Study of Changes in Public School Composition During the Pandemic and Their Relationship With NAEP Performance

Martin Hooper, Rebecca Shipan, David Bamat, and George Bohrnstedt
American Institutes for Research

June 2024
Commissioned by the NAEP Validity Studies (NVS) Panel

The NAEP Validity Studies Panel was formed by the American Institutes for Research under contract with the National Center for Education Statistics. Points of view or opinions expressed in this paper do not necessarily represent the official positions of the U.S. Department of Education or the American Institutes for Research.

The NAEP Validity Studies (NVS) Panel was formed in 1995 to provide a technical review of NAEP plans and products and to identify technical concerns and promising techniques worthy of further study and research. The members of the panel have been charged with writing focused studies and issue papers on the most salient of the identified issues.

Panel Members:

Keena Arbuthnot
Louisiana State University

Peter Behuniak
Criterion Consulting, LLC

Derek Briggs
University of Colorado

Jack Buckley
American Institutes for Research

Phil Daro
Strategic Education Research Partnership (SERP) Institute

Richard P. Durán
University of California, Santa Barbara

David Grissmer
University of Virginia

Gerunda Hughes
Howard University

Akisha Osei Sarfo
Council of the Great City Schools

James Pellegrino
University of Illinois at Chicago

Gary Phillips
Cambium Assessment

Jennifer Randall
University of Michigan

Lorrie Shepard
University of Colorado–Boulder

David Thissen
University of North Carolina, Chapel Hill

Gerald Tindal
University of Oregon

Sheila Valencia
University of Washington

Denny Way
Consultant

Project Director:
Sami Kitmitto
American Institutes for Research

Project Officer:
Grady Wilburn
National Center for Education Statistics

For Information:
NAEP Validity Studies (NVS) Panel
American Institutes for Research
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
Email: naepvalidityst udies@air.org

CONTENTS

EXHIBITS

EXECUTIVE SUMMARY

This study was motivated by the desire to better understand the major decline in eighth-grade mathematics achievement on the National Assessment of Educational Progress (NAEP) from 2019 to 2022. Given sharp losses in the enrollment of public school students during the same period, the study has two main purposes: (a) to examine changes in the demographic composition of students among eighth graders in U.S. public schools between 2019 and 2022 and (b) to estimate the extent to which the compositional changes relate to NAEP eighth-grade mathematics achievement scores in the same time period.

Overall, the study found that compositional changes between 2019 and 2022 generally correspond to trends in compositional changes observed before 2019. The changes in student and school characteristics between 2019 and 2022 are similar to observed trends in compositional changes in the years leading up to 2019. The exception is a moderate departure from the trend line for English learners (ELs), whose population growth increased more than would have been expected based on the previous trend in the population change of ELs.

In addition, a relatively small proportion of the national decrease in eighth-grade mathematics achievement is associated with compositional changes. Results from two separate analyses tell a similar overarching story: The observed decreases in mathematics achievement from 2019 to 2022 are best understood as the result of decreases in performance.

- The degree to which decreases are associated with compositional changes is more pronounced in certain states and districts.
- In most states and participating districts, compositional changes could have further accentuated the NAEP achievement declines between 2019 and 2022.

However, the findings are limited by changes in how demographic characteristics were measured between 2019 and 2022. For instance, in response to the pandemic, some jurisdictions increased access to free or reduced-price meals, which might have unduly influenced observed changes in students determined to be eligible for the National School Lunch Program.

BACKGROUND

The purpose of this study is to (a) examine changes in demographic composition among eighth graders in U.S. public schools between 2019 and 2022 and (b) estimate the extent to which these compositional changes relate to National Assessment of Educational Progress (NAEP) Grade 8 mathematics achievement scores in this time period, nationally and across states and districts participating in the Trial Urban District Assessments (TUDAs).

The 2022 Main NAEP results were the first results released following the start of the COVID-19 pandemic in 2020. The fourth- and eighth-grade mathematics and reading results provided evidence on how achievement changed since 2019, showing declines among national public school students of 5 points in Grade 4 mathematics, 3 points in Grade 4 reading, 8 points in Grade 8 mathematics, and 3 points in Grade 8 reading.

Most states and TUDAs also showed declines in their results. In Grade 8 mathematics, for example, 51 states/jurisdictions showed score declines, and only two states/jurisdictions showed no change. Changes in reading scores ranged from a 13-point decline in Oklahoma to no difference in Utah and the Department of Defense Education Activity schools. There was a similar pattern within TUDAs, with 22 TUDAs showing a decline in Grade 8 mathematics scores and only four TUDAs showing no significant change. The largest decline was in Shelby County (TN), with a loss of 13 points. Los Angeles, California; Miami-Dade, Florida; New York, New York; and Philadelphia, Pennsylvania, showed no change.

It may be tempting to attribute these results to COVID-19 policies and variations in policies across states, but a myriad of factors could be associated with the changes in results, including, potentially, changes in the composition of students within public schools. According to the National Center for Education Statistics (2022), enrollment in U.S. public schools shifted dramatically between fall 2019 and fall 2020, with the number of prekindergarten through Grade 8 public school students dropping by 1.5 million across the country, from 35.6 million students in 2019 to 34.1 million students in 2020.

The decrease in public school numbers during and following the pandemic was documented in *The New York Times* (Hubler, 2022). Providing numerous local examples, the article detailed that student enrollment in New York City public schools declined by 50,000 students; likewise, Michigan public schools lost 50,000 students. The article noted two potential causes:

> Some parents became so fed up with remote instruction or mask mandates that they started home-schooling their children or sending them to private or parochial schools that largely remained open during the pandemic. And other families were thrown into such turmoil by pandemic-related job losses, homelessness and school closures that their children simply dropped out.

A 2020 Household Pulse survey carried out by the U.S. Census Bureau (Eggleston & Fields, 2021) showed a significant increase between April 2020 and October 2020 in the proportion of students in homeschooling, and the size of this shift varied across racial/ethnic groups. By fall 2020, 16% of Black students were being homeschooled compared with 10% of White students, 12% of Hispanic students, and 9% of Asian students. The proportion of students moving to homeschooling across this time period represented a fivefold increase for Black students and about a twofold increase for other groups. A majority of states showed an increase in the

proportion of students being homeschooled during this time frame, but the magnitude of the increase varied considerably, with nine states showing increases of greater than 10%.

More recent results from the American Institutes for Research's COVID-19 and Equity in Education (CEE) Enrollment Explorer show heterogeneity in enrollment changes across the six states analyzed (California, Florida, New York, Tennessee, Texas, and Washington) between the 2019–20 and 2021–22 school years. The CEE found enrollment declines in all six focal states, ranging from a 5% decline in New York and Washington to a 1% decline in Florida. White student enrollment declined by 10% in California and Washington but by only 3% in Florida. Black student enrollment fell by 8% in California, but Latino student enrollment in California fell by only 3%. In Florida, however, Latino student enrollment increased by 2%. Such heterogeneity in population shifts raises the question: "To what extent are such population shifts also associated with changes in NAEP national, state, and TUDA scores?"

The COVID-19 era shifts away from public schooling compound an already declining public school enrollment that was attributed to a declining birth rate. As noted by Ladd and Huberlie (2023), "From 2008–20, birth rates in the US declined by ~1.3% annually, translating to ~700k total fewer students entering the schooling system from SY2012 up to SY2025."

This report explores the shifts in student enrollment patterns from public schools to private schools and/or homeschooling during the pandemic by analyzing changes in the NAEP public school population from 2019 to 2022. The report also estimates the degree to which changes in achievement are associated with changes in the demographic characteristics of the public school population (i.e., compositional changes).

RESEARCH QUESTIONS

This study examines two research questions (RQs):

How did the characteristics of the NAEP-assessed population change between 2019 and 2022?

Were observed compositional changes in the NAEP-assessed population between 2019 and 2022 related to changes in NAEP national, state, and TUDA scores?

The study uses results from eighth-grade mathematics, which exhibited the largest NAEP score declines nationally.

METHODOLOGIES

RQ1. How did the characteristics of the NAEP-assessed population change between 2019 and 2022?

To answer this question, we conducted a descriptive examination of trends in the demographic composition of the NAEP-assessed population. Of particular interest are the trend lines for the major NAEP reporting groups: race/ethnicity, students eligible for the National School Lunch Program (NSLP), English learners (ELs), and students with disabilities (SDs). Changes in the trajectory of the trend line between 2019 and 2022 compared with the trend between 1990 and 2019 provide evidence as to whether the change from 2019 to 2022 was in line with previous trends or diverged.

RQ2. Were observed compositional changes in the NAEP-assessed population between 2019 and 2022 related to changes in NAEP national, state, and TUDA scores?

To answer this question, we implemented two different techniques: (a) Beaton and Chromy's (2007) partitioning method and (b) a propensity weighting method.

Beaton and Chromy's Partitioning Method

Beaton and Chromy's 2007 NAEP Validity Studies (NVS) paper examined NAEP reading and mathematics performance changes from the 1970s through 2004 using a partitioning methodology. They partitioned changes in mean NAEP reading and mathematics scores into three components—the portion associated with changes in student racial/ethnic demographics, the portion associated with changes in student proficiency, and the interaction between these two—a technique similar to standardization traditionally applied in demography research (Mosteller & Tukey, 1977). Briefly, a traditional application of the method uses a "standard population" (in Beaton and Chromy's study, the baseline year of NAEP data) and represents the mean score for an alternative time point or sample in terms of what it would have been if the population remained the same as at baseline.

Beaton and Chromy (2007) showed that interpreting trends in NAEP's national mean scores using this partitioning method tells only part of the story because changes in the composition of the student population across assessments may mask performance changes for individual student groups when viewed in aggregate. In particular, they showed that from the 1970s through 2004, positive gains occurred in average NAEP scores across race/ethnicity groups, whereas national mean scores remained virtually unchanged—a phenomenon known as Simpson's paradox (Simpson, 1951). This result was explained by noting that traditionally lower

performing student groups comprised increasingly larger proportions of the NAEP population across time.

The current study uses data from the 2019 and 2022 Grade 8 NAEP mathematics assessments. The reporting sample was public school students nationally, in each state, and in each participating TUDA. The demographic variables used in the decompositions were as follows:

Student-level racial/ethnic data as provided by school records (NAEP variable SRACE10)

Student-level data on NSLP eligibility (NAEP variable SLUNCH1)

Student status as an EL (NAEP variable LEP)

Student disability status (NAEP variable IEP)

School locale (NAEP variable UTOL4)

Charter school designation (NAEP variable CHRTRPT)

The variable representing the percentage of students eligible for NSLP at the school level also was considered for analysis but ultimately discarded because of a large amount of missing data in 2022.

For the race/ethnicity analysis, average NAEP mathematics scores are decomposed for the aggregate student sample in each year of data, as well as group averages in each year for the seven racial/ethnic groups represented in NAEP's race/ethnicity variable: White, Black, Hispanic, Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, and Two or More Races. These groups are mutually exclusive and exhaustive. For example, the "Hispanic" group includes students of any racial background whose school records indicate Hispanic ethnicity; other race categories exclude Hispanic origin. Similar analyses are conducted with data on students' NSLP eligibility, EL status, and disability status, as well as the school variables for school locale and charter school designation.

The student variables for NSLP eligibility, EL status, and disability status included the option "information not available." To ensure use of the full sample, we retained "information not available" as an analysis variable, instead of treating it as missing data. Thus, fluctuations in the information-not-available category across time could appear as a population change in this analysis.

The average scores among all students and for each student group for each year were estimated following standard NAEP analysis procedures, including using each student's sampling weight and a set of 20 plausible achievement values. As detailed in Beaton and Chromy (2007), the difference in overall mean scores between two time points ($t = 1$ and $t = 2$) for students in the aggregate can be partitioned using the formula in Equation 1, in which each symbol is a vector of values for the seven racial/ethnic groups. In Equation 1, $\boldsymbol{X}_t$ is a vector of average NAEP scores for the component student groups at the given time points; P_t is a vector of the student groups' average proportion in each year; and $\bar{x_t}$ is the overall NAEP score as reported across all students, where $t = 2$ refers to 2022 and $t = 1$ refers to 2019.

$$\begin{aligned} \bar{x_2} - \bar{x_1} &= \boldsymbol{P}_2'\boldsymbol{X}_2 - \boldsymbol{P}_1'\boldsymbol{X}_1 \\ &= (\boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1)'(\boldsymbol{X}_1 + \boldsymbol{X}_2 - \boldsymbol{X}_1) - \boldsymbol{P}_1'\boldsymbol{X}_1 \\ &= \boldsymbol{P}_1'(\boldsymbol{X}_2 - \boldsymbol{X}_1) + (\boldsymbol{P}_2 - \boldsymbol{P}_1)'\boldsymbol{X}_1 + (\boldsymbol{P}_2 - \boldsymbol{P}_1)'(\boldsymbol{X}_2 - \boldsymbol{X}_1) \end{aligned} \quad (1)$$

The three terms shown in the final equation are as follows:

$\boldsymbol{P}_1'(\boldsymbol{X}_2 - \boldsymbol{X}_1)$ is the *performance effect*,[1] reflecting the mean gain or loss, assuming that the population composition had remained the same at $t = 2$ as it was at $t = 1$.

$(\boldsymbol{P}_2 - \boldsymbol{P}_1)'\boldsymbol{X}_1$ is the *population effect*, reflecting the gain or loss caused by changes in population, assuming that the group mean performance values had remained the same at $t = 2$ as they were at $t = 1$.

$(\boldsymbol{P}_2 - \boldsymbol{P}_1)'(\boldsymbol{X}_2 - \boldsymbol{X}_1)$ is the *joint effect residual*, reflecting the joint effect of performance and population on score change for the focal group being analyzed. This residual allows the performance and population effects to add to the score difference and should not be interpreted as substantively meaningful. This value is positive if the groups increasing in relative size also are increasing in relative performance (and those whose relative size is decreasing are the ones with relatively lower performance), and this value is negative when the opposite is true (i.e., group performance changes and population changes are in opposing directions).

It is possible in this methodology to have effects of different directionality—for example, a positive performance effect and a negative population effect. A more thorough explanation of the methodology is in Appendix D.

One limitation of the Beaton-Chromy method, as acknowledged by Beaton and Chromy themselves (2007), is that it lacks the calculation of standard errors and thus the opportunity to judge the statistical significance of the results. Another limitation of this methodology is that suspect effects can be found when a category has zero students in one year and a nonzero number of students in the other year. We also noticed large joint effect residuals when the category "information not available" increased dramatically across the assessment years for the NSLP eligibility, EL status, and disability status variables. If either of these scenarios resulted in a joint effect residual greater than ± 2—a value we judged to be atypically large compared with other residual values—we omitted the state or TUDA from the analysis.

The results of the partitioning analysis should not be interpreted to conclude that shifts in demographics cause changes in performance. Each demographic variable highly correlates with multiple unmeasured characteristics (e.g., wealth, opportunity to learn, discrimination, availability of food and health resources, quality schools), and these unmeasured characteristics may causally relate to both the measured demographic variables and NAEP performance.

One drawback of using the partitioning approach demographic variable (e.g., race/ethnicity) is that it ignores the multiplicities and intersections of student characteristics. We tried to ameliorate this concern by cross-tabulating multiple variables (race/ethnicity, NSLP), but this made the results difficult to interpret because it often led to cases with zero students in a cell in one year but not in another year.

Propensity Weighting Analysis

The Beaton-Chromy partitioning method aids in understanding the bivariate relationships between changes in student composition and differences in NAEP achievement across time; however, our goal also was to determine a more complex multivariate relationship between changes in composition and NAEP achievement. Thus, we used a multivariate approach—

[1] We use the word "effect" throughout this paper to talk about a statistical relationship such as correlations, not to imply a causal relationship.

propensity weighting analysis—to better account for students' multiple demographic characteristics.

Typically used in a quasi-experimental design to compare contemporaneous outcomes of treatment and comparison subjects, we expand on the propensity weighting framework to define 2019 students as the "treatment group" and 2022 students as the "comparison group." In this application of the method, student and school characteristics are used to predict the propensity (probability) that students with such characteristics tested in 2022 also could have been sampled for testing in 2019. We then used this propensity to reweight the 2022 sample so that it had characteristics similar to those of the 2019 sample. With propensity-weighted data, we can reexamine changes between 2019 and 2022 student scores to gauge the degree to which performance effects between testing years exist after reweighting, while simultaneously controlling for multiple student and school characteristics.

By accounting for the probability of being sampled in 2019, the theory guiding propensity weighting holds that differences in both observable characteristics (e.g., eligibility for NSLP) and unobservable characteristics (e.g., parental decisions to unenroll children from traditional public schools) are held constant. Therefore, the contribution of population effects is removed from the performance effect estimate (change in learning). In brief, we pose the following question through this analysis: "How much would mean achievement have changed from 2019 to 2022 had student characteristics stayed the same across assessment years?"

Of course, as with other approaches relying on the assumption of "ignorability," it is impossible to know whether the propensity weighting approach effectively captures all the unmeasured compositional changes occurring between assessment years (Raudenbush, 2004; Rosenbaum & Rubin, 1983). We assume that the observed predictor variables entered in the propensity score model fully account for all potential bias, and no hidden biases exist related to unmeasured factors. This assumption can never fully be met—an unavoidable limitation of propensity score analysis. However, we can reduce the threat of biases, including selection bias, by including variables in the propensity model that successfully predict the hypothetical likelihood of being sampled in 2019 and predict the outcome—performance in NAEP eighth-grade mathematics.

Nationally, there were 249,950 students sampled for Grade 8 mathematics across 2019 and 2022: 142,200 in 2019 (56.9%) and 107,740 in 2022 (43.1%).[2]

Here are the steps used to carry out the propensity weighting analysis.

Step 1. Conduct a Logistic Regression Analysis

The first step is to conduct a logistic regression analysis to estimate the probability that 2022 students with a set of eight characteristics would have been sampled in 2019. Six student- and two school-level demographic characteristics served as predictor variables (controls) in the logistic regression model.

Nationally, the student variables included are as follows:

- Student-level racial/ethnic data as provided by school records (NAEP variable SRACE10)
- Student-level data on NSLP eligibility (NAEP variable SLUNCH1)
- Student status as an EL (NAEP variable LEP)

[2] The total sample size does not equal the sum of samples across 2019 and 2022 because of rounding.

Student disability status (NAEP variable IEP)

Highest level of parental education (NAEP variable PARED)

Number of books at home (NAEP variable B013801)

Differences in the population composition between 2019 and 2022 also can vary by school type. The school variables included are as follows:

School locale (NAEP variable UTOL4)

Charter school designation (NAEP variable CHRTRPT)

Although the variable list is not exhaustive, the list represents most of the variables used in reporting NAEP results, including multiple socioeconomic status (SES) indicators—student-level data on NSLP eligibility, highest level of parental education, and books at home—because recent research shows that including multiple SES variables provides a more robust measure of SES than NSLP status alone and explains more of the variance in performance as well (Broer et al., 2023). Similar to our approach to the partitioning analysis, we did not include the school-level percentage of students eligible for NSLP because of large increases in missing data for this variable in 2022. All predictor variables were transformed into sets of indicator variables (e.g., dummy variables; see Exhibit E1 in Appendix E).

A separate propensity score model was run for each state and TUDA. To ensure reliable and stable estimation for coefficients in the model, a requirement was imposed in which each level of a variable needed to have at least 10 students in each assessment year to be included in the analysis. When a variable level had fewer than 10 students, no dummy variable was created for that level.

We also reviewed variable descriptive statistics to identify irregularities; for example, Oregon reported 100% of students (with rounding) as NSLP eligible in 2022; Los Angeles had a high percentage of information not available for NSLP in 2022. After identifying an irregular case, we deleted the variable from the analysis for that state or TUDA. Information on the excluded variables and the levels for each state and TUDA are in Appendix G.

Care also was taken in developing and testing the validity of the results. The checks included ensuring that certain variables did not overly influence the results. In addition, we examined the balance of the demographics in 2022 and 2019 after applying the propensity weights. See Appendix E for information on the results from the national covariate balance check analysis and Appendix F for an elaboration on the balancing effects of propensity weighting.

Step 2. Compute the Propensity Score Weights

To compute propensity score weights (*psw*) from propensity scores (*ps*), the formula in Equation 2 was applied to the propensity scores of the comparison group (i.e., 2022 sample; Austin & Stewart, 2015; Ridgeway et al., 2015).

$$psw_c = ps_{\text{comparison}}/(1 - ps_{\text{comparison}}) \quad (2)$$

To arrive at a final weight (*fw*) for estimating the treatment effect, the propensity score weights were multiplied by their original sampling weights (NAEP variable ORIGWT), an approach endorsed by Ridgeway et al. (2015) when computing propensity score weights with complex survey data (see Equation 3). Note that the final weight for treatment students equals the NAEP sampling weight.

$$fw = psw \times \text{NAEP sampling weight} \quad (3)$$

Similar to the population decomposition analysis, a major limitation of the propensity weight analytic approach is the plausibility that the measure or classification implementation procedures (e.g., classification to EL status or NSLP eligibility) were not static between 2019 and 2022, so shifting the weighting in 2022 to match that of 2019 may not have always reflected actual changes in the population.

A general limitation of this study, regardless of statistical methodology, is that some population shifts described are likely linked to changes in the measurement or classification of compositional characteristics rather than actual changes to the populations themselves. For example, NSLP eligibility is sometimes viewed as a proxy for low family income. However, policies expanding free or reduced-price lunch eligibility in certain states or TUDAs during the pandemic could have led to changes in the percentage of students reported as eligible. In such cases, it would be incorrect to infer that a state had a change in the percentage of students with low family income; rather, it reflects a change in the measure, or classification procedure, used.

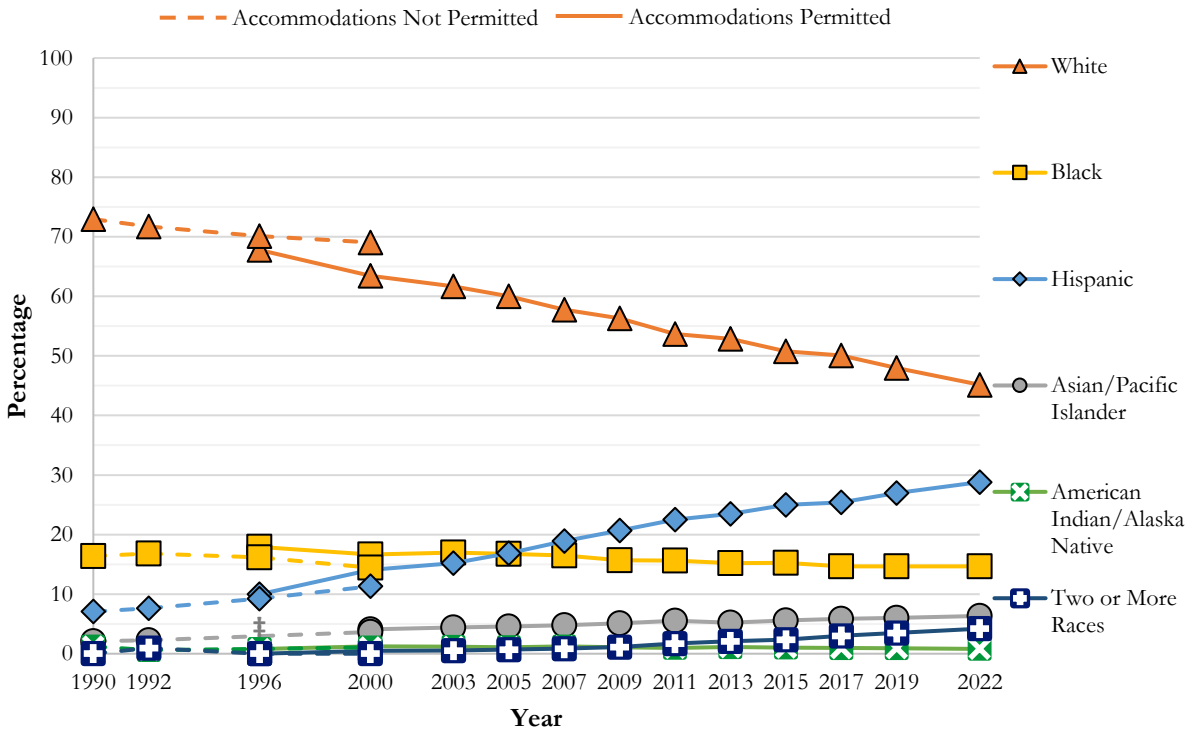
Another potential confounding factor could relate to the ability of schools to accurately classify students as ELs during the pandemic. Research evidence suggests that pandemic disruptions were associated with decreased rates in the reclassification of ELs as former ELs. Consequently, the sampled population of ELs in 2022 could have included a greater proportion of students who would have been classified as non-EL in previous testing years (Villegas & Garcia, 2022). In The School District of Philadelphia, for instance, pandemic-related disruptions were associated with a sharp decrease in the number of students administered the ACCESS for ELLs assessment. This test, which measures English-language proficiency, plays a role in determining the reclassification of students out of EL status. During the 2021–22 school year, ACCESS was administered to about 2,700 students in The School District of Philadelphia, down from about 16,000 students at the same time point in the 2019–20 school year (Bamat et al., 2023). Nationally, EL classification was further complicated by changes in immigration patterns during and after the pandemic. For example, immigration to the United States decreased dramatically during the pandemic (Schachter et al., 2021).

RQ1 RESULTS: How did the characteristics of the NAEP-assessed population change between 2019 and 2022?

Trends in the percentage of students in each population group were examined at the national level to explore to what extent the 2022 national public school student population aligns with previous population trends.

