is the z th indicator of interest for a ZCTA z , \bar{y} is the mean of this same indicator across all units in the sample, and $M_{zz'}$ is the matrix of influence that identifies units z

in close proximity from unit z . As described above, the rule to identify the neighbors in the ZCTA data structure was sharing a border or touching a given point. With this information, one can obtain y_z , or the mean value of neighboring units on the indicator of interest. This latter indicator is measured to assess the extent to y_z and y_z are systematically over or below the mean. If a unit and its neighbors are below the mean, the multiplication of their differences from the overall mean $\left[\left(y_z - \bar{y} \right) \left(y_z - \bar{y} \right) \right]$ will be positive. The same is true if these differences are consistently above the mean. Accordingly, positive Moran's I values indicate consistent neighboring under- or overperformance, therefore capturing negative or positive geographical dependence of the outcomes. Negative values indicate greater performance disparities than expected under random spatial distribution, and values of zero indicate no outcome association or dispersion. Finally, Moran's I is normalized to range from -1 to $+1$, with $+1$ indicating a perfect spatial autocorrelation, suggesting a 1-to-1 increase in a given outcome and the average outcomes of a unit's neighbors. Finally, although equation (2) relied on the matrix M , the same rationale applies to the matrix W discussed above, with the main difference being that the outcomes account for school mathematical performance to be compared to the performance of their neighboring schools (i.e., those identified to be located within the distance threshold of 1.5 kilometers, in this case).

To test whether spatial autocorrelation was present in the data analyzed, two sets of Moran's I tests were implemented. The first set tested whether nearby schools had more similar outcomes (this was achieved by using matrix W). The second set tested whether neighboring ZCTAs had more similar outcomes (using matrix M). The results (contained in Table 4) consistently corroborated the presence of both forms of spatial dependence in mathematical proficiency (Moran's $I = 0.401$ and 0.268 at the ZCTA and school levels, respectively). Because both cases reached statistical significance ($p < .001$), multilevel models should not only consider nesting but also *the spatial context where the nesting occurred* (Dong et al., 2015), as nesting and spatial context effects may have affected the outcome of interest.

Based on the multilevel and geographical structure of the data analyzed, the models presented in this study accounted for within and between group correlation (Dong & Harris, 2015; Dong et al., 2015). These models accounted for residual dependence in the outcome variable measured within schools and among neighboring zip codes through spatial simultaneous autoregressive processes. Conceptually, this modeling approach represented an important analytic advancement for schools, and geographical contexts may have simultaneously affected the outcomes of interest even after conditioning on higher and lower level covariates. This result implies that, in the presence of spatial dependence or when autocorrelation is corroborated, in addition to

controlling for observed indicators measured at the school and location levels, spatial analytic techniques should be used to account for neighborhood effects. Formally accounting for these latter effects is relevant given that the effect of location has been found to affect individuals' outcomes even among those with nearly identical personal attributes and socioeconomic characteristics but living in different areas (Jones & Duncan, 1995).

Multilevel Simultaneous Autoregressive Models. In the multilevel modeling with spatial interaction effects framework, the observed value of a given location is allowed to be potentially influenced by the values of surrounding or nearby locations at two levels, as captured by the matrices of influence W or M . This simultaneity issue is modeled in simultaneous autoregressive (SAR) processes (Dong et al., 2015), which enable the modeling of spatial spillovers across higher and lower level units.

As discussed, the matrices of influence M and W captured spatial dependence among ZCTAs and among schools that met the distance threshold. This MSAR framework also required the identification of the unit nesting grade-level outcomes to account for fixed group effects at the two levels of nesting as $(g \rightarrow s, z)$, as defined in equation (1). This matrix is identified as Δ in equation (3) and is a block diagonal design matrix that is traditionally referred to as an ascription matrix in the network analysis literature (Breiger, 1974; González Canché, 2019, 2023b). Therefore, Δ accounted for grade (g), school (s), and neighborhood (z) fixed effects to capture place-based unobserved heterogeneity within ZCTAs and within schools where grade-level outcomes were realized.¹³

These three matrices (M , W , and Δ) had the following dimensions: M was Z by Z , with Z being the number of ZCTAs to be included in the models; W was S by S , with S referring to the total number of units included in the entire network that, in this case, accounted for the number of grades g and schools s (in the form $g \rightarrow s$) represented in the data set. Finally, Δ had dimensions S by Z , which allowed identification of all grade-level outcomes observed within schools (captured by S or, more formally, $g \rightarrow s$), both of which were located within a given ZCTA. As indicated above, M and W were row standardized (i.e., the sum of rows added to 1), thus assuming similar simultaneous levels of influence or competition among neighbors (Bivand et al., 2013). In the case of the block diagonal design matrix Δ , the intersection between each unit in the rows (or $g \rightarrow s$) and its corresponding ZCTA of ascription (z) captured in the columns had a value of 1 when $g \rightarrow s$ was located in ZCTA z , or a value of 0 otherwise.

With this information, the MSAR model with spatial interaction effects is specified as

$$y = \rho W y + X \beta + Z \gamma + \Delta \theta + \varepsilon, \theta = \lambda M \theta + u \quad (3),$$

where y is the indicator of mathematical proficiency measured at the grade level, with the main parameter of interest being the expected changes in proficiency levels across Grades 3–8, after controlling for grade, school, and location indicators (see Figure 1). X is an N by K matrix denoting covariates measured at the grade and school levels, and Z is a Z by P matrix consisting of variables measured at the ZCTA level.¹⁴ The strength of spatial dependence of the outcome of interest at the lower level is captured by ρ via W ; ε captures the lower level residuals (after accounting for or modeling spatial dependence at this level). Residuals at the ZCTA level that also include lower level units' information are captured by θ by relying on the block diagonal design matrix Δ , representing random contextual effects. Following the multilevel framework, θ is modeled with λ capturing spatial interactions at the ZCTA level, given the matrix of influence M that identifies zip code contiguity. The residuals u are distributed as $N(0, \sigma_u^2)$; similarly, ε is also distributed $N(0, \sigma_\varepsilon^2)$, so that neither the higher level nor the lower level units have residuals that are spatially dependent. Moreover, u and ε are independent.

According to Dong and Harris (2015), if the variances associated with u and ε reached statistically significant levels, these findings further corroborated the need to model the variation captured by the matrices M and W , as identified in equation (3). They went on to state that ignoring these forms of dependence could result in underestimated standard errors for covariance effects. Table 4 corroborates the significance levels of u and ε across all model identifications (see Grade school [$g \rightarrow s$] variance [ρW] and zip code variance [z] [λM] in Table 4). The data (available at González Canché, 2023b) and R code required to replicate these models are available at <https://cutt.ly/N4zRstL>; note that the code also allows users to load the data directly when executing the procedures.

MSAR was implemented via a Bayesian Markov Chain Monte Carlo (BMCMC). This modeling framework draws samples sequentially from the conditional posterior distributions of each model unknown parameter. The inferences are based on the posterior distributions of model parameters based on three MCMC chains with 10,000 iterations each and a burn-in period of 5,000 (Dong & Harris, 2015).¹⁵ The prior distribution for ρ in equation (2) is obtained from the minimum and maximum eigenvalues of the spatial connectivity matrix W . The model also assumes a uniform prior distribution for λ over $1/(\text{minimum eigenvalue of } M)$, also shown in equation (3). For more details about this implementation, please see Dong and Harris (2015).

Feature Selection: A Machine Learning Strategy to Deal With Place-Based Multicollinearity. The tenets of geography of advantage/disadvantage suggest that the geographical indicators selected may be highly *correlated*—for example, zones with high crime are likely to have high poverty levels.

This correlation, which is typically observed in studies modeling environmental factors (Li et al., 2016), may affect the observed variable importance of the predictors. Following Li et al. (2016), before model estimation, variable inclusion criteria relied on a feature selection algorithm (Kursa & Rudnicki, 2010) to detect all nonredundant variables to predict mathematical proficiency variation via machine learning. This process effectively addressed multicollinearity issues by identifying and easing the exclusion of redundant features (González Canché, 2022, 2023b). This nonredundant feature selection was implemented by using the Boruta function, a random forest regression procedure. Boruta is a wrapper algorithm that subsets features—the X s and Z s depicted in equation (3)—and trains a model using them to try to capture all the relevant indicators with respect to an outcome variable.

As depicted by Kursa and Rudnicki (2010), relevance is identified when there is a subset of attributes in the data set among which a given indicator is not redundant when predicting the outcome of interest. Procedurally, Boruta duplicates the data set and shuffles the values in each column, referring to these shuffled indicators as shadow features. Then, a random forest algorithm is used to learn whether the actual feature performs better than its randomly generated shadow. The Boruta implementation relied on 1,000 iterations; however, the optimal result was consistently found in less than 50 iterations, indicating that each attribute had predictive relevance levels higher than its shadow attributes.¹⁶ The results of this process, shown in Figure 2 (see an interactive version at <https://cutt.ly/M4WNq7W>), indicate that all the features discussed in the data section were selected as relevant predictors of the percentage of students reaching mathematical proficiency. Accordingly, these features were included in the multilevel models with spatial interaction effects discussed above.¹⁷

Findings

Table 3 contains aggregated descriptions of outcomes and attributes (under the column labeled “Total”) and disaggregated results by ZCTA poverty-level tertiles.¹⁸ The column p value tests whether at least one of these tertile distributions is significantly different from the other two. Finally, given that the main comparison group of interest is Grade 3 outcomes, the table also shows summary statistics for Grade 3 students by economically disadvantaged status—that is, it presents mathematical proficiency by Grade 3 students classified as economically disadvantaged and by those not classified as economically disadvantaged, and, as indicated, all these estimates are disaggregated by the poverty tertile of the schools these students attended.

For context, before describing mathematical performance levels, the prevalence of economic disadvantage in the sample needs to be discussed. Overall, the indicator called “pct

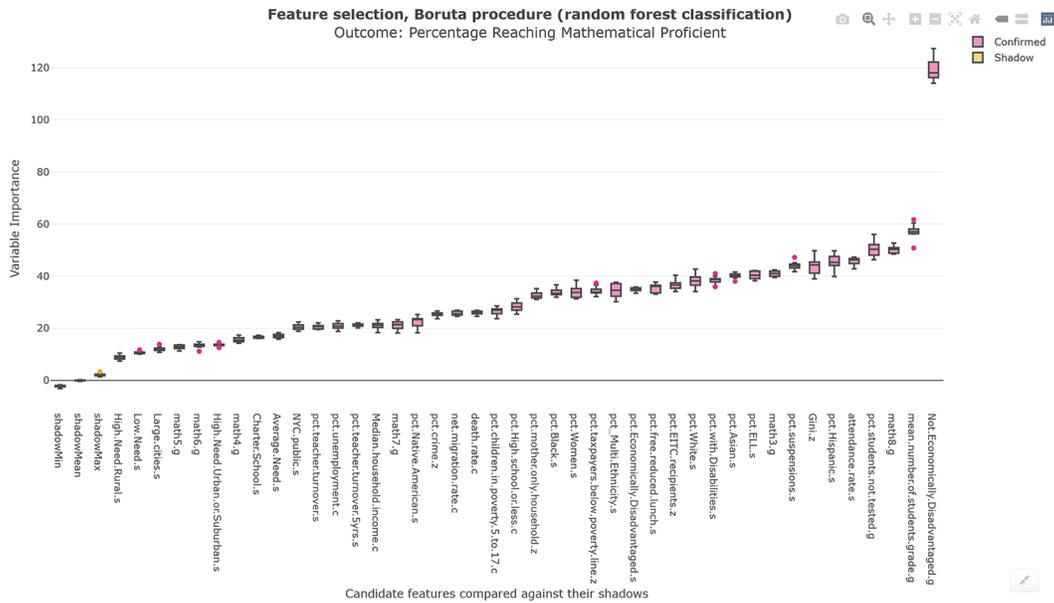


FIGURE 2. Feature selection results based on random forest classification. See interactive version here <https://cutt.ly/M4WNq7W>

Economically Disadvantaged s” shows that 57.91% of the 2.09 million test-takers represented in the data were classified as economically disadvantaged and that 79.41% of these economically challenged students also attended schools located in the highest poverty zones. More to the point, the lower concentration of economically disadvantaged students attended schools also located in the lowest poverty areas. Both of these findings agree with the notions of geography of disopportunity or disadvantage, which state that advantage begets advantage and disadvantage begets disadvantage (Elijah, 1990; Jargowsky & Tursi, 2015; Pastor, 2001; Tate, 2008).

Regarding mathematical proficiency, the average proficiency level across all grades (math proficiency all grades) was 46.91% ($SD = 24.61$); students attending schools in the lowest poverty areas attained the highest proficiency levels in mathematics (53.68%, $SD = 23.51$), around 10 percentage points higher than those of their peers located in the highest and the second-highest tertile classes for poverty zones (73.22%, $SD = 23.56$, and 43.99%, $SD = 18.30$, respectively).

