Do we need more adults in U. S. school districts? Achievement Effects of Increases in Teachers, Paraprofessionals, and Non-Teaching Staff

Carlos X. Lastra-Anadón
IE University
clastra@faculty.ie.edu

Paul E. Peterson
Harvard University
ppeterso@gov.harvard.edu

August 28, 2025

Abstract

According to a strain of public administration theory, services are delivered more efficiently when government jurisdictions consolidate, but organizational theorists expect returns to diminish as personnel increases in density or size. Empirical findings are scattered and inconsistent. In the first plausibly causal analysis of density and size effects that makes use of state representative data from twenty waves of the National Assessment of Educational Progress (NAEP), we estimate effects of changes in personnel density on biennial changes in 8th grade reading and math performances between 2003 and 2022. Models employ district fixed effects and state-by-wave fixed effects to estimate impacts of marginal shifts in the density of teaching, paraprofessional and non-teaching staff. No significant impacts are detected overall, but important heterogeneities are observed. Positive academic impacts are detected for districts below the median in enrollment size, but negative impacts are observed in above-median districts. Similar adverse effects are detected for in the most dense quartile. Disproportionately adverse effects are observed for black students in large districts, but positive ones are detected for Hispanic students. Salary increases have small positive effects. Findings are attributed to loosely coupled organizations.

1 Do we need more adults in our schools?

According to the law of diminishing returns, benefits can be expected to fade with increments in personnel density and size. If returns decline to the vanishing point, or point of satiety, benefits from additional staff will be no longer detectable or will acquire a negative sign. Organizational theorists interpret the vanishing point as the spot at which communication and co-ordination inefficiencies introduced by additional personnel outweigh benefits from increments in staff time devoted to organizational goals (March and Simon, 1958). A science of administration (SA) literature posits efficiencies from consolidating jurisdictions into larger aggregations, though it has not theorized as to whether they are infinite (Holzer et al., 2009; Cubberley, 1919).

A scattered, inconsistent literature has addressed these questions piecemeal, but no previous study has used a national data set to estimate causal estimates of effects of increased personnel density on student academic performance. Nor have researchers identified interactions between school district size and density or heterogeneities by student demographic characteristics or by changes in student enrollment. Estimates of test-score returns vary between as much as 0.2 standard deviations (sd) from a one-third reduction in class size (Krueger, 1999) to no detectable effects (Chingos, 2012; Hoxby, 2000; Grissmer et al., 2000). Academic returns from paraprofessional staff (teacher aides) usually show little or no return from the investment (Krueger, 1999). Effects of non-teaching personnel range from strong positive effects of principals (Brady, Fiorina and Wilkins, 2011; Hanushek et al., 2024) and professional medical personnel to little or no effects for other non-teaching staff (Clotfelter, Hemelt and Ladd, 2016).

The density of staff working in K-12 education rose during the first two decades of the 21st Century (Smith, Campbell and Barnard, 2024). n the district sample used for this study, the number of enrollees increased by a little over 2 percent between 2003 and 2022 and the number of teachers by 1.4 percent (3,004,000 to 3,046,000), but the number of paraprofessionals jumped by 27.8 percent (from 625,550 to 800,320) and the number of non-

teaching, school-based personnel by 57.4 percent (from 1,236,650 adults to 1,994,4700). See Figure 1 and Table 1.

In this paper, we estimate the returns from district staffing choices by linking data on student performances between 2003 and 2022 in 8th grade on biennial waves of the National Assessment of Educational Progress (NAEP), often called the nation's report card, to staffing information available from the Common Core of Data (CCD) assembled by the National Center for Education Statistics (NCES). Reading and math tests were administered every two years except for the interval between 2019 and 2022. The CCD contains information for each school district on the numbers of teaching, paraprofessional, and other school staff, average salary levels, student enrollment, and other information. We estimate plausibly causal effects with two-way fixed effects difference-in-difference models. The first difference is the cohort difference in math and reading achievement within each district between survey waves (every two years, except for the interval between 2019 and 2022; the second difference is between these district changes in each district and comparable changes statewide (statewide biannual trends). Heterogeneities by district size, district density, district growth, and student characteristics are estimated. NAEP-provided weights based on student background characteristics are used to estimate representativeness of the sample. By including state-by-wave fixed effects, the analysis controls for state economic growth rate, political and educational climate, economic policy, and other state-specific factors that may impact student learning. The main threat to a causal interpretation of estimated impacts are unobserved changes due to factors specific to the district correlated with outcomes but uncorrelated with control variables.

Our analysis shows that the vanishing point, or point of satiety, where marginal returns from additional personnel cannot be detected or turn negative, was reached in the United States in the first two decades of the 21st Century. Between 2003 and 2022, increases in the density of teachers, paraprofessionals, and non-teaching school staff showed no significant correlation with changes in cohort performance on 8th grade NAEP math and reading tests. Heterogeneities by size and density of the district are observed, however. An additional

teacher for every 100 students has, on the margins, a positive effect on student performance in small districts (enrollments below the national median of 1,057 students) but a negative impact on learning in districts with enrollments above the median, particularly in the largest quartile consisting of districts attended by 80 percent of the student population. Effects of nonteaching school personnel are also negative in the largest quartile of the distribution. Increments in the density of paraprofessional teaching staff has a curvilinear relationship with district size: Positive impacts are observed in the second and third quartiles. For all three types of staff, adverse effects are observed in the quartile that had the greatest personnel density at the beginning of the period.

Enhanced teacher density in the smallest quartile, by enrollment, is positively correlated with learning for students from higher income households, but it is negatively correlated with achievement in the largest quartile for students coming from low-income households. Adverse impacts are disproportionately large for black students in the largest quartile of the enrollment distribution. but for Hispanic students, positive impacts of increased density of teachers and paraprofessionals are observed. Information on staff salaries is imprecisely estimated, but results show small positive impacts on achievement. We attribute negative findings to the loosely-coupled organizational structure of large school districts. Positive findings for Hispanic students are attributed to the need for new staff in contexts where specialized skills are desirable. Negative impacts on disadvantaged students are attributed to standard operating procedures used in large school districts.

2 Theory and Literature

Two perspectives inform contemporary thinking about the value of increased organizational size and density. The first is associated with the seminal work of James March and Herbert Simon, as presented in their encyclopedic volume, *Organizations* (1958). The second is the science of public administration (SA) literature initiated during the progressive era, which

Edward Cubberly, a professor in the Stanford School of Education, applied to the school sector (Cubberley, 1927).

2.1 Organizational theory

Although economists have celebrated economies of scale, few, if any, argue they are infinite. On the contrary, organizational theorists identify problems of communication and co-ordination as reasons to expect diminishing returns from additional personnel once an organization has reached a certain size. In March and Simon's words: the "capacity of an organization to maintain a complex, highly interdependent pattern of activity is limited in part by its capacity to handle the communication required for coordination" (March and Simon, 1958, p.162).

Downs (1967, p.143) agrees: "The larger any organization becomes, the weaker is the control over its actions exercised by those at the top." and "as any bureaucratic organization grows larger the proportion of wasted activity [...] rises steadily." As organizations grow, leaders must adjust their internal structure either by adding new administrative tiers, making it more hierarchical, or by flattening the structure by giving existing tiers broader responsibility over more personnel. Either strategy removes those responsible for setting organizational goals (e.g., school boards and superintendents) further from those (e.g. teachers, paraprofessionals, and non-teaching personnel) who produce outputs (e.g., academic achievement). Whether one creates more tiers or flattens the structure, threats to goal attainment rise. In education, academic returns from additional personnel diminish, and, at some point, can be expected to turn negative, as the organizational size and density becomes dysfunctional. As one theorist put it, top administrators lose touch with and control over what is happening below, to the point where they may be "operating in purely imaginative worlds" (Williamson, 1967, p. 123).

Despite the law of diminishing returns, administrators often benefit personally from additional personnel, regardless of their contribution toward organizational goals. If an insti-

tution is growing, those in charge are able to enhance their own status by supervising more personnel, promote key aides, reduce internal conflict by assigning new duties to incoming personnel, and enhance institutional autonomy (Downs, 1967, p.16-17).