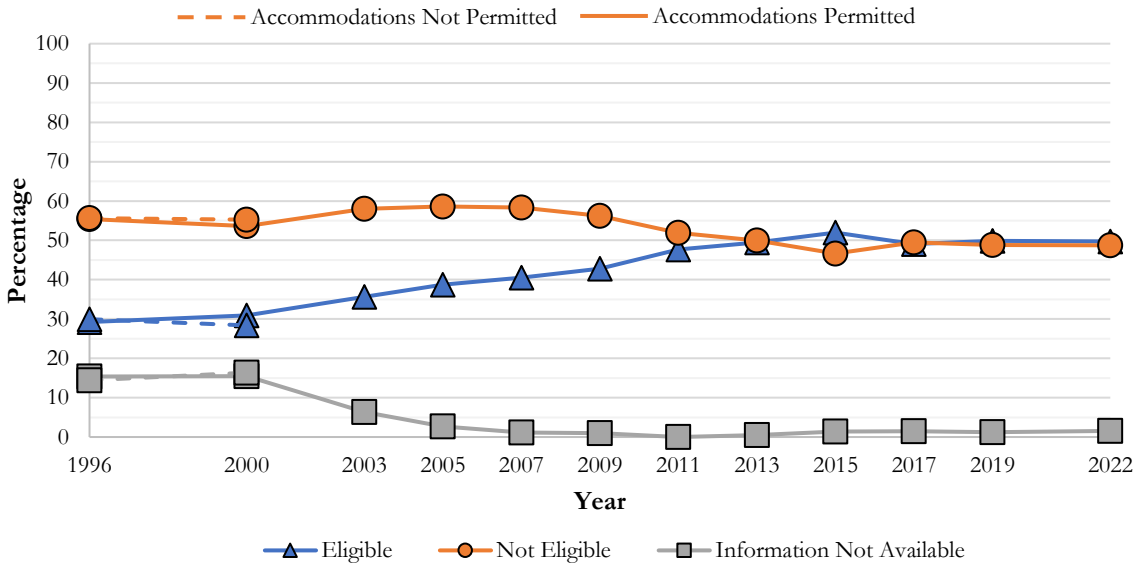
For race/ethnicity, NSLP eligibility, and SD status, the results show a linear trend line across recent years (Exhibits 1–3), without fluctuation in the trend line between 2019 and 2022. For example, the increase in the percentage of the Hispanic population between 2019 and 2022 seems in line with the change between 2017 and 2019.

Exhibit 1. Trends in the Percentage Distribution of Eighth-Grade Public School Students in NAEP Mathematics, by Race/Ethnicity: Various Years, 1990–2022



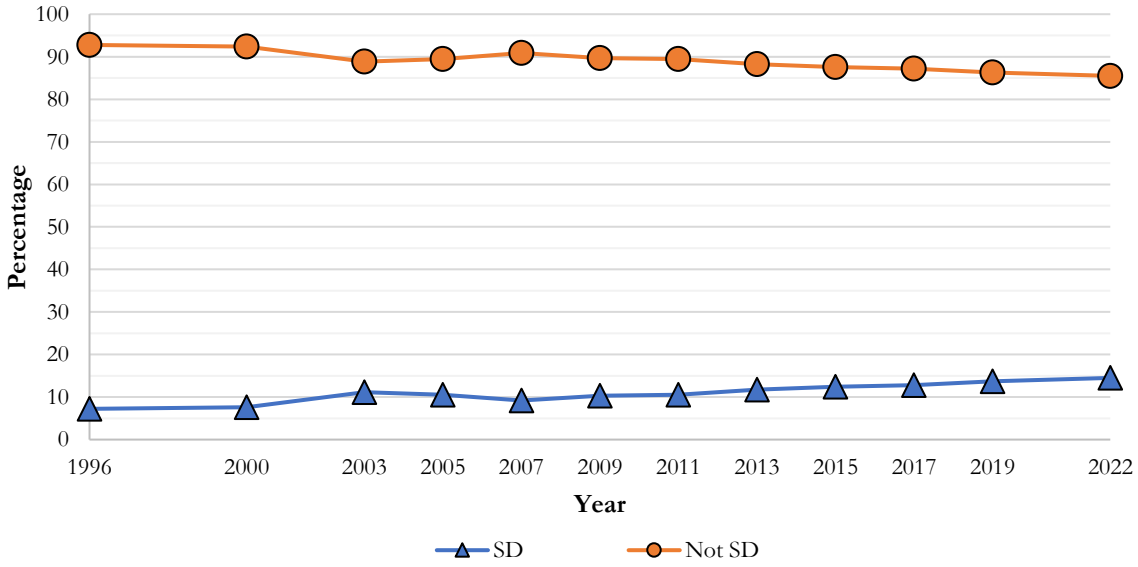
‡ Reporting standards not met.
NOTE: NAEP = National Assessment of Educational Progress. Dashed lines indicate that accommodations were not permitted; solid lines indicate that accommodations were permitted. Black includes African American, Hispanic includes Latino, and Pacific Islander includes Native Hawaiian. Race categories exclude Hispanic origin. Prior to 2011, students in the Two or More Races category were categorized as unclassified. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 1990, 1992, 1996, 2000, 2003, 2005, 2007, 2009, 2011, 2013, 2015, 2017, 2019, and 2022 Grade 8 Mathematics Assessments.

Exhibit 2. Trends in the Percentage Distribution of Eighth-Grade Public School Students in NAEP Mathematics, by NSLP Eligibility: Various Years, 1996–2022



NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. Dashed lines indicate that accommodations were not permitted; solid lines indicate that accommodations were permitted. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 1996, 2000, 2003, 2005, 2007, 2009, 2011, 2013, 2015, 2017, 2019, and 2022 Grade 8 Mathematics Assessments.

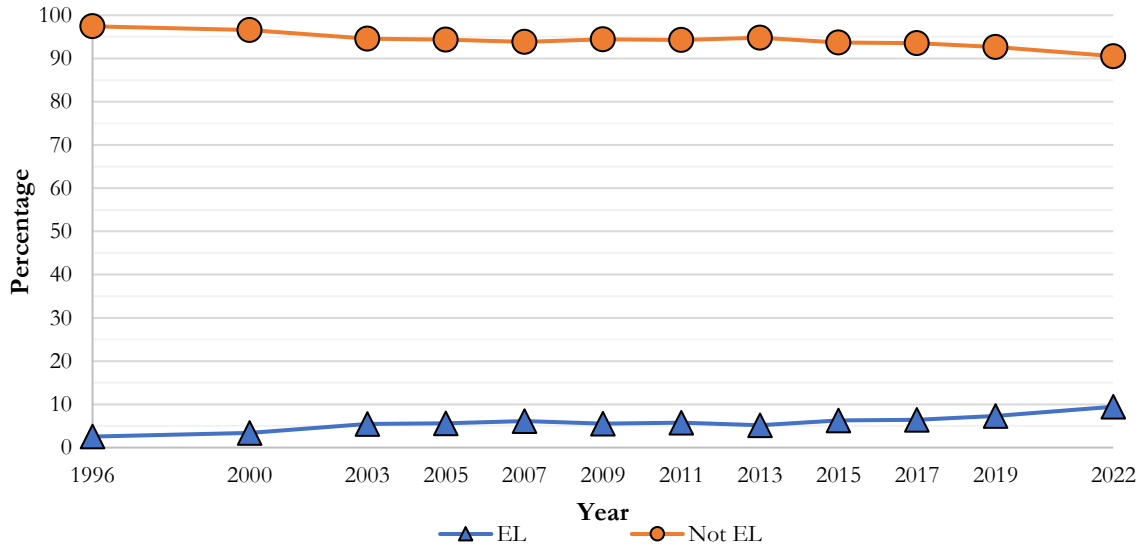
Exhibit 3. Trends in the Percentage Distribution of Eighth-Grade Public School Students in NAEP Mathematics, by SD Status: Various Years, 1996–2022



NOTE: NAEP = National Assessment of Educational Progress. SD = students with disabilities. Dashed lines indicate that accommodations were not permitted; solid lines indicate that accommodations were permitted. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 1996, 2000, 2003, 2005, 2007, 2009, 2011, 2013, 2015, 2017, 2019, and 2022 Grade 8 Mathematics Assessments.

However, the population of ELs increased by 2 percentage points between 2019 and 2022 compared with an increase of only 1 percentage point from 2017 to 2019 (Exhibit 4). The 2019–2022 growth in ELs marked an increase of nearly 30%, from 7.4% of the population of eighth graders in 2019 to 9.4% of the eighth-grade population in 2022.

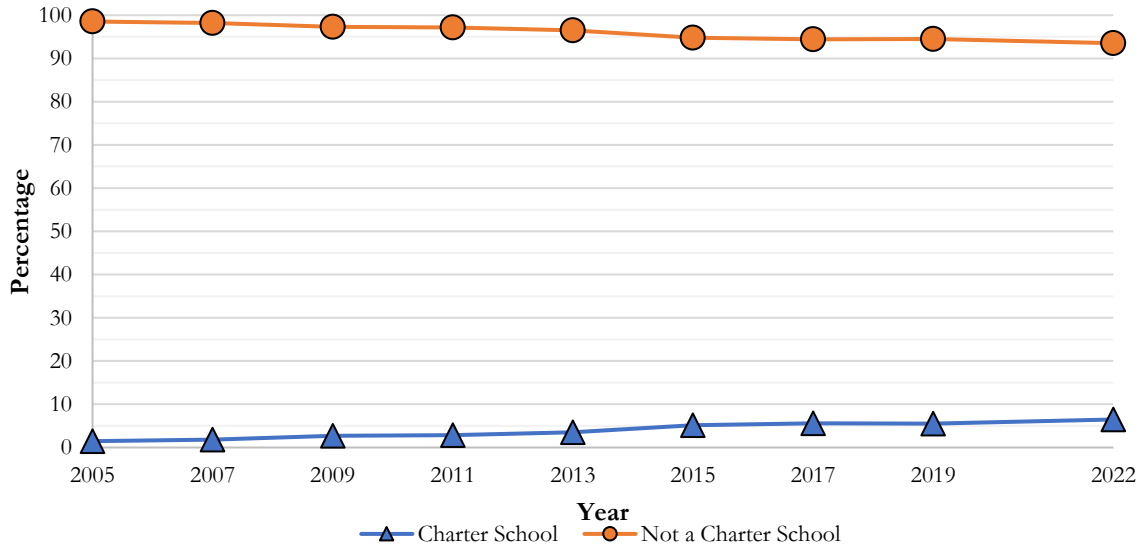
Exhibit 4. Trends in the Percentage Distribution of Eighth-Grade Public School Students in NAEP Mathematics, by EL Status: Various Years, 1996–2022



NOTE: EL = English learner. NAEP = National Assessment of Educational Progress. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 1996, 2000, 2003, 2005, 2007, 2009, 2011, 2013, 2015, 2017, 2019, and 2022 Grade 8 Mathematics Assessments.

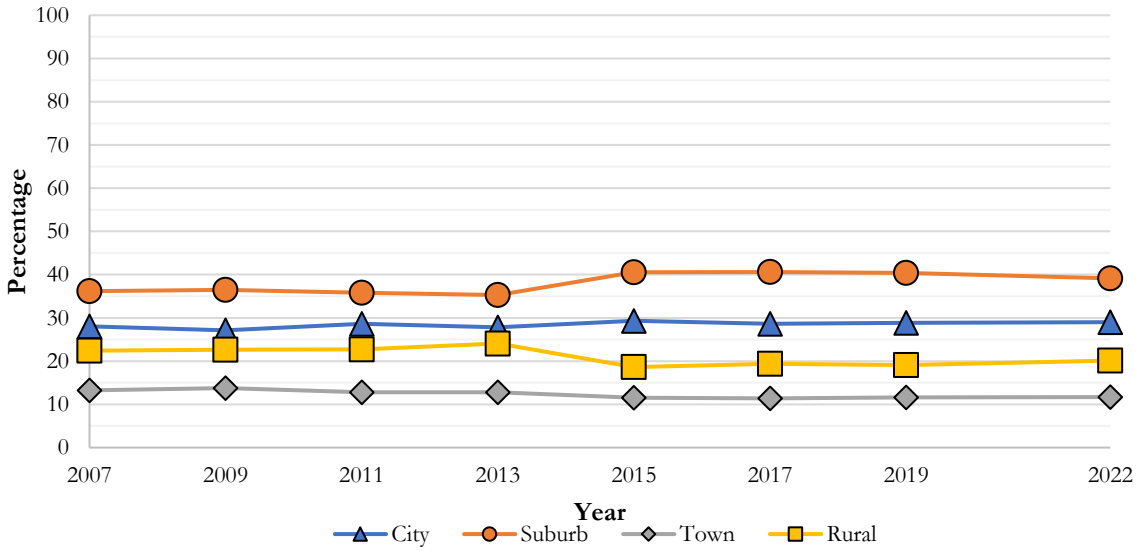
School characteristic trends between 2019 and 2022 were consistent with prior trends, with the percentage of students in charter schools and the percentage of students by school locale following approximately similar trends as they had from 2015 onward (Exhibits 5–6). The school locale analysis showed a sizable growth of students in suburban schools and a decrease of students in city schools from 2013 to 2015.

Exhibit 5. Trends in the Percentage Distribution of Eighth-Grade Students in NAEP Mathematics, by Charter School Status: Various Years, 2005–2022



NOTE: NAEP = National Assessment of Educational Progress. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2005, 2007, 2009, 2011, 2013, 2015, 2017, 2019, and 2022 Grade 8 Mathematics Assessments.

Exhibit 6. Trends in the Percentage Distribution of Eighth-Grade Students in NAEP Mathematics, by School Locale: Various Years, 2007–2022



NOTE: NAEP = National Assessment of Educational Progress. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2007, 2009, 2011, 2013, 2015, 2017, 2019, and 2022 Grade 8 Mathematics Assessments.

In summary, the national trends did not show sizable shifts in the student population and the public schools they attend. The lone trend that did show a shift was the growth of ELs, with a gain of nearly 30% from 2019 to 2022.

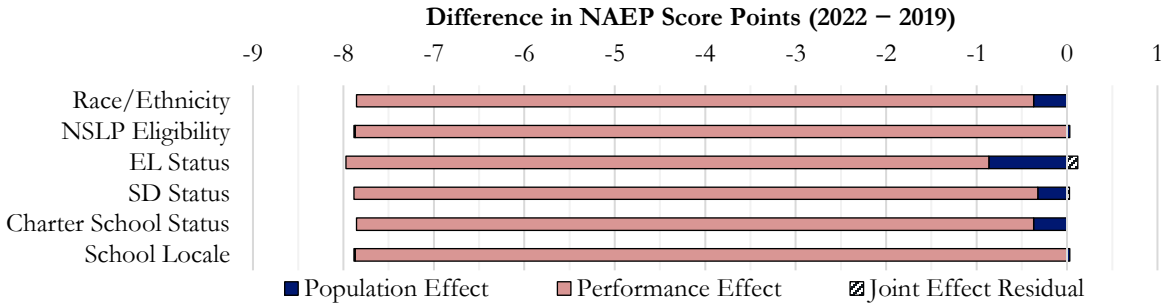
RQ2 PARTITIONING ANALYSIS RESULTS: Were observed compositional changes in the NAEP-assessed population between 2019 and 2022 related to changes in NAEP national, state, and TUDA scores?

In this section, we apply the Beaton-Chromy partitioning method at the national, state, and TUDA levels to examine the relationship between compositional changes and changes in NAEP performance.

National Results

The student-level national results show that less than 1 point in the nearly 8-point decline in Grade 8 mathematics performance relates to a population effect for any student or school factors: race/ethnicity, NSLP eligibility, EL status, SD status, charter school status, and school locale (Exhibit 7). The largest effect (nearly 1 point) is linked to changes in the EL population.

Exhibit 7. Partitioned Effects Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by Race/Ethnicity, NSLP Eligibility, EL Status, SD Status, Charter School Status, and School Locale



NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. EL = English learner. SD = students with disabilities. The results for SDs and ELs are based on students who were assessed and cannot be generalized to the total population of such students. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

State Results

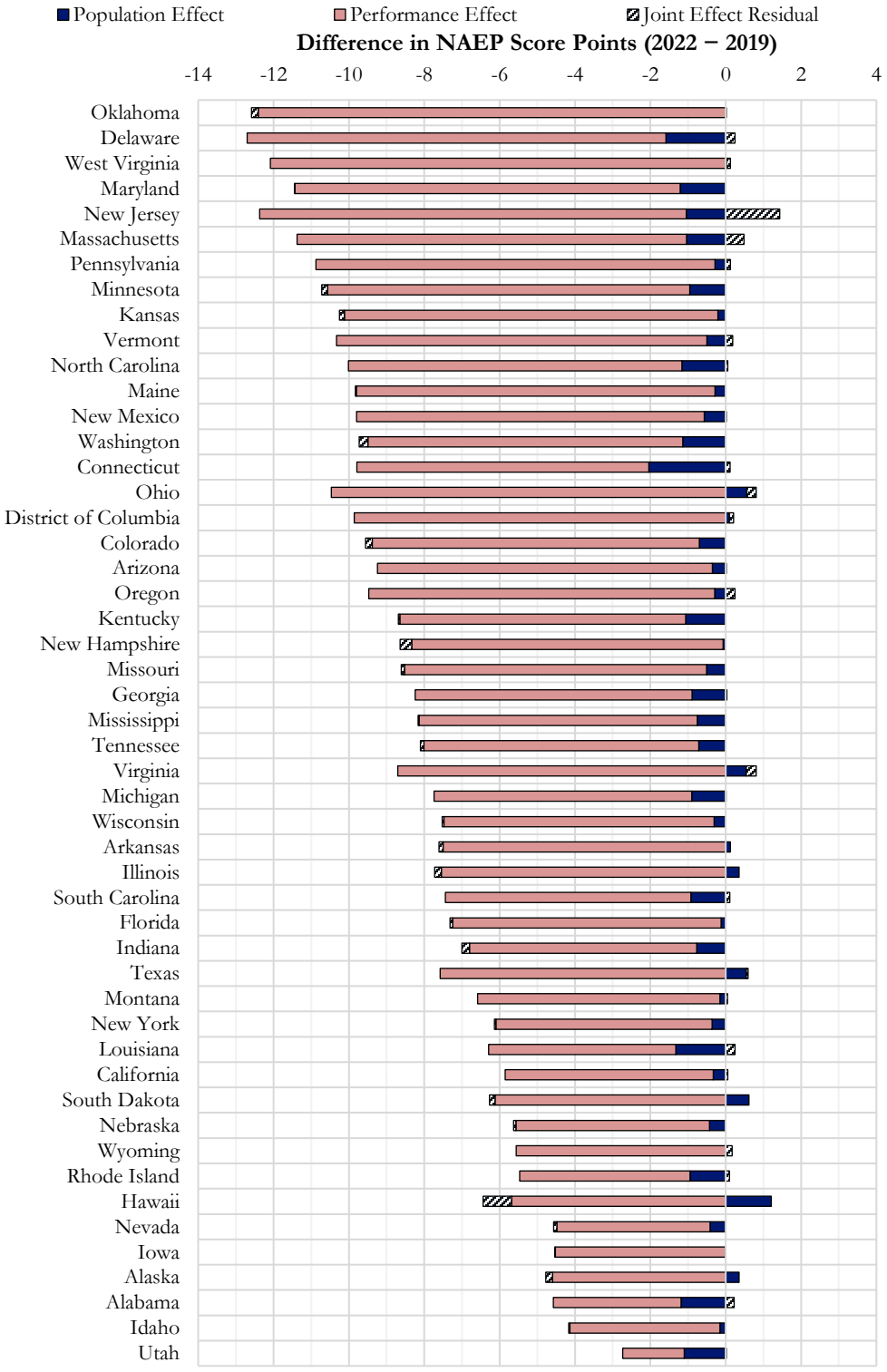
We also applied the partitioning analysis to the state results for all six analysis variables. Although arbitrary, some criterion of substantive or practical significance is necessary, and for these analyses we chose population effects of ±2 points.[3] We describe the results for states meeting this criterion in this section and present their results in separate exhibits.

By Change in Race/Ethnicity

Although the estimate of the population effect for the race/ethnicity variable equals −0.4 points nationally, state-level population effect estimates range from −2.0 points (Connecticut) to +1.2 points (Hawaii). Only Connecticut showed a shift of ±2 points or more (Exhibit 8). For the vast majority of states, the score decrease was almost entirely related to changes in NAEP performance, not to changes in the population by race/ethnicity.[4]

[3] As noted in the limitation section, we considered using a percentage of change criterion when highlighting states and TUDAs. Ultimately, we selected a score change criterion because it better captures substantively meaningful changes. For example, for Miami-Dade (Exhibit 19) race/ethnicity has population effect that equals nearly 40% of the score difference (0.8 of Miami-Dade's 1.9-point decline), yet the population effect may not be substantively meaningful at 0.8 of a NAEP score point.
[4] The numerical results used to generate the state figures in Exhibits 8, 10, 12, 14, 15, and 16 are in Appendix B.

Exhibit 8. Partitioned Effects for Race/Ethnicity Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by State



NOTE: NAEP = National Assessment of Educational Progress. States are ordered by the 2022 − 2019 difference in average scale score. North Dakota did not report data for students of Two or More Races in 2019 and thus was excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

The larger performance effect in Connecticut appears to be associated with a decrease in the enrollment of groups of students that are generally higher performing and an increase in the enrollment of groups that are generally lower performing (Exhibit 9).

Exhibit 9. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics in Connecticut, by Race/Ethnicity: 2019 and 2022

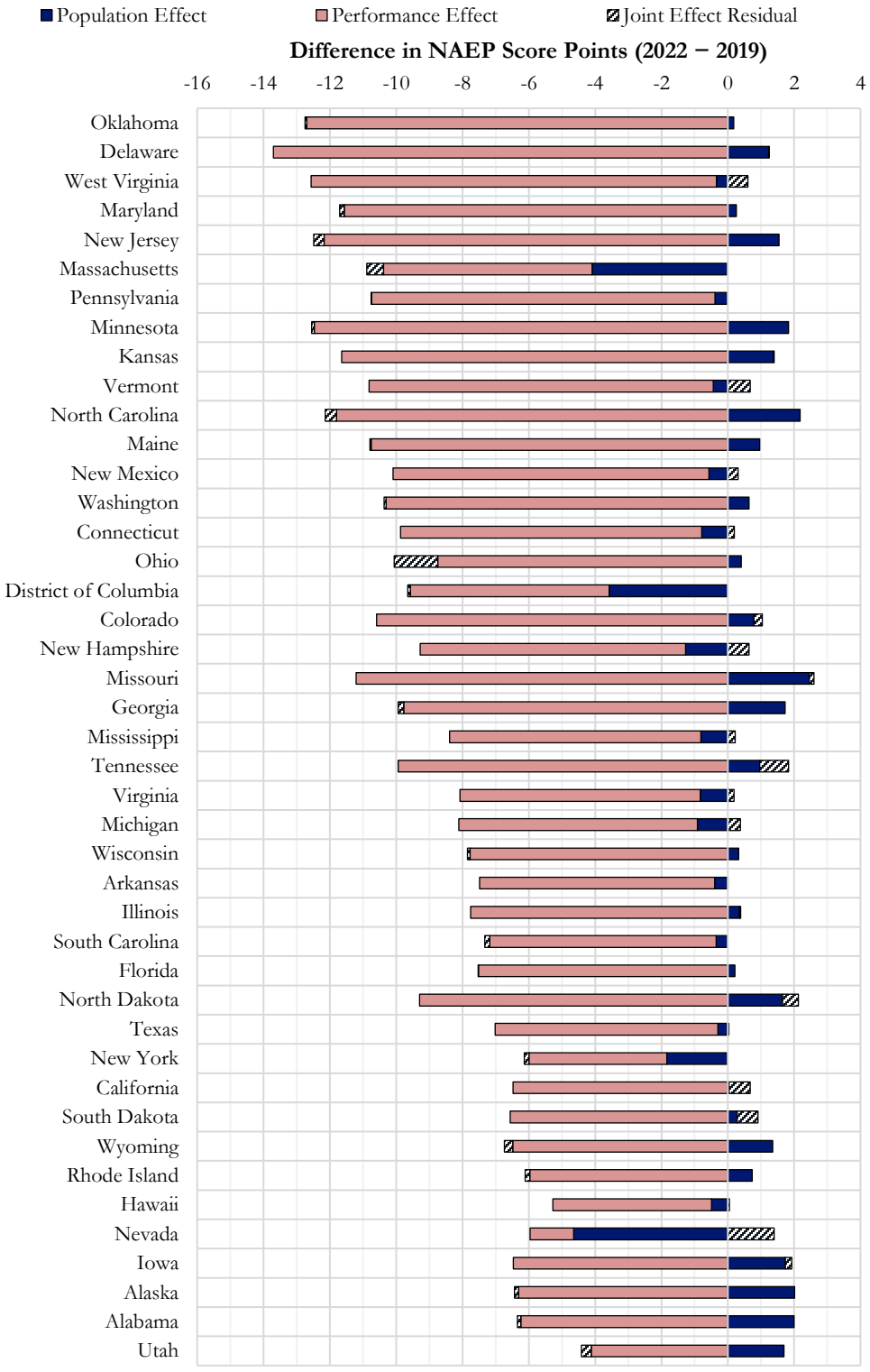
State	Race/ethnicity	2019	2022	2019		2022		2022 – 2019 average scale score difference	Partitioned effects		
		N	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
Connecticut	**Total**	**2,250**	**1,710**	**100.0**	**286.2**	**100.0**	**276.5**	**−9.7**	**−7.7**	**−2.0**	**0.1**
	White	1,260	900	56.4	298.5	50.8	290.2	−8.3			
	Black	260	220	11.6	256.4	13.6	248.9	−7.6			
	Hispanic	550	440	23.5	262.5	26.2	257.2	−5.3			
	Asian	110	90	5.1	328.9	4.9	315.3	−13.9			
	American Indian/ Alaska Native	10	10	0.3	275.9	0.3	276.4	0.5			
	Native Hawaiian/ Other Pacific Islander	‡	‡	‡	‡	‡	‡	‡			
	Two or More Races	70	60	3.0	283.9	4.0	275.5	−8.3			

‡ Reporting standards not met.
NOTE: NAEP = National Assessment of Educational Progress. Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

By Change in NSLP Eligibility

Changes in the NSLP-eligible population generally had small effects compared with changes in performance across states (Exhibit 10). However, three states/regions showed changes of more than 2 points in a negative direction: Massachusetts (−4.1), the District of Columbia (−3.6), and Nevada (−4.6). In Massachusetts's case, the percentage of NSLP-eligible students increased from 28% to 41%; in the District of Columbia, from 71% to 80%; and in Nevada, from 56% to 77% (Exhibit 11). In contrast, positive effects were associated with the change in NSLP status in three other states: North Carolina (2.2), Missouri (2.5), and Alaska (2.0). North Carolina saw a decrease in the percentage of NSLP-eligible students from 43% in 2019 to 35% in 2022, Missouri decreased from 50% to 41%, and Alaska decreased from 47% to 42%. In states in which positive effects were found, those effects partially offset the negative effects related to changes in student performance. Despite the positive effects associated with change in NSLP status, North Carolina, Missouri, and Alaska still had substantial declines in their NAEP mathematics scores.