When considering only Grade 3 outcomes, which is useful for baseline comparisons in the MSAR model interpretations below, non–economically disadvantaged Grade 3 students attending the lowest poverty zone schools reached mean of proficiency rates (72.27%, $SD = 15.95$) 23.32 percentage points above their *economically disadvantaged peers attending these same schools located in the lowest poverty zones* (48.95%, $SD = 20.57$). The performance of non–economically disadvantaged Grade 3 students attending schools located in low poverty areas could be also compared to those of their non–economically disadvantaged

peers attending schools located in higher poverty zones. In this case, non–economically disadvantaged Grade 3 students attending schools in mid- and high-poverty areas performed about 8 and 12 percentage points below their non–economically disadvantaged peers.

These same comparisons among economically disadvantaged Grade 3 students had lower magnitudes but did not follow the same pattern. Economically disadvantaged Grade 3 students attending schools in the highest poverty zones performed more similarly than their Grade 3 peers attending schools located in the lowest poverty area (46.62%, $SD = 22.47$, compared to 48.95%, $SD = 20.57$). In this case, the worst performance was observed among Grade 3 students attending schools located in mid-poverty neighborhoods (43.91%, $SD = 19.77$). Although purely descriptive, mathematical performance seems to have been more greatly affected by poverty location differences among non–economically disadvantaged Grade 3 students (in terms of performance reductions) compared to the mathematical performance of their economically disadvantaged Grade 3 peers.

With respect to other indicators, it should be noted that schools located in the highest poverty areas had the lowest attendance rates (93.62, $SD = 2.87$), highest teacher turnover rates (5.34%, $SD = 13.02$), and highest concentration of English language learners (13.59%, $SD = 12.32$) and African American (27.7%, $SD = 26.51$) and Hispanic (38.1%, $SD = 27.77$) students. At least 60% of students attending schools located in the lowest and mid-poverty zones were White; these students represented about 21% of the students attending schools located in the highest poverty zones. For comparison, note that although Black and Hispanic students only accounted for 17% and 24% of the

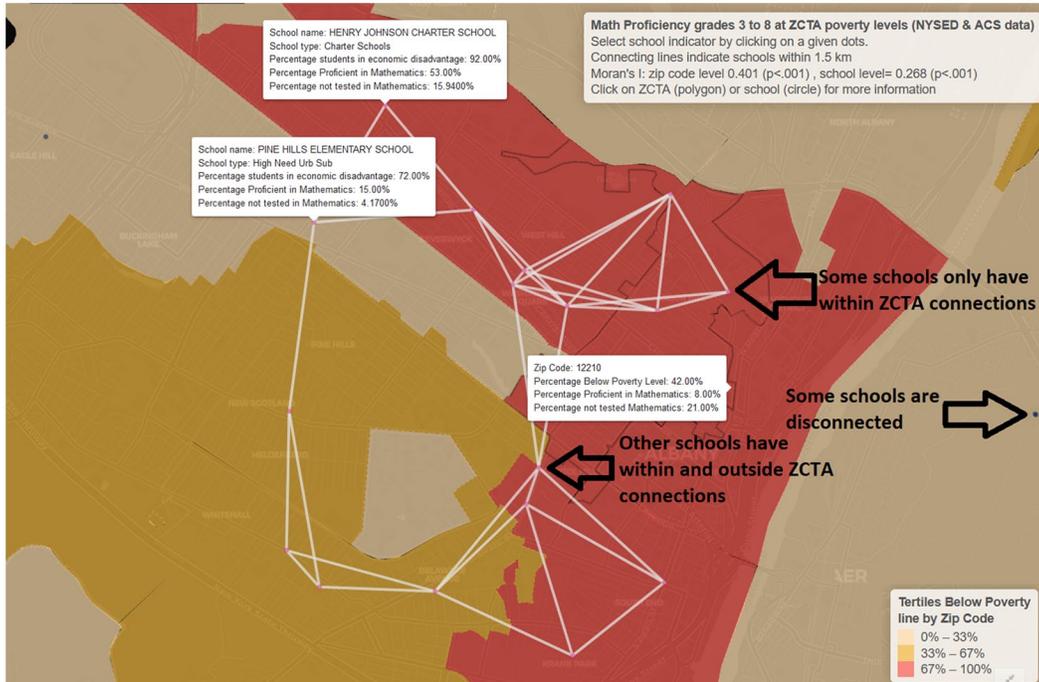


FIGURE 3. Example of identification of matrices of influence M and W to capture spatial dependence among zip codes and among schools that met the distance threshold (1.5 km), respectively. All dots represent a school. In this area of the state (Albany, New York), there are some schools without neighbors (dark blue dots) based on the radius-based threshold. Similarly, there are other schools with neighbors only within their ZCTA and other schools have neighbors both within and outside their ZCTAs. Interactive map available at <https://cutt.ly/54Odkyh>.

student population, these students accounted for about 28% and 38% of all students attending schools in the highest poverty areas, whereas their highest represented White peers (accounting for 47.53% of all students) were less represented than Black and Hispanic students in these same highest poverty zones (with about 21%).

Also in alignment with geography of disadvantage tenets, students participating in free and reduced-price lunch programs were concentrated in the highest poverty areas (75.93%, $SD = 14.38$). These schools also had the highest suspension rates, double those of the lowest poverty schools (3.82% vs. 1.58%).

When observing the categorical school-level features included in the models, also in Table 3, the needs index reported by the NYSED (2020b) shows that 51% of all schools located in NYC are also located in the highest poverty areas. The second most represented school type in the highest poverty areas is *charter schools*. Notably, although these schools represented only 8% of the analytic sample, 71% of these charter schools (2,159/3,054) were located in the highest poverty zones in our analytic sample. To put these descriptive findings into perspective, Figure 3 shows a visualization of these structures. This figure shows a charter school named Henry Johnson, located in a highest poverty zone in Albany, New York, and with 92% of the student body classified as economically disadvantaged. Despite this,

the mathematical proficiency observed in this school was 53% (i.e., over the average mean performance rate of 46.91%, $SD = 24.61$, discussed above). For comparison, the figure shows a neighboring school named Pine Hills Elementary, located in the lowest poverty tertile and with a lower proportion of its student body classified as economically disadvantaged (72%). In this latter school, only 15% of test-takers reached proficiency. In addition to validating the accuracy of this study’s identification strategy for MSAR, this representation aids with the identification of individual cases that may eventually help yield global statements about the expected performances when accounting for place-based, school-level, and students’ economically disadvantaged indicators. An interactive version of this visualization that includes all schools represented in the analytic sample can be accessed at <https://cutt.ly/54Odkyh>.

This brief discussion now transitions to a description of place-based indicators. Congruent with the geography of disadvantaged tenets, the Gini coefficient reflected the highest income inequality in schools located in the highest poverty zones; these schools also reflected the highest participation rates in the EITC program, with 48% ($SD = 7.68$) of inhabitants in these zip codes participating. Additionally, 42% ($SD = 17.3$) of households in these zones were classified as led by single mothers, which was almost triple this classification in zones with the lowest poverty rates (16%, $SD = 7.7$). Schools

in the highest poverty zones had the highest concentration of inhabitants ages 5–17 living in poverty (22.25%, $SD = 8.23$) and the highest percentage of inhabitants with a high school diploma as their highest education level (42.95%, $SD = 8.99$). Similarly, whereas the lowest poverty zones had median income levels that were 21% over the state’s median income, the highest poverty zones were 8.5% below the median state income.

With this brief description of the features that were found to be relevant predictors of outcome variation, it is time to discuss the main results of the MSAR models, presented in Table 4.

Multilevel Simultaneous Autoregressive Findings

Table 4 contains the results from the MSAR procedures. The posterior variances identified with the matrices of influence M and W discussed in equation (1) reached statistical significance across all models, corroborating the presence of significant variation at the zip code, school, and grade levels that needed to be modeled to avoid standard errors that were underestimated across the covariates included in the models (Dong & Harris, 2015). Moreover, as part of this analysis of spatial autocorrelation or dependence at different levels of nesting, Table 4 also presents the Moran’s I estimates at the ZCTA and at the school levels. All these Moran’s I s consistently reached statistical significance, indicating the presence of positive autocorrelation of the outcome of interest that needed to be modeled via MSAR. Once more, Figure 3 illustrates all these neighboring matrices of influence, where schools located in close proximity, regardless of whether they were within the same ZCTA, were considered neighbors (W matrix), and all contiguous ZCTAs (M matrix) were also considered sources of influence, as discussed in the methods section above.

To address the study’s research questions and to provide as comprehensive an analysis as possible, in addition to providing a model containing all schools and all students (“Aggregate” model in Table 4), six more models were estimated conditional on the intersection between students’ economic standing (i.e., economically disadvantaged and not economically disadvantaged) and their schools’ location poverty rates (i.e., low, middle, and high poverty rates of the ZCTAs where schools were physically located), as represented in Figure 1.

In accordance with the hypothetical relations and congruent with previous descriptive results, such as those presented by the NAEP (2020), the results corroborate that students from higher grades reached lower proficiency rates, with the largest gaps found in Grade 8. However, different from previous analyses on this topic, the results also test for and corroborate disparities by schools’ location poverty levels. That is, compared to the mathematical proficiency rates achieved in Grade 3, proficiency rates in Grade 4 decreased across all groups, yet the smallest magnitude of this decrease happened among students attending schools located in the lowest poverty zones. Specifically, Grade 4 students who attend

schools located in the lowest poverty areas had proficiency rates 3.44% (not economically disadvantaged) and 4.37% (economically disadvantaged) lower than the proficiency rates reached by their Grade 3 counterparts also attending schools located in the lowest poverty zones. Proficiency gaps in Grade 5 reached magnitudes over 10.6% in the highest poverty zones, with 11.41 and 10.67 percentage-point decreases among non–economically disadvantaged and economically disadvantaged Grade 5 students, respectively. These gaps continued to increase among Grade 6 students also in the highest poverty schools, reaching gaps of 13.05% and 10.99%.

The largest proficiency gaps so far were observed among non–economically disadvantaged Grade 7 students attending schools located in the highest poverty zones. These students were 17.79% less proficient than their non–economically disadvantaged Grade 3 counterparts attending schools located in high-poverty ZCTAs. For economically disadvantaged Grade 7 students also attending schools in high-poverty zones, this gap was 13.75%. Nonetheless, note that across Grades 4–7, students located in the lowest poverty zones realized the smallest proficiency drops. This trend was not true for Grade 8 students; in this case, the lowest gaps were found among economically disadvantaged and non–economically disadvantaged students attending the highest poverty zone schools, with 15.31% (economically disadvantaged) and 21.18% (not economically disadvantaged) reductions in proficiency rates compared to their Grade 3 counterparts. Overall, these findings suggest that, after accounting for two forms of spatial dependence, school-level indicators, and students’ own economic disadvantage status, there is a geography of mathematical disadvantage, wherein students attending lowest poverty areas tend to perform better than their peers attending schools in poorer zones, even though, as discussed below, these gaps represent different mathematical performance gains or drops (i.e., a gap of 5 percentage points is only meaningful when considering baselines of that group; hence the relevance of showing third performance rates by economic disadvantage status in Table 3).

In moving to the analyses of school-level indicators, the coefficients associated with percentage of ELLs consistently showed negative relationships across all models, with greater magnitudes realized in schools located in high poverty areas, reaching performance drops of at least 0.36 percentage points per each percentage point increase in ELL representation. Another consistently negative indicator was the percentage of students with disabilities attending a given school s . In this case, the highest coefficient magnitudes were found in schools located in the mid-poverty zones for non–economically disadvantaged students. In the case of economically disadvantaged students, this performance drop reached a magnitude of 0.5% for each increase in the percentage of students with disabilities in a given school.

With respect to the percentage of students who were not tested, results consistently indicated a negative relationship with mathematical proficiency. Regarding the racial/ethnic composition of the schools, Black students reached lower proficiency rates than their White counterparts. In the case of Asian students, the coefficients had positive signs, indicating that Asian students consistently reached higher proficiency rates than White students.

The indicator *percentage of free and reduced-price lunch* showed that students attending schools located in the highest poverty areas (economically disadvantaged and not economically disadvantaged) had higher performance rates to the extent that a higher proportion of the student body participated in this program. For example, if all students in these highest poverty area schools benefited from this program, the mathematical proficiency rates would have been expected to increase by at least 22 percentage points, which may serve to justify the current nationwide emphasis of making free lunch available to all students, regardless of need.¹⁹ Another consistently positive indicator was attendance rates, with greater magnitudes observed in high-poverty zones for economically disadvantaged and non-economically disadvantaged students (1.39% and 1.65%).

The school-level indicator of economic disadvantage was negative for the outcomes of all non-economically disadvantaged students. Notably, this indicator only affected economically disadvantaged students attending the schools in the highest poverty areas (0.46%)—that is, economically disadvantaged peers attending schools in the lowest and second lowest poverty zones were not affected by the presence of other economically disadvantaged peers attending those same schools.

The needs index classification considered schools located in NYC as the comparison group. These analyses revealed that schools located in other large cities and in high-poverty areas performed at least 10 percentage points higher in mathematical proficiency than their comparable NYC counterparts. Schools classified as average need performed statistically significantly worse than NYC schools in all but one instance. Non-economically disadvantaged test-takers attending average-need schools located in high-poverty areas performed statistically similarly to test-takers enrolled in these same types of schools but located in NYC. Finally, note that test-takers attending charter schools located in middle- and high-poverty areas performed better than those in similar schools in NYC. Charter school students in low-poverty areas performed similarly to their NYC peers, whether economically disadvantaged or not economically disadvantaged.