Unnecessary organizational growth occurs, because, according to March and Simon, organizations "satisfice" –they do not maximize—goal attainment. Since goals are multiple, none can be fully attained without compromising others, and the appropriate maximizing strategy requires information, predictive powers, and management flexibility beyond the leaders' capacities. Those in charge thus search for solutions that "satisfice" over the short run rather than maximize long-term objectives.

March and Simon do not distinguish between private and public organizations, but their theory seems especially applicable to public institutions in a democratic society, such as school boards and their principal agent, the school superintendent. These school leaders are likely to have multiple goals, short horizons, and multiple constraints on available alternatives, factors that encourage satisficing behavior. Loose-coupling theory provides a useful lens for understanding the perverse effects of additional personnel. Weick, (1976) Weick (1976) says many public organizations, including school systems, resemble trains that have loose joints, which give rail cars leeway to twist and turn on the bends and curves of the roadbed. The multiple joints leave professionals distant from those who deliver services.

2.2 SA theory

SA was proposed in the latter part of the 19th Century, when social scientists, journalists and social critics became increasingly concerned about governmental inefficiency and corruption (Callahan, 2010). To take politics out of administration, SA theorists favored, among other policies, a merit-based, politically neutral civil service, citywide rather than ward-based elections, and consolidation of small jurisdictions into more comprehensive ones (Burnham, 1965). Municipal governments were encouraged to annex outlying areas, cities and counties were urged to unify into a single metropolitan government (Grant, 1968) and special

districts were said to work best when they were given broad scope across a metropolitan area. In education, applications of SA theory included recruitment of professionally trained school administrators, city-wide school board elections, and consolidation of small, "little red schoolhouse" districts into larger districts with comprehensive high schools (Berry and West, 2010). Harvard President James Conant favored school district consolidation on the grounds that schools could operate more efficiently by offering college-preparatory classes to talented students in separate classrooms (Conant, 1959).

In sum, organizational and SA theories have somewhat different expectations for the marginal returns from increases in the number of personnel within organizations. Organizational theory expects returns to diminish and even turn in a negative direction once an organization achieves a size that complicates internal communication and co-ordination. SA theory expects greater efficiency from larger, more comprehensive organizations.

2.3 Municipal studies

A small empirical literature on municipalities and school districts, mainly focused on U. S. jurisdictions, has subjected propositions derived from these theories to empirical tests. In an excellent review of that literature, (Holzer et al., 2009, p.10) hypothesize the relationship could be U-shaped, with efficiency steadily declining with size until it turns upward when municipalities become very large. However, they draw a sharp distinction between the delivery of capital intensive services (utilities, roads, etc), for which positive returns from size are observed, and labor-intensive services, such as "public education, fire protection, police protection, refuse collection, etc" for which they find "consolidation ... [has] little, if any, significant effect on [the reduction in] per capita expenditures ... Beyond a point found in medium-sized communities, growth or consolidation will tend to produce per capita expenditure increases." Similarly, a quasi-experimental study by Blom-Hansen et al. (2016) finds no expenditure savings from municipal consolidation in Denmark. Berry (2009) finds that lower costs if services are provided by municipalities rather than by special districts responsible for

specific services but does not report any expenditure savings from jurisdictional size.

Most municipal studies focus on the relationship between size and municipal expenditure, not on the quality of the services rendered. However, a few studies of police services report "very small-to-medium-sized police departments consistently perform more effectively and frequently at less cost than do large police departments (Ostrom (1976), see also (Advisory Commission on Intergovernmental Relations, 1987)).[1] A recent causal study focusing on federal grants from the 2009 American Recovery and Reinvestment Act for increasing the number of police in municipalities (by, on average less than 1 police officer per 10,000 residents) resulted in a reduction in crime of all kinds, albeit the effects are small (Mello, 2019).

2.4 Education research

Education researchers have been able to use information on student test score performances and other educational outcomes to estimate connections among density, size and organizational outputs. Findings vary widely, with positive estimates observed more frequently when studies focus on specific classrooms rather than larger systems, such as districts, states and nations.

Krueger (1999) found the three-year effects of a reduction in class size from 24 to 16 children in grades K-3 to be no less than 0.2sd on standardized test scores.[2] Positive effects for black students persisted through college (Dynarski, Hyman and Schanzenbach, 2013). Comparable positive effects of class size on test-performance were observed in Israel (Angrist and Lavy, 1999), though a subsequent study on more recent data failed to confirm these results (Angrist et al., 2019). Similarly, Hoxby (2000)'s quasi-experimental research in Connecticut found no academic returns from class size reduction. District, state and nationwide reductions in the pupil-teacher ratio are less likely to discern beneficial consequences from

[1] That finding is disputed by Pachon and Lovrich Jr (1977) who claim the differences are due to population characteristics.

[2] This is equivalent in the to going from a 4.1 teachers per 100 students to 6.3 teachers per 100 students, in the metrics we will use.

increments in the size of the teaching staff. An early review of the class size literature found little connection between class size and student outcomes (Hanushek, 2011). In both Florida and California, statewide mandates to reduce class size throughout the state showed no benefits from class size reduction (Chingos, Henderson and West, 2012; Chingos, 2013; Jepsen and Rivkin, 2009). Brunner, Hyman and Ju (2020), exploiting variation in the unionization of states, show that learning gains are greater in states that use extra resources to increase teacher salaries (due to union strength through collective bargaining), rather than adding more teachers. Internationally, positive effects of a reduction in the pupil-teacher ratio in Iceland and Greece were observed but there were no more than minimal effects in thirteen other countries (Woessmann and West, 2006). A recent study using a school reform in Norway that funneled extra teachers to schools (in a country with already large teacher-student ratios) resulted in no gains in learning, though it improved student satisfaction (Borgen et al., 2025).

In our data set, paraprofessionals, also known as teacher aides, comprised 13 percent of the total elementary and secondary school labor force in 2020 (Table 1). The experimental study in Tennessee found significant impact of teacher aides on student performance (Krueger, 1999). A worldwide review of the literature finds inconsistent effects, with greater impacts in developing countries (Alborz et al., 2009). A statewide study of teacher aide impacts in grades three through five in North Carolina found positive effects in reading but none in math (Clotfelter, Hemelt and Ladd, 2016).

Of all school-level staff, principals are generally thought to be the most relevant for student academic performance, as they are typically responsible for selecting and guiding teachers and other staff, allocating resources within the school, and setting its educational climate (Grissom et al., 2021). A recent study of principal effectiveness finds a 0.03 and 0.01sd impact on math and reading performance, respectively (Bartanen, Husain and Liebowitz, 2024). A North Carolina team finds positive effects of health-care workers on student achievement (Clotfelter, Hemelt and Ladd, 2016). Other quasi-experimental or high-quality descriptive

research finds positive impacts of counsellors and health-care professionals on non-cognitive outcomes, such as classroom behavior and teen-age birth rates, but no causal effects on achievement (Carrell and Carrell, 2006; Lovenheim, Reback and Wedenoja, 2016; Reback, 2010*b*,*a*). Berry and West (2010) find negative returns to adult earnings from consolidation of school districts. Bloom, Thompson and Unterman (2010) found positive academic returns from the transformation of large New York City public schools into several smaller ones.

2.5 Summary

In sum, the empirical literature on school districts generally supports organizational theory's central proposition that returns diminish from increasing organizational size and personnel density. Nor does most of the SA research show efficiency gains from increases in jurisdictional size for the delivery of human services. But that literature is skimpy, inconsistent, and often small scale. No U. S. study has systematically addressed the question with a nationwide data set that includes indicators of organizational outcomes. That is the task to which we now turn.