Exhibit 10. Partitioned Effects for NSLP Eligibility Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by State



NOTE: NSLP = National School Lunch Program. NAEP = National Assessment of Educational Progress. States are ordered by the 2022 − 2019 difference in average scale score. Oregon reported 100% of students (with rounding) as NSLP eligible in 2022 and thus was excluded from this analysis. Arizona, Idaho, Indiana, and Louisiana were excluded from this analysis because of large changes in the percentage of students reported as information not available. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit 11. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics in Massachusetts, North Carolina, the District of Columbia, Missouri, Nevada, and Alaska, by NSLP Eligibility: 2019 and 2022

State	NSLP eligibility	2019	2022	2019		2022		2022 – 2019 average scale score difference	Partitioned effects		
		N	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
Massachusetts	**Total**	**3,420**	**2,490**	**100.0**	**294.5**	**100.0**	**283.6**	**−10.9**	**−6.3**	**−4.1**	**−0.5**
	Eligible	1,450	1,290	27.9	272.4	41.3	263.4	−9.0			
	Not eligible	1,970	1,200	71.7	303.2	58.4	297.9	−5.3			
	Information not available	10	10	0.4	267.9	0.3	272.7	4.9			
North Carolina	**Total**	**4,420**	**3,320**	**100.0**	**283.7**	**100.0**	**273.7**	**−10.0**	**−11.8**	**2.2**	**−0.3**
	Eligible	2,090	1,230	43.4	268.3	35.4	259.0	−9.3			
	Not eligible	2,330	2,090	56.5	295.5	64.6	281.8	−13.7			
	Information not available	‡	‡	‡	‡	‡	‡	‡			
District of Columbia	**Total**	**1,860**	**1,570**	**100.0**	**269.2**	**100.0**	**259.5**	**−9.7**	**−6.0**	**−3.6**	**−0.1**
	Eligible	1,290	1,260	71.0	257.0	79.6	250.2	−6.9			
	Not eligible	570	310	28.9	299.1	20.4	296.2	−2.9			
	Information not available	‡	0	‡	‡	0	†	†			
Missouri	**Total**	**2,420**	**1,800**	**100.0**	**280.8**	**100.0**	**272.2**	**−8.6**	**−11.2**	**2.5**	**0.1**
	Eligible	1,210	740	50.0	267.2	41.1	256.5	−10.8			
	Not eligible	1,210	1,060	49.9	294.4	58.9	283.1	−1.3			
	Information not available	‡	0	‡	‡	0	†	†			
Nevada	**Total**	**2,370**	**1,940**	**100.0**	**273.8**	**100.0**	**269.2**	**−4.6**	**−1.3**	**−4.6**	**1.4**
	Eligible	1,340	1,530	56.5	263.6	76.6	265.3	1.7			
	Not eligible	990	400	41.7	288.3	22.6	283.1	-5.2			
	Information not available	40	20	1.8	255.8	0.8	250.5	-5.3			
Alaska	**Total**	**2,090**	**1,540**	**100.0**	**274.3**	**100.0**	**269.9**	**−4.4**	**−6.3**	**2.0**	**−0.1**
	Eligible	950	640	47.2	260.0	41.6	254.3	−5.6			
	Not eligible	1,090	890	50.2	288.7	57.1	281.7	−7.1			
	Information not available	60	20	2.6	257.0	1.3	253.3	−3.7			

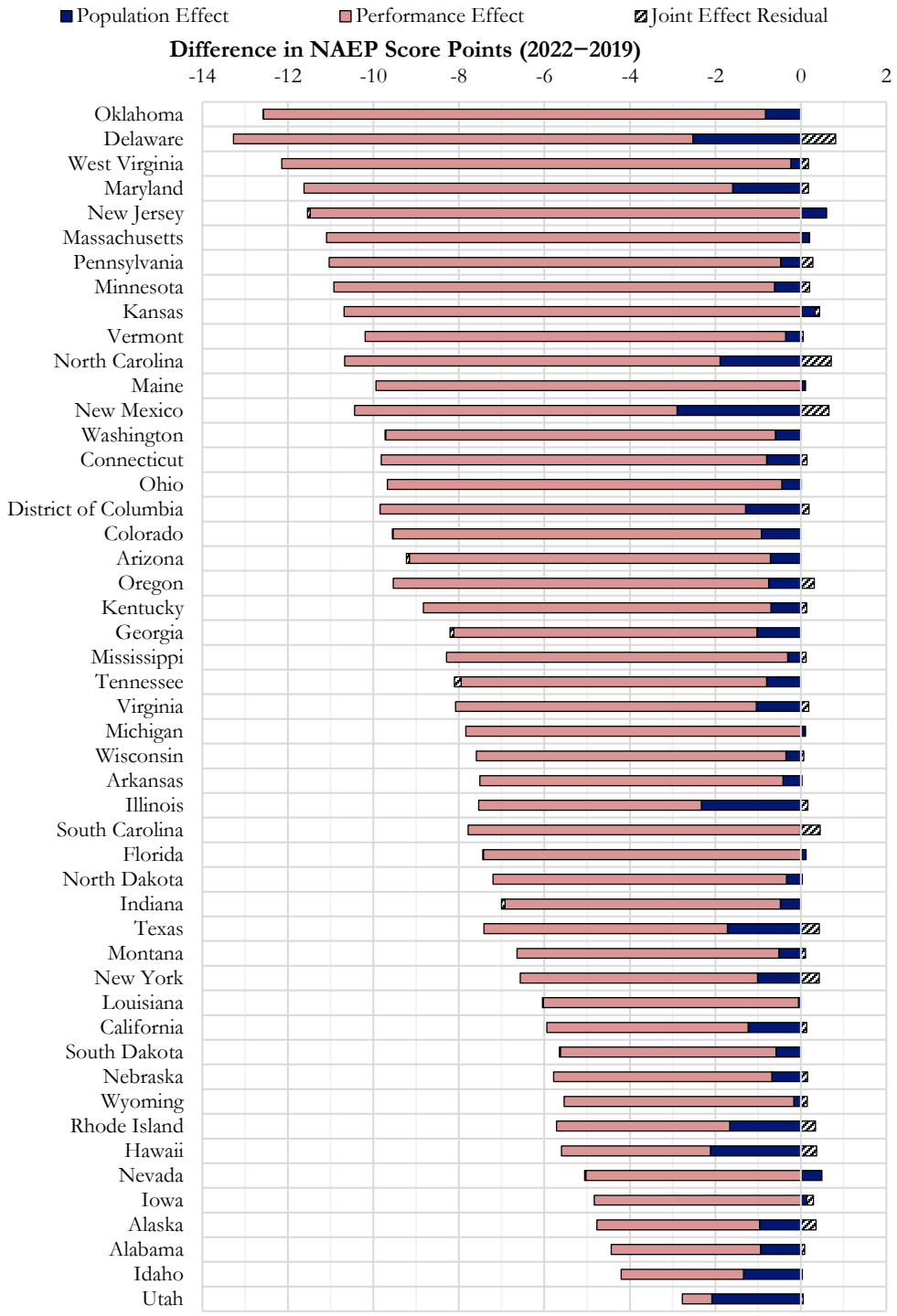
† = Not applicable. ‡ = Reporting standards not met.
NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. States are ordered by the 2022 − 2019 difference in average scale score. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

By Change in EL Status

As illustrated in Exhibit 12, negative effects of 2 or more points were related to EL population changes in five states: Delaware (−2.5), New Mexico (−2.9), Illinois (−2.3), Hawaii (−2.1), and Utah (−2.1). These states showed increases in their EL populations of at least 4 percentage points between 2019 and 2022. Delaware's EL population increased from 4% to 10%, New Mexico's from 12% to 20%, Illinois's from 6% to 12%, Hawaii's from 6% to 11%, and Utah's

from 6% to 11% (Exhibit 13). No states had a positive gain of 2 or more points related to changes in the population of ELs.

Exhibit 12. Partitioned Effects for EL Status Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by State



NOTE: EL = English learner. NAEP = National Assessment of Educational Progress. The results for ELs are based on students who were assessed and cannot be generalized to the total population of such students. States are ordered by the 2022 − 2019 difference in average scale score. New Hampshire had an increase in the percentage of omitted students and thus was excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit 13. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics in Delaware, New Mexico, Illinois, Hawaii, and Utah, by EL Status: 2019 and 2022

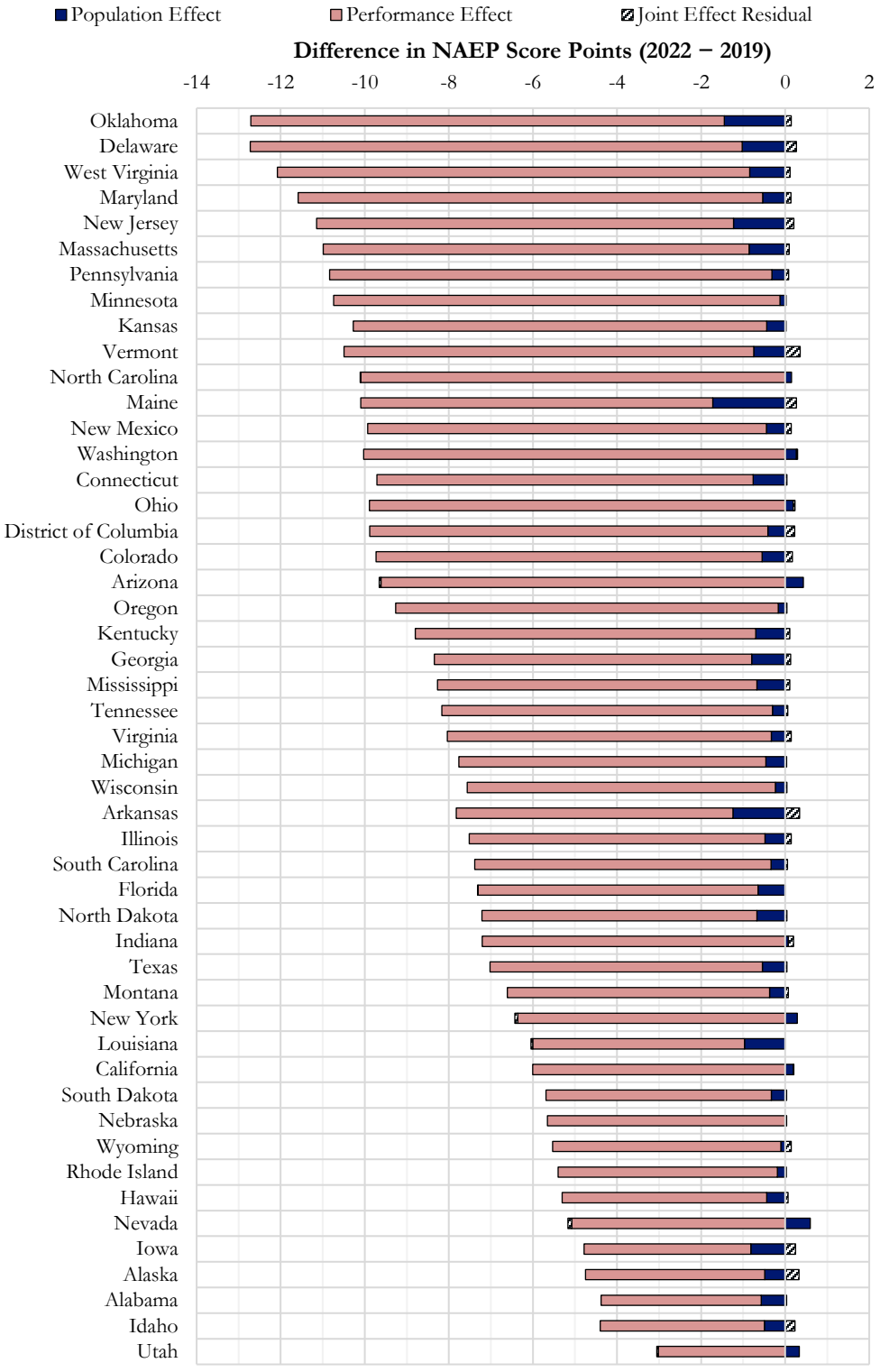
State	EL status	2019	2022	2019		2022		2022 − 2019 average scale score difference	Partitioned effects		
		N	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
Delaware	**Total**	**2,290**	**1,720**	**100.0**	**276.7**	**100.0**	**264.2**	**−12.5**	**−10.7**	**−2.5**	**0.8**
	EL	80	150	4.2	236.3	9.9	237.7	1.4			
	Not EL	2,210	1,570	95.8	278.5	90.0	267.2	−11.3			
	Omitted	0	‡	0	†	‡	‡	†			
New Mexico	**Total**	**2,770**	**2,180**	**100.0**	**268.8**	**100.0**	**259.0**	**−9.8**	**−7.5**	**−2.9**	**0.7**
	EL	340	430	11.8	238.2	19.7	236.8	−1.5			
	Not EL	2,430	1,740	88.2	272.9	80.2	264.5	−8.4			
	Omitted	0	‡	0	†	‡	‡	†			
Illinois	**Total**	**3,560**	**2,740**	**100.0**	**282.6**	**100.0**	**275.2**	**−7.4**	**−5.2**	**−2.3**	**0.2**
	EL	260	410	6.3	244.4	12.0	241.9	−2.5			
	Not EL	3290	2330	93.7	285.1	88.0	279.7	−5.4			
	Omitted	‡	‡	‡	‡	‡	‡	‡			
Hawaii	**Total**	**2,210**	**1,670**	**100.0**	**275.3**	**100.0**	**270.1**	**−5.2**	**−3.5**	**−2.1**	**0.4**
	EL	130	170	5.5	240.4	11.2	242.9	2.6			
	Not EL	2,090	1,500	94.5	277.4	88.8	273.5	−3.8			
	Omitted	0	0	0	†	0	†	†			
Utah	**Total**	**2,450**	**1,720**	**100.0**	**284.9**	**100.0**	**282.2**	**−2.7**	**−0.7**	**−2.1**	**0.1**
	EL	160	180	5.5	246.1	10.6	246.5	0.3			
	Not EL	2,300	1,540	94.5	287.2	89.4	286.5	−0.8			
	Omitted	0	0	0	†	0	†	†			

† = Not applicable. ‡ = Reporting standards not met.
NOTE: NAEP = National Assessment of Educational Progress. EL = English learner. The results for ELs are based on students who were assessed and cannot be generalized to the total population of such students. States are ordered by the 2022 − 2019 difference in average scale score. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

By Change in SD Status

As illustrated in Exhibit 14, no states had shifts of ±2 points or more associated with changes in the population of SDs.

Exhibit 14. Partitioned Effects for SD Status Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by State

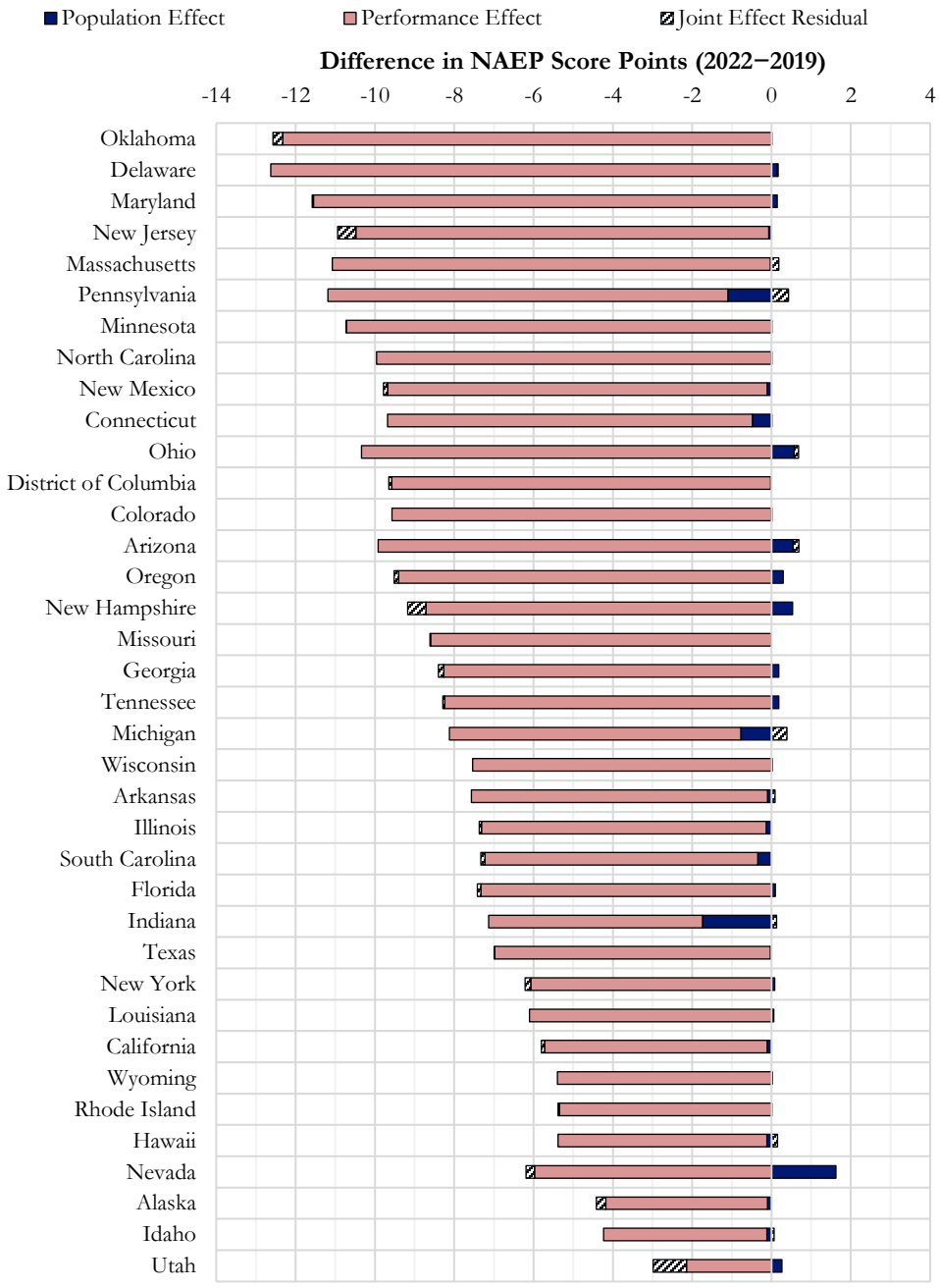


NOTE: SD = students with disabilities. NAEP = National Assessment of Educational Progress. The results for SDs are based on students who were assessed and cannot be generalized to the total population of such students. States are ordered by the 2022 − 2019 difference in average scale score. New Hampshire had an increase in the percentage of omitted students and thus was excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

By Change in Charter School Status

As illustrated in Exhibit 15, among the 37 states with NAEP data on public charter schools, no state had an effect of ±2 points or more related to changes in charter school status.

Exhibit 15. Partitioned Effects for Charter School Status Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by State

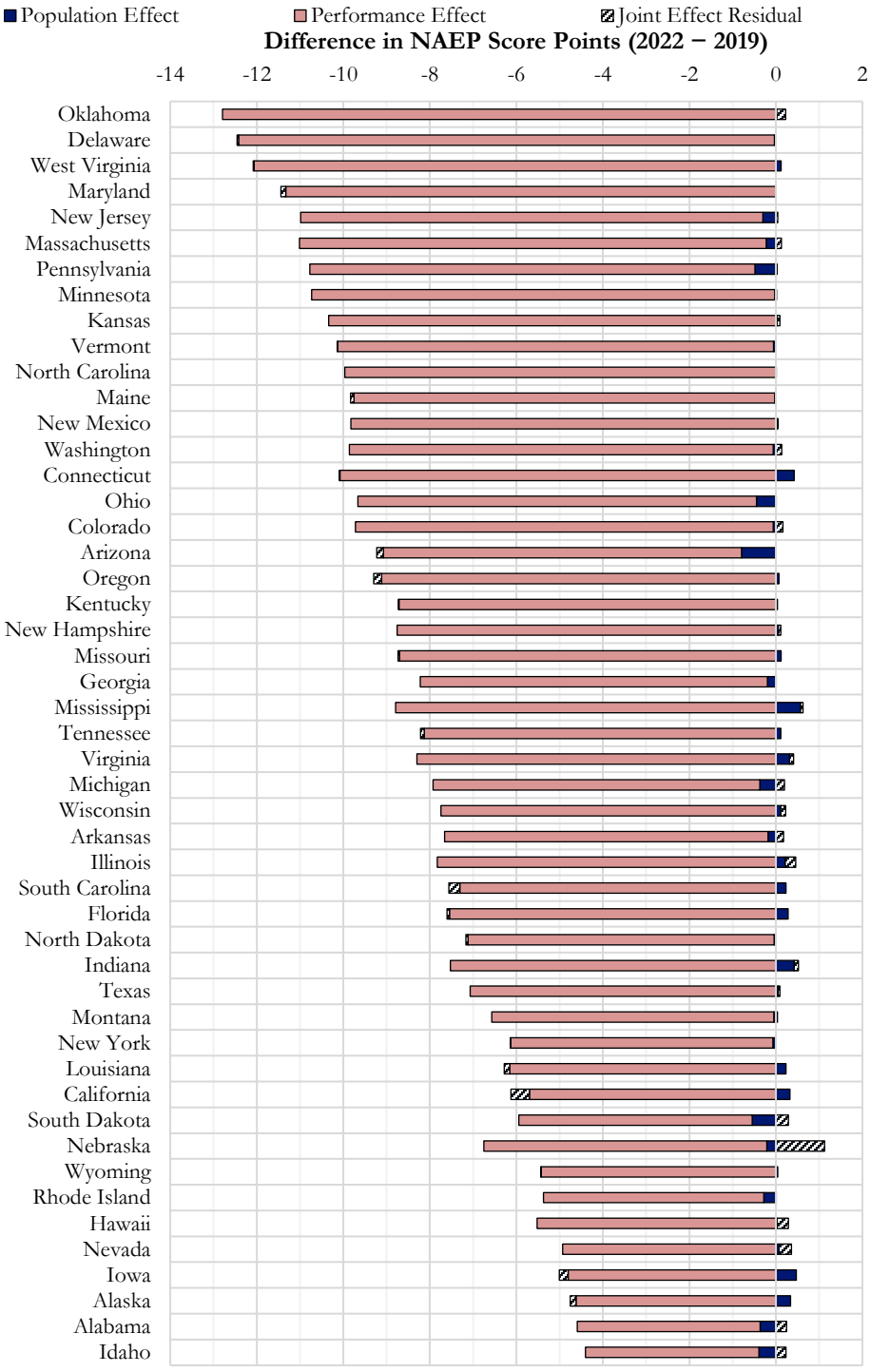


NOTE: NAEP = National Assessment of Educational Progress. States are ordered by the 2022 − 2019 difference in average scale score. Alabama, Iowa, Kansas, Kentucky, Maine, Mississippi, Montana, Nebraska, North Dakota, South Dakota, Vermont, Virginia, Washington, and West Virginia did not report charter school status data in 2019 or 2022 and thus were excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

By Change in School Locale

Finally, as illustrated in Exhibit 16, there were no population effects of ±2 points or more related to changes in the distribution of school locale for any state.

Exhibit 16. Partitioned Effects for School Locale Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by State



NOTE: NAEP = National Assessment of Educational Progress. States are ordered by the 2022 − 2019 difference in average scale score. The District of Columbia did not report data for the categories of suburb, town, and rural and thus was excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

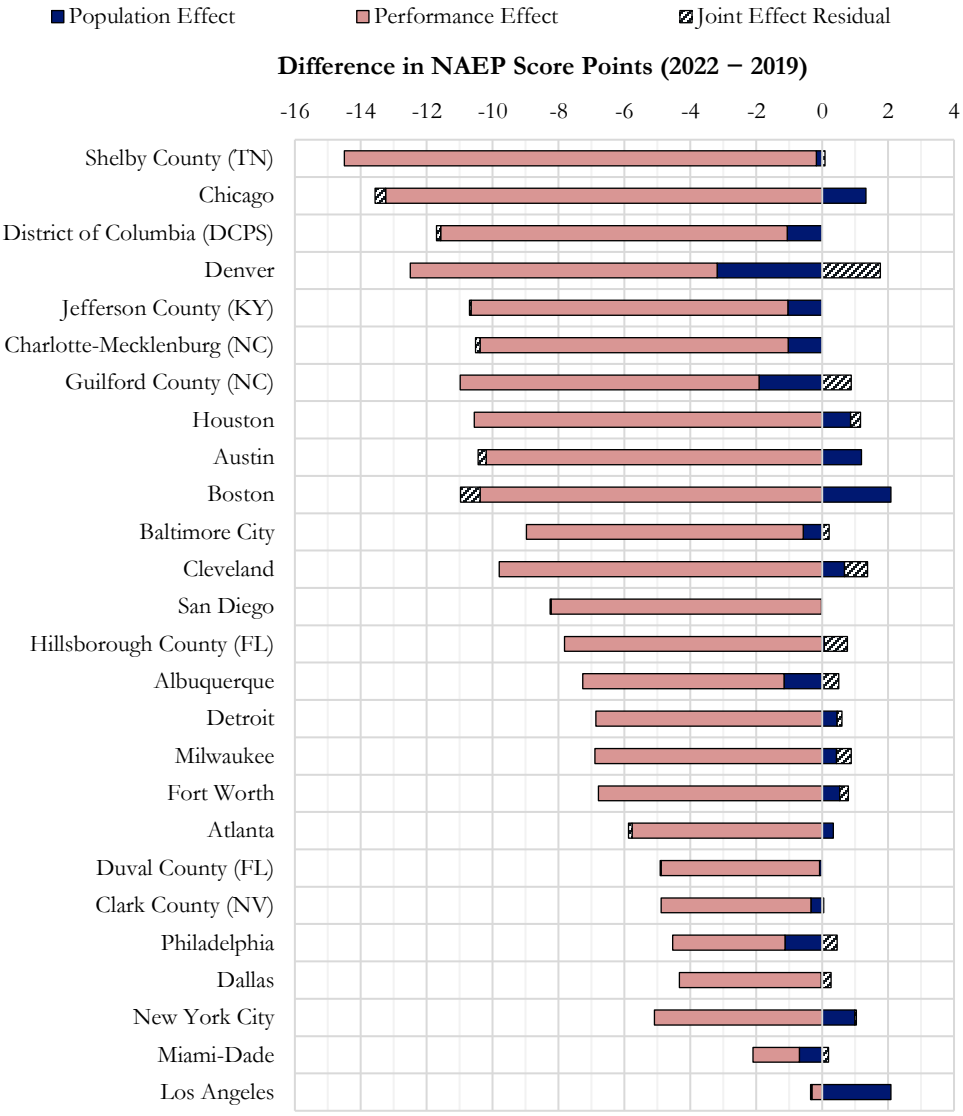
TUDA Results

We applied the partitioning analysis to the TUDA results as well. Similar to the state analyses, we examine the results for TUDAs with a population effect of ±2 points or more.