With respect to place-based features and despite Moran's *I* and Boruta feature relevance analyses, note that after this comprehensive level of analysis, after accounting for spatial interaction effects, few place-based indicators reached significant levels. For example, the Gini coefficient was statistically significant only among economically disadvantaged

and non-economically disadvantaged students enrolled in schools located in the lowest poverty zones but with different signs. Non-economically disadvantaged test-takers attending schools located in low-poverty areas showed increases in mathematical performance to the extent that more economic disparities were observed in those locations. For economically disadvantaged test-takers attending schools located in low-poverty areas, this index of economic disparities was negatively associated with their mathematical performance. The EITC indicator reached statistical significance levels only in the highest poverty areas for economically disadvantaged and non-economically disadvantaged students, and this coefficient was consistently negatively associated with mathematical proficiency.

At the county level, the estimate of median household income was positively associated with mathematical proficiency in all but the models accounting for schools located in high-poverty areas, and these findings were consistent for economically disadvantaged and non-economically disadvantaged test-takers. The death rate indicator, also measured at the county level, was negative and significant only for economically disadvantaged and non-economically disadvantaged students attending schools located in the highest poverty zones. Note that death rates are not necessarily crime- or violence-related but may capture systematic exposure to lower levels of access to health care services, infrastructure, and use (Baltrus et al., 2019) in different zones within the state. A counterintuitive finding was that the percentage of inhabitants ages 5–17 living in poverty in a given county had positive mathematical proficiency coefficients that reached statistically significant levels in five of the disaggregated models. In looking at these impacts, note that the highest positive effects were observed among test-takers attending schools located in the highest poverty areas. This result may indicate that such schools may potentially be (a) receiving more services; (b) more likely to be charter schools, which, as shown in the descriptive statistics, tended to be located in high-poverty zones; or (c) a combination of points (a) and (b).

Robustness and Sensitivity Checks

As indicated above, the data set analyzed accounts for the 2017–2018 and 2018–2019 academic years. However, all models and summary statistics described so far were reestimated by academic year, corroborating the findings just described. For transparency, in addition to these disaggregated tables for the 2017–2018 and 2018–2019 analyses in Tables 5–8, with Tables 5 and 7 showing summary statistics and Tables 6 and 8 containing MSAR models, three sets of replication codes are available (see “Replication_data_code_math_2018.R” at <https://cutt.ly/H4jsMe4>, “Replication_data_code_math_2019.R” at <https://cutt.ly/U4js6pk>, and “Replication_data_code_math_2018-2019.R” at <https://cutt.ly/l4jdpfY>).

TABLE 5

Summary Statistics 2017–2018 Analytic Sample

	Total (3,335 schools)	Lowest pov. tertile (1,152 schools)	Mid pov. tertile (1,048 schools)	Highest pov. tertile (1,135 schools)	<i>p</i> value
Outcomes: Mean (SD)					
math proficiency all grades	45.87 (24.78)	52.69 (23.76)	43.28 (23.16)	41.77 (25.88)	<0.001
math3 g Not Econ Dis g	65.88 (20.41)	72.09 (16.48)	64.59 (18.59)	60.13 (24.09)	<0.001
math3 g Econ Dis g	46.03 (21.13)	48.25 (20.45)	43.97 (19.92)	45.78 (22.79)	<0.001
Continuous or numeric features: Mean (SD)					
pct students not tested g	18.01 (20.94)	27.46 (23.85)	18.66 (19.65)	8.20 (13.42)	<0.001
pct Economically Disadvantaged s*	58.16 (24.98)	35.21 (20.42)	58.79 (16.61)	79.83 (13.54)	<0.001
attendance rate s	94.52 (2.29)	95.42 (1.55)	94.58 (1.78)	93.59 (2.91)	<0.001
pct teacher turnover s	2.55 (9.60)	0.72 (5.45)	1.66 (8.18)	5.19 (12.91)	<0.001
pct teacher turnover 5 years s	2.55 (9.86)	0.68 (5.18)	1.61 (8.08)	5.29 (13.59)	<0.001
pct Native American s	0.67 (2.83)	0.21 (0.98)	0.61 (2.03)	1.18 (4.27)	<0.001
pct Black s	16.96 (23.54)	8.50 (13.02)	14.68 (24.57)	27.39 (26.38)	<0.001
pct Hispanic s	23.80 (23.75)	18.34 (15.85)	14.99 (18.93)	37.69 (27.76)	<0.001
pct Asian s	7.77 (13.34)	8.35 (11.26)	4.98 (10.61)	9.91 (16.71)	<0.001
pct White s	47.99 (36.19)	61.22 (26.10)	61.67 (35.61)	21.80 (30.35)	<0.001
pct Multi Ethnicity s	2.75 (2.91)	3.30 (2.32)	3.02 (3.00)	1.95 (3.15)	<0.001
pct Women s	49.06 (4.89)	48.87 (3.47)	48.95 (3.42)	49.36 (6.88)	<0.001
pct with Disabilities s	17.33 (6.45)	15.97 (6.42)	16.77 (5.63)	19.21 (6.79)	<0.001
pct free-reduced lunch s	54.43 (24.76)	31.91 (20.05)	55.20 (16.62)	75.55 (14.37)	<0.001
pct suspensions s	2.90 (4.72)	1.60 (3.12)	3.28 (4.57)	3.78 (5.77)	<0.001
Gini z	0.45 (0.06)	0.43 (0.06)	0.43 (0.05)	0.47 (0.05)	<0.001
pct EITC recipients z	34.89 (12.59)	22.84 (6.99)	33.58 (6.96)	47.86 (7.80)	<0.001
pct mother only household z	28.16 (16.85)	16.02 (7.76)	26.33 (12.07)	41.72 (17.48)	<0.001
pct crime z	3.27 (1.49)	2.90 (1.33)	3.71 (1.13)	3.19 (1.81)	<0.001
pct children in poverty 5 to 17 c	17.83 (7.61)	13.87 (5.99)	17.23 (5.80)	22.27 (8.23)	<0.001
death rate c	8.50 (1.53)	8.40 (1.13)	9.10 (1.63)	8.03 (1.58)	<0.001
net migration rate c	-5.93 (5.85)	-3.50 (4.35)	-5.26 (5.87)	-8.93 (5.81)	<0.001
pct High school or less c	40.54 (8.03)	35.69 (5.85)	42.76 (6.58)	43.07 (8.98)	<0.001
pct unemployment c	4.27 (0.67)	3.92 (0.38)	4.43 (0.65)	4.47 (0.77)	<0.001
Median household income c	1.30 (29.72)	20.60 (29.93)	-7.60 (23.24)	-8.77 (25.80)	<0.001
pct taxpayers below poverty line z	0.36 (0.09)	0.26 (0.04)	0.35 (0.02)	0.46 (0.06)	<0.001
mean number of students grade g	55.36 (57.61)	62.55 (62.94)	47.86 (43.92)	55.67 (62.80)	<0.001
Categorical features: n (%)					
math3 g**	3,969.00 (21.08%)	1,389.00 (22.38%)	1,318.00 (21.15%)	1,262.00 (19.74%)	0.001
math4 g**	3,893.00 (20.67%)	1,365.00 (21.99%)	1,307.00 (20.98%)	1,221.00 (19.10%)	<0.001
math5 g**	3,747.00 (19.90%)	1,284.00 (20.69%)	1,203.00 (19.31%)	1,260.00 (19.71%)	0.14
math6 g**	2,680.00 (14.23%)	846.00 (13.63%)	889.00 (14.27%)	945.00 (14.78%)	0.2
math7 g**	2,432.00 (12.91%)	730.00 (11.76%)	795.00 (12.76%)	907.00 (14.19%)	<0.001
math8 g**	2,111.00 (11.21%)	593.00 (9.55%)	719.00 (11.54%)	799.00 (12.50%)	<0.001
Not Econ Disadvantaged***	9,372.00 (49.77%)	3,139.00 (50.57%)	3,105.00 (49.83%)	3,128.00 (48.92%)	0.2
Large cities s	771 (4.09%)	149 (2.40%)	139.00 (2.23%)	483.00 (7.55%)	<0.001
High Need Urban or Suburban s	1,416 (7.52%)	136 (2.19%)	576.00 (9.24%)	704.00 (11.01%)	<0.001
NYC public s	5,588 (29.67%)	1,160 (18.69%)	1,180.00 (18.94%)	3,248.00 (50.80%)	<0.001
High Need Rural s	1,803 (9.57%)	43 (0.69%)	1,072.00 (17.20%)	688.00 (10.76%)	<0.001
Average Need s	5,902 (31.34%)	2,846 (45.85%)	2,814.00 (45.16%)	242.00 (3.78%)	<0.001
Low Need s	1,883 (10.00%)	1,738 (28.00%)	145.00 (2.33%)	0.00 (0.00%)	<0.001
Charter School s	1,469 (7.80%)	135 (2.17%)	305.00 (4.89%)	1,029.00 (16.09%)	<0.001
N of grades (g)	18,832	6,207 (32.96%)	6,231 (33.09%)	6,394 (33.95%)	Not applicable
Total no. of students	1,042,452	388,242	298,233	355,977	Not applicable

Note. *SD* = standard deviation.

*In addition to this school level indicator, each grade contains proficiency levels disaggregated by students economically disadvantaged status.

**number of grades that are classified in 3, 4, 5, 6, 7, or 8 grades in the analytic sample

***number of analytic units that are classified as not being economically disadvantaged

TABLE 6

Multilevel SAR Models for Posterior Means of Mathematical Proficiency Rates 2017–2018 Analytic Sample

	Aggregate	Not economically disadvantaged			Economically disadvantaged		
		Low pov	Mid pov	High pov	Low pov	Mid pov	High pov
Intercept	-78.389*** (7.054)	-11.676 (9.411)	-18.903* (9.383)	-18.834* (9.235)	-8.992 (9.651)	-30.341*** (9.314)	-25.128*** (8.892)
math4 g	-5.493*** (0.377)	-4.186*** (0.782)	-6.001*** (0.907)	-4.748*** (1.116)	-5.210*** (0.917)	-7.081*** (0.812)	-6.317*** (0.858)
math5 g	-10.329*** (0.375)	-7.248*** (0.795)	-11.522*** (0.927)	-12.365*** (1.095)	-7.938*** (0.961)	-12.126*** (0.816)	-12.052*** (0.861)
math6 g	-9.712*** (0.443)	-5.112*** (0.983)	-8.974*** (1.071)	-13.948*** (1.235)	-6.203*** (1.138)	-8.783*** (0.939)	-12.167*** (0.962)
math7 g	-13.606*** (0.466)	-8.200*** (1.086)	-13.906*** (1.140)	-18.311*** (1.248)	-9.032*** (1.234)	-13.888*** (1.020)	-14.326*** (0.990)
math8 g	-22.349*** (0.533)	-27.373*** (1.339)	-27.801*** (1.329)	-22.587*** (1.390)	-20.748*** (1.431)	-19.696*** (1.118)	-16.356*** (1.074)
not Economically Disadvantaged g	18.305*** (0.249)	NA (NA)	NA (NA)	NA (NA)	NA (NA)	NA (NA)	NA (NA)
attendance rate s	1.311*** (0.069)	0.840*** (0.127)	0.882*** (0.116)	1.483*** (0.134)	0.616*** (0.135)	0.852*** (0.108)	1.244*** (0.111)
pct teacher turnover s	0.09 (0.060)	0.264 (0.221)	-0.12 (0.200)	0.213 (0.121)	-0.016 (0.265)	0.175 (0.176)	0.223*** (0.093)
pct teacher turnover 5 years s	0.08 (0.052)	0.1 (0.231)	0.24 (0.185)	0.021 (0.097)	0.296 (0.273)	0.081 (0.164)	0 (0.075)
pct ELL s	-0.318*** (0.024)	-0.222*** (0.073)	-0.186*** (0.073)	-0.360*** (0.051)	-0.347*** (0.087)	-0.267*** (0.063)	-0.309*** (0.040)
pct Native American s^	-0.05 (0.053)	-0.685** (0.312)	0.059 (0.166)	-0.095 (0.093)	-0.755* (0.364)	0.233 (0.143)	-0.031 (0.078)
pct Black s^	-0.135*** (0.014)	-0.123*** (0.038)	-0.169*** (0.030)	-0.208*** (0.043)	-0.116** (0.046)	-0.083*** (0.026)	-0.071* (0.036)
pct Hispanic s^	-0.035** (0.015)	0.005 (0.041)	-0.069 (0.038)	-0.048 (0.043)	-0.04 (0.048)	-0.006 (0.034)	0.034 (0.035)
pct Asian s^	0.237*** (0.018)	0.163*** (0.035)	0.242*** (0.050)	0.217*** (0.053)	0.373*** (0.043)	0.334*** (0.044)	0.348*** (0.043)
pct Multi Ethnicity s^	-0.08 (0.064)	0.047 (0.150)	0.261* (0.129)	-0.11 (0.189)	-0.472*** (0.181)	-0.01 (0.112)	-0.378*** (0.154)
pct Women s	0.105*** (0.027)	-0.052 (0.083)	0.123 (0.095)	0.101 (0.055)	0.18 (0.093)	0.113 (0.080)	0.04 (0.043)
pct Economically Disadvantaged s	-0.345*** (0.037)	-0.304*** (0.096)	-0.248*** (0.094)	-0.665*** (0.093)	-0.075 (0.114)	-0.095 (0.081)	-0.612*** (0.075)