3 Data

Our measure of student achievement comes from 2,425,000 anonymized student-level test scores in math and reading administered to 8th grade students as part of the main NAEP in ten waves in each subject between 2002 and 2023. We exclude NAEP 4th and 12th grade data, as 4th grade performances are heavily shaped by early childhood factors (Shakeel and Peterson, 2022), and scholars have expressed doubts as to whether 12th grade students are sufficiently motivated to perform well on low stakes tests (Brophy and Ames, 2005).[3] District-level test scores are linked to information by district on numbers of teachers, para-professionals and non-teaching school staff, salary expenditures by personnel category, and

[3]Researchers with restricted-use data licenses may obtain NAEP individual-level micro-data from the NCES, which is expected to approve for disclosure all results reported in this paper.

other information available in the CCD of the NCES. Observations use post-stratification weights to be district representative (Little, 1993). All outcomes are reported in standard deviations, calculated over the sample within subject (math or reading).[4] The following background characteristics of individual test-takers are used as control variables and for estimations of interactions: parental education , gender, race (white, black, Asian), Hispanic identity, special education status, English learner status, eligibility for participation in free-and-reduced-lunch programs.[5] For use of this indicator, see Michelmore and Dynarski (2017); Koedel and Parsons (2020); Chingos (2016). Parental education status, as reported by test-taker, is dichotomized between with and without one or more parents with a four-year college degree.[6]

Personnel data come from the NCES Common Core of Data (NCES, 2024). Number of students in each district refers to those in K-12 grades in public schools. Personnel are divided into three categories: teachers, paraprofessionals, and school staff (which include school counselors, other guidance counselors, school psychologists, instructional coordinators, student support services staff, other support services staff, librarians and media specialists, school administrative support, and school-level administrators). District-level administrators and staff are excluded.

Average salaries are calculated at the district level by dividing the total salary for the type of personnel category (excluding fringe benefits and pensions), as given in the CCD [from LEA Finance survey F-33]) by its total full-time equivalent personnel. Given the available information, estimates of the effect of changes in salary levels on achievement can only be indicative. CCD reports at the district level total expenditures on salaries for just two categories: 1) the combined salaries of teachers and paraprofessionals and 2) the salaries of all other school-level personnel. Each of these subtotals is divided by the number of employees

[4]The scores throughout the period are centered at their mean within-subject value, normalized to be 0. To give a sense of the magnitudes, the average standardized score for math in 2003 was -.05 sd, while it was -.14 in 2022. It was -.01 for reading in 2003 and -.06 in 2022.

[5]We report estimates of effects on those with and without an Hispanic identity in separate tables from race, as this characteristic overlaps NAEP racial classification.

[6]Effects by English learner status are shown in the Appendix.

in each of the two categories. An estimate of salary effects can be made if one makes the strong assumption that all districts distribute salaries in the same way for subcategories. For example, we make the strong assumption that paraprofessional salaries have the same ratio to teacher salaries in all districts. Salaries are not available by specific category for other school staff. Effects of district-level staff, estimated separately, have neither a positive nor negative effect (See Table A.3).

4 Analytical strategies

We estimate impacts on student learning in combined math and reading achievement of additions to each of three components of the U. S. public school work force: teachers, paraprofessionals, and non-teaching, school-based personnel. In our preferred density estimates, we use a combined math-reading index as the outcome variable but include a subject fixed effects indicator to adjust for inter-subject differences.[7] Estimates of a change in a district of one additional staff member for every 100 students are reported in standard deviations (sd).

Our basic analytical strategy is given by the following regression specification, estimated separately for each of the three staff categories:

$$Y_{idsy} = \beta S_{dy} + \gamma W_{dy} + \Sigma D_d + \Sigma A_{sy} + R_i + \delta \mathbf{X}_i + G_{dy} + \epsilon_{idsy} \quad (1)$$

where the outcome Y is the NAEP test score for individual i in district d, in state s and year y, measured in standard deviations; S is the staff number per 100 students in the category in the district; W are the salary in the district-year for the staff category; D are district fixed effects; A are state-by-year (of the NAEP wave) fixed effects. R is an indicator for whether the test is a reading or math test. $\mathbf{X}$ is a vector of the individual-level covariates described above, and G is an indicator for whether student enrollment number in the district have grown by more than 10% in the previous ten years. Finally, ϵ is the error term. We use

[7] Results do not change materially when each subject is estimated separately.

the NAEP post-stratification weights in our estimations throughout.

We explore differential effects by size of district enrollment. We delete the smallest 10 percent of the districts in the NAEP sample from the analysis because observations are too few to be reliable. Even so, the smallest district in our NAEP analytic sample has 83 students. For ease of interpretation while accounting for possible nonlinear trends in size we divide district-years by quartile of size. We add three interaction terms $Q_{qd} \times S_{dy}$ (as well as quartile main effects) to our basic model in (1):

$$Y_{iqdsy} = \Sigma(\zeta_q Q_{qd} \times S_{dy}) + \beta S_{dy} + \Sigma\eta_q Q_{qd} + \gamma W_{dy} + \Sigma D_d + \Sigma A_{sy} + R_i + \delta \mathbf{X}_i + G_{dy} + \epsilon_{iqdsy} \quad (2)$$

Q are three quartile indicator variables, with the reference category being the lowest quartile by enrollment size (relative to the NCES Common Core of Data 2003 sample, see cutoffs in Table 1).

The number of districts in the United States is currently 16,092. The smallest district remaining in the sample has 84 students and the average enrollment in the smallest quartile of the distribution is 161 students. The minimum in the second quartile is 84 students, with an average of 161 students. The minimum in the third quartile (median value) is 1,057 enrollees, with an average of 1,748 students; the largest quartile has an average of 10,541, its minimum is 2,742. The largest district (New York City) has 996,495 pre-K to 12th grade students in 2003. Only 1.8 percent of students attend schools in the smallest quartile and 5.1 percent in the second quartile, but 13.2 percent attend schools in the third quartile, and no less than 79.9 percent in the largest. For policy purposes, special attention is to be given to impacts observed in the largest enrollment quartile.

5 Results

The academic returns from increasing teacher density within U. S. school districts has been at the vanishing point (also known as the point of satiety) during the first two decades of the 21st Century. For districts nationwide, we are unable to detect positive impacts on average 8th grade achievement of increments in teacher density, paraprofessional density, or other personnel density. No significant differences in patterns of relationships are observed between the first half and the second half of the period (Appendix Table A.8), nor is the pattern significantly different for the final period (2019-2023) when many schools were offering courses online (though the paucity of observations precludes a precise estimate for this period).

Other heterogeneities are observed, however. For those below the median in enrollment size, positive impacts are sometimes observed; for those above the median, especially in the largest quartile, effects are often significantly negative. Heterogeneous effects are reported relative to the base category. To ascertain total effects for any particular category, add the total of that category to the base category (or categories, when required). In Table 2 , for example, the disproportionate effect for students in the largest quartile relative to the base (smallest) quartile is -0.027sd but the total effect is -0.022sd, as the base category has a value of 0.005sd.

Teacher density effects in smaller and larger districts. Those in the analytic sample of districts identified as teachers in the CCD increased by 1.4 percent between 2003 and 2022 (Table 1), less than the 2.5 percent increase in the number of enrolled students. Upward ticks in teacher density had no statistically significant effect on cohort math and reading performances in 8th grade, nationally (Table 2). The average effect blurs heterogeneities by quartile of enrollment size. In districts below the median in enrollment, increased density has a positive impact; for those in the second quartile, it is 0.01sd between NAEP waves (Table 2). However, estimated effects turn negative for those districts in the largest quartile (-0.027sd). As mentioned, 80 percent of all student enrollments are in districts that fall

within the largest quartile.[8] The finding does not show students learn less if they attend schools in large districts. Other things being equal, scores are higher in larger than smaller school districts, perhaps because schools in larger districts have greater resources, teachers have higher compensation levels, and families are more supportive of their children's education. But whatever advantages students enjoy by attending schools in the largest districts, learning falls when teacher-pupil ratios climb.

Paraprofessional density in small and large districts. School districts have invested considerably more resources in the recruitment of paraprofessional teachers during the 21st Century. The number of teacher aides in the districts within our analytic sample increased from 626,000 to 800,000, a 28 percent increment over the period. A significantly positive impact from additional paraprofessionals is identified for students in the second (0.01sd) and third (0.02sd) quartiles of the enrollment distribution but a large (0.04), though insignificant, adverse effect is observed for those in the largest quartile (Table 2).

Non-teaching staff density in small and large districts. The size of the school level staff not directly engaged in teaching grew by 57 percent during the period. Overall, no statistically significant academic impacts are observed from this major investment in non-teaching staff. In the largest of the quartiles, a large (though statistically insignificant) negative effect is observed (-0.03sd).