By Change in Race/Ethnicity

Of the 26 TUDAs, Denver (−3.2) had a negative population effect of greater than 2 points related to the change in the distribution of race/ethnicity, and two TUDAs—Boston (2.1) and Los Angeles (2.1)—had positive population effects of 2 or more points (Exhibits 17–18).[5]

Exhibit 17. Partitioned Effects for Race/Ethnicity Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by TUDA



NOTE: NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. TUDAs are ordered by the 2022 − 2019 difference in average scale score. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

[5] The numerical results used to generate the TUDA figures in Exhibits 17, 19, 21, 23, and 24 are in Appendix C.

Exhibit 18. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics in Denver, Boston, and Los Angeles, by Race/Ethnicity: 2019 and 2022

TUDA	Race/ethnicity	2019	2022	2019		2022		2022 − 2019 average scale score difference	Partitioned effects		
		N	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
Denver	**Total**	**1,010**	**860**	**100.0**	**275.3**	**100.0**	**264.6**	**−10.7**	**−9.3**	**−3.2**	**1.8**
	White	250	170	24.7	307.8	22.6	296.8	−11.1			
	Black	130	130	13.3	254.0	13.8	255.3	1.3			
	Hispanic	540	500	53.5	263.0	55.7	251.3	−11.7			
	Asian	30	30	3.2	290.9	3.2	287.3	−3.5			
	American Indian/ Alaska Native	10	‡	1.2	280.6	‡	‡	‡			
	Native Hawaiian/Other Pacific Islander	0	10	0	†	0.7	223.8	†			
	Two or More Races	40	30	4.2	294.4	3.8	289.4	−5.1			
Boston	**Total**	**1,290**	**790**	**100.0**	**278.8**	**100.0**	**269.9**	**−8.9**	**−10.4**	**2.1**	**−0.6**
	White	210	110	15.0	315.2	14.3	304.6	−10.5			
	Black	380	250	29.6	262.8	27.8	253.5	−9.3			
	Hispanic	540	330	43.1	264.5	42.6	255.4	−9.1			
	Asian	140	80	11.1	323.7	11.8	313.6	−10.1			
	American Indian/ Alaska Native	10	0	0.2	227.0	0	†	†			
	Native Hawaiian/Other Pacific Islander	0	0	0	†	0	†	†			
	Two or More Races	10	30	0.9	337.4	3.6	287.0	−50.5			
Los Angeles	**Total**	**1,720**	**1,310**	**100.0**	**260.7**	**100.0**	**262.4**	**1.7**	**−0.3**	**2.1**	**0.0**
	White	130	150	7.8	287.0	11.9	285.3	−1.7			
	Black	90	80	5.5	248.1	5.7	246.8	−1.3			
	Hispanic	1,350	950	78.1	255.1	72.0	254.3	−0.8			
	Asian	130	100	6.6	297.1	7.8	303.5	6.4			
	American Indian/ Alaska Native	10	10	0.2	307.8	0.3	235.5	−72.2			
	Native Hawaiian/Other Pacific Islander	10	‡	0.4	264.5	‡	‡	‡			
	Two or More Races	30	30	1.5	290.5	2.2	296.1	5.5			

† = Not applicable. ‡ = Reporting standards not met.

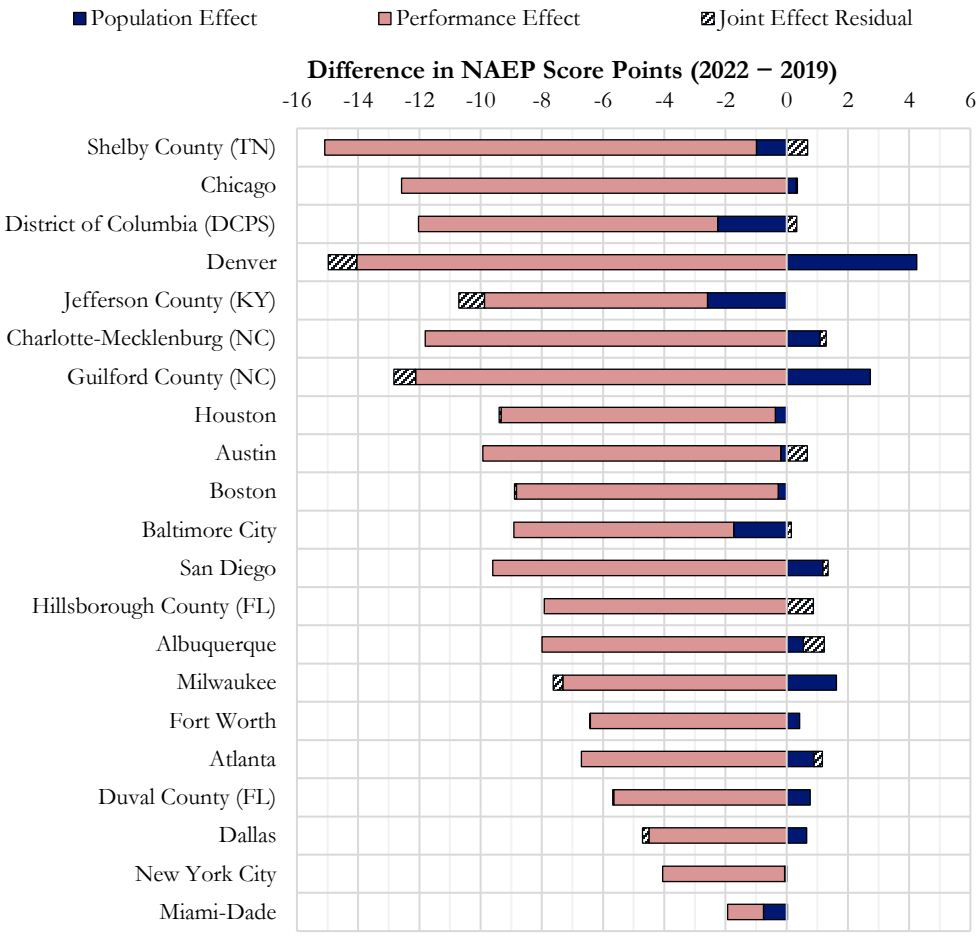
NOTE: NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. TUDAs are ordered by the 2022 − 2019 difference in average scale score. Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

By Change in NSLP Eligibility

As demonstrated in Exhibit 19, four TUDAs had population effects of more than ±2 points associated with changes in the population of NSLP-eligible students. For District of Columbia Public Schools (DCPS) and Jefferson County (KY), these changes were associated with negative population effects (−2.3 and −2.6 points, respectively), but for Denver and Guilford County (NC), the population effects related to NSLP were positive (4.2 and 2.7 points, respectively). In examining what might account for these results, both DCPS and Jefferson County showed increases in the percentage of NSLP-eligible students (by 4 percentage points and nearly 10 percentage points, respectively), whereas Denver and Guilford County showed decreases in the NSLP-eligible percentage (by 12 percentage points and 8 percentage points, respectively; Exhibit 20).

Exhibit 19. Partitioned Effects for NSLP Eligibility Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by TUDA



NOTE: NSLP = National School Lunch Program. NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. TUDAs are ordered by the 2022 − 2019 difference in average scale score. Clark County (NV) and Cleveland reported 100% of students (with rounding) as NSLP eligible in 2022 and thus were excluded from this analysis. Detroit, Los Angeles, and Philadelphia were excluded from this analysis because of a large change in the percentage of students reported as information not available. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit 20. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics in the District of Columbia, Denver, Jefferson County, and Guilford County, by NSLP Eligibility: 2019 and 2022

TUDA	NSLP eligibility	2019	2022	2019		2022		2022 − 2019 average scale score difference	Partitioned effects		
		N	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
District of Columbia (DCPS)	**Total**	**950**	**750**	**100.0**	**268.6**	**100.0**	**256.9**	**−11.7**	**−9.8**	**−2.3**	**0.3**
	Eligible	640	550	69.6	252.5	74.0	243.4	−9.1			
	Not eligible	310	200	30.3	305.6	26.0	295.1	−10.5			
	Information not available	‡	0	‡	‡	0	†	†			
Denver	**Total**	**1,010**	**860**	**100.0**	**275.3**	**100.0**	**264.6**	**−10.7**	**−14.0**	**4.2**	**−0.9**
	Eligible	650	460	64.0	260.2	51.8	251.2	−9.0			
	Not eligible	360	390	36.0	302.1	47.9	279.2	−22.9			
	Information not available	0	10	0	†	0.3	241.2	†			
Jefferson County (KY)	**Total**	**1,080**	**930**	**100.0**	**273.3**	**100.0**	**262.6**	**−10.7**	**−7.3**	**−2.6**	**−0.8**
	Eligible	660	650	60.8	262.8	70.5	252.2	−10.7			
	Not eligible	420	280	39.2	289.6	29.5	287.5	−2.1			
	Information not available	0	0	0	†	0	†	†			
Guilford County (NC)	**Total**	**1,120**	**870**	**100.0**	**280.3**	**100.0**	**270.2**	**−10.1**	**−12.1**	**2.7**	**−0.7**
	Eligible	610	390	52.1	264.2	43.9	256.3	−7.9			
	Not eligible	510	480	47.9	297.8	56.1	281.1	−16.7			
	Information not available	0	0	0	†	0	†	†			

† = Not applicable. ‡ = Reporting standards not met.

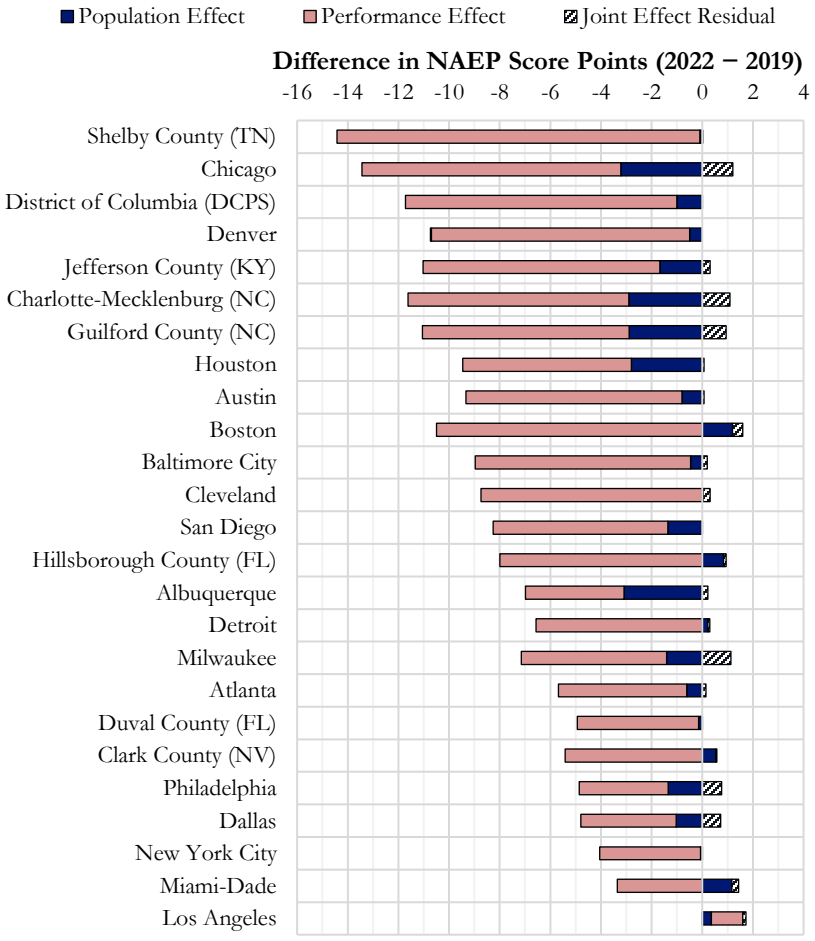
NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. TUDA = Trial Urban District Assessment. TUDAs are ordered by the 2022 − 2019 difference in average scale score. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

By Change in EL Status

For change in EL status, Chicago, Charlotte-Mecklenburg (NC), Guilford County (NC), Houston, and Albuquerque all showed negative population effects of more than 2 points (Exhibit 21). Concomitant with these changes, these TUDAs showed an increase of 5 or more percentage points in the EL population (Exhibit 22). For example, the EL population in Chicago increased from 11% to 20%, the EL population in Houston increased from 23% to 32%, and the EL population in Albuquerque increased from 14% to 23%.

Exhibit 21. Partitioned Effects for EL Status Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by TUDA



NOTE: EL = English learner. NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. The results for ELs are based on students who were assessed and cannot be generalized to the total population of such students. TUDAs are ordered by the 2022 − 2019 difference in average scale score. Fort Worth was excluded from this analysis because the joint effect residual was greater than 2 points. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit 22. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics in Chicago, Charlotte-Mecklenburg (NC), Guilford County (NC), Houston, and Albuquerque, by EL status: 2019 and 2022

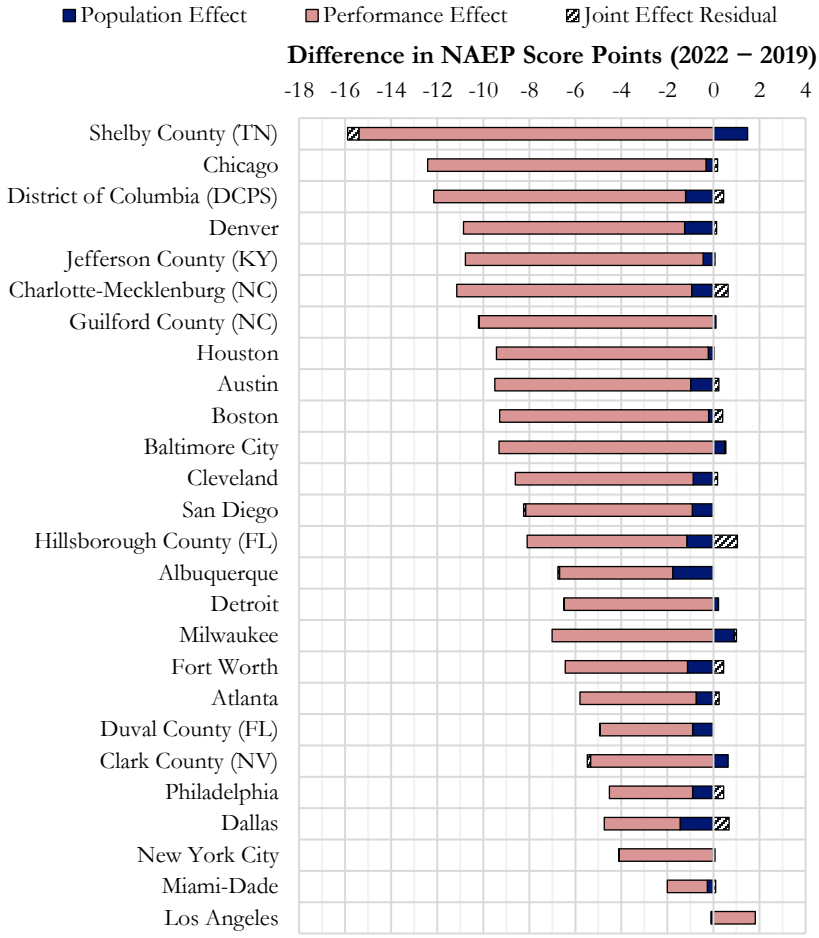
TUDA	EL status	2019	2022	2019		2022		2022 − 2019 average scale score difference	Partitioned effects		
		N	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
Chicago	**Total**	**1,670**	**1,340**	**100.0**	**275.3**	**100.0**	**263.0**	**−12.2**	**−10.2**	**−3.2**	**1.2**
	EL	170	260	11.2	244.2	20.4	245.7	1.5			
	Not EL	1,500	1,080	88.8	279.2	79.6	267.5	−11.7			
	Omitted	‡	‡	‡	‡	‡	‡	‡			
Charlotte-Mecklenburg (NC)	**Total**	**1,070**	**920**	**100.0**	**288.1**	**100.0**	**277.6**	**−10.5**	**−8.7**	**−2.9**	**1.1**
	EL	90	110	8.1	238.7	13.4	244.0	5.3			
	Not EL	980	810	91.8	292.4	86.6	282.8	−9.6			
	Omitted	‡	0	‡	‡	0	†	†			
Guilford County (NC)	**Total**	**1,120**	**870**	**100.0**	**280.3**	**100.0**	**270.2**	**−10.1**	**−8.2**	**−2.9**	**0.9**
	EL	50	90	4.7	236.5	10.9	242.6	6.2			
	Not EL	1,070	770	95.3	282.5	89.1	273.6	−8.9			
	Omitted	0	0	0	†	0	†	†			
Houston	**Total**	**1,560**	**1,270**	**100.0**	**274.1**	**100.0**	**264.7**	**−9.4**	**−6.7**	**−2.8**	**0.1**
	EL	360	380	22.7	251.6	32.3	245.5	−6.1			
	Not EL	1,200	890	77.3	280.7	67.7	273.8	−6.8			
	Omitted	0	0	0	†	0	†	†			
Albuquerque	**Total**	**1,120**	**860**	**100.0**	**266.8**	**100.0**	**260.0**	**−6.8**	**−3.9**	**−3.1**	**0.2**
	EL	170	180	14.3	235.8	22.9	234.1	−1.7			
	Not EL	950	680	85.7	272.0	77.1	267.7	−4.3			
	Omitted	0	0	0	†	0	†	†			

† = Not applicable. ‡ = Reporting standards not met.
NOTE: NAEP = National Assessment of Educational Progress. EL = English learner. TUDA = Trial Urban District Assessment. The results for ELs are based on students who were assessed and cannot be generalized to the total population of such students. TUDAs are ordered by the 2022 − 2019 difference in average scale score. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

By Change in SD Status

As demonstrated in Exhibit 23, none of the TUDAs showed effects of ±2 points or more linked to changes in the SD population.

Exhibit 23. Partitioned Effects for SD Status Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by TUDA

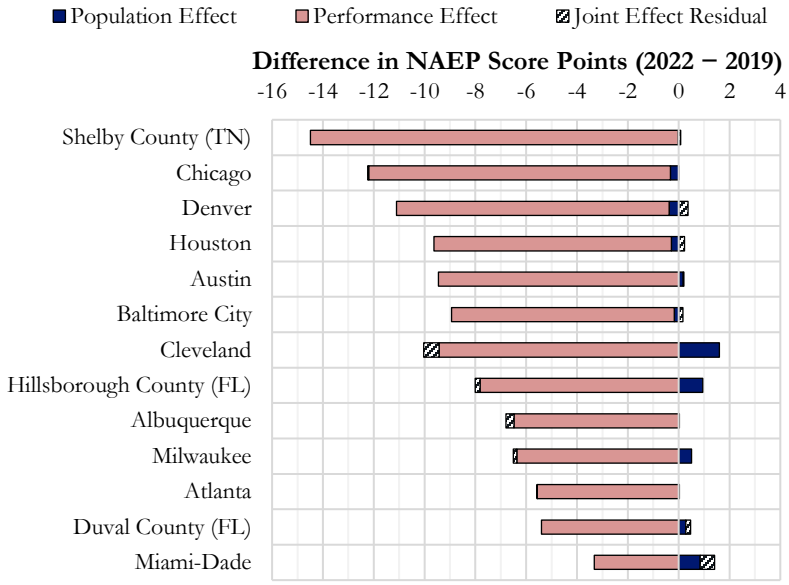


NOTE: SD = students with disabilities. NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. The results for SDs are based on students who were assessed and cannot be generalized to the total population of such students. TUDAs are ordered by the 2022 − 2019 difference in average scale score. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

By Change in Charter School Status

None of the TUDAs had effects of ±2 points or more linked to changes in charter school status (Exhibit 24).

Exhibit 24. Partitioned Effects for Charter School Status Between 2019 and 2022 for Public School Students in Grade 8 NAEP Mathematics, by TUDA



NOTE: NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. TUDAs are ordered by the 2022 − 2019 difference in average scale score. Boston, Charlotte-Mecklenburg, Clark County, the District of Columbia, Dallas, Detroit, Fort Worth, Guilford County, Jefferson County, New York City, Philadelphia, and San Diego did not report charter school status data in 2019 or 2022 and thus were excluded from this analysis. Los Angeles reported no charter schools in 2019 and thus was excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Because most TUDAs include only urban schools, analyses on the effects of change in school locale were not conducted for the TUDAs.

In summary, the partitioning analyses revealed small population effects, with a national population effect near 1 point only for EL status. Almost all score changes appear to relate to changes in student performance between 2019 and 2022. However, the results were more heterogeneous at the state and TUDA levels, especially for race/ethnicity, NSLP eligibility, and EL status. For these variables, some states or TUDAs exhibited population effects of 2 or more points in either a positive or a negative direction.

RQ2 PROPENSITY WEIGHTING ANALYSIS RESULTS: Were observed compositional changes in the NAEP-assessed population between 2019 and 2022 related to changes in NAEP national, state, and TUDA scores?

National Results

In our analysis, the difference in mean achievement estimates between 2019 and 2022 for Grade 8 mathematics was 7.9 points without accounting for differences in the hypothetical probability of being sampled in 2019 (Exhibit 25). The mean achievement estimate for students in 2022 (273.1 points) was 7.9 points lower than the mean achievement estimate for students in 2019 (280.9 points).

When applying propensity weights that account for differences in the probability of being sampled in 2019, the difference in mean achievement estimates decreased from 7.9 points to 6.2 points. The adjusted estimate for 2022 students was 274.8 points. These results indicate that about one fifth of the total difference in scores across years related to differences in the hypothetical likelihood of being sampled in 2019. Alternatively, these results suggest that about four fifths of the observed decrease in achievement between 2019 and 2022 were associated with changes in NAEP performance.

Exhibit 25. Unadjusted and Adjusted Mean Achievement Estimates of 2019 and 2022 Sampled Students in Grade 8 NAEP Mathematics

Unadjusted average scale score			Adjusted average scale score (on propensity weights)	
2019 (n = 142,200)	2022 (n = 107,740)	2022 − 2019 score difference	2022 (n = 107,740)	2022 − 2019 score difference
280.9 (0.26)	273.1 (0.31)	−7.9	274.8 (0.31)	−6.2

NOTE: NAEP = National Assessment of Educational Progress. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Standard errors are shown in parentheses. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

State Results

A similar analysis was run for each state.[6] The results show that nearly all states experienced a smaller decline between 2019 and 2022 based on the adjusted difference than on the unadjusted difference (Exhibit 26). (See Exhibit F7 in Appendix F for full details.) The results showing unadjusted differences as larger than adjusted differences imply that, for these states, compositional changes explained part of the declines in NAEP scores between 2019 and 2022.

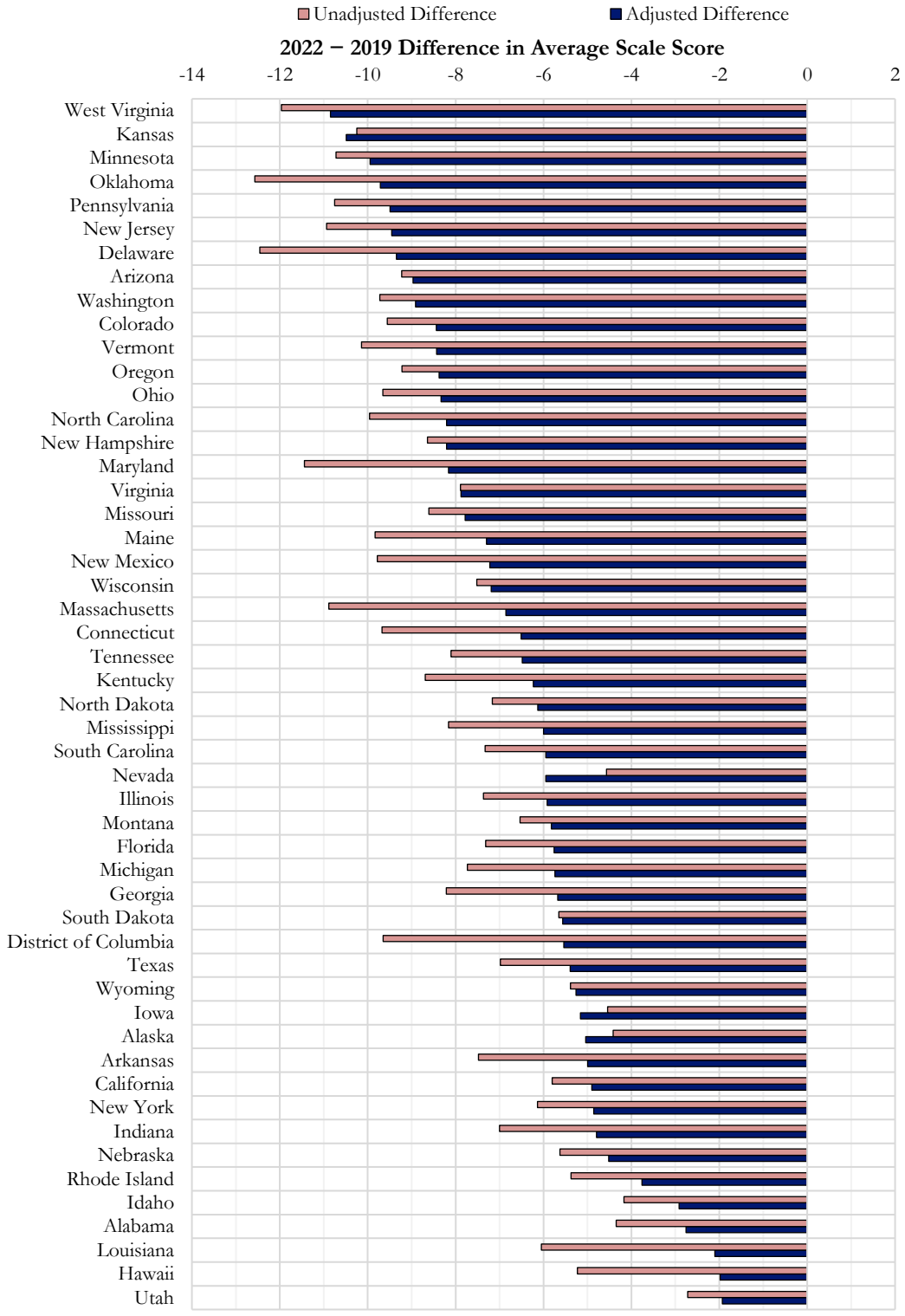
Compositional changes were most highly associated with score declines in The District of Columbia and Massachusetts. In both jurisdictions, unadjusted score declines were more than 4 points greater than adjusted declines. In five other states—Delaware, Maryland, Connecticut, Louisiana, and Hawaii—unadjusted score declines were more than 3 points greater than adjusted declines.