(continued)

TABLE 6. (CONTINUED)

	Not economically disadvantaged			Economically disadvantaged			
	Aggregate	Low pov	Mid pov	High pov	Low pov	Mid pov	High pov
pct with Disabilities s	-0.374*** (0.027)	-0.428*** (0.058)	-0.457*** (0.069)	-0.324*** (0.074)	-0.316*** (0.072)	-0.464*** (0.058)	-0.473*** (0.059)
Large cities s^^	1.66 (1.148)	-7.030*** (2.727)	1.131 (2.825)	7.145*** (2.360)	-10.711*** (3.127)	-2.325 (2.523)	7.189*** (2.070)
High-Need Urban/Suburban s^^	-3.442*** (1.047)	-9.899*** (2.784)	-5.591*** (2.302)	3.128 (2.368)	-7.145*** (3.187)	-6.654*** (2.015)	2.374 (1.954)
High-Need Rural s^^	-6.042*** (1.085)	-1.624 (3.972)	-6.935*** (2.286)	-4.024 (2.871)	-7.064 (4.552)	-8.708*** (2.039)	-3.646 (2.468)
Average Need s^^	-5.567*** (0.984)	-6.019*** (1.851)	-5.614*** (2.224)	-3.829 (2.966)	-11.914*** (2.084)	-8.334*** (1.982)	-4.22 (2.609)
Low Need s^^	-7.209*** (1.077)	-8.109*** (1.858)	-3.043 (3.032)	-0.061 (9.991)	-12.890*** (2.103)	-8.636*** (2.736)	-0.083 (10.043)
Charter School s^^	9.152*** (1.309)	-2.313 (4.093)	12.077*** (3.279)	8.695*** (2.674)	7.704 (4.527)	7.973*** (2.920)	10.050*** (2.171)
pct free-reduced lunch s	0.213*** (0.035)	0.117 (0.095)	0.06 (0.092)	0.417*** (0.074)	-0.05 (0.112)	-0.002 (0.080)	0.345*** (0.060)
pct suspensions s	-0.133*** (0.037)	-0.531*** (0.172)	-0.215*** (0.089)	-0.197*** (0.087)	0.04 (0.117)	-0.111 (0.073)	-0.160*** (0.068)
mean number of students grade g	0.028*** (0.002)	0.013*** (0.005)	0.019 (0.011)	0.024 (0.020)	-0.018 (0.009)	-0.012 (0.007)	-0.002 (0.005)
Gini z	-1.04 (3.729)	11.750* (5.845)	6.025 (6.218)	-10.852 (8.137)	-0.014 (6.427)	-9.53 (5.579)	-9.708 (7.889)
pct EITC recipients z	0.00 (0.031)	-0.06 (0.079)	-0.027 (0.076)	-0.174 (0.100)	-0.092 (0.095)	-0.038 (0.067)	-0.156 (0.090)
pct mother only household z	-0.03 (0.019)	-0.016 (0.054)	-0.021 (0.036)	-0.049 (0.044)	-0.109 (0.069)	-0.069*** (0.031)	-0.038 (0.040)
pct crime z	-0.06 (0.190)	0.194 (0.488)	0.396 (0.410)	0.153 (0.350)	0.319 (0.550)	-0.071 (0.357)	-0.152 (0.326)
pct children in poverty 5-17 c	0.383*** (0.062)	0.324*** (0.122)	0.382*** (0.124)	0.738*** (0.166)	0.102 (0.143)	0.371*** (0.108)	0.572*** (0.153)
death rate c	-0.30 (0.237)	0.635 (0.606)	-0.012 (0.374)	-2.798*** (0.817)	-0.335 (0.682)	0.214 (0.327)	-2.304*** (0.701)
net migration rate c	-0.09 (0.058)	-0.163 (0.143)	-0.119 (0.097)	0.123 (0.193)	-0.142 (0.160)	-0.146 (0.086)	0.117 (0.168)
pct High school or less c	-0.04 (0.047)	-0.187 (0.115)	-0.15 (0.086)	-0.116 (0.133)	-0.248 (0.132)	-0.029 (0.074)	-0.028 (0.120)

(continued)

TABLE 6. (CONTINUED)

	Aggregate	Not economically disadvantaged			Economically disadvantaged		
		Low pov	Mid pov	High pov	Low pov	Mid pov	High pov
pct unemployment c	0.27 (0.557)	2.036 (1.649)	1.807 (0.985)	-2.672 (1.522)	4.085** (1.839)	0.237 (0.874)	-0.61 (1.340)
Median household income c	0.107*** (0.017)	0.142*** (0.033)	0.158*** (0.037)	0.055 (0.060)	0.081** (0.038)	0.097*** (0.032)	0.061 (0.053)
pct students not tested g	-0.151*** (0.009)	-0.206*** (0.019)	-0.174*** (0.021)	-0.182*** (0.033)	-0.066*** (0.021)	-0.069*** (0.018)	-0.117*** (0.028)
Pseudo R ²	0.542	0.525	0.451	0.464	0.373	0.454	0.575
Lambda M	13.63*** (1.405)	12.042*** (3.009)	7.597** (3.494)	10.606*** (4.596)	6.104*** (4.638)	5.243** (2.315)	16.462*** (3.851)
Rho W	282.349*** (3.521)	209.663 (6.749)	271.646 (8.281)	390.568*** (11.176)	290.836*** (9.068)	210.139*** (6.060)	242.688*** (7.852)
Moran's I Regions^^^	0.369*** (-0.001)	0.281*** (-0.002)	0.054 (0.002)	0.397*** (0.005)	0.267** (-0.003)	0.228*** (-0.002)	0.528** (-0.005)
Expectation	0.248*** (-0.00005)	0.315*** (-0.0003)	0.271*** (-0.0003)	0.21*** (-0.0003)	0.3181*** (-0.0003)	0.363*** (-0.0003)	0.29*** (-0.0003)
N of schools and grades	18,832	3,139	3,105	3,128	3,068	3,126	3,266
No. of students	1,042,452	259,379	120,030	68,994	128,863	178,203	286,983
Δ	[18832, 1019]	[3139, 385]	[3105, 415]	[3128, 220]	[3068, 376]	[3126, 415]	[3266, 220]
M	[1019, 1019]	[385, 385]	[415, 415]	[220, 220]	[376, 376]	[415, 415]	[220, 220]
W	[18832, 18832]	[3139, 3139]	[3105, 3105]	[3128, 3128]	[3068, 3068]	[3126, 3126]	[3266, 3266]

Note. EITC = Eamed Income Tax Credit; ELL = English language learner; MSAR = multilevel simultaneous autoregressive; NYC = New York City.

^ Comparison group is White students.

^^ Comparison is NYC schools.

^^^ The parameter space for these variables is to be (-1,1), and these tests were conducted independently from MSAR, as can be seen in the replication code.

*p < .05. **p < .01. ***p < .0001.

TABLE 7
Summary Statistics 2018–2019 Analytic Sample

	Total (3,372 schools)	Lowest pov. tertile (1,168 schools)	Mid pov. tertile (1,068 schools)	Highest pov. tertile (1,136 schools)	<i>p</i> value
Outcomes: Mean (SD)					
math proficiency all grades	47.92 (24.39)	54.65 (23.21)	44.83 (22.96)	44.40 (25.52)	<0.001
math3 g Not Econ Dis g	66.45 (19.44)	72.44 (15.44)	64.47 (18.07)	61.52 (22.99)	<0.001
math3 g Econ Dis g	47.03 (20.95)	49.62 (20.67)	44.21 (19.54)	47.15 (22.31)	<0.001
Continuous or Numeric Features: Mean (SD)					
pct students not tested g	15.59 (19.64)	25.29 (23.44)	15.22 (17.58)	6.50 (11.28)	<0.001
pct Economically Disadvantaged s*	57.66 (24.55)	34.83 (19.56)	58.62 (16.38)	79.00 (13.53)	<0.001
attendance rate s	94.54 (2.26)	95.43 (1.53)	94.58 (1.77)	93.64 (2.85)	<0.001
pct teacher turnover s	2.63 (9.63)	0.67 (5.07)	1.64 (7.99)	5.53 (13.17)	<0.001
pct teacher turnover 5 years s	2.61 (9.83)	0.63 (4.82)	1.57 (7.89)	5.57 (13.73)	<0.001
pct ELL s	7.81 (9.95)	4.81 (5.80)	5.00 (8.08)	13.50 (12.12)	<0.001
pct Native American s	0.71 (3.21)	0.24 (1.76)	0.65 (2.10)	1.24 (4.76)	<0.001
pct Black s	17.10 (23.66)	8.35 (12.93)	14.45 (24.25)	28.25 (26.64)	<0.001
pct Hispanic s	24.22 (23.90)	18.66 (16.00)	15.61 (19.48)	38.14 (27.77)	<0.001
pct Asian s	7.90 (13.46)	8.67 (11.73)	5.05 (10.62)	9.94 (16.69)	<0.001
pct White s	47.08 (36.07)	60.49 (26.08)	60.97 (35.58)	20.31 (29.34)	<0.001
pct Multi Ethnicity s	2.93 (3.03)	3.54 (2.40)	3.21 (3.16)	2.05 (3.24)	<0.001
pct Women s	49.07 (4.90)	48.88 (3.26)	48.99 (3.62)	49.33 (6.90)	<0.001
pct with Disabilities s	17.67 (6.49)	16.19 (6.19)	17.01 (5.71)	19.76 (6.95)	<0.001
pct free-reduced lunch s	54.70 (25.06)	31.68 (20.10)	55.63 (16.70)	76.25 (14.43)	<0.001
pct suspensions s	2.93 (4.70)	1.56 (2.58)	3.29 (4.64)	3.91 (5.89)	<0.001
Gini z	0.45 (0.06)	0.43 (0.06)	0.43 (0.05)	0.48 (0.05)	<0.001
pct EITC recipients z	34.99 (12.59)	22.81 (6.99)	33.78 (6.85)	48.08 (7.59)	<0.001
pct mother only household z	28.35 (16.92)	15.94 (7.72)	26.43 (12.00)	42.36 (17.24)	<0.001
pct crime z	3.27 (1.50)	2.91 (1.34)	3.71 (1.13)	3.19 (1.83)	<0.001
pct children in poverty 5 to 17 c	17.84 (7.60)	13.83 (5.98)	17.25 (5.73)	22.33 (8.23)	<0.001
death rate c	8.49 (1.53)	8.41 (1.13)	9.09 (1.65)	7.99 (1.56)	<0.001
net migration rate c	-5.97 (5.86)	-3.48 (4.33)	-5.43 (5.93)	-8.95 (5.78)	<0.001
pct High school or less c	40.53 (8.01)	35.74 (5.84)	42.92 (6.52)	42.84 (9.04)	<0.001
pct unemployment c	4.27 (0.67)	3.92 (0.39)	4.41 (0.65)	4.47 (0.76)	<0.001
Median household income c	1.36 (29.67)	20.70 (30.05)	-7.81 (22.80)	-8.47 (25.88)	<0.001
pct taxpayers below poverty line z	0.36 (0.09)	0.26 (0.04)	0.35 (0.02)	0.46 (0.06)	<0.001
mean number of students grade g	54.39 (56.12)	61.87 (61.34)	47.39 (43.95)	53.99 (60.43)	<0.001
Categorical features: n (%)					
math3 g**	4,050.00 (21.02%)	1,434.00 (22.59%)	1,347.00 (20.99%)	1,269.00 (19.51%)	<0.001
math4 g**	3,971.00 (20.61%)	1,378.00 (21.71%)	1,329.00 (20.71%)	1,264.00 (19.43%)	0.006
math5 g**	3,842.00 (19.94%)	1,311.00 (20.65%)	1,260.00 (19.64%)	1,271.00 (19.54%)	0.2
math6 g**	2,824.00 (14.66%)	882.00 (13.89%)	917.00 (14.29%)	1,025.00 (15.76%)	0.007
math7 g**	2,496.00 (12.95%)	753.00 (11.86%)	828.00 (12.91%)	915.00 (14.07%)	<0.001
math8 g**	2,086.00 (10.83%)	590.00 (9.29%)	735.00 (11.46%)	761.00 (11.70%)	<0.001
Not Econ Disadvantaged***	9,597.00 (49.81%)	3,214.00 (50.63%)	3,194.00 (49.78%)	3,189.00 (49.02%)	0.2
Large cities s	808.00 (4.19%)	154.00 (2.43%)	132.00 (2.06%)	522.00 (8.02%)	<0.001
High-Need Urban or Suburban s	1,450.00 (7.53%)	142.00 (2.24%)	579.00 (9.02%)	729.00 (11.21%)	<0.001
NYC public s	5,701.00 (29.59%)	1,170.00 (18.43%)	1,283.00 (20.00%)	3,248.00 (49.93%)	<0.001
High-Need Rural s	1,827.00 (9.48%)	46.00 (0.72%)	1,154.00 (17.99%)	627.00 (9.64%)	<0.001
Average Need s	5,930.00 (30.77%)	2,895.00 (45.60%)	2,794.00 (43.55%)	241.00 (3.70%)	<0.001
Low Need s	1,960.00 (10.17%)	1,806.00 (28.45%)	154.00 (2.40%)	0.00 (0.00%)	<0.001
Charter School s	1,585.00 (8.23%)	135.00 (2.13%)	320.00 (4.99%)	1,130.00 (17.37%)	<0.001
N of grades (g)	19,269	6,348 (32.94%)	6,416 (33.30%)	6,505 (33.76%)	Not Applicable
Total no. of students	1,048,061	392,760	304,075	351,226	Not Applicable

Note. EITC = Earned Income Tax Credit; ELL = English language learner; MSAR = multilevel simultaneous autoregressive; NYC = New York City; SD = standard deviation.