Personnel Increments by Density Quartile. 35 percent of students live in districts with the least personnel density at the beginning of the period, 26 percent of students live in the second quartile, 24 in the third quartile, and 13 in the largest teacher density quartile (in 2003). If a district is in the quartile with the greatest personnel density at the beginning of the period (2003), increments in personnel density have a significantly negative effect on student achievement. Effects are 0.03sd for teachers and 0.02sd for both paraprofessionals and for other school staff.

Demographic Factors. Heterogeneities are observed by race and Hispanic identity

[8] Very similar effects, not controlling for growing district are shown in Appendix Table A.1.

as well as by household income and parental education. Forty-three percent of the students in our analytic sample were eligible for free-and-reduced lunch, with a low of 40 percent in the smallest quartile and a high of 48 percent in the third quartile, by enrollment size (Table 1). For students from high-income households, positive achievement returns from increased teacher density are detected for students attending school in small districts (particularly the second quartile), but adverse effects are observed for low-income students in the three larger quartiles (Table 3). In other words, the negative impacts of increasing density tend to be concentrated on students from low-income households. This inference also seems applicable to paraprofessional staffing in larger districts. Nationwide, no detectable impacts are observed for students from households in either the high or low income category, but in the third quartile of the enrollment distribution, marginally positive returns from paraprofessional density are observed for high-income students. (Table 3). Meanwhile effects of enhanced paraprofessional density on low-income students in the largest quartile are both negative and large (-0.045sd), though not statistically significant.

Half the analytical sample consists of students from households where neither parent has a college degree (Table 1). Overall, no significant heterogeneities by parental education are detected for teachers and other school personnel (Table 4). However, we do observe a positive impact of paraprofessionals (0.01sd) on the performance of students from more educated backgrounds in districts in the second and third enrollment quartiles, a contrast from the negative, if still insignificant, effect of increased paraprofessional density on students from less educated backgrounds in the second quartile (-0.01sd).

The distribution of the analytic sample by race and ethnicity is as follows: 55 percent white, 15 percent black, 22 percent Hispanic, and 5 percent Asian. Those who are identified as Hispanic by NAEP are also given a racial identity, so that category is not mutually exclusive of the others. Effects of density and size on achievement do not vary among students identified as white or Asian (Table 5), but among those classified as black, a disproportionately negative impact (is detected for those in the largest quartile (-.03). For Hispanic

students, a different pattern of results emerges: Enhanced teacher density has a positive differential impact on Hispanic student learning (as compared to that of non-Hispanic students) in districts above the size median (a significant 0.06sd in the third quartile and 0.05sd in the largest quartile). See Table 6. The differential effect of paraprofessional density on achievement is also significantly positive (0.07sd) in the third quartile of the size distribution, but, oddly, the sign is negative (though insignificant) in the largest quartile of the enrollment distribution (-0.03sd).[9]

District Growth. A district is identified as a no growth district if enrollments did not expand by 10 percent or more over the decade prior to the test-score observation. By this definition, 48 percent of district observations experienced no enrollment growth or declining growth over the period. Growth was more likely in districts above the median in enrollment. In the largest districts by enrollment, no or negative growth occurs in only 24 percent of the observations, 51 percent of those in the third quartile show no significant increase in students, as do 66 percent and 74 percent of the districts in the second and smallest quartiles, respectively. See Table 1. If a district is below median in size, a lack of growth has no detectable consequence for the density-achievement relationship, but in districts above the median, no growth adversely affects the relationship between density and achievement (Table 7). In these large, no-growth districts, increasing density has a negative impact of -0.01sd on achievement in the third quartile, and 0.04sd in the largest quartile. In the largest quartile, achievement is also negatively affected by intensifying density of other staff (-0.05sd) and of paraprofessionals (-0.09sd, though not significant). (See Table 8). These relationships between no growth and density impacts are to be interpreted cautiously, as the connection may be endogenous. If a district is declining in size, that could negatively affect the density-achievement connection, but the opposite may also happen: When a district's academic productivity slips, it may encourage more family exits than arrivals.

Salaries. Given the limited information available from the CCD (see discussion above),

[9]Similarly, increased teacher density has a positive impact on the performance of students identified as English Language Learners (ELL). See Appendix Table A.2.

estimates of the effects of salary increments on achievement can only be indicative. A $10,000 increase in average teacher salary levels has an impact on achievement of 0.004sd. That effect is concentrated in districts above the median in enrollment, where it ranges slightly from a significant 0.006sd in the third quartile to an insignificant to 0.012sd in the largest (See Table 8). Positive salary effects on achievement are concentrated on students from higher-income households, those with college-educated parents, and Asian students (See Tables Appendix tables A.4, A.5, A.6). The substantial increments in salaries have only small effects on student test scores, but at least they do not have perverse effects often observed when marginal dollars are used to recruit additional staff.

Regional, Collective Bargaining and Partisan differences. In a search for meaningful differences across states, we substituted a nationwide year-fixed effects model for the state-year fixed model used elsewhere. We felt comfortable with this relaxation of the two-way fixed effects model, because the main findings reported in Table 2 are not sensitive to the use of a state or national fixed effects models. The revised model is used to estimate personnel effects by census region, state collective bargaining laws, or partisan propensity (see Appendix Tables A.9, A.10 and A.11). Only one of the differences across states has a differential effect on the density-achievement connection. Density impacts are particularly adverse in the three continental states bordering the Pacific Ocean. We hesitate to interpret that finding, however, as adverse effects by state show no discernible pattern. It is hard to draw strong conclusions when Texas and California both show particularly adverse effects of enhanced teacher density (Figure 2).

6 Discussion and limitations

Although the analysis we have presented has certain limitations, the findings warrant summation and interpretation.

6.1 Summary

Major investments in personnel density were made in the first two decades of the 21st Century. Student enrollments increased by 2 percent and the number of teachers by 1.6 percent, but the number of paraprofessionals grew by 28 percent and non-teaching school staff by 57 percent. On average, those investments in personnel had no detectable effect on 8th grade math and reading performances on NAEP. However, heterogeneities are observed. Academic benefits from additional teachers are identified for those attending schools in small districts and for Hispanic students. However, increased teacher and other staff density have negative effects on student performance in large districts, particularly in those that have not grown over the previous decade. Negative effects are also observed for increments in teaching, paraprofessional, and other staff in the quartile of districts that had the highest personnel densities at the beginning of the period. About 80 percent of all students attend schools in districts in the largest quartile of the distribution. The negative effects of teacher density are disproportionately concentrated on black students, those from low-income households, and those born of parents without a four-year college degree. We do not observe variation by state voting behavior in the 2000 presidential election, which preceded the beginning of the period. Nor do we observe differences between states that do or do not require districts to bargain collectively with teachers, nor by region of the country, except for adverse effects observed in continental states bordering the Pacific Ocean. Effects do not differ significantly between the first and second half of the period or for the final three-years of the period, during which many schools were offering instruction online rather than in the classroom.

6.2 Limitations

Like other studies that rely upon NAEP and NCES data, this analysis assumes test-scores are valid indicators of academic learning, the numbers and salaries of employees are accurately reported to NCES, test data are representative of students within a state, observables capture any unobserved trends specific to a district, and district classification error is random.

NAEP test performances are for biannual cross-sections of cohorts of 8th grade students, not the progress of individual students over time. Though the ten waves in each subject are temporally linked and designed to be state and nationally representative, non-compliance and sampling error could bias estimates. NAEP administers low-stakes tests. Performances of individual students, classes, schools, or districts are not made publicly available. This has the advantage of minimizing incentives to cheat, but low-stakes tests may not be taken seriously by test-takers. However, NAEP tests of 8th grade students correlate well with high-stakes accountability tests administered by states (Lastra-Anadón and Peterson, 2023). The relationship between staff density and achievement growth in growing districts should be interpreted with greater caution than other findings. Growth could enhance the impact of staff density, but effective schools could attract families with children to districts. NCES provides only broad categories for personnel salaries by district, reducing the precision of the salary estimates used here. Finally, a two-way fixed analysis estimates differences between trends across NAEP waves within districts and differences in trends statewide. Causal relationships may be cautiously inferred from differences-in-differences analyses, but they do not meet the gold standard of a well-designed randomized field trial. However, this analysis, unlike most randomized trials, may, if valid, be generalized to U. S. schools nationwide.