[6] Reference categories for the state and TUDA propensity score analyses are White; not eligible for the NSLP; not EL; not SD; parent(s) graduated from college; 26–100 books at home; not a charter school; and city (school location).

Kansas, Nevada, Iowa, and Alaska were the only states in which the adjusted score decline was larger than the unadjusted score decline, suggesting that the declines for these states might have been larger between the 2019 and 2022 results if the student composition had remained the same.

The states with the largest adjusted performance declines from 2019 to 2022 were West Virginia, Kansas, and Oklahoma, and those with the smallest adjusted performance declines were Louisiana, Hawaii, and Utah. If demographic composition were held constant from 2019 to 2022, West Virginia, Kansas, and Oklahoma would have had the largest score declines in that time period, and Louisiana, Hawaii, and Utah would have had the smallest score declines.

Exhibit 26. Unadjusted and Adjusted 2022 − 2019 Differences in Average Scale Score in Grade 8 NAEP Mathematics, by State



NOTE: NAEP = National Assessment of Educational Progress. States are ordered by the 2022 − 2019 adjusted difference in average scale score. Oregon reported 100% of students (with rounding) as eligible for the National School Lunch Program (NSLP) in 2022; thus, the NSLP eligibility variable was excluded from the state's propensity score model. Variables and levels were excluded from a state propensity score model if there was an insufficient sample size in one or both years. For a full list of variables and levels that were excluded from state propensity score models, refer to Exhibit G1 in Appendix G. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

TUDA Results

Propensity weighting analyses also were conducted for each TUDA. The results showed that most TUDAs had a smaller decline based on the adjusted differences than on the unadjusted differences (Exhibit 27), providing evidence that changes in demographic composition accentuated at least some of the decline in NAEP scores from 2019 to 2022 among the TUDAs.

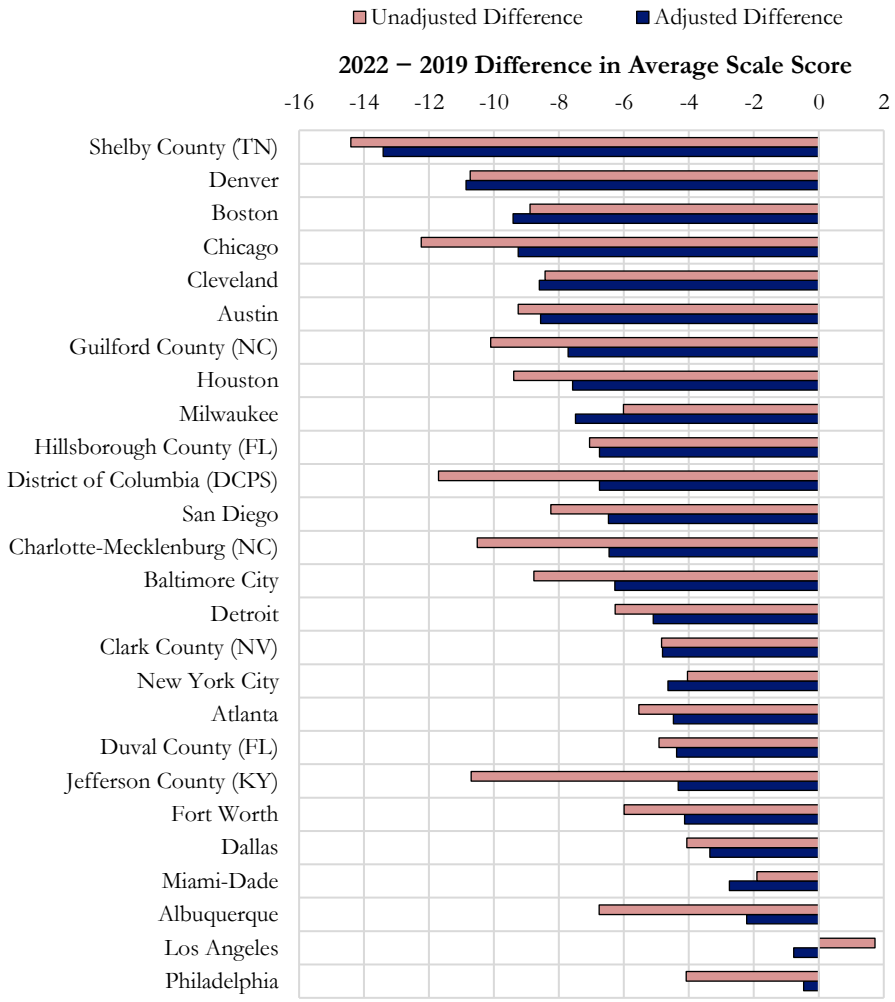
Compositional changes were most strongly associated with score declines in DCPS, Charlotte-Mecklenburg (NC), Jefferson County (KY), Albuquerque, and Philadelphia, all of which had unadjusted declines that were more than 4 points larger in magnitude than their adjusted declines.

Denver, Boston, Cleveland, Milwaukee, New York City, Miami-Dade, and Los Angeles showed larger score declines based on their adjusted score differences than on their unadjusted differences, suggesting that the decline in scores between 2019 and 2022 would have been larger without compositional changes.

The TUDAs with the largest adjusted performance declines from 2019 to 2022 were Shelby County (TN), Denver, and Boston. If demographic composition based on the variables in these models were held constant from 2019 to 2022, these three TUDAs would have had the largest declines. Shelby County also had the largest unadjusted decline.

The TUDAs with the smallest adjusted performance declines were Albuquerque, Los Angeles, and Philadelphia, meaning that if demographic composition were held constant from 2019 to 2022, these three TUDAs would have had the smallest performance declines. Los Angeles is a unique case because it was the only TUDA that showed an unadjusted increase in scores between 2019 and 2022. However, with adjustment, Los Angeles's scores also showed a small decline.

Exhibit 27. Unadjusted and Adjusted 2022 − 2019 Differences in Average Scale Score in Grade 8 NAEP Mathematics, by TUDA



NOTE: NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. TUDAs are ordered by the 2022 − 2019 unadjusted difference in average scale score. National School Lunch Program (NSLP) eligibility was excluded from the propensity score model for Los Angeles because of a large change in the percentage of students reported as information not available. Variables and levels were excluded from a TUDA propensity score model if there was an insufficient sample size in one or both years. For a full list of variables and levels that were excluded from TUDA propensity score models, refer to Exhibit G2 in Appendix G. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

The propensity weighting analysis, which is a multivariate methodology, in contrast to the bivariate Beaton-Chromy methodology, indicated the extent to which compositional changes may have been related to changes in NAEP scores between 2019 and 2022. The national results indicated a small but meaningful relationship between changes in composition and NAEP scores. State and TUDA results showed substantial variation, with many states and TUDAs showing differences of more than 3 points when comparing their adjusted results to their unadjusted results. The results at all three levels (national, state, and TUDA) suggest that the declines in achievement would have been smaller if student composition had remained the same across the two administrations. This finding can be understood as related to minor population decreases across jurisdictions among students who generally have higher average scores (e.g., students not eligible for NSLP) and population increases among students who generally have lower average scores (e.g., students eligible for NSLP).

DISCUSSION

This study addressed two questions:

1. How did the characteristics of the NAEP-assessed population change between 2019 and 2022?
2. Were observed compositional changes in the NAEP-assessed population between 2019 and 2022 related to changes in NAEP national, state, and TUDA scores?

For the first research question, our initial trend analysis rarely showed noteworthy demographic shifts at the national level. Most demographic and school characteristic trends showed a similar pattern across time, except perhaps for ELs, which showed a greater rate of growth from 2019 to 2022.

For the second research question, the national results generally showed small population effects related to changes in NAEP performance. The Beaton-Chromy (2007) partitioning method showed a national population effect near 1 point for EL status. National results from the propensity weighting analysis, in comparison, indicated small but stronger relationships between changes in demographic composition and NAEP scores. The adjusted decrease in mean achievement from 2019 to 2022 was 6.2 points versus the actual decrease of 7.9 points. To help place this finding in context, the results imply that about 20% of the total difference in mean achievement between 2019 and 2022 related to differences in the hypothetical probability of being included in the 2019 sample.

Compared with findings from the Beaton-Chromy (2007) partitioning method, in which population effects were evaluated on a variable-by-variable basis, the findings from the multivariate propensity weighting analysis indicated that when we simultaneously account for multiple changes in the characteristics of the 2022 student sample, the population effect is more pronounced. Although the population effect estimated from the propensity weighting analysis represented about 20% of the total mean difference in achievement across testing years, the population effects estimated from the bivariate partitioning analyses ranged from 0.4% (NSLP) to 11.0% (EL) of the total mean difference in achievement. Given that each variable used in the Beaton-Chromy partitioning method explained some of the score difference, it is perhaps not surprising that simultaneously accounting for all variables in the propensity weighting analysis resulted in a larger estimated effect.

Nationally, both methods showed that the observed changes in demographic composition between 2019 and 2022 had a small yet nonnegligible relationship with changes in NAEP results. The observed decrease in mean national achievement from 2019 to 2022 should mainly be understood as related to a general decrease in performance. In states and TUDAs, the degree to which compositional changes were associated with changes in achievement varied. Although the results from both methodological approaches showed that the relationships between compositional and achievement changes were negligible in some states, in other states they were much more pronounced. For example, results from the propensity score analysis showed that the District of Columbia, Massachusetts, Jefferson County (KY), and Albuquerque all had differences of more than 4 points between their adjusted and unadjusted declines. One potential explanation for these diverging results could be that internal migration between particular states and/or TUDAs led to demographic shifts associated with changes in state or TUDA NAEP results.

Limitations

A few limitations are important to consider when interpreting the results from this study. First, the classification criteria for defining the demographic variables in this study may have changed in some states and districts from 2019 to 2022. For example, in districts and states that began providing free or reduced-price lunch during COVID-19, the percentage of students reported to be NSLP eligible may have increased between 2019 and 2022 without an actual measured increase in the percentage of families meeting previously established income criteria. This caveat is important to note because the validity of results from this type of analysis rests on the assumption of measurement invariance of compositional characteristics across time.

Second, the Beaton-Chromy method does not produce results with standard errors. Rather, the results are point estimates that measure the degree to which changes in average scores can be partitioned into performance, population, and joint (residual) effects. Without standard errors, we cannot draw inferences about whether the effects are real or different from zero. Because of this limitation, we needed some way to determine the practical significance of differences, so we chose to focus on states and TUDAs with population effects of at least ±2 NAEP points. However, without a statistical or theoretical justification for this criterion, it also needs to be seen as having limitations. For example, in a jurisdiction with an unadjusted difference of 10 points between 2019 and 2022, an adjusted difference of 2.0 points would account for 20% of the difference, whereas in another jurisdiction with an unadjusted difference of 6 points, the same adjusted difference would account for 33% of the difference.

Also, the use of different sets of predictor variables in propensity score models across jurisdictions could be viewed as a limitation. This arises because many states and TUDAs have very few or zero sampled students from certain demographic groups, preventing the application of the same set of predictor variables used in the national analysis across all state and TUDA models. Consequently, mean achievement scores were not adjusted using a uniform set of inputs (variables) across jurisdictions.

Finally, the list of variables included in the model that generated the propensity weights was not exhaustive. The list represented many of the NAEP reporting variables, including multiple SES indicators that are commonly used in NAEP reporting or as control variables to explain variation in achievement. The results may be sensitive to the input variables included or not included in the model.

REFERENCES

Austin, P. C., & Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine*, *34*, 3661–3679. https://doi.org/10.1002/sim.6607

Bamat, D., Vannata, S., Liu, L, & Schlesinger, M. (2023). *Exploring the diversity and supports for School District of Philadelphia English learners as they prepare to transition to high school*. Philadelphia Education Research Consortium. https://phledresearch.org/the-diversity-and-supports-for-els-transitioning-to-high-school/

Beaton, A. E., & Chromy, J. R. (2007). *Partitioning NAEP trend data.* American Institutes for Research. http://www.air.org/sites/default/files/downloads/report/PartitioningNAEPdata_BeatonChromy_2007_0.pdf

Broer, M., Bai, Y., & Shipan, R. (2023). *Proposal for a NAEP SES index: Background, findings, and remaining challenges* (AIR-NAEP Working Paper No. 2023-02). American Institutes for Research. https://www.air.org/sites/default/files/2023-09/NAEP-SES-Index-Working-Paper-2023-02.pdf

Eggleston, C., & Fields, J. (2021, March 22). *Census Bureau's Household Pulse Survey shows significant increase in homeschooling rates in fall 2020*. U.S. Census Bureau. https://www.census.gov/library/stories/2021/03/homeschooling-on-the-rise-during-covid-19-pandemic.html

Hubler, S. (2022, May 17). With plunging enrollment, a 'seismic hit' to public schools. *The New York Times*. https://www.nytimes.com/2022/05/17/us/public-schools-falling-enrollment.html

Ladd, H., & Huberlie, A. (2023). Where have all the students gone? *EY Parthenon.* https://www.ey.com/en_us/strategy/declining-enrollment-in-public-K–12-schools

Mosteller, F., & Tukey, J. W. (1977). *Data analysis and regression*. Addison-Wesley.

National Center for Education Statistics. (2022). *Condition of Education: Public school enrollment*. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/programs/coe/indicator/cga

Olmos, A., & Govindasamy, P. (2015). A practical guide for using propensity score weighting in R. *Practical Assessment, Research & Evaluation*, *20*(13). https://www.math.umd.edu/~slud/s818M-MissingData/PropensityScoreWeightingR.pdf

Raudenbush, S. W. (2004). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics*, *29*(1), 121–129. https://www.jstor.org/stable/3701310

Ridgeway, G., Kovalchik, S. A., Griffin, B. A., & Kabeto, M. U. (2015). Propensity score analysis with survey weighted data. *Journal of Causal Inference*, *3*(2), 237–249. https://doi.org/10.1515/jci-2014-0039

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.2307/2335942

Schachter, J., Borsella, P., & Knapp, A. (2021, December 21). *New population estimates show COVID-19 pandemic significantly disrupted migration across borders*. U.S. Census Bureau. https://www.census.gov/library/stories/2021/12/net-international-migration-at-lowest-levels-in-decades.html

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, *13*(2), 238–241. https://doi.org/10.1111/j.2517-6161.1951.tb00088.x

Villegas, L., & Garcia, A. (2022). *Educating English learners during the pandemic*. New America. https://www.newamerica.org/education-policy/reports/educating-english-learners-during-the-pandemic/

APPENDICES

Appendix A. Partitioned Effects for National Public

Exhibit A1. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics, by Race/Ethnicity: 2019 and 2022

Race/ethnicity	2019 *N*	2022 *N*	2019 % of total	2019 Avg. scale score	2022 % of total	2022 Avg. scale score	2022 − 2019 average scale score difference	Partitioned effects: Performance effect	Partitioned effects: Population effect	Partitioned effects: Joint effect residual
Total	**142,200**	**107,740**	**100.0**	**281.0**	**100.0**	**273.1**	**−7.9**	**−7.5**	**−0.4**	**0.0**
White	70,350	50,910	48.0	291.5	45.2	283.6	−7.9			
Black	24,970	19,140	14.7	259.2	14.7	252.1	−7.1			
Hispanic	31,400	25,550	27.0	268.0	28.8	260.8	−7.2			
Asian	6,350	4,920	5.6	312.6	6.0	306.3	−6.3			
American Indian/ Alaska Native	2,990	1,740	0.9	262.8	0.8	256.7	−6.0			
Native Hawaiian/ Other Pacific Islander	1,110	690	0.4	262.7	0.3	261.6	−1.1			
Two or More Races	5,030	4,790	3.5	285.0	4.2	275.4	−9.5			

NOTE: NAEP = National Assessment of Educational Progress. Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit A2. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics, by NSLP Eligibility: 2019 and 2022

NSLP eligibility	2019 *N*	2022 *N*	2019 % of total	2019 Avg. scale score	2022 % of total	2022 Avg. scale score	2022 − 2019 average scale score difference	Partitioned effects: Performance effect	Partitioned effects: Population effect	Partitioned effects: Joint effect residual
Total	**142,200**	**107,740**	**100.0**	**281.0**	**100.0**	**273.1**	**−7.9**	**−7.9**	**0.0**	**0.0**
Eligible	72,050	54,160	49.9	266.1	49.7	259.5	−6.6			
Not eligible	68,700	52,130	48.8	296.0	48.7	287.0	−9.0			
Information not available	1,450	1,460	1.3	286.2	1.6	274.1	−12.1			

NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit A3. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics, by EL Status: 2019 and 2022

	2019	2022	2019		2022		2022 − 2019 average scale score difference	Partitioned effects		
EL status	*N*	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
Total	**142,200**	**107,740**	**100.0**	**281.0**	**100.0**	**273.1**	**−7.9**	**−7.1**	**−0.9**	**0.1**
EL	9,560	9,130	7.3	242.8	9.4	241.0	−1.9			
Not EL	132,630	98,570	92.6	284.0	90.5	276.5	−7.5			
Omitted	20	40	0.0	269.4	0.0	265.0	−4.4			

NOTE: NAEP = National Assessment of Educational Progress. EL = English learner. The results for ELs are based on students who were assessed and cannot be generalized to the total population of such students. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit A4. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics, by SD Status: 2019 and 2022

	2019	2022	2019		2022		2022 − 2019 average scale score difference	Partitioned effects		
SD status	*N*	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
Total	**142,200**	**107,740**	**100.0**	**281.0**	**100.0**	**273.1**	**−7.9**	**−7.6**	**−0.3**	**0.0**
SD	18,810	14,960	13.7	246.8	14.5	242.3	−4.5			
Not SD	123,370	92,750	86.3	286.4	85.5	278.4	−8.1			
Omitted	20	40	0.0	281.5	0.0	269.3	−12.2			

NOTE: NAEP = National Assessment of Educational Progress. SD = students with disabilities. The results for SDs are based on students who were assessed and cannot be generalized to the total population of such students. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit A5. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics, by Charter School Status: 2019 and 2022

	2019	2022	2019		2022		2022 − 2019 average scale score difference	Partitioned effects		
Charter school status	*N*	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
Total	**142,200**	**107,740**	**100.0**	**281.0**	**100.0**	**273.1**	**−7.9**	**−7.8**	**0.0**	**0.0**
Charter school	7,990	7,170	5.5	276.8	6.5	267.9	−8.9			
Not a charter school	134,220	100,580	94.5	281.2	93.5	273.5	−7.7			

NOTE: NAEP = National Assessment of Educational Progress. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit A6. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics, by School Locale: 2019 and 2022

School locale	2019	2022	2019		2022		2022 − 2019 average scale score difference	Partitioned effects		
	N	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
Total	**142,200**	**107,740**	**100.0**	**281.0**	**100.0**	**273.1**	**−7.9**	**−7.8**	**−0.1**	**0.0**
City	52,690	40,250	28.9	275.5	29.0	268.5	−7.1			
Suburb	39,800	30,090	40.4	285.8	39.1	277.5	−8.3			
Town	19,280	14,280	11.6	276.1	11.7	268.4	−7.7			
Rural	30,440	23,120	19.1	282.0	20.2	274.1	−8.0			

NOTE: NAEP = National Assessment of Educational Progress. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Appendix B. Partitioned Effects by State

Exhibit B1. Average Scale Score Differences and Partitioned Effects for Race/Ethnicity for Public School Students in Grade 8 NAEP Mathematics, by State: 2019 to 2022

State	2022 − 2019 average scale score difference	Partitioned effects for race/ethnicity		
		Performance effect	Population effect	Joint effect residual
Oklahoma	−12.6	−12.4	0.0	−0.2
Delaware	−12.5	−11.1	−1.6	0.2
West Virginia	−12.0	−12.1	0.0	0.1
Maryland	−11.4	−10.2	−1.2	0.0
New Jersey	−10.9	−11.3	−1.0	1.4
Massachusetts	−10.9	−10.3	−1.0	0.5
Pennsylvania	−10.7	−10.6	−0.3	0.1
Minnesota	−10.7	−9.6	−1.0	−0.2
Kansas	−10.2	−9.9	−0.2	−0.1
Vermont	−10.1	−9.8	−0.5	0.2
North Carolina	−10.0	−8.9	−1.2	0.1
Maine	−9.8	−9.5	−0.3	0.0
New Mexico	−9.8	−9.2	−0.6	0.0
Washington	−9.7	−8.4	−1.1	−0.2
Connecticut	−9.7	−7.7	−2.0	0.1
Ohio	−9.7	−10.5	0.6	0.2
District of Columbia	−9.7	−9.9	0.1	0.1
Colorado	−9.6	−8.7	−0.7	−0.2
Arizona	−9.2	−8.9	−0.4	0.0
Oregon	−9.2	−9.2	−0.3	0.2
Kentucky	−8.7	−7.6	−1.1	0.0
New Hampshire	−8.6	−8.3	−0.1	−0.3
Missouri	−8.6	−8.0	−0.5	−0.1
Georgia	−8.2	−7.4	−0.9	0.0
Mississippi	−8.2	−7.4	−0.8	0.0
Tennessee	−8.1	−7.3	−0.7	−0.1
Virginia	−7.9	−8.7	0.6	0.3
Michigan	−7.7	−6.8	−0.9	0.0
Wisconsin	−7.5	−7.2	−0.3	0.0
Arkansas	−7.5	−7.5	0.1	−0.1
Illinois	−7.4	−7.5	0.4	−0.2
South Carolina	−7.3	−6.5	−0.9	0.1
Florida	−7.3	−7.1	−0.1	−0.1
Indiana	−7.0	−6.0	−0.8	−0.2
Texas	−7.0	−7.6	0.6	0.0
Montana	−6.5	−6.4	−0.2	0.1
New York	−6.1	−5.7	−0.4	0.0
Louisiana	−6.0	−5.0	−1.3	0.2
California	−5.8	−5.5	−0.3	0.1
South Dakota	−5.7	−6.1	0.6	−0.1
Nebraska	−5.6	−5.1	−0.4	−0.1
Wyoming	−5.4	−5.6	0.0	0.2
Rhode Island	−5.4	−4.5	−0.9	0.1
Hawaii	−5.2	−5.7	1.2	−0.8
Nevada	−4.6	−4.1	−0.4	−0.1
Iowa	−4.5	−4.5	0.0	0.0
Alaska	−4.4	−4.6	0.4	−0.2
Alabama	−4.3	−3.4	−1.2	0.2

State	2022 − 2019 average scale score difference	Partitioned effects for race/ethnicity		
		Performance effect	Population effect	Joint effect residual
Idaho	−4.2	−4.0	−0.2	0.0
Utah	−2.7	−1.6	−1.1	0.0

NOTE: NAEP = National Assessment of Educational Progress. States are ordered by the 2022 − 2019 average scale score difference. North Dakota did not report data for students of Two or More Races in 2019 and thus was excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit B2. Average Scale Score Differences and Partitioned Effects for NSLP Eligibility for Public School Students in Grade 8 NAEP Mathematics, by State: 2019 to 2022

State	2022 − 2019 average scale score difference	Partitioned effects for NSLP eligibility		
		Performance effect	Population effect	Joint effect residual
Oklahoma	−12.6	−12.7	0.2	−0.1
Delaware	−12.5	−13.7	1.2	0.0
West Virginia	−12.0	−12.2	−0.3	0.6
Maryland	−11.4	−11.6	0.3	−0.1
New Jersey	−10.9	−12.2	1.5	−0.3
Massachusetts	−10.9	−6.3	−4.1	−0.5
Pennsylvania	−10.7	−10.4	−0.4	0.0
Minnesota	−10.7	−12.5	1.8	−0.1
Kansas	−10.2	−11.6	1.4	0.0
Vermont	−10.1	−10.4	−0.4	0.7
North Carolina	−10.0	−11.8	2.2	−0.3
Maine	−9.8	−10.8	1.0	0.0
New Mexico	−9.8	−9.5	−0.6	0.3
Washington	−9.7	−10.3	0.6	−0.1
Connecticut	−9.7	−9.1	−0.8	0.2
Ohio	−9.7	−8.7	0.4	−1.3
District of Columbia	−9.7	−6.0	−3.6	−0.1
Colorado	−9.6	−10.6	0.8	0.2
Kentucky	−8.7	−8.5	−0.3	0.1
New Hampshire	−8.6	−8.0	−1.3	0.6
Missouri	−8.6	−11.2	2.5	0.1
Georgia	−8.2	−9.8	1.7	−0.2
Mississippi	−8.2	−7.6	−0.8	0.2
Tennessee	−8.1	−9.9	1.0	0.9
Virginia	−7.9	−7.2	−0.8	0.2
Michigan	−7.7	−7.2	−0.9	0.4
Wisconsin	−7.5	−7.8	0.3	−0.1
Arkansas	−7.5	−7.1	−0.4	0.0
Illinois	−7.4	−7.8	0.3	0.0
South Carolina	−7.3	−6.8	−0.3	−0.2
Florida	−7.3	−7.5	0.2	0.0
North Dakota	−7.2	−9.3	1.7	0.5
Texas	−7.0	−6.7	−0.3	0.0
Montana	−6.5	−6.5	−1.4	1.4
New York	−6.1	−4.2	−1.8	−0.1
California	−5.8	−6.5	0.0	0.6
South Dakota	−5.7	−6.6	0.3	0.6
Nebraska	−5.6	−6.2	0.5	0.1
Wyoming	−5.4	−6.5	1.3	−0.3
Rhode Island	−5.4	−6.0	0.7	−0.1
Hawaii	−5.2	−4.8	−0.5	0.0