*In addition to this school-level indicator, each grade contains proficiency levels disaggregated by students' economically disadvantaged status.

**Number of grades that are classified in 3, 4, 5, 6, 7, or 8 grades in the analytic sample

***Number of analytic units that are classified as not being economically disadvantaged

TABLE 8

Multilevel SAR Models for Posterior Means of Mathematical Proficiency Rates Analytic Sample 2018–2019

	Not economically disadvantaged				Economically disadvantaged		
	Aggregate	Low pov	Mid pov	High pov	Low pov	Mid pov	High pov
Intercept	-82.602*** (7.092)	-11.679 (9.479)	-16.169 (9.333)	-17.83 (9.510)	-11.072 (9.462)	-29.149*** (9.197)	-23.726*** (9.032)
math4 g	-4.581*** (0.373)	-2.865*** (0.745)	-5.462*** (0.891)	-5.986*** (1.129)	-3.545*** (0.903)	-5.180*** (0.809)	-5.448*** (0.881)
math5 g	-8.802*** (0.382)	-6.318*** (0.757)	-10.740*** (0.920)	-10.575*** (1.131)	-7.459*** (0.943)	-9.756*** (0.817)	-9.578*** (0.873)
math6 g	-7.966*** (0.425)	-3.917*** (0.951)	-7.893*** (1.037)	-11.943*** (1.242)	-3.988*** (1.117)	-7.046*** (0.948)	-9.947*** (0.945)
math7 g	-12.351*** (0.465)	-8.118*** (1.077)	-13.122*** (1.110)	-17.221*** (1.252)	-6.831*** (1.204)	-12.238*** (1.015)	-13.280*** (1.013)
math8 g	-19.684*** (0.528)	-23.890*** (1.298)	-24.469*** (1.316)	-19.744*** (1.448)	-16.597*** (1.425)	-18.213*** (1.124)	-14.335*** (1.116)
not Economically Disadvantaged g	17.707*** (0.243)	NA (NA)	NA (NA)	NA (NA)	NA (NA)	NA (NA)	NA (NA)
attendance rate s	1.323*** (0.070)	0.734*** (0.125)	0.794*** (0.116)	1.695*** (0.142)	0.528*** (0.139)	0.831*** (0.111)	1.298*** (0.111)
pct teacher turnover s	0.133*** (0.056)	-0.398 (0.222)	0.012 (0.204)	0.277*** (0.113)	-0.198 (0.262)	0.437*** (0.175)	0.143 (0.085)
pct teacher turnover 5 years s	0.029*** (0.048)	0.486** (0.226)	0.133 (0.182)	-0.081 (0.093)	0.649*** (0.272)	-0.089 (0.160)	0.038 (0.071)
pct ELL s	-0.388*** (0.024)	-0.291*** (0.075)	-0.371*** (0.072)	-0.484*** (0.055)	-0.339*** (0.088)	-0.319*** (0.064)	-0.379*** (0.041)
pct Native American s^	-0.048*** (0.044)	-0.076 (0.158)	0.091 (0.161)	-0.142 (0.090)	0.068 (0.185)	0.178 (0.144)	-0.072 (0.066)
pct Black s^	-0.136*** (0.014)	-0.113*** (0.038)	-0.163*** (0.031)	-0.246*** (0.045)	-0.086 (0.046)	-0.100*** (0.027)	-0.116*** (0.036)
pct Hispanic s^	-0.036*** (0.015)	-0.001 (0.040)	-0.021 (0.040)	-0.096*** (0.045)	-0.031 (0.048)	-0.005 (0.036)	-0.023 (0.037)
pct Asian s^	0.232*** (0.018)	0.220*** (0.033)	0.222*** (0.050)	0.121*** (0.055)	0.372*** (0.042)	0.346*** (0.047)	0.299*** (0.045)
pct Multi Ethnicity s^	-0.065*** (0.062)	0.065 (0.144)	0.191 (0.129)	-0.248 (0.197)	-0.163 (0.176)	-0.008 (0.113)	-0.387*** (0.157)
pct Women s	0.089*** (0.026)	0.105 (0.083)	0.197** (0.090)	0.012 (0.056)	0.215** (0.098)	0.174** (0.076)	-0.008 (0.044)

(continued)

TABLE 8. (CONTINUED)

	Not economically disadvantaged			Economically disadvantaged			
	Aggregate	Low pov	Mid pov	High pov	Low pov	Mid pov	High pov
pct Economically Disadvantaged s	-0.226*** (0.028)	-0.161*** (0.070)	-0.141** (0.067)	-0.354*** (0.077)	-0.147 (0.083)	-0.114 (0.061)	-0.352*** (0.062)
pct with Disabilities s	-0.394*** (0.026)	-0.309*** (0.059)	-0.466*** (0.070)	-0.423*** (0.076)	-0.334*** (0.068)	-0.488*** (0.057)	-0.515*** (0.058)
Large cities s^^	3.419*** (1.171)	-6.445*** (2.653)	-6.369*** (2.988)	10.244*** (2.398)	-7.723*** (3.078)	-4.276 (2.722)	9.184*** (1.950)
High-Need Urban/Suburban s^^	-2.688*** (1.056)	-9.923*** (2.639)	-8.444*** (2.331)	5.674*** (2.366)	-7.652*** (3.095)	-8.576*** (2.120)	3.898*** (1.916)
High-Need Rural s^^	-5.972*** (1.092)	-4.755 (3.717)	-8.808*** (2.295)	-4.674 (2.991)	-6.288 (4.347)	-9.805*** (2.120)	-3.365 (2.465)
Average Need s^^	-5.252*** (0.975)	-5.487*** (1.757)	-6.556*** (2.308)	-3.594 (3.060)	-13.789*** (2.059)	-8.858*** (2.070)	-5.043 (2.522)
Low Need s^^	-6.212*** (1.085)	-6.245*** (1.776)	4.685 (3.063)	0.046 (9.839)	-14.334*** (2.063)	-7.370*** (2.831)	0.015 (9.960)
Charter School s^^	9.175*** (1.268)	6.244 (3.999)	8.843*** (3.473)	7.877*** (2.554)	3.475 (4.504)	4.801 (2.835)	10.596*** (2.066)
pct free-reduced lunch s	0.112*** (0.026)	-0.002 (0.067)	0.006 (0.067)	0.136*** (0.061)	-0.018 (0.078)	0.046 (0.060)	0.153*** (0.048)
pct suspensions s	-0.193*** (0.037)	-0.393*** (0.168)	-0.099 (0.093)	-0.311*** (0.087)	-0.139 (0.155)	-0.025 (0.073)	-0.240*** (0.067)
mean number of students grade g	0.027*** (0.002)	0.015*** (0.005)	0.021 (0.011)	0.004 (0.019)	-0.011 (0.009)	-0.017*** (0.007)	-0.006 (0.005)
Gini z	2.425*** (3.674)	16.736*** (5.805)	7.675 (6.277)	-8.878 (8.326)	-1.325 (6.404)	-7.124 (5.879)	-7.685 (7.962)
pct EITC recipients z	0.016*** (0.032)	-0.053 (0.079)	-0.038 (0.080)	-0.123 (0.110)	-0.127 (0.094)	-0.013 (0.073)	-0.121 (0.089)
pct mother-only household z	-0.032*** (0.019)	0.005 (0.054)	-0.04 (0.036)	-0.053 (0.049)	-0.118 (0.067)	-0.056 (0.033)	-0.024 (0.041)
pct crime z	-0.035*** (0.192)	0.42 (0.444)	0.079 (0.414)	-0.177 (0.390)	0.438 (0.553)	-0.047 (0.391)	0.004 (0.334)
pct children in poverty 5-17 c	0.412*** (0.061)	0.106 (0.122)	0.472*** (0.131)	0.739*** (0.179)	0.153 (0.140)	0.277*** (0.121)	0.580*** (0.151)
death rate c	-0.508*** (0.239)	0.19 (0.572)	0.204 (0.382)	-4.475*** (0.885)	0.166 (0.661)	-0.26 (0.357)	-3.685*** (0.716)

(continued)

TABLE 8. (CONTINUED)

	Not economically disadvantaged			Economically disadvantaged			
	Aggregate	Low pov	Mid pov	High pov	Low pov	Mid pov	High pov
net migration rate c	-0.077*** (0.059)	-0.26 (0.138)	-0.142 (0.099)	0.404 (0.208)	0.063 (0.161)	-0.14 (0.092)	0.241 (0.170)
pct High school or less c	-0.019*** (0.046)	-0.237*** (0.111)	-0.239*** (0.088)	0.003 (0.147)	-0.107 (0.133)	-0.048 (0.081)	0.037 (0.123)
pct unemployment c	0.316*** (0.570)	2.725 (1.553)	1.673 (0.963)	-3.578** (1.659)	4.323*** (1.917)	0.775 (0.914)	-0.409 (1.352)
Median household income c	0.123*** (0.017)	0.092*** (0.033)	0.192*** (0.041)	0.016 (0.067)	0.128*** (0.038)	0.101*** (0.036)	0.074 (0.053)
pct students not tested g	-0.150*** (0.009)	-0.166*** (0.019)	-0.191*** (0.024)	-0.235*** (0.036)	-0.085*** (0.021)	-0.073*** (0.021)	-0.134*** (0.031)
Pseudo R ²	0.528	0.483	0.418	0.451	0.398	0.449	0.556
Lambda M	15.127*** (1.608)	12.581*** (2.934)	10.077*** (3.559)	20.204*** (5.931)	17.399*** (4.656)	10.759*** (2.734)	16.126*** (4.070)
Rho W	279.93*** (3.495)	204.147*** (6.352)	276.766*** (7.924)	397.572*** (12.649)	285.289*** (9.234)	219.431*** (6.568)	252.277*** (8.499)
Moran's I Regions ^{^^^}	0.398*** (-0.0009)	0.316*** (-0.002)	0.104** (-0.002)	0.365*** (-0.005)	0.292*** (-0.002)	0.227*** (-0.002)	0.548*** (-0.005)
Expectation	0.261*** (-0.00005)	0.335*** (-0.00003)	0.303*** (-0.00003)	0.200*** (-0.00003)	0.368*** (-0.00003)	0.363*** (-0.00003)	0.282*** (-0.00003)
N of schools and grades	19,269	3,214	3,194	3,189	3,134	3,222	3,316
No. of students	1,048,061	261,756	122,094	72,198	131,004	181,981	279,028
Δ	[19269, 1021]	[3214, 393]	[3194, 418]	[3189, 213]	[3134, 383]	[3222, 417]	[3316, 213]
M	[1021, 1021]	[393, 393]	[418, 418]	[213, 213]	[383, 383]	[417, 417]	[213, 213]
W	[19269, 19269]	[3214, 3214]	[3194, 3194]	[3189, 3189]	[3134, 3134]	[3222, 3222]	[3316, 3316]

Note. EITC = Earned Income Tax Credit; ELL = English language learner; MSAR = multilevel simultaneous autoregressive; NYC = New York City; SAR = simultaneous autoregressive.

[^] Comparison group is White students.

^{^^} Comparison is NYC schools.

^{^^^} The parameter space for these variables was to be (-1,1), and these tests were conducted independently from MSAR, as can be seen in the replication code.

p* < .05. *p* < .01. ****p* < .0001.

Having said this, although all these results are included as robustness checks, my preferred specifications still remain those of the models that include both years of data. Conceptually and econometrically speaking, these time fixed-effects models allow the incorporation of *time* into the MSAR specification, an approach that the developers of MSAR recommended and implemented (see Dong et al., 2015). This methodological decision is important, for MSAR models are primarily cross-sectional, but I have demonstrated how researchers may include these temporal elements into their specification strategies. Having said that, in this case, it should be highlighted that the cross-sectional and time-effects models were consistent, but consistent with complex systems tenets that model life itself (see Gonzalez Canche, 2021), this may not necessarily be the case in models estimated with other data sources and that address different research questions.

Limitations: How Would Student Composition and Parents' Wealth Affect the Findings?

The analytic sample accounted for *all public schools at the state level*, including charter schools, which are public schools yet also different from public schools (see Fischler, 2021). Although the models included all available data from public schools, it is possible that families *who could afford* to move, looking for better public schooling options, may have done so for the betterment of their students' academic prospects or may have simply placed them in private schools. It could also be true that families with kids would have preemptively selected to live in "good public-school districts" rather than move to them, given the high costs associated with moving (Lindenfeld Hall, 2022; Pogol, 2019; Sahadi, 2015). Nonetheless, it is also possible that these parents may not have needed to move; instead, if they were located in high-poverty zones, they could have potentially enrolled their students at charter schools, which, as described above, tended to be overwhelmingly located in the highest poverty areas (with 71% of them located in such zones, as described in the findings section—see Table 3). Finally, another arguably less expensive strategy may have consisted of relying on private tutoring (online or in person) and/or on after-school programs.