6.3 Interpretation

In large districts a one-percent increment in the number of teachers will have the unfortunate consequence of adversely affecting average student performance by a little more than one month every biennium.[10] A similar adverse effect is observed for districts in the highest quartile of the personnel density distribution at the start of the period. Since 80 percent

[10]The findings are reported in standard deviations for technical accuracy, but they can be translated into months of learning for discussion within the larger policy community, though any translation from a technical to a vernacular language is necessarily imprecise. The average 8th grader scores about one standard deviation higher on the math and reading tests than does the average student in 4th grade. If learning accumulates at a linear pace from one grade to the next during this four year interval, then the average student learns about 0.25sd per year. Estimates are given for marginal changes in staff density, not average effects for all teachers or other personnel.

of enrolled students fall within the largest enrollment quartile, the academic learning of a large portion of U. S. students is suffering from recruitment of excess personnel. Increased teacher density nonetheless has a positive impact on learning in small districts. In these districts, superintendents and principals seem to be less constrained by rules and routines, which gives them more capacity to direct extra personnel to situations where need exists. When enrollments increase, administrators seem to find it easier to circumvent standard operating procedures (e. g., student-staff ratios; compensation rules; seniority privileges) that otherwise pre-determine staff allocation. Additional personnel can also be put to work effectively if the organization is assuming new tasks that require personnel with specialized skills. Such is the case with Hispanic and ELL students, many of whom require bilingual instruction or instruction in English as a Second Language. The Hispanic population grew rapidly in schools during the period but, overall, the number of teachers did not (Table 1). Those districts able to hire extra teachers and paraprofessionals may have acquired new personnel with language skills needed to lift the reading and math performance of Hispanic students.

Loose-coupling theory provides a useful lens for understanding the perverse effects of additional personnel (Weick, 1976). Teachers exercise wide discretion over classroom management and instructional practice. Principals have extensive latitude in the recruitment of personnel and allocation of resources within schools. Multiple administrative tiers are required to manage large districts, and, at lower tiers, the range of responsibility spreads widely over dozens of schools. The professional superintendent in charge of administrating the vast enterprise reports to an elected school board or one appointed by an elected official. The professional does not have sole discretion to set goals but is guided by a democratically chosen set of supervisors. Adding to the complexity are state regulations, court decisions, and federal guidelines. The multiple joints, barely connecting the semi-autonomous units, leave superintendents distant from teaching, paraprofessional, and non-teaching personnel.

Unless the district is led by a cohesive board, working in harmony with a skillful, deter-

mined superintendent (Vogel, 2021), a large school system may become dysfunctional. Years after a new superintendent arrived in Atlanta, one recalled she had found "the system was in crisis. I was the fifth superintendent in 10 years, and so the central office reflected the total disarray that existed in the organization. There were clearly people there who were working really hard, but they were not sure of the vision, they were not sure of the direction." (Atlanta Public Schools, 2009).

The work of a superintendent of a large system differs sharply from that of a small one. Those in charge of small districts say most of their time is occupied with budgets, school facility operations, and "gathering and consuming data about their schools" (Diliberti and Schwartz, 2024). Those tasks seem close to the ground, in rough proximity to the specifics of the tasks of the organization. By contrast, superintendents of large districts report "spending the largest shares of their time on external and internal communications, including with school board members, parents, and central office staff." Communication, not content, preoccupies those responsible for the cosmic social space encompassed by large school districts.

Yet new superintendents are expected by their boards to propose new policies, alternative approaches, and novel strategies to lift student performance (Hess, 2011). They hire additional staff to implement the proposed changes, who then encounter resistance from pre-existing personnel. Old-timers have tenure, precluding dismissal other than through a prolonged quasi-legal process. In the private sector, organizational heads address these issues by periodically imposing wholesale staff reductions across the firm (Todd, 2022). That is often found objectionable in the public sector, so school superintendents muddle through rather than take the political risk of releasing employees without specific justification. Nothing is more controversial than school closures, even when under-enrollment would seem to require such an action (Kogan, 2025).

Given the loose connections between administrative tiers, school districts rely heavily upon standard operating procedures. Boards enact standard salary schedules that vary compensation not by competence but by university degree and years of service; they offer vacant

positions first to senior teachers; they grant tenure within three years of initial hiring; and they devolve to principals the responsibility for recruiting new teachers, paraprofessionals, and other staff. These standard procedures facilitate co-operation but tend to give junior personnel the responsibility for educating students of low-income. Yet beginning teachers are least effective in the classroom (Kane, Rockoff and Staiger, 2008; Chetty, Friedman and Rockoff, 2014). "Schools serving at-risk children" find it difficult to attract and retain effective teachers, say Guarino and colleagues. "At-risk schools not only attract lower percentages of entering [likely effective] teachers . . . but lose these teachers to other schools at higher rates." (Guarino, Brown and Wyse, 2011). As many as 7 percent of teachers change schools within a district each year, and a similar share move to another district (Luekens, 2004; Boe, Cook and Sunderland, 2008, p.8). Taken together, these standard operating procedures disproportionately burden schools with higher concentrations of disadvantaged students when personnel density is on the rise. As staffing becomes increasingly dense, opportunities for experienced teachers expand, but the quality of the work force responsible for the education of the disadvantaged shifts in the opposite direction.

In large districts, marginal dollars might better be put to salary increments. Gains in student performance are positively correlated with increments in personnel salaries. Even larger effects might be estimated if more precise data could be obtained. At the very least salary increments appear cause no harm.

If a district is not expanding, or, even worse, if it is trying to adjust to falling enrollments, downward adjustments are likely to aggravate staff (Cavin, Murnane and Brown, 1985) and enflame communities (Boyd, 1979). Superintendents in such places seem willing to forego compulsory departures even if staff/pupil ratios rise. Yet when downsizing is voluntary, the first to leave are those capable enough to find alternatives. Density increases may ease organizational conflict, but the results are not necessarily to the advantage of students. Enrollment decline marks the country's educational system in the 21st Century's third decade. Fertility rates are sliding, immigration rates are crashing, and alternatives to public schooling are on

the increase. The federal government projects that by 2030 public schools will enroll just 46.9 million students, a downward shift of about 5 percent from the level in 2021 (National Center for Education Statistics, 2022). If projections are correct, an increasing number of districts will find themselves grappling with the logistical and political challenges posed by school closures and demands for reductions-in-force. Increasingly applicable could become the familiar adage, "too many cooks spoils the broth."

References

Advisory Commission on Intergovernmental Relations. 1987. "The Organization of Local Public Economies." pp. 1–56.

Alborz, A., D. Pearson, P. T. Farrell and A. J. Howes. 2009. "The Impact of Adult Support Staff on Pupils and Mainstream Schools." *Institute of Education* . https://eppi.ioe.ac.uk/cms/Portals/0/PDF

Angrist, Joshua D and Victor Lavy. 1999. "Using Maimonides' rule to estimate the effect of class size on scholastic achievement." *The Quarterly journal of economics* 114(2):533–575.

Angrist, Joshua D, Victor Lavy, Jetson Leder-Luis and Adi Shany. 2019. "Maimonides rule redux." *American Economic Review: Insights* 1(3):309–324.

Atlanta Public Schools. 2009. "Atlanta Public Schools Report, 2009." Quoted in Wallace Foundation PDF (2023 edition). No publisher (n.p.), as cited in Meredith I. Honig et al., *Central Office Transformation for District-Wide Teaching and Learning Improvement*, Wallace Foundation, 2010.

Bartanen, Brendan, Aliza N Husain and David D Liebowitz. 2024. "Rethinking principal effects on student outcomes." *Journal of Public Economics* 234:105115.

Berry, Christopher R. 2009. *Imperfect union: Representation and taxation in multilevel governments*. Cambridge University Press.