State	2022 − 2019 average scale score difference	Partitioned effects for NSLP eligibility		
		Performance effect	Population effect	Joint effect residual
Nevada	−4.6	−1.3	−4.6	1.4
Iowa	−4.5	−6.5	1.7	0.2
Alaska	−4.4	−6.3	2.0	−0.1
Alabama	−4.3	−6.2	2.0	−0.1
Utah	−2.7	−4.1	1.7	−0.3

NOTE: NSLP = National School Lunch Program. NAEP = National Assessment of Educational Progress. States are ordered by the 2022 − 2019 average scale score difference. Oregon reported 100% of students (with rounding) as NSLP eligible in 2022 and thus was excluded from this analysis. Arizona, Idaho, Indiana, and Louisiana were excluded from this analysis because of a large change in the percentage of students reported as information not available. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit B3. Average Scale Score Differences and Partitioned Effects for EL Status for Public School Students in Grade 8 NAEP Mathematics, by State: 2019 to 2022

State	2022 − 2019 average scale score difference	Partitioned effects for EL status		
		Performance effect	Population effect	Joint effect residual
Oklahoma	−12.6	−11.7	−0.8	0.0
Delaware	−12.5	−10.7	−2.5	0.8
West Virginia	−12.0	−11.9	−0.2	0.2
Maryland	−11.4	−10.0	−1.6	0.2
New Jersey	−10.9	−11.5	0.6	−0.1
Massachusetts	−10.9	−11.1	0.2	0.0
Pennsylvania	−10.7	−10.6	−0.5	0.3
Minnesota	−10.7	−10.3	−0.6	0.2
Kansas	−10.2	−10.7	0.3	0.1
Vermont	−10.1	−9.8	−0.4	0.1
North Carolina	−10.0	−8.8	−1.9	0.7
Maine	−9.8	−9.9	0.1	0.0
New Mexico	−9.8	−7.5	−2.9	0.7
Washington	−9.7	−9.1	−0.6	0.0
Connecticut	−9.7	−9.0	−0.8	0.1
Ohio	−9.7	−9.2	−0.4	0.0
District of Columbia	−9.7	−8.5	−1.3	0.2
Colorado	−9.6	−8.6	−0.9	0.0
Arizona	−9.2	−8.4	−0.7	−0.1
Oregon	−9.2	−8.8	−0.8	0.3
Kentucky	−8.7	−8.1	−0.7	0.1
Missouri	−8.6	−7.8	−0.7	−0.1
Georgia	−8.2	−7.1	−1.0	−0.1
Mississippi	−8.2	−8.0	−0.3	0.1
Tennessee	−8.1	−7.1	−0.8	−0.2
Virginia	−7.9	−7.0	−1.0	0.2
Michigan	−7.7	−7.8	0.1	0.0
Wisconsin	−7.5	−7.2	−0.3	0.1
Arkansas	−7.5	−7.1	−0.4	0.0
Illinois	−7.4	−5.2	−2.3	0.2
South Carolina	−7.3	−7.8	0.0	0.5
Florida	−7.3	−7.4	0.1	0.0
North Dakota	−7.2	−6.9	−0.3	0.0
Indiana	−7.0	−6.4	−0.5	−0.1
Texas	−7.0	−5.7	−1.7	0.4
Montana	−6.5	−6.1	−0.5	0.1

State	2022 − 2019 average scale score difference	Partitioned effects for EL status		
		Performance effect	Population effect	Joint effect residual
New York	−6.1	−5.6	−1.0	0.4
Louisiana	−6.0	−6.0	−0.1	0.0
California	−5.8	−4.7	−1.2	0.1
South Dakota	−5.7	−5.0	−0.6	0.0
Nebraska	−5.6	−5.1	−0.7	0.2
Wyoming	−5.4	−5.4	−0.2	0.1
Rhode Island	−5.4	−4.1	−1.7	0.3
Hawaii	−5.2	−3.5	−2.1	0.4
Nevada	−4.6	−5.0	0.5	0.0
Iowa	−4.5	−4.8	0.1	0.2
Alaska	−4.4	−3.8	−1.0	0.4
Alabama	−4.3	−3.5	−0.9	0.1
Idaho	−4.2	−2.9	−1.3	0.0
Utah	−2.7	−0.7	−2.1	0.1

NOTE: EL = English learner. NAEP = National Assessment of Educational Progress. The results for ELs are based on students who were assessed and cannot be generalized to the total population of such students. States are ordered by the 2022 − 2019 average scale score difference. New Hampshire had an increase in the percentage of omitted students and thus was excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit B4. Average Scale Score Differences and Partitioned Effects for SD Status for Public School Students in Grade 8 NAEP Mathematics, by State: 2019 to 2022

State	2022 − 2019 average scale score difference	Partitioned effects for SD status		
		Performance effect	Population effect	Joint effect residual
Oklahoma	−12.6	−11.3	−1.4	0.1
Delaware	−12.5	−11.7	−1.0	0.3
West Virginia	−12.0	−11.2	−0.8	0.1
Maryland	−11.4	−11.0	−0.5	0.1
New Jersey	−10.9	−9.9	−1.2	0.2
Massachusetts	−10.9	−10.1	−0.9	0.1
Pennsylvania	−10.7	−10.5	−0.3	0.1
Minnesota	−10.7	−10.6	−0.1	0.0
Kansas	−10.2	−9.8	−0.4	0.0
Vermont	−10.1	−9.8	−0.7	0.4
North Carolina	−10.0	−10.1	0.2	0.0
Maine	−9.8	−8.4	−1.7	0.3
New Mexico	−9.8	−9.5	−0.4	0.1
Washington	−9.7	−10.0	0.3	0.0
Connecticut	−9.7	−9.0	−0.8	0.0
Ohio	−9.7	−9.9	0.2	0.0
District of Columbia	−9.7	−9.5	−0.4	0.2
Colorado	−9.6	−9.2	−0.6	0.2
Arizona	−9.2	−9.6	0.4	0.0
Oregon	−9.2	−9.1	−0.2	0.0
Kentucky	−8.7	−8.1	−0.7	0.1
Missouri	−8.6	−8.3	−0.5	0.2
Georgia	−8.2	−7.5	−0.8	0.1
Mississippi	−8.2	−7.6	−0.7	0.1
Tennessee	−8.1	−7.9	−0.3	0.1
Virginia	−7.9	−7.7	−0.3	0.1
Michigan	−7.7	−7.3	−0.5	0.0
Wisconsin	−7.5	−7.3	−0.2	0.0

State	2022 − 2019 average scale score difference	Partitioned effects for SD status		
		Performance effect	Population effect	Joint effect residual
Arkansas	−7.5	−6.6	−1.2	0.3
Illinois	−7.4	−7.0	−0.5	0.1
South Carolina	−7.3	−7.0	−0.3	0.1
Florida	−7.3	−6.7	−0.6	0.0
North Dakota	−7.2	−6.5	−0.7	0.0
Indiana	−7.0	−7.2	0.1	0.1
Texas	−7.0	−6.5	−0.5	0.0
Montana	−6.5	−6.2	−0.4	0.1
New York	−6.1	−6.4	0.3	−0.1
Louisiana	−6.0	−5.0	−1.0	0.0
California	−5.8	−6.0	0.2	0.0
South Dakota	−5.7	−5.4	−0.3	0.0
Nebraska	−5.6	−5.7	0.0	0.0
Wyoming	−5.4	−5.4	−0.1	0.1
Rhode Island	−5.4	−5.2	−0.2	0.0
Hawaii	−5.2	−4.9	−0.4	0.1
Nevada	−4.6	−5.1	0.6	−0.1
Iowa	−4.5	−4.0	−0.8	0.2
Alaska	−4.4	−4.3	−0.5	0.3
Alabama	−4.3	−3.8	−0.6	0.0
Idaho	−4.2	−3.9	−0.5	0.2
Utah	−2.7	−3.0	0.3	0.0

NOTE: SD = students with disabilities. NAEP = National Assessment of Educational Progress. The results for SDs are based on students who were assessed and cannot be generalized to the total population of such students. States are ordered by the 2022 − 2019 average scale score difference. New Hampshire had an increase in the percentage of omitted students and thus was excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit B5. Average Scale Score Differences and Partitioned Effects for Charter School Status for Public School Students in Grade 8 NAEP Mathematics, by State: 2019 to 2022

State	2022 − 2019 average scale score difference	Partitioned effects for charter school status		
		Performance effect	Population effect	Joint effect residual
Oklahoma	−12.6	−12.3	0.0	−0.2
Delaware	−12.5	−12.6	0.1	0.0
Maryland	−11.4	−11.5	0.1	0.0
New Jersey	−10.9	−10.4	−0.1	−0.5
Massachusetts	−10.9	−11.0	0.0	0.2
Pennsylvania	−10.7	−10.1	−1.1	0.4
Minnesota	−10.7	−10.7	0.0	0.0
North Carolina	−10.0	−9.9	0.0	0.0
New Mexico	−9.8	−9.6	−0.1	−0.1
Connecticut	−9.7	−9.2	−0.5	0.0
Ohio	−9.7	−10.3	0.6	0.1
District of Columbia	−9.7	−9.6	0.0	−0.1
Colorado	−9.6	−9.6	0.0	0.0
Arizona	−9.2	−9.9	0.5	0.1
Oregon	−9.2	−9.4	0.3	−0.1
New Hampshire	−8.6	−8.7	0.5	−0.5
Missouri	−8.6	−8.6	0.0	0.0
Georgia	−8.2	−8.3	0.2	−0.1
Tennessee	−8.1	−8.2	0.2	0.0
Michigan	−7.7	−7.3	−0.8	0.4

State	2022 − 2019 average scale score difference	Partitioned effects for charter school status		
		Performance effect	Population effect	Joint effect residual
Wisconsin	−7.5	−7.5	0.0	0.0
Arkansas	−7.5	−7.5	−0.1	0.1
Illinois	−7.4	−7.2	−0.1	−0.1
South Carolina	−7.3	−6.9	−0.3	−0.1
Florida	−7.3	−7.3	0.1	−0.1
Indiana	−7.0	−5.4	−1.7	0.1
Texas	−7.0	−6.9	0.0	0.0
New York	−6.1	−6.1	0.1	−0.1
Louisiana	−6.0	−6.1	0.0	0.0
California	−5.8	−5.6	−0.1	−0.1
Wyoming	−5.4	−5.4	0.0	0.0
Rhode Island	−5.4	−5.3	0.0	0.0
Hawaii	−5.2	−5.3	−0.1	0.2
Nevada	−4.6	−6.0	1.6	−0.2
Alaska	−4.4	−4.1	−0.1	−0.2
Idaho	−4.2	−4.1	−0.1	0.1
Utah	−2.7	−2.1	0.3	−0.8

NOTE: NAEP = National Assessment of Educational Progress. States are ordered by the 2022 − 2019 average scale score difference. Alabama, Iowa, Kansas, Kentucky, Maine, Mississippi, Montana, Nebraska, North Dakota, South Dakota, Vermont, Virginia, Washington, and West Virginia did not report charter school status data in 2019 or 2022 and thus were excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit B6. Average Scale Score Differences and Partitioned Effects for School Locale for Public School Students in Grade 8 NAEP Mathematics, by State: 2019 to 2022

State	2022 − 2019 average scale score difference	Partitioned effects for school locale		
		Performance effect	Population effect	Joint effect residual
Oklahoma	−12.6	−12.8	0.0	0.2
Delaware	−12.5	−12.4	0.0	0.0
West Virginia	−12.0	−12.1	0.1	0.0
Maryland	−11.4	−11.3	0.0	−0.1
New Jersey	−10.9	−10.7	−0.3	0.0
Massachusetts	−10.9	−10.8	−0.2	0.1
Pennsylvania	−10.7	−10.3	−0.5	0.0
Minnesota	−10.7	−10.7	0.0	0.0
Kansas	−10.2	−10.3	0.0	0.1
Vermont	−10.1	−10.1	−0.1	0.0
North Carolina	−10.0	−10.0	0.0	0.0
Maine	−9.8	−9.7	0.0	−0.1
New Mexico	−9.8	−9.8	0.0	0.0
Washington	−9.7	−9.8	−0.1	0.1
Connecticut	−9.7	−10.1	0.4	0.0
Ohio	−9.7	−9.2	−0.4	0.0
Colorado	−9.6	−9.7	−0.1	0.2
Arizona	−9.2	−8.3	−0.8	−0.2
Oregon	−9.2	−9.1	0.1	−0.2
Kentucky	−8.7	−8.7	0.0	0.0
New Hampshire	−8.6	−8.8	0.1	0.1
Missouri	−8.6	−8.7	0.1	0.0
Georgia	−8.2	−8.0	−0.2	0.0
Mississippi	−8.2	−8.8	0.6	0.1
Tennessee	−8.1	−8.1	0.1	−0.1

State	2022 − 2019 average scale score difference	Partitioned effects for school locale		
		Performance effect	Population effect	Joint effect residual
Virginia	−7.9	−8.3	0.3	0.1
Michigan	−7.7	−7.6	−0.4	0.2
Wisconsin	−7.5	−7.7	0.1	0.1
Arkansas	−7.5	−7.5	−0.2	0.2
Illinois	−7.4	−7.8	0.2	0.2
South Carolina	−7.3	−7.3	0.2	−0.3
Florida	−7.3	−7.5	0.3	−0.1
North Dakota	−7.2	−7.1	0.0	0.0
Indiana	−7.0	−7.5	0.4	0.1
Texas	−7.0	−7.1	0.0	0.0
Montana	−6.5	−6.5	0.0	0.0
New York	−6.1	−6.1	−0.1	0.0
Louisiana	−6.0	−6.2	0.2	−0.1
California	−5.8	−5.7	0.3	−0.4
South Dakota	−5.7	−5.4	−0.5	0.3
Nebraska	−5.6	−6.5	−0.2	1.1
Wyoming	−5.4	−5.4	0.0	0.0
Rhode Island	−5.4	−5.1	−0.3	0.0
Hawaii	−5.2	−5.5	0.0	0.3
Nevada	−4.6	−4.9	0.1	0.3
Iowa	−4.5	−4.8	0.5	−0.2
Alaska	−4.4	−4.6	0.3	−0.1
Alabama	−4.3	−4.2	−0.4	0.2
Idaho	−4.2	−4.0	−0.4	0.2
Utah	−2.7	−3.3	−0.2	0.7

NOTE: NAEP = National Assessment of Educational Progress. States are ordered by the 2022 − 2019 average scale score difference. The District of Columbia did not report data for the categories of suburb, town, and rural and thus was excluded from this analysis. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Appendix C. Partitioned Effects by TUDA

Exhibit C1. Average Scale Score Differences and Partitioned Effects for Race/Ethnicity for Public School Students in Grade 8 NAEP Mathematics, by TUDA: 2019 to 2022

TUDA	2022 − 2019 average scale score difference	Partitioned effects for race/ethnicity		
		Performance effect	Population effect	Joint effect residual
Shelby County (TN)	−14.4	−14.3	−0.2	0.1
Chicago	−12.2	−13.3	1.3	−0.3
District of Columbia	−11.7	−10.5	−1.1	−0.1
Denver	−10.7	−9.3	−3.2	1.8
Jefferson County (KY)	−10.7	−9.6	−1.0	−0.1
Charlotte-Mecklenburg (NC)	−10.5	−9.4	−1.0	−0.1
Guilford County (NC)	−10.1	−9.1	−1.9	0.9
Houston	−9.4	−10.6	0.9	0.3
Austin	−9.3	−10.2	1.2	−0.2
Boston	−8.9	−10.4	2.1	−0.6
Baltimore City	−8.8	−8.4	−0.6	0.2
Cleveland	−8.4	−9.8	0.7	0.7
San Diego	−8.3	−8.2	0.0	0.0
Hillsborough County (FL)	−7.1	−7.8	0.1	0.7
Albuquerque	−6.8	−6.1	−1.2	0.5
Detroit	−6.3	−6.9	0.5	0.1
Milwaukee	−6.0	−6.9	0.4	0.4
Fort Worth	−6.0	−6.8	0.5	0.3
Atlanta	−5.5	−5.8	0.3	−0.1
Duval County (FL)	−4.9	−4.8	−0.1	0.0
Clark County (NV)	−4.8	−4.5	−0.3	0.0
Philadelphia	−4.1	−3.4	−1.1	0.5
Dallas	−4.1	−4.3	0.0	0.3
New York City	−4.0	−5.1	1.0	0.0
Miami-Dade	−1.9	−1.4	−0.7	0.2
Los Angeles	1.7	−0.3	2.1	0.0

NOTE: NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. TUDAs are ordered by the 2022 − 2019 average scale score difference. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit C2. Average Scale Score Differences and Partitioned Effects for NSLP Eligibility for Public School Students in Grade 8 NAEP Mathematics, by TUDA: 2019 to 2022

TUDA	2022 − 2019 average scale score difference	Partitioned effects for NSLP eligibility		
		Performance effect	Population effect	Joint effect residual
Shelby County (TN)	−14.4	−14.1	−1.0	0.7
Chicago	−12.2	−12.6	0.3	0.0
District of Columbia	−11.7	−9.8	−2.3	0.3
Denver	−10.7	−14.0	4.2	−0.9
Jefferson County (KY)	−10.7	−7.3	−2.6	−0.8
Charlotte-Mecklenburg (NC)	−10.5	−11.8	1.1	0.2
Guilford County (NC)	−10.1	−12.1	2.7	−0.7
Houston	−9.4	−9.0	−0.4	−0.1
Austin	−9.3	−9.7	−0.2	0.7
Boston	−8.9	−8.5	−0.3	−0.1
Baltimore City	−8.8	−7.2	−1.7	0.1
San Diego	−8.3	−9.6	1.2	0.2
Hillsborough County (FL)	−7.1	−7.9	0.0	0.9
Albuquerque	−6.8	−8.0	0.5	0.7
Milwaukee	−6.0	−7.3	1.6	−0.3
Fort Worth	−6.0	−6.4	0.4	0.0
Atlanta	−5.5	−6.7	0.9	0.3
Duval County (FL)	−4.9	−5.6	0.8	0.0
Dallas	−4.1	−4.5	0.6	−0.2
New York City	−4.0	−4.0	−0.1	0.0
Miami-Dade	−1.9	−1.2	−0.8	0.0

NOTE: NSLP = National School Lunch Program. NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. TUDAs are ordered by the 2022 − 2019 average scale score difference. Clark County and Cleveland reported 100% of students (with rounding) as NSLP eligible in 2022 and thus were excluded from this analysis. Detroit, Los Angeles, and Philadelphia were excluded from this analysis because of a large change in the percentage of students reported as information not available. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit C3. Average Scale Score Differences and Partitioned Effects for EL Status for Public School Students in Grade 8 NAEP Mathematics, by TUDA: 2019 to 2022

TUDA	2022 − 2019 average scale score difference	Partitioned effects for EL status		
		Performance effect	Population effect	Joint effect residual
Shelby County (TN)	−14.4	−14.3	−0.1	0.0
Chicago	−12.2	−10.2	−3.2	1.2
District of Columbia	−11.7	−10.7	−1.0	0.0
Denver	−10.7	−10.2	−0.5	0.0
Jefferson County (KY)	−10.7	−9.3	−1.7	0.3
Charlotte-Mecklenburg (NC)	−10.5	−8.7	−2.9	1.1
Guilford County (NC)	−10.1	−8.2	−2.9	0.9
Houston	−9.4	−6.7	−2.8	0.1
Austin	−9.3	−8.5	−0.8	0.1
Boston	−8.9	−10.5	1.2	0.4
Baltimore City	−8.8	−8.5	−0.5	0.2
Cleveland	−8.4	−8.7	0.0	0.3
San Diego	−8.3	−6.9	−1.4	0.0
Hillsborough County (FL)	−7.1	−8.0	0.8	0.1
Albuquerque	−6.8	−3.9	−3.1	0.2
Detroit	−6.3	−6.6	0.2	0.1
Milwaukee	−6.0	−5.7	−1.4	1.1
Fort Worth	−6.0	−3.9	−4.4	2.3
Atlanta	−5.5	−5.1	−0.6	0.1
Duval County (FL)	−4.9	−4.8	−0.2	0.0
Clark County (NV)	−4.8	−5.4	0.5	0.0
Philadelphia	−4.1	−3.5	−1.3	0.8
Dallas	−4.1	−3.8	−1.0	0.7
New York City	−4.0	−4.0	−0.1	0.0
Miami-Dade	−1.9	−3.3	1.2	0.3
Los Angeles	1.7	1.2	0.4	0.1

NOTE: EL = English learner. NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. The results for ELs are based on students who were assessed and cannot be generalized to the total population of such students. TUDAs are ordered by the 2022 − 2019 average scale score difference. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit C4. Average Scale Score Differences and Partitioned Effects for SD Status for Public School Students in Grade 8 NAEP Mathematics, by TUDA: 2019 to 2022

TUDA	2022 − 2019 average scale score difference	Partitioned effects for SD status		
		Performance effect	Population effect	Joint effect residual
Shelby County (TN)	−14.4	−15.4	1.5	−0.5
Chicago	−12.2	−12.1	−0.3	0.2
District of Columbia	−11.7	−10.9	−1.2	0.4
Denver	−10.7	−9.6	−1.3	0.1
Jefferson County (KY)	−10.7	−10.3	−0.4	0.1
Charlotte-Mecklenburg (NC)	−10.5	−10.2	−0.9	0.6
Guilford County (NC)	−10.1	−10.2	0.1	0.0
Houston	−9.4	−9.2	−0.2	0.0
Austin	−9.3	−8.5	−1.0	0.2
Boston	−8.9	−9.1	−0.2	0.4
Baltimore City	−8.8	−9.3	0.5	0.1
Cleveland	−8.4	−7.7	−0.9	0.2
San Diego	−8.3	−7.2	−0.9	−0.1
Hillsborough County (FL)	−7.1	−6.9	−1.2	1.0
Albuquerque	−6.8	−4.9	−1.8	−0.1
Detroit	−6.3	−6.5	0.2	0.0
Milwaukee	−6.0	−7.0	0.9	0.1
Fort Worth	−6.0	−5.3	−1.1	0.4
Atlanta	−5.5	−5.0	−0.8	0.3
Duval County (FL)	−4.9	−4.0	−0.9	0.0
Clark County (NV)	−4.8	−5.3	0.6	−0.1
Philadelphia	−4.1	−3.6	−0.9	0.4
Dallas	−4.1	−3.3	−1.5	0.7
New York City	−4.0	−4.1	0.1	0.0
Miami-Dade	−1.9	−1.7	−0.3	0.1
Los Angeles	1.7	1.8	−0.1	0.0

NOTE: SD = students with disabilities. NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. The results for SDs are based on students who were assessed and cannot be generalized to the total population of such students. TUDAs are ordered by the 2022 − 2019 average scale score difference. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit C5. Average Scale Score Differences and Partitioned Effects for Charter School Status for Public School Students in Grade 8 NAEP Mathematics, by TUDA: 2019 to 2022

TUDA	2022 − 2019 average scale score difference	Partitioned effects for charter school status		
		Performance effect	Population effect	Joint effect residual
Shelby County (TN)	−14.4	−14.5	0.0	0.1
Chicago	−12.2	−11.9	−0.3	0.0
Denver	−10.7	−10.7	−0.4	0.4
Houston	−9.4	−9.3	−0.3	0.2
Austin	−9.3	−9.5	0.2	0.0
Baltimore City	−8.8	−8.8	−0.2	0.2
Cleveland	−8.4	−9.4	1.6	−0.6
Hillsborough County (FL)	−7.1	−7.8	1.0	−0.2
Albuquerque	−6.8	−6.5	0.0	−0.3
Milwaukee	−6.0	−6.4	0.5	−0.1
Atlanta	−5.5	−5.6	0.0	0.0
Duval County (FL)	−4.9	−5.4	0.3	0.2
Miami-Dade	−1.9	−3.3	0.9	0.6

NOTE: NAEP = National Assessment of Educational Progress. TUDA = Trial Urban District Assessment. Boston, Charlotte-Mecklenburg, Clark County, the District of Columbia, Dallas, Detroit, Fort Worth, Guilford County, Jefferson County, New York City, Philadelphia, and San Diego did not report charter school status data in 2019 or 2022 and thus were excluded from this analysis. Los Angeles reported no charter schools in 2019 and thus was excluded from this analysis. TUDAs are ordered by the 2022 − 2019 average scale score difference. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Appendix D. Demonstrating the Calculations for Partitioning Results

This appendix provides a more detailed demonstration of the calculation of the Beaton-Chromy results. As noted in the Beaton and Chromy's Partitioning Method section in this report, the total in scores for the nation across years can be partitioned using the formula in Equation 4. Here, $\boldsymbol{X}_t$ is a vector of average NAEP scores for the component student groups at the given time points; $\boldsymbol{P}_t$ is a vector of the student groups' average proportion in each year; and $\bar{x_t}$ is the overall NAEP score as reported across all students, where $t = 2$ refers to 2022 and $t = 1$ refers to 2019.

$$
\begin{aligned}
\bar{x_2} - \bar{x_1} &= \boldsymbol{P}_2'\boldsymbol{X}_2 - \boldsymbol{P}_1'\boldsymbol{X}_1 \qquad (4)\\
&= (\boldsymbol{P}_1 + \boldsymbol{P}_2 - \boldsymbol{P}_1)'(\boldsymbol{X}_1 + \boldsymbol{X}_2 - \boldsymbol{X}_1) - \boldsymbol{P}_1'\boldsymbol{X}_1 \\
&= \boldsymbol{P}_1'(\boldsymbol{X}_2 - \boldsymbol{X}_1) && \text{the performance effect}\\
&+ (\boldsymbol{P}_2 - \boldsymbol{P}_1)'\boldsymbol{X}_1 && \text{the population effect}\\
&+ (\boldsymbol{P}_2 - \boldsymbol{P}_1)'(\boldsymbol{X}_2 - \boldsymbol{X}_1) && \text{the joint effect residual}
\end{aligned}
$$

To demonstrate the method, we examine in Exhibit D1 how NSLP-eligible students' scores were partitioned.