If mobility is assumed to be the primarily driver that may ultimately influence students' performance as a source of unobserved heterogeneity, note that moving, attending a private school, and/or buying a house in "good" public-school districts is expensive, which may not reflect the decisions that the vast majority of parents in our analytic samples may have made—as has been widely reported (Lindenfeld Hall, 2022; Pogol, 2019; Sahadi, 2015). Nonetheless, following the rationale of compositional changes based on parents' decisions to *move* their students, given such parents' socioeconomic standing, the worst-case scenario that may have

translated into heightened compositional changes was supposed to affect grades and schools located in more affluent regions. This implies, then, that the set of models that should be "distrusted" the most are those measuring mathematical performance in the wealthiest neighborhoods in the analytic samples—after all, the types of students attending those areas would, *in theory*, have been coming from households that may have been able to afford to move into those areas and schools, precisely based on their known higher level of school quality. Alternatively, the mathematical performance in these wealthier zones may have been influenced by parents' decisions to send their students to private schools, which means that the models would not have captured these students' mathematical performance—that is, these wealthiest kids would likely not have been part of the analytic sample.

Despite these possibilities, the results of the models of students attending wealthier school districts are consistent with the results found among students attending schools located in less economically privileged neighborhoods. These models' agreements alleviate concerns that student compositional changes may have resulted in misleading findings or grossly erroneous conclusions. Nonetheless, unfortunately, I do not have access to data that would have captured the specific movement of students from poor areas to enroll at schools located in wealthier zones (or movement from public non-charter to charter schools or participation in after-school programs or private tutoring). Indeed, I am not familiar with any available data set that is detailed enough to capture such a movement (or such parental strategies to improve their students' mathematical ability). Having said this, although this movement analysis was not a goal in this study, it motivated me to account for net outmigration as part of the control indicators included in the models. Moreover, despite those idiosyncratic decisions made by families, I am confident that because moving and attending private school are expensive, the models shown in this study captured *most of the experiences of typical students* rather than those experiences of privileged families who were embarking in this strategic movement and school choice. As with any statistical model, outlier cases may be present, and the experiences of these atypical and privileged cases may surely have diverged from the overwhelming majority reflected in the analytic samples, which should not disqualify this paper or my inferences. Having said this, despite the fact that I tested and accounted for two different forms of spatial dependence present in the analytic samples, which, if left unmodeled, may have been sources of bias I am not claiming causality or absolute lack of bias. As a matter of fact, given the untestable set of assumptions required to claim causality, I strongly refrain from making causal claims in education and in social science research.

Before listing another limitation of the study, it should be noted that researchers interested in mapping and estimating

the impact of this student mobility (and/or participation in private tutoring or after-school programs) could do so with individual-level panel data access. However, I do not have such data sources, despite having requested such access. Nonetheless, reliance on publicly available administrative data has allowed me to showcase how applied researchers may leverage these data to offer innovative and nuanced estimates of factors affecting mathematical proficiency that are fully reproducible; all data and code can be accessed at <https://cutt.ly/N4zRstL>.

Another limitation of this study is that although economic disadvantage encompasses a wide array of factors that may capture students' economic challenges, and although students may experience more than one of these indicators, the NYSED data do not provide levels of economic disadvantage. From this view, rather than having a single dichotomic indicator, when possible, future studies should provide more nuanced measurements (perhaps categorical or even numeric) of the economic conditions that may influence students' prospects of remaining mathematically proficient. As depicted in this study, even though the analyses and findings are informative in capturing performance gaps based on this binary indicator, to the extent that the geography of mathematical disadvantage is prevalent—that is, to the extent that students experience a concentration of these economic challenges at the individual, school, and geographical levels—their opportunities for upward mobility will surely decline.

Methodological Discussion and Implications: How Specific Should Nesting Structures Be When Not Relying on Multilevel Simultaneous Autoregressive to Address Residual Autocorrelation?

As discussed in the methods section, multilevel models are the standard to address nesting structures that may affect outcome variation over and above individual-level attributes. From this perspective, González Canché (2014, 2021, 2023b) and Wolf et al. (2021) demonstrated that when the nesting structure is *specific enough*, model residuals obtained from multilevel models (i.e., not spatial multilevel models) will successfully address outcome dependence that may then result in spatial outcome dependence or autocorrelation (see González Canché, 2023b). However, as these nesting structures become less specific (i.e., moving from nesting at the classroom, to nesting at the school, to nesting at the school district, to nesting at the county or state level), spatial outcome dependence or autocorrelation may result in the presence of residual dependence—that is, in not identically and independently distributed (iid) residuals.

With this understanding, one of the study's reviewers encouraged me to demonstrate whether these patterns were realized when analyzing the public administrative data employed in this study, with an emphasis on comparing not

only whether model residuals became iid but also on how model coefficients differed or remained congruent with the model coefficients found with MSAR. Table 9 addresses this recommendation, including the aggregate model shown in Table 4 as well as two nonspatial multilevel models: one nesting classrooms at the school level and another where the nesting structure was conducted at the ZCTA level, hence matching the two levels used in every MSAR specification employed in this study.

Table 9 shows that if the nesting of grade levels was identified or operationalized at the school level, residual dependence became iid. On the other hand, when this nesting structure was specified at the ZCTA level, the models once again suffered from residual dependence—that is, the resulting model residuals were not iid. These findings and patterns were then congruent with past research (González Canché, 2014, 2021, 2023b; Wolf et al., 2021): When the nesting structures were less specific, model residuals would remain an issue in nonspatial multilevel models.

With respect to model coefficients, although the less specific nesting structure suffered from residual dependence (nested ZCTA model in Table 9—see also González Canché, 2014), some of its coefficients were closer in magnitude to the coefficients found via MSAR. For example, gaps between Grades 3 and 8 math had a magnitude of 20.87 and 20.83 percentage points in the MSAR and the nested ZCTA models, respectively, but this magnitude in the Nested School model indicated a gap between students in Grades 3 and 8 of only 17.31 percentage points. Other coefficients, however, resulted in notable differences. For example, the percentage of single-mother households was statistically significant in both multilevel models, but not in the MSAR specification, which may have been a function of the former not incorporating any spatial information as part of the model estimation. From this perspective, when considering that MSAR models were capable of (a) accounting for two forms of nesting structures, (b) accounting for fixed effects via the matrix Δ that captured schools' ascription to ZCTAs, and (c) rendering residuals that were iid, I consider the MSAR estimates to have been more robust to multiple sources of bias.

Based on this brief discussion, multilevel-alone models (i.e., not spatial multilevel models) may have either suffered from residual dependence (like the ZCTA nesting structure model) or rendered coefficients that were arguably downwardly biased (i.e., like the school nesting structure model). Having said this, it should be noted that despite the statistical benefits associated with MSAR, these models remained cumbersome to be operationalized and executed. This complexity was based on the need to create three matrices of influence (Δ , M , and W), which required advanced feature engineering. Although all code for the complete model replication is provided (see <https://cutt.ly/N4zRstL>), note also that González Canché (2023b) developed low-code functions²⁰ that automatically created all matrices

TABLE 9
Comparison Coefficient Magnitudes and Residual Dependence Issue of MSAR and Multilevel-Alone Models With Divergent Nesting Structures

	MSAR aggregate [§]	Nested schools 3,484	Nested ZCTA 1,028
Intercept	-107.190*** (6.007)	-133.068*** (13.102)	-141.455*** (8.020)
math4 g	-4.999*** (0.266)	-4.804*** (0.217)	-4.948*** (0.248)
math5 g	-9.483*** (0.271)	-9.165*** (0.224)	-9.455*** (0.253)
math6 g	-8.735*** (0.307)	-7.150*** (0.287)	-9.024*** (0.291)
math7 g	-12.795*** (0.328)	-10.408*** (0.319)	-13.237*** (0.311)
math8 g	-20.869*** (0.372)	-17.308*** (0.367)	-20.833*** (0.364)
not Economically Disadvantaged g	17.789*** (0.174)	17.499*** (0.141)	17.382*** (0.162)
Academic year 2018–2019_s	1.580*** (0.174)	1.623*** (0.143)	1.518*** (0.163)
attendance rate s	1.502*** (0.057)	1.861*** (0.123)	2.045*** (0.067)
pct teacher turnover s	0.084*** (0.040)	0.089 (0.090)	0.07 (0.039)
pct teacher turnover 5 years s	0.065 (0.035)	0.056 (0.077)	0.052 (0.034)
pct ELL s	-0.362*** (0.018)	-0.294*** (0.030)	-0.387*** (0.018)
pct Native American s [^]	-0.03 (0.038)	-0.066 (0.057)	-0.06 (0.044)
pct Black s [^]	-0.137*** (0.011)	-0.138*** (0.017)	-0.167*** (0.013)
pct Hispanic s [^]	-0.042*** (0.012)	-0.063*** (0.018)	-0.070*** (0.014)
pct Asian s [^]	0.219*** (0.014)	0.268*** (0.022)	0.269*** (0.016)
pct Multi Ethnicity s [^]	-0.022 (0.048)	-0.029 (0.071)	-0.025 (0.052)
pct Women s	0.108*** (0.019)	0.104*** (0.035)	0.137*** (0.019)

(continued)

TABLE 9. (CONTINUED)

	MSAR aggregate ^s	Nested schools 3,484	Nested ZCTA 1,028
pct Economically Disadvantaged s	-0.244*** (0.022)	-0.083** (0.029)	-0.273*** (0.023)
pct with Disabilities s	-0.376*** (0.020)	-0.252*** (0.033)	-0.366*** (0.020)
Large cities s^^	5.248*** (0.894)	-3.907** (1.414)	4.027*** (0.982)
High-Need Urban/Suburban s^^	-1.026 (0.828)	-9.343*** (1.226)	-3.152** (0.987)
High-Need Rural s^^	-4.829*** (0.878)	-11.615*** (1.296)	-6.789*** (1.155)
Average Need s^^	-4.440*** (0.761)	-9.009*** (1.123)	-4.403*** (0.984)
Low Need s^^	-6.101*** (0.857)	-9.212*** (1.260)	-3.927*** (1.099)
Charter School s^^	9.106*** (0.902)	7.754*** (1.941)	7.739*** (0.911)
pct free-reduced lunch s	0.146*** (0.021)	-0.058 (0.031)	0.164*** (0.021)
pct suspensions s	-0.136*** (0.026)	-0.251*** (0.048)	-0.140*** (0.026)
mean number of students grade g	0.028*** (0.001)	0.036*** (0.002)	0.025*** (0.002)
Gini z	-0.189 (3.264)	15.349*** (4.508)	14.829*** (5.080)
pct EITC recipients z	0.013 (0.027)	-0.057 (0.037)	-0.042 (0.038)
pct mother-only household z	-0.011 (0.016)	-0.067** (0.022)	-0.071** (0.024)
pct crime z	-0.086 (0.173)	-0.034 (0.187)	-0.38 (0.262)
pct children in poverty 5-17 c	0.407*** (0.051)	0.432*** (0.066)	0.541*** (0.077)
death rate c	-0.523*** (0.199)	0.412 (0.277)	-0.582* (0.290)
net migration rate c	-0.033 (0.051)	-0.259*** (0.067)	-0.320*** (0.073)

(continued)

TABLE 9. (CONTINUED)

	MSAR aggregate [§]	Nested schools 3,484	Nested ZCTA 1,028
pct High school or less c	0.022 (0.040)	-0.071 (0.053)	-0.006 (0.058)
pct unemployment c	0.053 (0.495)	1.617* (0.643)	0.038 (0.709)
Median household income c	0.114*** (0.014)	0.177*** (0.018)	0.145*** (0.021)
pct students not tested g	-0.142*** (0.007)	-0.191*** (0.007)	-0.159*** (0.007)
Lambda <i>M</i>	16.492*** (1.288)	NA	NA
Rho <i>W</i>	274.441*** (2.390)	NA	NA
Moran's <i>I</i> Regions ^{^^^}	0.401*** (-0.0009)	NA	NA
Expectation	0.268*** (-0.00002)	NA	NA
Moran's <i>I</i> Schools ^{^^^}	38,101	38,101	38,101
Expectation	2,090,513	2,090,513	2,090,513
<i>N</i> of schools and grades	[38101, 1028]	NA	NA
No. of students	[1028, 1028]	NA	NA
Δ	[38101, 38101]	NA	NA
<i>M</i>	Modeled with	-0.033	0.022***
<i>W</i>	MSAR		
Residual dependence			

Note. EITC = Eamed Income Tax Credit; ELL = English language learner; MSAR = multilevel simultaneous autoregressive; NYC = New York City; ZCTA = zip code tabulated area.

[^] Comparison group is White students.

^{^^} Comparison is NYC schools.

^{^^^} The parameter space for these variables was to be (-1,1), and these tests were conducted independently from MSAR, as can be seen in the replication code.