Berry, Christopher R and Martin R West. 2010. "Growing pains: The school consolidation movement and student outcomes." *The Journal of Law, Economics, & Organization* 26(1):1–29.

Blom-Hansen, Jens, Kurt Houlberg, Søren Serritzlew and Daniel Treisman. 2016. "Jurisdiction size and local government policy expenditure: Assessing the effect of municipal amalgamation." *American political science review* 110(4):812–831.

Bloom, Howard S, Saskia Levy Thompson and Rebecca Unterman. 2010. "Transforming the High School Experience: How New York City's New Small Schools Are Boosting Student Achievement and Graduation Rates." *MDRC* .

Boe, Erling E, Lynne H Cook and Robert J Sunderland. 2008. "Teacher turnover: Examining exit attrition, teaching area transfer, and school migration." *Exceptional children* 75(1):7–31.

Borgen, Nicolai T., Lars J. Kirkebøen, Andreas Kotsadam and Oddbjørn Raaum. 2025. "Do Funds for More Teachers Improve Student Outcomes?" *Journal of Human Resources* . **URL:** *https://jhr.uwpress.org/content/early/2025/02/28/jhr.0523-12892R2*

Boyd, William Lowe. 1979. "Introduction: The changing politics of changing communities." *Education and Urban Society* 11(3):275–284.

Brady, David W., Morris P. Fiorina and Arjun S. Wilkins. 2011. "The 2010 elections: Why did political science forecasts go awry?" *PS: Political Science & Politics* 44(2):247–250.

Brophy, Jere and Carole Ames. 2005. "NAEP Testing for Twelfth Graders: Motivational Issues." *National Assessment Governing Board* .

Brunner, Eric, Joshua Hyman and Andrew Ju. 2020. "School finance reforms, teachers' unions, and the allocation of school resources." *Review of Economics and Statistics* 102(3):473–489.

Burnham, Walter Dean. 1965. "The changing shape of the American political universe." *American political science review* 59(1):7–28.

Callahan, Raymond E. 2010. *Education and the Cult of Efficiency: A Study of the Social Forces That Have Shaped the Adminstration of the Public Schools*. University of Chicago Press.

Carrell, Scott E and Susan A Carrell. 2006. "Do lower student to counselor ratios reduce school disciplinary problems?" *The BE Journal of Economic Analysis & Policy* 5(1):000010151515380645 1463.

Cavin, Edward S, Richard J Murnane and Randall S Brown. 1985. "School District Responses to Enrollment Changes: The Direction of Change Matters!" *Journal of Education Finance* 10(4):426–440.

Chetty, Raj, John N Friedman and Jonah E Rockoff. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American economic review* 104(9):2633–2679.

Chingos, Matthew M. 2012. "The impact of a universal class-size reduction policy: Evidence from Florida's statewide mandate." *Economics of Education Review* 31(5):543–562.

Chingos, Matthew M. 2013. "Class size and student outcomes: Research and policy implications." *Journal of Policy Analysis and Management* pp. 411–438.

Chingos, Matthew M. 2016. "Instructional quality and student learning in higher education: Evidence from developmental algebra courses." *The Journal of Higher Education* 87(1):84–114.

Chingos, Matthew M, Michael Henderson and Martin R West. 2012. "Citizen Perceptions of Government Service Quality: Evidence from Public Schools." *Quarterly Journal of Political Science* 7:411–445.

Clotfelter, Charles T, Steven W Hemelt and Helen F Ladd. 2016. "Teaching Assistants and Nonteaching Staff: Do They Improve Student Outcomes? Working Paper 169." *National Center for Analysis of Longitudinal Data in Education Research (CALDER)* .

Conant, James Bryant. 1959. *The American high school today: A first report to interested citizens*. McGraw-Hill Book Company.

Cubberley, Ellwood Patterson. 1919. *Public education in the United States: A study and interpretation of American educational history; an introductory textbook dealing with the larger problems of present-day education in the light of their historical development.* Houghton Mifflin.

Cubberley, Ellwood Patterson. 1927. *State school administration: A textbook of principles.* Houghton Mifflin.

Diliberti, Melissa Kay and Heather L. Schwartz. 2024. Staffing, Budget, Politics, and Academic Recovery in Districts: Selected Findings from the Fall 2023 American School District Panel Survey. Technical Report RR-A956-19 RAND Corporation. Accessed January 23, 2025.
URL: *https://www.rand.org/pubs/research*$_r$*eports/RRA*$956-19.html$

Downs, Anthony. 1967. *Inside bureaucracy.* Little, Brown & Company.

Dynarski, Susan, Joshua Hyman and Diane Whitmore Schanzenbach. 2013. "Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion." *Journal of policy Analysis and management* 32(4):692–717.

García, Emma and Eunice Han. 2021. "The Impact of Changes in Public-Sector Bargaining Laws on Districts' Spending on Teacher Compensation." *Economic Policy Institute* .

Grant, Daniel R. 1968. "The Metropolitan Government Approach: Should, Can, and Will It Prevail?" *Urban Affairs Quarterly* 3(3):103–110.

Grissmer, David W, Ann Flanagan, Jennifer H Kawata, Stephanie Williamson and Tom LaTourrette. 2000. *Improving student achievement: What state NAEP test scores tell us.* Rand Corporation.

Grissom, Jason A, Anna J Egalite, Constance A Lindsay et al. 2021. "How principals affect students and schools." *Wallace Foundation* 2(1):30–41.

Guarino, Cassandra M, Abigail B Brown and Adam E Wyse. 2011. "Can districts keep good teachers in the schools that need them most?" *Economics of Education Review* 30(5):962–979.

Hanushek, Eric. 2011. *Does money matter?: The effect of school resources on student achievement and adult success*. Brookings Institution Press.

Hanushek, Eric A, Andrew J Morgan, Steven G Rivkin, Jeffrey C Schiman, Ayman Shakeel and Lauren Sartain. 2024. The Lasting Impacts of Middle School Principals. Technical report National Bureau of Economic Research.

Hess, Frederick M. 2011. *Spinning wheels: The politics of urban school reform*. Brookings Institution Press.

Holzer, Marc, John Fry, Etienne Charbonneau, Robert Shick, Eileen Burnash, Atta Ceesay, Sunjoo Kwak, Weiwei Lin, Gautam Nayer and Alicia Schatteman. 2009. *Literature review and analysis related to costs and benefits of service delivery consolidation among municipalities*. Rutgers–Newark, School of Public Affairs and Administration.

Hoxby, Caroline M. 2000. "The effects of class size on student achievement: New evidence from population variation." *The quarterly journal of economics* 115(4):1239–1285.

Jepsen, Christopher and Steven Rivkin. 2009. "Class size reduction and student achievement: The potential tradeoff between teacher quality and class size." *Journal of human resources* 44(1):223–250.

Kane, Thomas J, Jonah E Rockoff and Douglas O Staiger. 2008. "What does certification tell us about teacher effectiveness? Evidence from New York City." *Economics of Education review* 27(6):615–631.

Koedel, Cory and Eric Parsons. 2020. "The Effect of the Community Eligibility Provision on the Ability of Free and Reduced-Price Meal Data to Identify Disadvantaged Students.

Working Paper No. 234-0320." *National Center for Analysis of Longitudinal Data in Education Research (CALDER)* .

Kogan, Vladimir. 2025. *No Adult Left Behind: How Politics Hijacks Education Policy and Hurts Kids*. Cambridge University Press.

Krueger, Alan B. 1999. "Experimental estimates of education production functions." *The Quarterly Journal of Economics* 114(2):497–532.

Lastra-Anadón, Carlos X and Paul E Peterson. 2023. "The efficiency-equity trade-off in a federal system: Local financing of schools and student achievement." *Publius: The Journal of Federalism* 53(2):174–200.

Little, Roderick JA. 1993. "Post-stratification: a modeler's perspective." *Journal of the American Statistical Association* 88(423):1001–1012.

Lovenheim, Michael F and Alexander Willén. 2019. "The long-run effects of teacher collective bargaining." *American Economic Journal: Economic Policy* 11(3):292–324.

Lovenheim, Michael F, Randall Reback and Leigh Wedenoja. 2016. How does access to health care affect teen fertility and high school dropout rates? Evidence from school-based health centers. Technical report National Bureau of Economic Research.