Exhibit D1. Sample Sizes, Percentage Distribution, Average Scale Scores, Changes in Average Scale Scores, and Partitioned Effects for Public School Students in Grade 8 NAEP Mathematics, by NSLP Eligibility: 2019 and 2022

	2019	2022	2019		2022		2022 − 2019 average scale score difference	Partitioned effects		
NSLP eligibility	*N*	*N*	% of total	Avg. scale score	% of total	Avg. scale score		Performance effect	Population effect	Joint effect residual
Total	**142,200**	**107,740**	**100.0**	**281.0**	**100.0**	**273.1**	**−7.9**	**−7.9**	**0.0**	**0.0**
Eligible	72,050	54,160	49.9	266.1	49.7	259.5	−6.6	−3.3	−0.6	0.0
Not eligible	68,700	52,130	48.8	296.0	48.7	287.0	−9.0	−4.4	−0.3	0.0
Information not available	1,450	1,460	1.3	286.2	1.6	274.1	−12.1	−0.2	0.9	0.0

NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences and partitioned effects were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Plugging in the corresponding numbers for NSLP-eligible students from Exhibit D1, with the first time point being the 2019 data and the second time point being the 2022 data:[7]

$$
\begin{aligned}
&= 0.499 \times (259.5 - 266.1) && \text{the performance effect} \qquad (5)\\
&+ (0.497 - 0.499) \times 266.1 && \text{the population effect}\\
&+ (0.497 - 0.499) \times (259.5 - 266.1) && \text{the joint effect residual}
\end{aligned}
$$

Carrying out the operation, the partitioned effects for the NSLP-eligible row are replicated:

$$
\begin{aligned}
&= -3.3 && \text{the performance effect} \qquad (6)\\
&+ -0.6 && \text{the population effect}\\
&+ 0.0 && \text{the joint effect residual}
\end{aligned}
$$

The effects can then be summed across all NSLP eligibility groups to calculate the total partitioned effects for 2019 to 2022. For example, the total population effect (0.0) is the sum of the partitioned population effects for students who are eligible (−0.6), students who are not eligible (−0.3), and students with information not available (0.9).

[7] With rounding, the results of this calculation vary slightly from those in Exhibit D1 and Equation 6, which were calculated using unrounded values.

Appendix E. Detailed Statistical Results From the National Propensity Weighting Analysis

This appendix provides details for the national public propensity score model, including coefficient estimates from the propensity score model and results of the covariate balance check.

Exhibit E1. Coefficient Estimates From Propensity Score (Logistic Regression) Model for Public School Students in Grade 8 NAEP Mathematics

Covariate	Coefficient estimate	*SE*	*p*	Odds
(Intercept)	−0.2	0.24	0.39	0.81
Race and ethnicity (Reference = White)				
Black	0.0	0.03	0.65	0.99
Hispanic	0.0	0.03	0.33	0.97
Asian	−0.1	0.05	0.12	0.92
American Indian/Alaska Native	0.2	0.06	0.01	1.17
Native Hawaiian/Other Pacific Islander	0.2	0.09	0.11	1.17
Two or More Races	−0.2	0.04	< 0.01	0.80
National School Lunch Program (NSLP) (Reference = Information not available)				
Eligible	0.3	0.25	0.26	1.233
Not eligible	0.2	0.25	0.47	1.20
English learner (EL)	−0.2	0.05	< 0.01	0.81
Students with disabilities (SD)	0.0	0.02	0.01	0.96
Highest parental level of education (Reference = Omitted, Missing)				
Did not finish high school	0.2	0.06	< 0.01	1.20
Graduated high school	0.1	0.06	0.18	1.09
Some education after high school	0.2	0.06	0.01	1.16
Graduated college	0.1	0.06	0.07	1.11
"I don't know"	0.0	0.06	0.76	1.02
Books at home (Reference = Omitted, Missing)				
0–10 books	−0.2	0.06	0.01	0.86
11–25 books	−0.1	0.06	< 0.07	0.89
26–100 books	0.0	0.06	0.77	1.08
More than 100 books	0.1	0.07	0.27	1.08
Charter school	−0.2	0.13	0.19	0.84
School locale (Reference = City)				
Suburb	0.0	0.03	0.99	1.00
Town	−0.1	0.06	0.24	0.94
Rural	−0.1	0.04	< 0.01	0.89

NOTE: Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Covariate Balance Check

In this appendix, we present results from fitting simple logistic regressions, whereby each predictor variable from the propensity score model is regressed on the year indicator (2019 = 1; 2022 = 0) when the propensity score weights are applied. The purpose of the propensity weighting analysis was to ensure that the 2022 sample was statistically similar to the 2019 sample on the demographic characteristics.

We performed these regressions to assess the balance of the covariates after reweighting (Olmos & Govindasamy, 2015). The null results (*p* values greater than 0.05) from Exhibits E2–E24 indicate that balance between the 2019 and 2022 samples was achieved across predictors (covariates).

Exhibit E2. Covariate Balance Check: *Black* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−1.8	0.02	< 0.01
Treatment indicator (2019 sample)	0.0	0.03	0.94

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E3. Covariate Balance Check: *Hispanic* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−1.0	0.02	< 0.01
Treatment indicator (2019 sample)	0.0	0.03	0.92

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E4. Covariate Balance Check: *Asian* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−2.8	0.05	< 0.01
Treatment indicator (2019 sample)	0.0	0.00	0.96

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E5. Covariate Balance Check: *American Indian/Alaska Native* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−4.7	0.05	< 0.01
Treatment indicator (2019 sample)	0.0	0.07	0.98

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E6. Covariate Balance Check: *Native Hawaiian/Other Pacific Islander* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−5.6	0.08	< 0.01
Treatment indicator (2019 sample)	0.0	0.11	0.98

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E7. Covariate Balance Check: *Two or More Races* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−3.3	0.02	< 0.01
Treatment indicator (2019 sample)	0.0	0.03	0.96

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E8. Covariate Balance Check: *NSLP Eligible* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	0.0	0.02	0.99
Treatment indicator (2019 sample)	0.0	0.03	0.89

NOTE: NSLP = National School Lunch Program.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E9. Covariate Balance Check: *Not NSLP Eligible* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−0.1	0.03	0.01
Treatment indicator (2019 sample)	0.0	0.03	0.87

NOTE: NSLP = National School Lunch Program.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E10. Covariate Balance Check: *English Learner* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−2.5	0.00	< 0.01
Treatment indicator (2019 sample)	0.0	0.00	0.98

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E11. Covariate Balance Check: *Individualized Education Program* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−1.8	0.01	< 0.01
Treatment indicator (2019 sample)	0.0	0.00	0.80

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E12. Covariate Balance Check: *Did Not Finish High School (Parental Education)* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−2.7	0.02	< 0.01
Treatment indicator (2019 sample)	0.0	0.04	0.85

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E13. Covariate Balance Check: *Graduated High School (Parental Education)* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−1.8	0.02	< 0.01
Treatment indicator (2019 sample)	0.0	0.00	0.96

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E14. Covariate Balance Check: *Some Education After High School (Parental Education)* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−1.9	0.01	< 0.01
Treatment indicator (2019 sample)	0.0	0.02	0.95

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E15. Covariate Balance Check: *Graduated College (Parental Education)* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	0.0	0.06	0.03
Treatment indicator (2019 sample)	0.0	0.02	0.91

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E16. Covariate Balance Check: *I Don't Know (Parental Education)* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−2.0	0.02	< 0.01
Treatment indicator (2019 sample)	0.0	0.03	0.93

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E17. Covariate Balance Check: *0–10 Books* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−1.4	0.02	< 0.01
Treatment indicator (2019 sample)	0.0	0.02	0.94

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E18. Covariate Balance Check: *11–25 Books* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−1.1	0.01	< 0.01
Treatment indicator (2019 sample)	0.0	0.02	0.92

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E19. Covariate Balance Check: *26–100 Books* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−0.8	0.01	< 0.01
Treatment indicator (2019 sample)	0.0	0.02	0.97

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E20. Covariate Balance Check: *More Than 100 Books* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−1.4	0.02	< 0.01
Treatment indicator (2019 sample)	0.0	0.02	0.94

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E21. Covariate Balance Check: *Charter School* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−2.8	0.09	< 0.01

Treatment indicator (2019 sample)	0.0	0.12	0.98

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E22. Covariate Balance Check: *Suburb* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−0.4	0.02	< 0.01
Treatment indicator (2019 sample)	0.0	0.02	0.92

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E23. Covariate Balance Check: *Town* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−2.0	0.05	< 0.01
Treatment indicator (2019 sample)	0.0	0.06	1.00

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit E24. Covariate Balance Check: *Rural* Regressed on Treatment Indicator

Term	Coefficient estimate	*SE*	*p*
Intercept	−1.5	0.03	< 0.01
Treatment indicator (2019 sample)	0.0	0.04	0. 91

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Appendix F. A Deeper Dive Into the Propensity Weighting Analysis Results

In this appendix, we unpack the results for selected states and TUDAs to further demonstrate how the propensity weighting analysis operates. Specifically, we juxtapose changes in demographic characteristics from 2019 to 2022 for selected jurisdictions with relatively small and large differences between their adjusted and unadjusted achievement results. This exercise serves to elucidate how greater changes in demographic characteristics of students tend to be associated with greater differences between adjusted and unadjusted achievement results.

For states, we contrasted changes in demographic characteristics between Massachusetts, which had a relatively large difference of −4.03 points between its adjusted and unadjusted results, and Virginia, which had a relatively small difference of −0.01 points. In Massachusetts, the percentage of students eligible for NSLP increased by more than 13 percentage points from 2019 to 2022. In Virginia, the NSLP-eligible percentage increased by just 0.4 percentage points. Similarly, Massachusetts experienced more change in its composition of students by highest level of parental education. For instance, the percentage of students whose parent(s) graduated college decreased by 3.8 percentage points in Massachusetts but increased by just 0.9 percentage points in Virginia.

Exhibit F1. Percentage Distribution for Public School Students in Grade 8 NAEP Mathematics in Massachusetts and Virginia, by Select Demographic Variables: 2019 and 2022

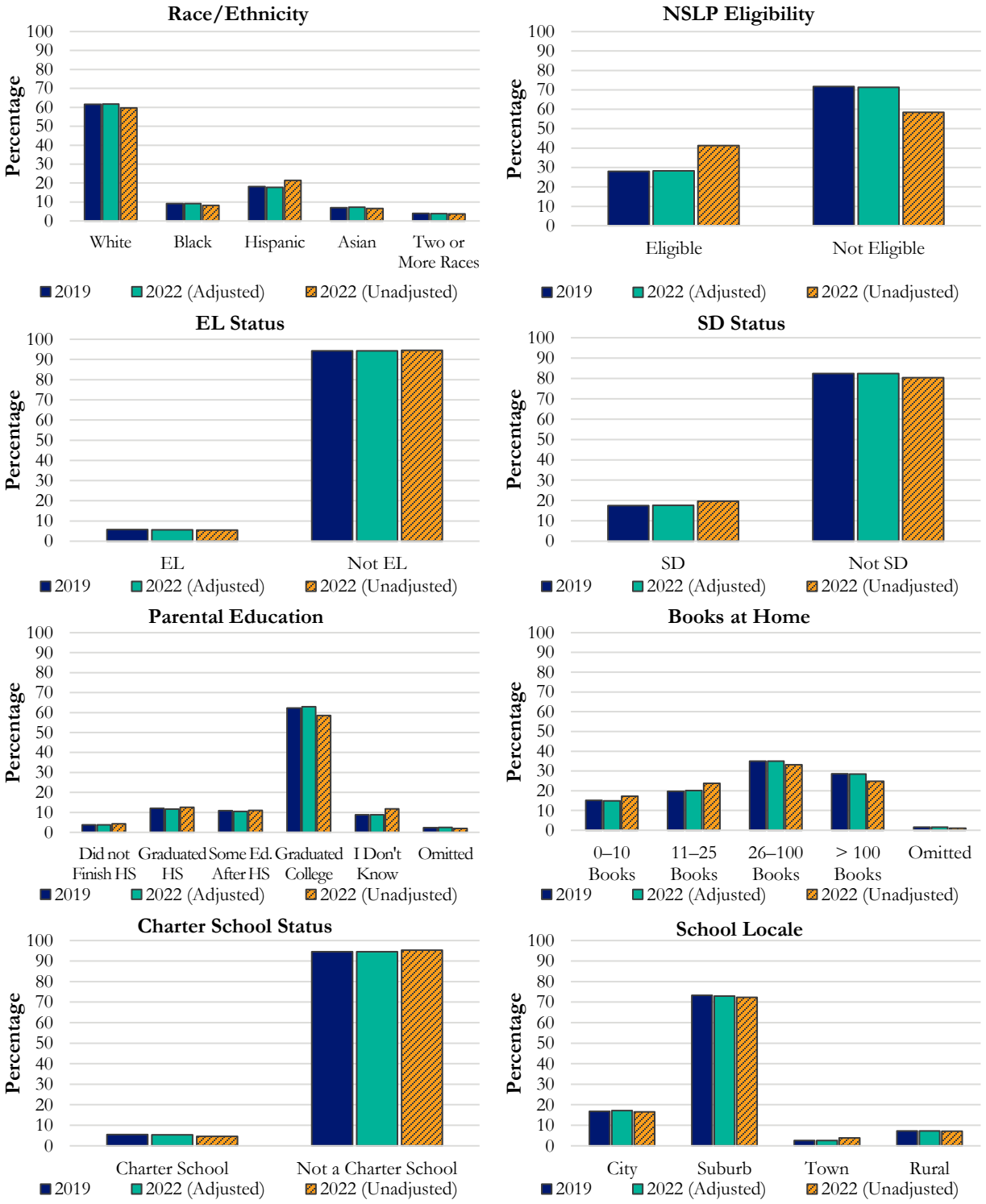
Variable	Massachusetts			Virginia		
	2019 percentage	2022 percentage	2022 − 2019 difference in percentage	2019 percentage	2022 percentage	2022 − 2019 difference in percentage
Race/ethnicity						
White	61.6	59.7	−1.9	49.8	45.5	−4.4
Black	9.1	8.2	−0.9	21.9	20.4	−1.5
Hispanic	18.1	21.4	3.3	15.4	17.7	2.3
Asian	7.0	6.6	−0.4	7.0	10.2	3.2
American Indian/Alaska Native	‡	‡	‡	‡	‡	‡
Native Hawaiian/Other Pacific Islander	‡	0.0	‡	0.0	‡	‡
Two or More Races	4.0	3.8	−0.2	5.5	6.1	0.7
NSLP eligibility						
Eligible	27.9	41.3	13.4	37.1	37.6	0.4
Not eligible	71.7	58.4	−13.3	62.5	60.5	−2.0
Information not available	‡	‡	‡	‡	1.9	‡
EL status						
EL	5.8	5.4	−0.3	4.8	6.6	1.8
Not EL	94.2	94.6	0.3	95.2	93.3	−1.8
Omitted	‡	0.0	‡	0.0	‡	‡
SD status						
SD	17.5	19.6	2.1	13.1	13.6	0.5
Not SD	82.5	80.4	−2.1	86.9	86.3	−0.5
Omitted	‡	0.0	‡	0.0	‡	‡
Parental education						
Did not finish high school	3.8	4.3	0.5	3.7	4.0	0.3
Graduated from high school	12.0	12.5	0.5	14.3	11.9	−2.4
Some education after high school	10.8	11.0	0.2	12.1	9.6	−2.5
Graduated college	62.3	58.5	−3.8	58.2	59.0	0.9
I don't know	8.8	11.8	3.0	10.5	13.1	2.6
Omitted	2.3	1.9	−0.4	1.2	2.3	1.2

Variable	Massachusetts			Virginia		
	2019 percentage	2022 percentage	2022 − 2019 difference in percentage	2019 percentage	2022 percentage	2022 − 2019 difference in percentage
Books at home						
0–10 books	15.1	17.3	2.1	16.7	18.4	1.8
11–25 books	19.7	23.8	4.1	21.5	22.1	0.6
26–100 books	35.0	33.1	−1.9	33.0	32.4	−0.6
More than 100 books	28.6	24.8	−3.8	28.5	25.7	−2.8
Omitted	1.6	1.0	−0.6	‡	1.4	‡
Charter school status						
Charter school	5.5	4.7	−0.8	100.0	100.0	0.0
Not a charter school	94.5	95.3	0.8	0.0	0.0	0.0
School location						
City	16.9	16.6	−0.3	20.4	19.6	-0.8
Suburb	73.3	72.3	−1.0	47.3	49.3	2.1
Town	2.6	3.9	1.3	8.5	6.2	−2.3
Rural	7.3	7.2	−0.1	23.7	24.8	1.1

‡ Reporting standards not met.
NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. EL = English learner. SD = students with disabilities. Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin. The results for ELs and SDs are based on students who were assessed and cannot be generalized to the total population of such students. The following levels were excluded from the propensity score models because of an insufficient sample size in one or both years: American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, and NSLP information not available for Massachusetts; and American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, NSLP information not available, books at home omitted, and charter school status for Virginia. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

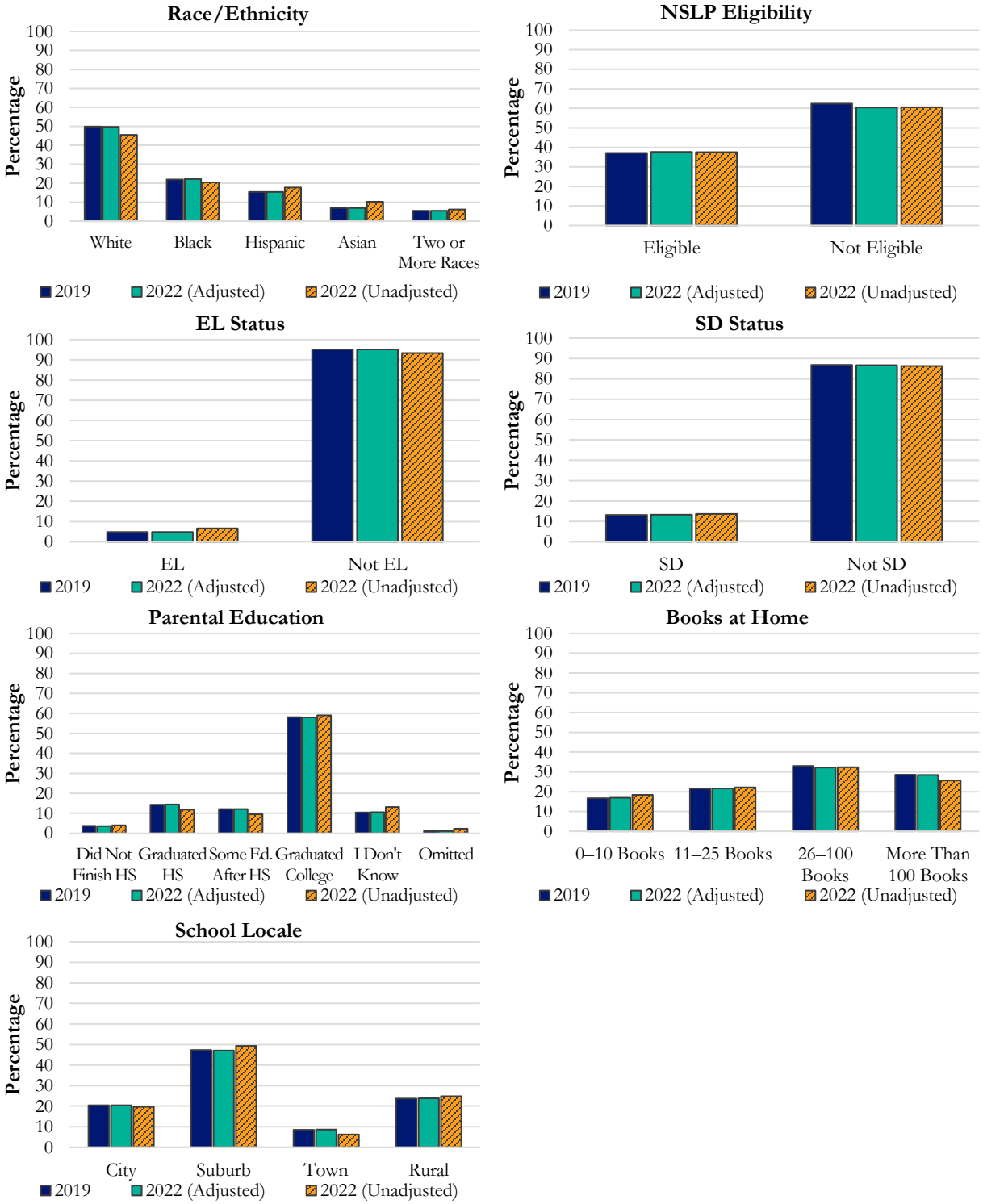
The sets of bar charts for Massachusetts and Virginia in Exhibits F2 and F3, respectively, demonstrate how propensity score weighting had the practical effect of making the demographic characteristics of students in 2019 and 2022 more similar for analysis. The charts illustrate how the demographic composition of 2019 and *adjusted* 2022 groups were more similar than 2019 and *unadjusted* 2022 groups. Consider, for instance, the example of students eligible for NSLP in Massachusetts in Exhibit F2. Although NSLP-eligible students represented about 30% of the students in the 2019 population and *adjusted* 2022 populations, they represented about 40% of the students in the *unadjusted* 2022 population. Importantly, weighting cases as a function of their propensity scores, so that the demographic composition of students is more similar for analysis, lends itself to addressing the second major research question from this study—whether compositional changes between 2019 and 2022 related to changes in NAEP achievement.

Exhibit F2. Unadjusted and Adjusted Percentage Distribution for Public School Students in Grade 8 NAEP Mathematics in Massachusetts, by Selected Demographic Variables: 2019 and 2022



NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. EL = English learner. SD = students with disabilities. HS = high school. Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin. The following levels were excluded from the propensity score model for Massachusetts because of an insufficient sample size in one or both years: American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, and NSLP information not available. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit F3. Unadjusted and Adjusted Percentage Distribution for Public School Students in Grade 8 NAEP Mathematics in Virginia, by Selected Demographic Variables: 2019 and 2022



NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. EL = English learner. SD = students with disabilities. HS = high school. Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin. The following levels were excluded from the propensity score model for Virginia due to insufficient sample size in one or both years: American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, NSLP information not available, Books at home omitted, and charter school status. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

For TUDAs, we contrasted changes in demographic characteristics between Jefferson County (KY), which had a relatively large difference of −6.4 points between its adjusted and unadjusted results, and Cleveland, which had a relatively small difference of 0.2 points. In Jefferson County, the percentage of students eligible for NSLP increased by nearly 10 percentage points from 2019 to 2022. In Cleveland, this percentage of students increased by about 3 percentage points. Similarly, Jefferson County experienced greater shifts in its composition of students by highest level of parental education. For instance, the percentage of students whose parent(s) graduated college decreased by 9.1 percentage points in Jefferson County but by just 3.8 percentage points in Cleveland.