[§] This is the aggregated model shown in Table 4. Replication code starts on line 394 in the file called "Replication_data_code_math_2018-2019.R" located at <https://cutt.ly/N4zRstL>.

p* < .05. *p* < .01. ****p* < .0001.

required for these models and offered Moran's I tests. The goal of developing these low-code functions was to remove the inherent complexity of estimating MSAR models. Moreover, in addition to MSAR, González Canché (2023b) offered access to other low-code functions capable of testing for nongeographical stationarity in coefficient estimates. That is, it may be possible that some areas were more prone to being affected by local attributes than others, which could be empirically tested via geographically weighted regressions. This test, however, is beyond what can be accommodated in single study; accordingly, future research may address whether mathematical proficiency is prone to spatial heterogeneity or whether a single global estimate may suffice to capture the relationships observed in this study. Another possibility is to estimate spatio-temporal studies in the form of difference-in-differences that may compare how mathematical proficiency may have changed before and after the pandemic across low-income and non-low-income participants. In this respect, González Canché (2023b) also offered low-code functions to estimate these models, an example of which may also be found at González Canché (2018).

Discussion: The Geography of Mathematical (Dis) Advantage

The study's purpose was to offer nuanced estimates of proficiency rates in mathematics across grades while *simultaneously* considering test-takers' own economic standing and schools' contextual and place-based observed poverty levels. The results consistently indicate that proficiency rates decreased as cohorts of students advanced through elementary and middle schools, with the highest decreases observed in Grade 8. Note that although these drops in proficiency rates were congruent with the gaps reported by the NAEP (2020), this study's estimates show a grimmer picture and also provide evidence of a geography of mathematical disopportunity, wherein the interaction of economic disadvantage and place-based poverty results in the lowest change of remaining mathematical proficiency across grades, but with interesting and more nuanced patterns. For example, Grade 8 students who were economically disadvantaged and attended schools located in the highest poverty zones realized the lowest gaps among all students, with changes in proficiency rates of -15.31% with respect to their Grade 3 test-taker counterparts.

These findings convey mixed results. On the one hand, students who arguably were the most vulnerable (Pearman, 2019) had the lowest gaps with respect to their Grade 3 peers. On the other hand, these findings are not positive when observing that the mean mathematical proficiency rates of Grade 3 students attending schools in the highest poverty zones (see Table 3) were the second lowest among all Grade 3 students (46.62% , $SD = 22.47$). When considering these starting points, the 26.38 percentage-point decrease

in mathematical proficiency (standard error 0.912) observed among non-economically disadvantaged Grade 8 students attending schools in the lowest poverty zones (see Table 4), when compared to the 15.31 percentage-point decrease realized by Grade 8 students in high-poverty schools (standard error 0.768), implied that the former test-takers had proficiency rates that still remained 14.92 percentage points higher than those of economically disadvantaged Grade 8 students attending schools located in high-poverty zones. For the sake of clarity, let us elaborate on this 14.92 percentage-point gap among Grade 8 test-takers. Note that non-economically disadvantaged Grade 3 test-takers had mathematical proficiency levels of 72.27%, whereas economically disadvantaged Grade 3 test-takers reached mathematical proficiency levels of 46.62%. Considering this, the 26.38 percentage-point decrease observed among Grade 8 test-takers attending schools located in low-poverty areas led to a performance of 46.23% for these non-economically disadvantaged students (or $72.27 - 26.038 = 46.232$). In the case of their economically disadvantaged peers attending schools located in high-poverty areas, when considering that the Grade 3 performance was 46.62%, the lower magnitude observed in the Grade 8 coefficient (-15.31% , as shown in Table 4) indicates that these Grade 8 students reached mathematical performance rates of 31.31% (or $46.62 - 15.307 = 31.313$). Taking the differences of these two differences yields a performance gap among Grade 8 students of 14.92 percentage points (or $46.23 - 31.31$). That is, although the magnitude of the coefficients shown in Table 4 seemed to be worse for non-economically disadvantaged Grade 8 test-takers attending schools located in low-poverty zones, the ultimate mathematical performance of their economically disadvantaged Grade 8 peers attending schools located in high-poverty areas was still much more dire.

In sum, although proficiency gaps between economically disadvantaged students in Grades 3 and 8 attending schools located in high-poverty zones were the smallest across all groups, performance gaps between economically disadvantaged and non-economically disadvantaged Grade 8 students favored the latter. In going back to the hypothetical relationships, these findings partially corroborated the hypothesis that "economically disadvantaged students would likely experience greater losses in mathematical proficiency levels compared to their non-economically disadvantaged peers." As just depicted, although these losses were greater in magnitude for non-economically disadvantaged Grade 8 test-takers, the end result was the same: Economically disadvantaged students remained less proficient than their non-economically disadvantaged peers.

The consistent finding across model specifications between opting out from testing and mathematical proficiency reflected no evidence of purposeful exclusion of students from testing. That is, had these results reflected a positive association between percentage of students who

were not tested and performance in math, this could have been perceived as students who may not have been identified/self-identified as nonproficient deciding not to participate or as school administrators preferring that they not participate in this standardized testing process. Also aligned with the notion of the geography of mathematical disopportunity, overall, schools located in the lowest poverty zones (see Table 3) had many advantages: They had the lowest turnover rates, lowest suspension levels, lowest percentage of students classified as economically disadvantaged, and the highest median income at the county level, which may have translated into higher levels of funding based on property taxes. Their students also performed the best.

Considering the positive spillover effects of mathematical proficiency, programs and policies should strive for all schools, but particularly the most socioeconomically vulnerable schools and at younger ages, to increase or maintain their students' proficiency rates. In this respect, note that participation in free and reduced-price lunch programs was consistently associated with increases in mathematical proficiency levels in these poorest areas; as indicated above, current national movements are pushing for the adoption of healthy school lunches at public and not-for-profit public schools. Recent in-person class interruptions due to COVID-19 highlighted that these programs often represented the main source of food for students living in the highest poverty zones. Accordingly, this study offers evidence to reinforce the positive impact that these programs may have in mathematical proficiency prospects, regardless of students' economic health.

The comparison of NYC schools and charter schools reflected that proficiency rates in the latter were at least 8.7 percentage points higher in high-poverty areas for economically disadvantaged and non-economically disadvantaged students, with the highest gains observed among economically disadvantaged test-takers attending schools in high-poverty areas. Although charter schools are public schools, they are also different from public schools (Fischler, 2021). In this respect, it should be noted that non-charter public schools could potentially benefit from collaborating with or implementing programs similar to those implemented by their charter counterparts, particularly because 71% of these charter schools were located in the highest poverty areas and were succeeding at higher rates.

To close, because 80% of students attending the highest poverty level zones were also classified as economically disadvantaged, positive spillovers of mathematical proficiency indicate that actions geared toward increasing their mathematical competence may lead to increasing their performance in other academic areas (Duncan et al., 2007; Farran et al., 2006; Lerkkanen et al., 2005; Pagani et al., 2010). The intrinsic relationship between mathematics and STEM, along with the relevance of these fields in the present and future economies, suggests that to the extent that more students, particularly those from economically disadvantaged

backgrounds, become and remain proficient in mathematics, their prospects of success in life will increase (González Canché, 2017). Mathematical proficiency is a worthwhile pursuit with equitizing, gap-closing implications.

In terms of areas for future research, researchers may build from the concepts and data set provided in this study to assess the impact of the COVID-19 pandemic on mathematical proficiency for economically disadvantaged and non-economically disadvantaged students. This endeavor is worthwhile, for the NAEP (2022) reported that although mathematical proficiency rates nationwide between students in Grades 4 and 8 had widened by 10 percentage points, favoring the former, in New York State, there was no gap, with a 28% proficiency score in 2022 in Grades 4 and 8. However, these depictions may be misleading by failing to account for test-takers' economically disadvantaged statuses along with the poverty rates and locations of their schools. Accordingly, future studies may leverage MSAR difference-in-differences by incorporating data on performance observed after the pandemic and compare these performance rates between economically disadvantaged and non-economically disadvantaged test-takers. In addition to being beyond the scope of this study, due to data availability constraints—there has been no post-pandemic data release by the NYSED—I am unable to offer these estimates but such analyses can be implemented following the code provided with this study. Finally, considering that charter students in high-poverty areas are doing better than their district peers, future research may study whether certain families in high-poverty areas are seeking better educational alternatives for their children and document who these families are to more clearly explain the mechanism behind this outstanding performance in the highest poverty zones.

Acknowledgments

I am grateful for the research support and conceptual and methodological discussions provided by Chelsea Zhang (PhD candidate and master's in statistics candidate at the University of Pennsylvania; email: chelszh@upenn.edu) and Kaiwen Zheng (undergraduate student at the University of International Business and Economics in Beijing, China; email: 201921025@uibe.edu.cn).

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by grants offered by the Spencer Foundation and the National Academy of Education Postdoctoral Fellowship as well as by a grant on education data analytics sponsored by the National Science of Education.

ORCID iD

Manuel S. Gonzalez Canche  <https://orcid.org/0000-0001-8926-7614>

Notes

1. Based on the New York State Education Department (NYSED, 2020a), students scoring 3 or 4 in a 1–4 Common Core Learning Standardized test administered at public elementary and middle schools in the state of New York were considered proficient in mathematics.

2. Students classified as “economically disadvantaged” participate in assistance programs, such as free or reduced-price lunch, Social Security Insurance, food stamps, foster care, refugee assistance, EITC, Home Energy Assistance Program, Safety Net Assistance, Bureau of Indian Affairs, or Family Assistance: Temporary Assistance for Needy Families (NYSED, 2020a).

3. Please note that lunch eligibility as a proxy for poverty measure is becoming less useful, as there has been a recent trend in district-level participation in the national lunch program (see National School Lunch Program by the U.S. Department of Agriculture at <https://www.fns.usda.gov/nslp>). I am thankful for this comment, which was raised by one of this study’s reviewers. However, the summary statistics in the analytic samples still showed a high degree of variation at the aggregated levels and when disaggregating the analytic samples by poverty levels of the zones where schools were located. Specifically, these disaggregated estimates ranged from 31.76% in the least poor areas to 75.93% in the highest poverty zones. These variation levels, along with the identification of this indicator as a relevant feature via Boruta (see Figure 2), as described in the methods section, justified its inclusion in the models.

4. This estimate was obtained as the ratio of taxpayers in a given zip code with adjusted income levels up to \$25,000 over the total number of taxpayers in that same zip code. Considering that the state of New York indicated that the poverty line was \$32,402 as of 2016 (Chatterjee et al., 2019), inhabitants with incomes up to \$25,000 were considered to be at least \$7,500 below this poverty line.

5. Hence, participation in assistance programs, such as free or reduced-price lunch, Social Security Insurance, food stamps, foster care, refugee assistance, EITC, Home Energy Assistance Program, Safety Net Assistance, Bureau of Indian Affairs, or Family Assistance: Temporary Assistance for Needy Families (NYSED, 2020a).

6. All data and *R* codes required to fully replicate all models are available at <https://cutt.ly/N4zRstL>.

7. See https://www.huduser.gov/portal/datasets/usps_cross-walk.html.

8. The NYSED counted the number of students per grade who took these Common Core Learning Standardized tests.

9. All definitions of grade- and school-level indicators were retrieved from the NYSED and can be accessed at <https://data.nysed.gov/glossary.php?report=reportcards>.

10. Including students with intellectual disability, hearing impairment, speech or language impairment, visual impairment, serious emotional disturbance, orthopedic impairment, autism, traumatic brain injury, developmental delay, other health impairment,

specific learning disability, deaf-blindness, or multiple disabilities and who, by reason thereof, receive special education and related services.

11. Higher level covariates captured at the county level could be added, but the main nesting structure was zip code. Spatial dependence at the county level indicated no spatial clustering, relying on Moran’s *I*, discussed below.

12. To test for sensitivity of this 1.5-kilometer specification, the models were replicated with matrices *W* with distance thresholds of 2 and 2.5 kilometers rendering similar estimates. Another specification relied on a 0.6-mile threshold, as discussed by Chetty et al. (2020), who did not rely on matrices of influence, as in the case of Dong et al. (2015). Similarly, no differences in the estimates were found (because this specification represented a distance of .33 mile, because 1.5 kilometers = 0.93 mile), accordingly, the findings were assumed to not be sensitive to the variations discussed herein. All these specifications can be tested with the replication code provided at <https://cutt.ly/N4zRstL>.

13. The replication code identified $g \rightarrow s, z$, relying on network principles by creating a composite key identifier of grades in schools that then were ascribed to ZCTAs to extract the block diagonal design matrix Δ .

14. County-level indicators could be identified and would remain the same across zip codes that were located in a given county.

15. Note that Raftery and Lewis (1991) found that after 500 burn-in samples, models were stable. In this study, following Dong and Harris (2015), I increased the burn-in samples to 5,000, or 10 times the recommended value in Raftery and Lewis’s seminal paper.

16. This feature selection was implemented with the Boruta package in *R* (see Kursa & Rudnicki, 2010).

17. Robustness checks conducted with data disaggregated for the 2017–2018 and 2018–2019 academic years corroborated the relevance of all these indicators. These HTML Boruta figures may be accessed at <https://cutt.ly/l4DSvLT> and at <https://cutt.ly/r4DSUjy> for the 2017–2018 and 2018–2019 academic years, respectively.