Luekens, Michael T. 2004. *Teacher attrition and mobility: Results from the teacher follow-up survey, 2000-01*. National Center for Education Statistics.

March, James G and Herbert A Simon. 1958. *Organizations*. John Wiley & sons.

Mello, Steven. 2019. "More COPS, less crime." *Journal of Public Economics* 172:174–200. **URL:** *https://www.sciencedirect.com/science/article/pii/S0047272718302305*

Michelmore, Katherine and Susan Dynarski. 2017. "The gap within the gap: Using longitudinal data to understand income differences in educational outcomes." *AERA Open* 3(1):2332858417692958.

National Center for Education Statistics. 2022. Digest of Education Statistics 2022: Table 203.10 Enrollment in Public Elementary and Secondary Schools, by Level and Grade: Selected Years, Fall 1980 through Fall 2031. Technical Report Digest Table 203.10 U.S. Department of Education, Institute of Education Sciences. Retrieved from NCES Digest of Education Statistics 2022.

URL: *https://nces.ed.gov/programs/digest/d22/tables/dt22*$_2$03.10.*asp*

NCES. 2024. "Common Core of Data (CCD) via the ElementarySecondary Information System (ElSi)." https://nces.ed.gov/ccd/elsi/. Accessed 1 July 2024.

Ostrom, Elinor. 1976. "Size and performance in a federal system." *Publius* 6(2):33–73.

Pachon, Harry P and Nicholas P Lovrich Jr. 1977. "The consolidation of urban public services: A focus on the police." *Public Administration Review* pp. 38–47.

Reback, Randall. 2010*a*. "Noninstructional spending improves noncognitive outcomes: Discontinuity evidence from a unique elementary school counselor financing system." *Education Finance and Policy* 5(2):105–137.

Reback, Randall. 2010*b*. "Schools' mental health services and young children's emotions, behavior, and learning." *Journal of Policy Analysis and Management* 29(4):698–725.

Shakeel, M Danish and Paul E Peterson. 2022. "A half century of progress in US student achievement: Agency and Flynn effects, ethnic and SES differences." *Educational Psychology Review* 34(3):1255–1342.

Smith, Aaron Garth, Jordan Campbell and Christian Barnard. 2024. Public Education at a Crossroads: A Comprehensive Look at K-12 Resources and Outcomes. Technical report Reason Foundation. Retrieved [August 10, 2025], from Reason Foundation.

URL: *https://reason.org/wp-content/uploads/public-education-crossroads.pdf*

Todd, Sarah. 2022. "The Short but Destructive History of Mass Layoffs." *Quartz at Work* . Updated July 21, 2022.

URL: *https://qz.com/work/1663731/mass-layoffs-a-history-of-cost-cuts-and-psychological-tolls*

Vogel, Mike. 2021. "Superintendent of Choice." *Florida Trend* 64(5):78–82. Accessed via digital publication.
URL: *https://www.mydigitalpublication.com/publication/?m=28897i=714077p=80pp=1ver=html5*

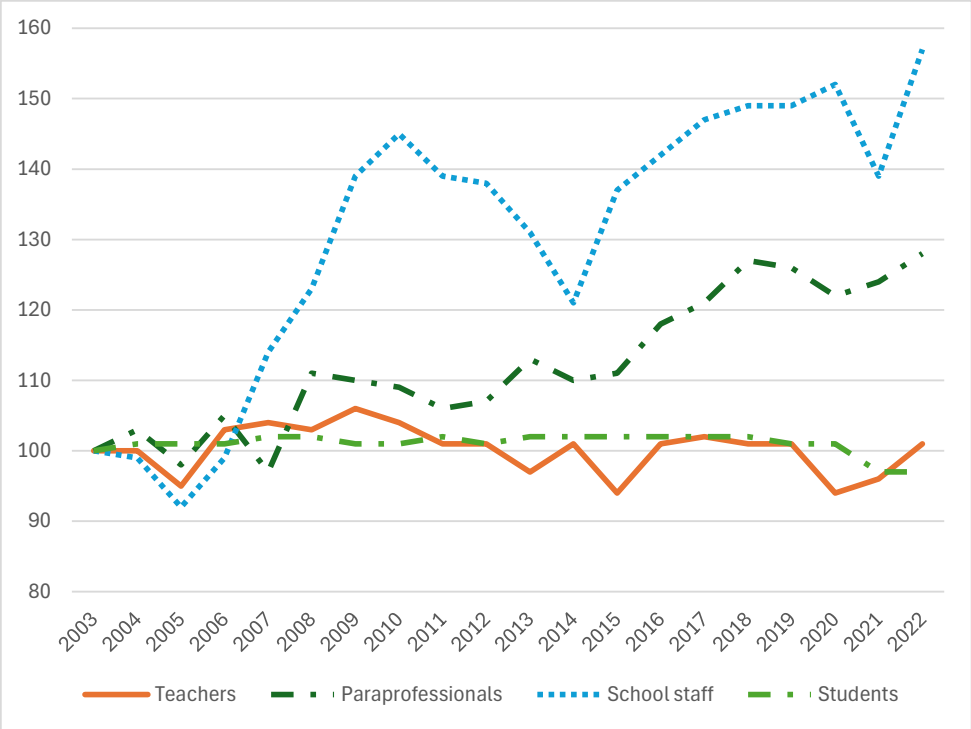
Weick, Karl E. 1976. "Educational Organizations as Loosely Coupled Systems." *Administrative Science Quarterly* 21(1):1–19.

Williamson, Oliver E. 1967. "Hierarchical control and optimum firm size." *Journal of political economy* 75(2):123–138.

Woessmann, Ludger and Martin West. 2006. "Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS." *European Economic Review* 50(3):695–736.