Exhibit F4. Percentage Distribution for Public School Students in Grade 8 NAEP Mathematics in Jefferson County (KY) and Cleveland, by Selected Demographic Variables: 2019 and 2022

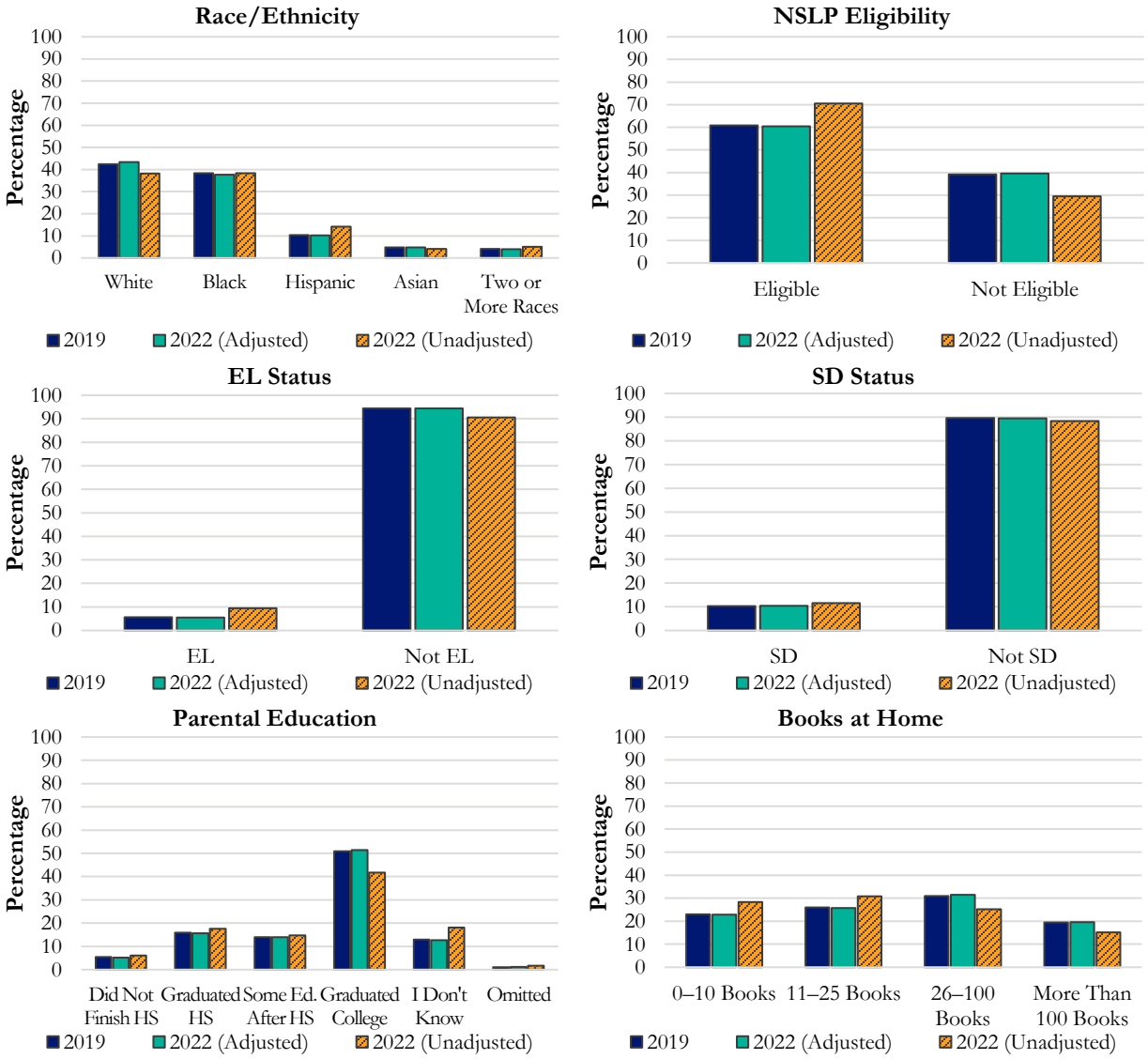
Variable	Jefferson County (KY)			Cleveland		
	2019 percentage	2022 percentage	2022 − 2019 difference in percentage	2019 percentage	2022 percentage	2022 − 2019 difference in percentage
Race/ethnicity						
White	42.4	38.2	−4.2	14.2	15.4	1.2
Black	38.4	38.3	−0.1	64.9	62.7	−2.2
Hispanic	10.3	14.2	3.9	16.9	17.3	0.4
Asian	4.7	4.1	−0.7	‡	2.1	‡
American Indian/Alaska Native	‡	‡	‡	‡	0.0	‡
Native Hawaiian/Other Pacific Islander	‡	‡	‡	‡	‡	‡
Two or More Races	4.0	5.1	1.0	2.7	2.3	−0.4
NSLP eligibility						
Eligible	60.8	70.5	9.7	96.8	100.0	3.2
Not eligible	39.2	29.5	−9.7	3.0	0.0	−3.0
Information not available	0.0	0.0	0.0	‡	0.0	‡
EL status						
EL	5.6	9.4	3.8	11.3	9.6	−1.7
Not EL	94.4	90.6	−3.8	88.7	90.2	1.5
Omitted	0.0	0.0	0.0	0.0	‡	‡
SD status						
SD	10.3	11.6	1.3	19.4	21.2	1.8
Not SD	89.7	88.4	−1.3	80.6	78.6	−2.0
Omitted	0.0	0.0	0.0	0.0	‡	‡
Parental education						
Did not finish high school	5.4	6.1	0.7	11.8	9.0	−2.8
Graduated from high school	15.9	17.5	1.7	21.0	23.0	2.0
Some education after high school	13.9	14.7	0.8	15.6	13.7	−1.9
Graduated college	50.9	41.8	−9.1	36.3	32.5	−3.8
I don't know	12.9	18.1	5.3	14.0	17.3	3.3
Omitted	1.1	1.7	0.7	1.4	4.5	3.2
Books at home						
0–10 books	23.0	28.3	5.3	31.0	34.7	3.7
11–25 books	26.0	30.9	4.9	34.0	32.9	−1.2
26–100 books	31.0	25.2	−5.8	22.1	19.4	−2.7
More than 100 books	19.4	15.1	−4.3	12.0	10.7	−1.3
Omitted	‡	‡	‡	‡	2.3	‡

Variable	Jefferson County (KY)			Cleveland		
	2019 percentage	2022 percentage	2022 − 2019 difference in percentage	2019 percentage	2022 percentage	2022 − 2019 difference in percentage
Charter school status						
Charter school	0.0	0.0	0.0	14.9	21.5	6.6
Not a charter school	100.0	100.0	0.0	85.1	78.5	−6.6

‡ = Reporting standards not met.
NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. EL = English learner. SD = students with disabilities. Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin. The following levels were excluded from the propensity score models because of an insufficient sample size in one or both years: American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, NSLP information not available, books at home omitted, and charter school status for Jefferson County; and Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, NSLP eligibility (all levels), and books at home omitted for Cleveland. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

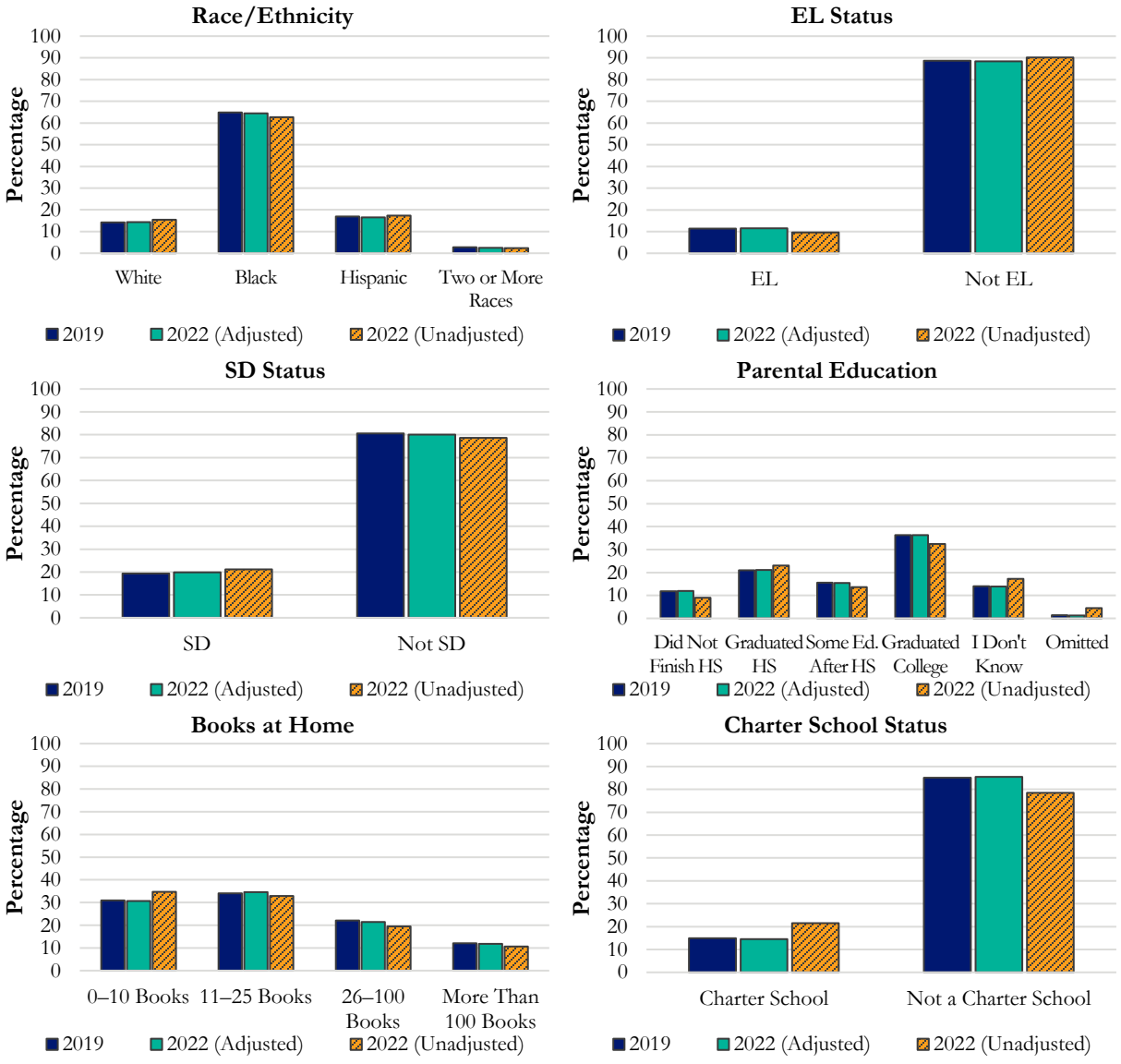
As with the sets of bar charts in Exhibits F2 and F3, the sets of bar charts for Jefferson County and Cleveland presented in Exhibits F5 and F6, respectively, demonstrate how the demographic characteristics of students in 2019 and 2022 became more similar for analysis when cases were weighted as a function of their propensity scores. Consider, for instance, the example of students eligible for NSLP in Jefferson County in Exhibit F5. Although NSLP-eligible students represented about 60% of the students in the 2019 population and the *adjusted* 2022 population, they represented about 70% of the students in the unadjusted 2022 population.

Exhibit F5. Unadjusted and Adjusted Percentage Distribution for Public School Students in Grade 8 NAEP Mathematics in Jefferson County (KY), by Selected Demographic Variables: 2019 and 2022



NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. EL = English learner. SD = students with disabilities. HS = high school. Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin. The following levels were excluded from the propensity score model for Jefferson County because of an insufficient sample size in one or both years: American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, NSLP information not available, books at home omitted, and charter school status. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Exhibit F6. Unadjusted and Adjusted Percentage Distribution for Public School Students in Grade 8 NAEP Mathematics in Cleveland, by Selected Demographic Variables: 2019 and 2022



NOTE: NAEP = National Assessment of Educational Progress. EL = English learner. SD = students with disabilities. HS = high school. Black includes African American, and Hispanic includes Latino. Race categories exclude Hispanic origin. The following levels were excluded from the propensity score model for Cleveland because of an insufficient sample size in one or both years: Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, National School Lunch Program eligibility, and books at home omitted. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Examining the Sensitivity of Results to Variables Included in Propensity Score Models

We followed an iterative model-building process to verify that the results were not unduly driven by particular variables included in the propensity score models. For example, when model building, we examined whether the results would have substantially differed if we had excluded select sets of variables related to socioeconomic status. Findings from this sensitivity analysis suggested that results were not overly sensitive to particular variables, lending support to the idea that using a multivariate analysis, in addition to the sets of bivariate analyses from the Beaton-Chromy (2007) method, offered a distinct and alternative way to address whether changes in the demographic composition of students related to changes in NAEP achievement.

In Exhibit F7, we present state results when including student NSLP eligibility (Model 1) and excluding NSLP eligibility (Model 2). We selected this variable because it was among the variables with the most explanatory power included in the models. Yet, as demonstrated through the results presented in Exhibit F7, whether this particular variable was included did not result in substantive differences in our calculations of adjusted achievement scores. The difference between unadjusted and adjusted scores changed by more than ±1 point in only six states: Arizona (−0.3 in Model 1; −1.4 in Model 2), Missouri (−0.8 in Model 1; −2.2 in Model 2), Massachusetts (−4.0 in Model 1; −2.5 in Model 2), the District of Columbia (−4.1 in Model 1; −3.1 in Model 2), Alaska (0.6 in Model 1; −0.5 in Model 2), and Alabama (−1.6 in Model 1; −2.7 in Model 2).

Exhibit F7. Unadjusted and Adjusted 2022 − 2019 State Differences in Average Scale Score in Grade 8 NAEP Mathematics, by Model

State	Unadjusted difference	Model 1 (with student NSLP eligibility)			Model 2 (without student NSLP eligibility)		
		Adjusted difference	Unadjusted minus adjusted difference	% difference	Adjusted difference	Unadjusted minus adjusted difference	% difference
West Virginia	−12.0	−10.8	−1.1	−9%	−11.0	−1.0	−8%
Kansas	−10.2	−10.5	0.2	2%	−9.8	−0.5	−5%
Minnesota	−10.7	−9.9	−0.8	−7%	−9.2	−1.5	−14%
Oklahoma	−12.6	−9.7	−2.9	−23%	−9.6	−2.9	−23%
Pennsylvania	−10.7	−9.5	−1.3	−12%	−9.5	−1.2	−11%
New Jersey	−10.9	−9.5	−1.5	−14%	−9.0	−1.9	−18%
Delaware	−12.5	−9.3	−3.1	−25%	−8.8	−3.7	−30%
Arizona	−9.2	−9.0	−0.3	−3%	−7.8	−1.4	−15%
Washington	−9.7	−8.9	−0.8	−8%	−8.6	−1.1	−12%
Colorado	−9.6	−8.4	−1.1	−12%	−7.5	−2.0	−21%
Vermont	−10.1	−8.4	−1.7	−17%	−7.7	−2.5	−24%
Oregon	−9.2	−8.4	−0.8	−9%	−8.4	−0.8	−9%
Ohio	−9.7	−8.3	−1.3	−14%	−8.7	−0.9	−10%
North Carolina	−10.0	−8.2	−1.7	−18%	−7.3	−2.7	−27%
New Hampshire	−8.6	−8.2	−0.4	−5%	−8.4	−0.2	−3%
Maryland	−11.4	−8.2	−3.3	−29%	−7.9	−3.5	−31%
Virginia	−7.9	−7.9	0.0	0%	−7.9	0.0	0%
Missouri	−8.6	−7.8	−0.8	−10%	−6.4	−2.2	−26%
Maine	−9.8	−7.3	−2.5	−26%	−6.8	−3.1	−31%
New Mexico	−9.8	−7.2	−2.6	−26%	−7.2	−2.6	−26%

State	Unadjusted difference	Model 1 (with student NSLP eligibility)			Model 2 (without student NSLP eligibility)		
		Adjusted difference	Unadjusted minus adjusted difference	% difference	Adjusted difference	Unadjusted minus adjusted difference	% difference
Wisconsin	−7.5	−7.2	−0.3	−4%	−7.1	−0.4	−5%
Massachusetts	−10.9	−6.9	−4.0	−37%	−8.4	−2.5	−23%
Connecticut	−9.7	−6.5	−3.2	−33%	−6.3	−3.3	−34%
Tennessee	−8.1	−6.5	−1.6	−20%	−5.8	−2.3	−28%
Kentucky	−8.7	−6.2	−2.5	−28%	−5.9	−2.8	−32%
North Dakota	−7.2	−6.1	−1.0	−14%	−5.7	−1.5	−20%
Mississippi	−8.2	−6.0	−2.2	−26%	−6.1	−2.0	−25%
South Carolina	−7.3	−5.9	−1.4	−19%	−5.8	−1.5	−20%
Nevada	−4.6	−5.9	1.4	30%	−5.2	0.6	13%
Illinois	−7.4	−5.9	−1.4	−20%	−5.7	−1.7	−23%
Montana	−6.5	−5.8	−0.7	−11%	−5.6	−0.9	−14%
Florida	−7.3	−5.8	−1.6	−21%	−5.3	−2.0	−27%
Michigan	−7.7	−5.7	−2.0	−26%	−5.9	−1.9	−24%
Georgia	−8.2	−5.7	−2.5	−31%	−4.7	−3.5	−42%
South Dakota	−5.7	−5.6	−0.1	−2%	−5.2	−0.5	−8%
District of Columbia	−9.7	−5.5	−4.1	−43%	−6.6	−3.1	−32%
Texas	−7.0	−5.4	−1.6	−23%	−5.4	−1.6	−23%
Wyoming	−5.4	−5.3	−0.1	−2%	−4.9	−0.5	−9%
Iowa	−4.5	−5.2	0.6	14%	−4.4	−0.1	−3%
Alaska	−4.4	−5.0	0.6	14%	−3.9	−0.5	−12%
Arkansas	−7.5	−5.0	−2.5	−33%	−4.9	−2.6	−35%
California	−5.8	−4.9	−0.9	−15%	−4.5	−1.3	−22%
New York	−6.1	−4.9	−1.3	−21%	−5.3	−0.8	−13%
Indiana	−7.0	−4.8	−2.2	−32%	−4.8	−2.2	−32%
Nebraska	−5.6	−4.5	−1.1	−20%	−4.0	−1.7	−29%
Rhode Island	−5.4	−3.8	−1.6	−30%	−3.3	−2.1	−39%
Idaho	−4.2	−2.9	−1.3	−30%	−2.5	−1.7	−40%
Alabama	−4.3	−2.8	−1.6	−37%	−1.7	−2.7	−62%
Louisiana	−6.0	−2.1	−3.9	−65%	−2.5	−3.5	−58%
Hawaii	−5.2	−2.0	−3.2	−62%	−1.9	−3.3	−63%
Utah	−2.7	−1.9	−0.8	−29%	−1.1	−1.6	−59%

NOTE: NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. States are ordered by the 2022 − 2019 adjusted difference in average scale score. Oregon reported 100% of students (with rounding) as eligible for the NSLP in 2022; thus, the NSLP eligibility variable was excluded from the state's propensity score model. Variables and levels were excluded from a state propensity score model if there was an insufficient sample size in one or both years. For a full list of excluded variables and levels for states, refer to Exhibit G1 in Appendix G. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Results for TUDAs also did not appear to be particularly sensitive to selected variables included in the propensity score models. For example, only Denver and Guilford County (NC) had differences between adjusted and unadjusted scores that changed by more than 1 point when including student NSLP eligibility in the model (Model 1, Exhibit F8) compared with excluding student NSLP eligibility from the model (Model 2, Exhibit F8).

Exhibit F8. Unadjusted and Adjusted 2022 − 2019 TUDA Differences in Average Scale Score in Grade 8 NAEP Mathematics, by Model

TUDA	Unadjusted difference	Model 1 (with student NSLP eligibility)			Model 2 (without student NSLP eligibility)		
		Adjusted difference	Unadjusted minus adjusted difference	% difference	Adjusted difference	Unadjusted minus adjusted difference	% difference
Shelby County (TN)	−14.4	−13.4	−1.0	−7%	−13.8	−0.6	−4%
Denver	−10.7	−10.9	0.1	1%	−9.1	−1.6	−15%
Boston	−8.9	−9.4	0.5	6%	−9.5	0.6	7%
Chicago	−12.2	−9.3	−3.0	−24%	−9.0	−3.2	−26%
Cleveland	−8.4	−8.6	0.2	2%	−8.6	0.2	2%
Austin	−9.3	−8.6	−0.7	−7%	−8.8	−0.5	−5%
Guilford County (NC)	−10.1	−7.7	−2.4	−24%	−6.7	−3.4	−34%
Houston	−9.4	−7.6	−1.8	−19%	−7.7	−1.7	−18%
Milwaukee	−6.0	−7.5	1.5	25%	−6.8	0.8	13%
Hillsborough County (FL)	−7.1	−6.8	−0.3	−4%	−6.8	−0.2	−3%
District of Columbia (DCPS)	−11.7	−6.7	−5.0	−42%	−6.8	−5.0	−42%
San Diego	−8.3	−6.5	−1.8	−21%	−6.0	−2.3	−27%
Charlotte-Mecklenburg (NC)	−10.5	−6.5	−4.1	−39%	−6.0	−4.5	−43%
Baltimore City	−8.8	−6.3	−2.5	−28%	−7.3	−1.5	−17%
Detroit	−6.3	−5.1	−1.2	−19%	−4.7	−1.6	−26%
Clark County (NV)	−4.8	−4.8	0.0	−1%	−4.8	0.0	−1%
New York City	−4.0	−4.6	0.6	15%	−4.6	0.5	13%
Atlanta	−5.5	−4.5	−1.1	−19%	−4.3	−1.3	−23%
Duval County (FL)	−4.9	−4.4	−0.5	−11%	−3.9	−1.0	−20%
Jefferson County (KY)	−10.7	−4.3	−6.4	−60%	−5.3	−5.4	−51%
Fort Worth	−6.0	−4.1	−1.9	−31%	−4.0	−2.0	−34%
Dallas	−4.1	−3.4	−0.7	−18%	−3.0	−1.0	−25%
Miami-Dade	−1.9	−2.8	0.8	44%	−3.0	1.1	59%
Albuquerque	−6.8	−2.2	−4.5	−67%	−1.7	−5.1	−75%
Los Angeles	1.7	−0.8	2.5	145%	−0.8	2.5	145%
Philadelphia	−4.1	−0.5	−3.6	−89%	−1.1	−3.0	−73%

NOTE: TUDA = Trial Urban District Assessment. NAEP = National Assessment of Educational Progress. NSLP = National School Lunch Program. TUDAs are ordered by the 2022 − 2019 unadjusted difference in average scale score. NSLP eligibility was excluded from the propensity score model for Los Angeles because of a large change in the percentage of students reported as information not available. Variables and levels were excluded from a TUDA propensity score model if there was an insufficient sample size in one or both years. For a full list of excluded variables and levels for TUDAs, refer to Exhibit G2 in Appendix G. The NAEP mathematics scale ranges from 0 to 500 at Grade 8. Differences were calculated using unrounded values. Some apparent differences between estimates may not be statistically significant.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress, 2019 and 2022 Grade 8 Mathematics Assessments.

Appendix G. Variables and Levels Excluded From Propensity Models

Exhibit G1. Variables and Levels Excluded From Propensity Models, by State

State	Levels excluded from propensity models						
	Race/ethnicity	NSLP eligibility	EL status	Parental education	Books at home	Charter school status	School locale
Alabama	Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	All levels	—
Alaska	—	—	—	All levels	All levels	—	—
Arizona	Native Hawaiian/Other Pacific Islander	—	—	—	—	—	—
Arkansas	—	—	—	—	Omitted	—	—
California	American Indian/Alaska Native	—	—	—	—	—	—
Colorado	Native Hawaiian/Other Pacific Islander	—	—	All levels	All levels	—	—
Connecticut	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
District of Columbia	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	All levels
Delaware	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
Florida	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
Georgia	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	—	—	—	—	—	—
Hawaii	American Indian/Alaska Native	Information not available	—	—	—	—	—
Idaho	Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
Illinois	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	—	—	—	—	—	—
Indiana	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
Iowa	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	—	—	—	—	All levels	—
Kansas	Native Hawaiian/Other Pacific Islander	—	—	—	Omitted	All levels	—
Kentucky	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	All levels	—
Louisiana	Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
Maine	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	All levels	—
Maryland	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	—	—	—	—	—	—
Massachusetts	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
Michigan	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
Minnesota	Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
Mississippi	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	All levels	—
Missouri	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	Omitted	—	—
Montana	Black, Asian, Native Hawaiian/Other Pacific Islander	Information not available	—	All levels	All levels	All levels	—
Nebraska	Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	All levels	—
Nevada	—	—	—	—	—	—	—

State	Levels excluded from propensity models						
	Race/ethnicity	NSLP eligibility	EL status	Parental education	Books at home	Charter school status	School locale
New Hampshire	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	—	—	All levels	All levels	—	—
New Jersey	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	—	—	—	—	—	—
New Mexico	Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
New York	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	—	—	—	—	—	—
North Carolina	Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
North Dakota	Native Hawaiian/Other Pacific Islander, Two or More Races	Information not available	—	—	—	All levels	—
Ohio	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	—	—	—	—	—	—
Oklahoma	Native Hawaiian/Other Pacific Islander	Information not available	—	—	Omitted	—	—
Oregon	—	All levels	—	—	—	—	—
Pennsylvania	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	—	—	—	—	—	—
Rhode Island	Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	Town
South Carolina	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
South Dakota	Native Hawaiian/Other Pacific Islander	Information not available	—	All levels	All levels	All levels	—
Tennessee	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
Texas	Native Hawaiian/Other Pacific Islander	—	—	—	—	—	—
Utah	—	—	—	All levels	All levels	—	—
Vermont	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	—	—	—	—	All levels	—
Virginia	American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	—	Omitted	All levels	—
Washington	—	Information not available	—	—	—	All levels	—
West Virginia	Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	All levels	—	—	All levels	—
Wisconsin	Native Hawaiian/Other Pacific Islander	Information not available	—	—	—	—	—
Wyoming	Asian, Native Hawaiian/Other Pacific Islander	—	—	—	—	All levels	Suburb

— Not applicable.

NOTE: NSLP = National School Lunch Program. EL = English learner. Reference categories for the state propensity score analysis were White; not eligible for NSLP; not English learner; not students with disabilities; parent(s) graduated from college; 26–100 books at home; not a charter school; and city (school location). No variables or levels were excluded for Nevada. Parental education and books at home were not reported and thus were excluded from the propensity score models for Colorado, New Hampshire, Montana, South Dakota, Alaska, and Utah. Oregon reported 100% of students (with rounding) as NSLP eligible in 2022; thus, the student NSLP variable was excluded from the state's propensity score model. The District of Columbia did not report data on school location; thus, this variable was excluded from the jurisdiction's propensity score model. Variables and levels were excluded from a state's propensity score model if there was an insufficient sample size in one or both years.

Exhibit G2. Variables and Levels Excluded From Propensity Models, by TUDA

TUDA	Levels excluded from propensity models				
	Race/ethnicity	NSLP eligibility	Parental education	Books at home	Charter school status
Albuquerque	Native Hawaiian/Other Pacific Islander	Information not available	—	Omitted	—
Atlanta	Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	Omitted	—
Austin	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	Omitted	—
Baltimore City	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander, Two or More Races	Information not available	—	—	—
Boston	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	Omitted	All levels
Charlotte-Mecklenburg (NC)	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	Omitted	All levels
Chicago	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander, Two or More Races	—	—	—	—
Clark County (NV)	American Indian/Alaska Native	All levels	—	—	All levels
Cleveland	Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	All levels	—	Omitted	—
Dallas	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	Omitted	All levels
Denver	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	All levels	All levels	—
Detroit	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander, Two or More Races	Information not available	—	Omitted	All levels
District of Columbia (DCPS)	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	—	All levels
Duval County (FL)	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	Omitted	Omitted	—
Fort Worth	Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	Omitted	All levels
Guilford County (NC)	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	Omitted	All levels
Hillsborough County (FL)	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	Omitted	—
Houston	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander, Two or More Races	Information not available	—	—	—
Jefferson County (KY)	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	Omitted	All levels
Los Angeles	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	All levels	—	—	All levels
Miami-Dade	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander, Two or More Races	Information not available	—	—	—
Milwaukee	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	—	—
New York City	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	—	All levels
Philadelphia	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	Omitted	All levels
San Diego	American Indian/Alaska Native, Native Hawaiian/ Other Pacific Islander	Information not available	—	Omitted	All levels
Shelby County (TN)	Asian, American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander	Information not available	—	Omitted	—

— Not applicable.
NOTE: TUDA = Trial Urban District Assessment. NSLP = National School Lunch Program. Reference categories for the TUDA propensity score analysis were White; not eligible for the National School Lunch Program; not English learner; not students with disabilities; parent(s) graduated from college; 26–100 books at home; and not a charter school. Parental education and books at home were not reported and thus were excluded from the propensity score model for Denver. Variables and levels were excluded from a TUDA propensity score model if there was an insufficient sample size in one or both years.