18. Highest: 66% lived below poverty levels. Lowest: 33% lived below poverty levels.

19. See National School Lunch Program by the U.S. Department of Agriculture at <https://www.fns.usda.gov/nslp>.

20. Minimal or low-code tools ascribe to data science democratizing efforts by removing vast amounts of statistical and computer programming proficiency requirements (or barriers) to apply sophisticated and rigorous state-of-the-art analytic and data interactive data visualization techniques (Gonzalez Canche, 2023b).

References

- Aber, J., Grannis, K., Owen, S., & Sawhill, I. (2013). *Middle childhood success and economic mobility*. Washington, DC: Centers on Children and Families at Brookings Institute. <https://www.brookings.edu/wp-content/uploads/2016/06/15-education-success-economic-mobility-aber-grannis-owen-sawhill.pdf>
- Anderson, S., Leventhal, T., & Dupere, V. (2014). Exposure to neighborhood affluence and poverty in childhood and adolescence and academic achievement and behavior. *Applied Developmental Science, 18*(3), 123–138.

- Baltrus, P., Malhotra, K., Rust, G., Levine, R., Li, C., & Gaglioti, A. H. (2019). Identifying county-level all-cause mortality rate trajectories and their spatial distribution across the United States. *Preventing Chronic Disease, 16*, 180486. <http://dx.doi.org/10.5888/pcd16.180486>
- Bivand, R. S., Pebesma, E. J., & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R*. Springer.
- Breiger, R. L. (1974). The duality of persons and groups. *Social forces, 53*(2), 181–190.
- Carlson, D., & Cowen, J. M. (2015). Student neighborhoods, schools, and test score growth: Evidence from Milwaukee, Wisconsin. *Sociology of Education, 88*(1), 38–55.
- Case, R., Griffin, S., & Kelly, W. M. (1999). Socioeconomic gradients in mathematical ability and their responsiveness to intervention during early childhood. In D. P. Keating, & C. Hertzman (Eds.), *Developmental health and the wealth of nations: Social, biological, and educational dynamics* (pp. 125–152). Guilford Press.
- Chatterjee, D., Krampner, J., Shin, J., & Virgin, V. (2019). New York City government poverty measure: 2005–2016. https://www1.nyc.gov/assets/opportunity/pdf/18_poverty_measure_report.pdf
- Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R., & Porter, S. R. (2020). *The opportunity atlas: Mapping the childhood roots of social mobility* (No. w25147). National Bureau of Economic Research. https://opportunityinsights.org/wp-content/uploads/2018/10/atlas_paper.pdf
- Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review, 106*(4), 855–902.
- Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics, 129*(4), 1553–1623.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2009). The academic achievement gap in Grades 3 to 8. *Review of Economics and Statistics, 91*(2), 398–419. <https://www.mitpressjournals.org/doi/pdf/10.1162/rest.91.2.398>
- Considine, G., & Zappalà, G. (2002). The influence of social and economic disadvantage in the academic performance of school students in Australia. *Journal of Sociology, 38*(2), 129–148.
- Denton, K., & West, J. (2002). *Children's reading and mathematics achievement in kindergarten and first grade*. National Center for Education Statistics, Office of Educational Research and Improvement, U.S. Department of Education.
- Dong, G., & Harris, R. (2015). Spatial autoregressive models for geographically hierarchical data structures. *Geographical Analysis, 47*(2), 173–191.
- Dong, G., Harris, R., Jones, K., & Yu, J. (2015). Multilevel modelling with spatial interaction effects with application to an emerging land market in Beijing, China. *PloS ONE, 10*(6).
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446. <https://doi.org/10.1037/0012-1649.43.6.1428>
- Elijah, A. (1990). *Streetwise: Race, class, and change in an urban community*. University of Chicago.
- Farran, D. C., Aydogan, C., Kang, S. J., & Lipsey, M. W. (2006). Preschool classroom environments and the quantity and quality of children's literacy and language behaviors. *Handbook of Early Literacy Research, 2*, 257–268.
- Federal Bureau of Investigation (FBI) (2019). Crime in the United States 1995: 2019 Uniform crime reports. U.S. Department of Justice. <https://www.fbi.gov/how-we-can-help-you/more-fbi-services-and-information/ucr/publications>
- Fischler, J. (2021, October 19). Understanding charter schools vs. Public schools. *U.S. News and World Report*. <https://www.usnews.com/education/k12/articles/understanding-charter-schools-vs-public-schools>
- Fogarty, R. J., Kerns, G. M., & Pete, B. M. (2018). *Unlocking student talent: The new science of developing expertise*. Teachers College Press.
- Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement, 8*, 369–395.
- González Canché, M.S. (2014). Localized competition in the non-resident student market. *Economics of Education Review, 43*, 21–35.
- González Canché, M. S. (2017). Community college scientists and salary gap: Navigating socioeconomic and academic stratification in the U.S. higher education system. *Journal of Higher Education, 88*(1), 1–32.
- González Canché, M. S. (2018). The statistical power of “zooming in”: Applying geographically based difference in differences using spatio-temporal analysis to the study of college aid and access. *New Directions for Institutional Research, 2018*(180), 85–107.
- González Canché, M. S. (2019). Geographical, statistical, and qualitative network analysis: A multifaceted method-bridging tool to reveal and model meaningful structures in education research. *Higher Education: Handbook of Theory and Research, 34*, 535–634.
- González Canché, M.S. (2021). A complex systems network approach to assessing classroom/teacher-level baseline outcome dependence and peer effects in clustered randomized control trials. In A. Bowers (Ed.), *Data visualization, dashboards, and evidence use in schools: Data collaborative workshop* (p. 417). Columbia University Libraries.
- González Canché, M. S. (2022). Post-purchase federal financial aid: How (in)effective is the IRS's student loan interest deduction (SLID) in reaching lower-income taxpayers and students? *Research in Higher Education, 63*(6), 933. <https://link.springer.com/article/10.1007/s11162-021-09672-6>
- González Canché, M. S. (2023a). Data to replicate: The geography of mathematical (dis)opportunity. *Mendeley Data, 1*. <https://cutt.ly/5wsdo808>
- González Canché, M. S. (2023b). *Spatial socio-econometric modeling (SSEM): A low-code toolkit for spatial data science and interactive visualizations using R*. Springer Nature. <https://link.springer.com/book/10.1007/978-3-031-24857-3>
- Hanushek, E. A., & Rivkin, S. G. (2006). *School quality and the Black-White achievement gap* (NBER Working Paper No. w12651). National Bureau of Economic Research. <https://www.nber.org/papers/w12651>

- Internal Revenue Service. (2020). *SOI tax stats-individual income tax statistics-ZIP code data (SOI)*. <https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-statistics-zip-code-data-soi>
- Iriti, J., Page, L. C., & Bickel, W. E. (2018). Place-based scholarships: Catalysts for systems reform to improve postsecondary attainment. *International Journal of Educational Development*, 58, 137–148.
- Jargowsky, P. A., & El Komi, M. (2011). Before or after the bell? School context and neighborhood effects on student achievement. In S. Wachter, E. L. Birch, & H. Newberger (Eds.), *Neighborhood and life chances: How place matters in modern America* (pp. 50–72). University of Pennsylvania Press.
- Jargowsky, P. A., & Tursi, N. O. (2015). Concentrated disadvantage. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (2nd ed, pp. 525–530). Elsevier.
- Jones, K., & Duncan, C. (1995). Individuals and their ecologies: Analysing the geography of chronic illness within a multilevel modelling framework. *Health and Place*, 1(1), 27–40.
- Konstantopoulos, S., & Borman, G. D. (2011). Family background and school effects on student achievement: A multi-level analysis of the coleman data. *Teachers College Record*, 113, 97–132.
- Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the Boruta package. *Journal of Statistical Software*, 36(11), 1–13. <https://www.jstatsoft.org/article/view/v036i11>
- Lerikainen, M. K., Rasku-Puttonen, H., Aunola, K., & Nurmi, J. E. (2005). Mathematical performance predicts progress in reading comprehension among 7-year-olds. *European Journal of Psychology of Education*, 20(2), 121–137.
- Li, J., Tran, M., & Siwabessy, J. (2016). Selecting optimal random forest predictive models: A case study on predicting the spatial distribution of seabed hardness. *PLoS ONE*, 11(2). <https://doi.org/10.1371/journal.pone.0149089>
- Lindenfeld Hall, S. (2022). How ethical is it to move to a neighborhood for its school rating? *Parents*. <https://www.parents.com/kids/education/elementary-school/how-ethical-is-it-to-move-to-a-neighborhood-for-its-school-rating/>
- Lloyd, C. (2010). *Spatial data analysis: an introduction for GIS users*. Oxford University Press.
- Manoli, D. S., & Turner, N. (2018). Cash-on-hand and college enrollment: Evidence from population tax data and policy nonlinearities. *American Economic Journal: Economic Policy*, 10, 242–271.
- Mervosh, S., & Wu, A. (2022, October 24). Math scores fell in nearly every state, and reading dipped on national exam. *New York Times*. <https://www.nytimes.com/2022/10/24/us/math-reading-scores-pandemic.html>
- National Assessment of Educational Progress (NAEP). (2020). National achievement-level results in mathematics. *The Nation's Report Card*. <https://www.nationsreportcard.gov/mathematics/nation/achievement/?grade=8>
- National Assessment of Educational Progress (NAEP). (2022). National achievement-level results in mathematics. *The Nation's Report Card*. <https://www.nationsreportcard.gov/profiles/stateprofile?chart=1&sub=MAT&sj=AL&sfj=NP&st=MN&year=2019R3>
- New York State Education Department (NYSED). (2020a). Definition of performance levels. <http://www.p12.nysed.gov/irs/pressRelease/20190822/documents/math-2019-scale-score-performance-level-conversion-charts.pdf>
- New York State Education Department (NYSED). (2020b). New York State elementary and middle school educational data. <https://data.nysed.gov/downloads.php>
- Pacione, M. (1997). The geography of educational disadvantage in Glasgow. *Applied Geography*, 17(3), 169–192.
- Pagani, L. S., Fitzpatrick, C., Archambault, I., & Janosz, M. (2010). School readiness and later achievement: A French-Canadian replication and extension. *Developmental Psychology*, 46(5), 984.
- Pastor, M., Jr. (2001). Geography and opportunity. In N. J. Smelser, W. J. Wilson, & F. Mitchell (Eds.), *America becoming: Racial trends and their consequences* (Vol. 1, pp. 435–467). National Academy Press.
- Pearman, F. A. (2019). The effect of neighborhood poverty on math achievement: Evidence from a value-added design. *Education and Urban Society*, 51(2), 289–307.
- Phillips, M., & Chin, T. (2004). School inequality: What do we know? In K. Neckerman (Ed.), *Social inequality* (pp. 1–68). Russell Sage Foundation.
- Pogol, G. (2019). *Are homes in better school districts worth the extra money?* HSH: The trusted mortgage resource since 1979. <https://www.hsh.com/homebuyer/should-you-spend-more-for-home-with-better-schools.html>
- Reardon, S. F., & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal*, 46(3), 853–891. <https://doi.org/10.3102/0002831209333184>
- Rosen, H. S. (1985). Housing subsidies: Effects on housing decisions, efficiency, and equity. *Handbook of Public Economics*, 1(1), 375–420.
- Rothwell, J. (2013). *The hidden STEM economy*. Metropolitan Policy Program at Brookings. <https://www.brookings.edu/research/the-hidden-stem-economy/>
- Sahadi, J. (2015). You probably can't afford to live near good schools. *CNN Money*. https://money.cnn.com/2015/11/19/real_estate/neighborhoods-good-schools-affordable/
- Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, 23(7), 691–697.
- Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly*, 19, 99–120.
- Stevenson, H. W., & Newman, R. S. (1986). Long-term prediction of achievement and attitudes in mathematics and reading. *Child Development*, 57, 646–659.
- Tate, W. F., IV. (2008). "Geography of opportunity": Poverty, place, and educational outcomes. *Educational Researcher*, 37(7), 397–411.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(sup1), 234–240.

- Weicher, J. C. (1979). Urban housing policy. In P. M. Mieszkowski, & M. R. Straszheim (Eds.), *Current issues in urban economics* (pp. 469–508). Johns Hopkins Press.
- Wolf, S., Gonzalez Canche, M. S., & Coe, K. (2021). A complex systems network approach to quantifying peer effects: Evidence from Ghanaian preprimary classrooms. *Child Development, 92*(6), e1242–e1259.
- Wright, T. S., Parks, A. N., Wilinski, B., Domke, L. M., & Hopkins, L. J. (2020). Examining certification requirements in early math and literacy: What do states expect prekindergarten teachers to know? *Journal of Teacher Education, 0022487120905514*.

Author

MANUEL S. GONZÁLEZ CANCHÉ is an associate professor at the University of Pennsylvania, 208 S. 37th St., Room 307, Philadelphia, PA 19104; *email*: msgc@upenn.edu. His research focuses on democratizing or expanding access to data science via no-code or low-code software tools. For no-code tools, see <https://doi.org/10.1016/j.eswa.2022.119265> and <https://doi.org/10.1177/15586898231166968>. For low-code tools, see <https://link.springer.com/book/10.1007/978-3-031-24857-3>.