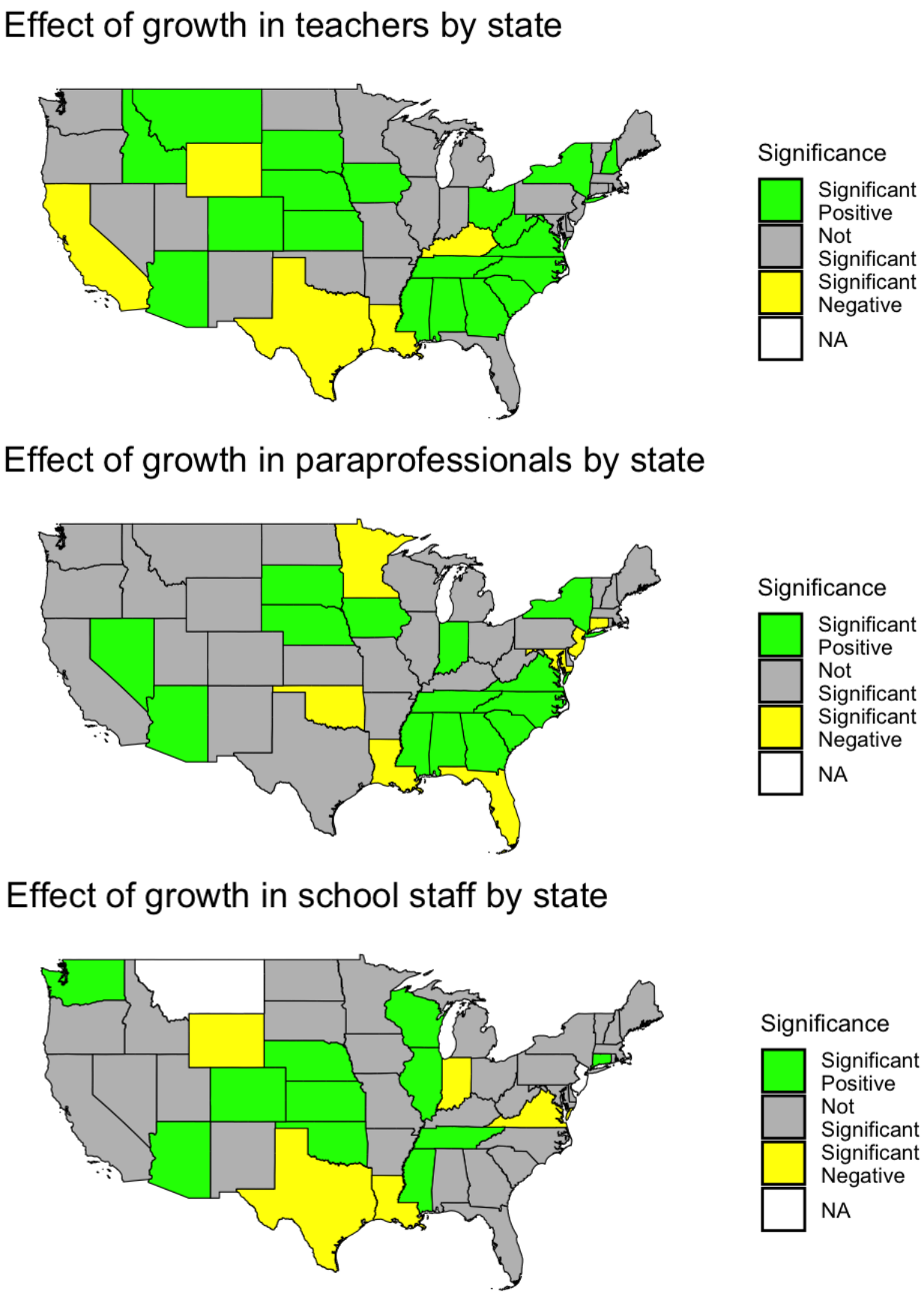
7 Figures

Figure 1: Evolution since 2003 of total school district staff by category, and students.



Rebased relative to 2003 (set to 100). Excludes smallest 10% of districts (less than 83 students). Source: NCES Common Core of Data.

Figure 2: Effect of teachers and paraprofessionals by state.



8 Tables

Table 1: Summary statistics of selected variables

Panel A: National total (thousands)

Variable	**Total 2003**	**Total 2022**	**Growth 2003–22**	**CAGR 2003–22**
Teachers	3,003.84	3,045.86	1.40%	0.07%
Paraprofessionals	625.55	800.32	27.94%	1.31%
School staff	1,236.65	1947.71	57.42%	2.42%
Students K-12	48,113.21	49,323.40	2.52%	0.13%

Panel A Notes: Excludes smallest 10% of districts (less than 83 students). CAGR: Compound annual growth rate. Source: NCES Common Core of Data.

Panel B: Summary Statistics in NAEP analytical sample, 2003–2022, by District Enrollment Size Quartile

Variable	**All Observations**	**Smallest size Quartile**	**2nd size Quartile**	**3rd size Quartile**	**Largest size Quartile**
White	0.547 (0.498)	0.749 (0.434)	0.665 (0.472)	0.480 (0.500)	0.443 (0.497)
Black	0.151 (0.358)	0.088 (0.283)	0.119 (0.324)	0.163 (0.370)	0.186 (0.389)
Asian	0.054 (0.226)	0.017 (0.131)	0.038 (0.192)	0.067 (0.249)	0.070 (0.254)
Native Am.	0.011 (0.103)	0.023 (0.150)	0.009 (0.096)	0.008 (0.091)	0.008 (0.088)
Hispanic	0.217 (0.412)	0.108 (0.311)	0.152 (0.359)	0.261 (0.439)	0.267 (0.442)
Free or reduced lunch	0.427 (0.495)	0.416 (0.493)	0.392 (0.488)	0.467 (0.499)	0.425 (0.494)
College parent	0.502 (0.500)	0.483 (0.500)	0.511 (0.500)	0.468 (0.499)	0.527 (0.499)
Teachers/100 stud.	6.27 (1.91)	7.05 (2.73)	6.37 (1.36)	5.94 (1.41)	6.08 (2.01)
Paraprof./100 stud.	1.38 (0.95)	1.63 (1.47)	1.49 (0.90)	1.24 (0.70)	1.30 (0.81)
School staff/100 stud.	3.01 (1.79)	3.06 (2.32)	2.91 (1.64)	2.89 (1.48)	3.18 (1.80)
Growing dist.	0.522 (0.500)	0.263 (0.440)	0.340 (0.474)	0.482 (0.500)	0.758 (0.428)
Salaries: teachers	69386 (21385)	61874 (20074)	69515 (20460)	69738 (20903)	73208 (22151)
Salaries: teachers, paraprof.	56332 (19424)	49491 (21711)	56146 (17713)	56927 (16948)	59736 (20517)
Salaries: school staff	53851 (38959)	45076 (32015)	51136 (32952)	55001 (40144)	59774 (44182)
Current exp. per student	11905 (28598)	13353 (63859)	12440 (4642)	10963 (3863)	11347 (3323)
Students in district (average)	3295 (14326)	161 (2)	678 (3)	1748 (8)	10541 (456)
Students in district (min.)	84	84	388	1057	2742
Share students per quartile		*1.78*	*5.13*	*13.23*	*79.85*
No. of districts		*3610*	*3611*	*3607*	*3611*
Math score (norm. avg.)	0.050 (1.008)	0.093 (0.938)	0.137 (0.985)	0.017 (1.017)	0.004 (1.039)
Reading score (norm. avg.)	0.055 (1.012)	0.100 (0.940)	0.121 (0.978)	-0.027 (1.016)	0.056 (1.051)

Panel B Notes: Mean and standard deviations (in parentheses) for all observations and observations in each size quartile. Observations weighted by NAEP-provided analytical weights. Excludes 1,603 smallest 10% of districts (less than 83 students). Whites are non-Hispanic whites. "Growing districts" are those that grew by 10% students or more over the previous 10 years. Quartiles are defined by share of all districts in 2003, by the

Table 2: Effect of Staffing Change on Math and Reading Achievement, by Enrollment Quartile and controlling for Growth, 2003-2022.

	Teachers	Paraprofessionals	School staff
Average effects	-0.00205	0.00285	-0.00470
	(0.00825)	(0.00613)	(0.00570)
Personnel effects:			
Smallest district quartile	0.00474	-0.00144	0.00178
(Base category)	(0.00437)	(0.00345)	(0.00266)
Base × Q2 size	0.0101*	0.0109*	0.000800
	(0.00429)	(0.00510)	(0.00261)
Base × Q3 size	-0.00599	0.0171**	-0.00466
	(0.00478)	(0.00618)	(0.00479)
Base × Largest district quartile	-0.0270*	-0.0321	-0.0316
	(0.0122)	(0.0422)	(0.0168)
Control variables:			
Growing district	-0.000538	-0.000793	-0.00118
	(0.00441)	(0.00532)	(0.00566)
Salaries	0.000469**	0.000192	0.0000955
	(0.000147)	(0.000307)	(0.0000789)
Q2 size	-0.0609	-0.00295	0.00658
	(0.0337)	(0.0171)	(0.0180)
Q3 size	0.00882	-0.0298	0.0173
	(0.0336)	(0.0232)	(0.0218)
Largest q. size	0.148	0.0383	0.101
	(0.0884)	(0.0533)	(0.0522)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. For reference, shows the coefficients for control variables. Base category is the smallest district quartile by student size. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, salary controls, as well as a subject fixed effects and individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3: Effect of Staffing Change on Math and Reading Achievement by household income status (free or reduced lunch eligibility) and by Enrollment quartile, 2003-2022.

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of staff by household income			
High income	-0.000312	0.00306	-0.00316
(Base category)	(0.00742)	(0.00681)	(0.00498)
Base × Low income	-0.000927	-0.000941	-0.00339
	(0.00377)	(0.00611)	(0.00297)
Panel B: Effect of staff by household income and district size			
High income , smallest quartile	0.00342	0.000116	0.000692
(Base category)	(0.00353)	(0.00341)	(0.00257)
Base × 2nd quartile	0.0128**	0.0130*	0.00237
	(0.00467)	(0.00545)	(0.00263)
Base × 3rd quartile	0.00437	0.0290**	-0.00187
	(0.00551)	(0.00913)	(0.00545)
Base × Largest quartile	-0.00674	-0.0104	-0.0240
	(0.0160)	(0.0398)	(0.0139)
Low income × Base	0.00350	-0.00357	0.000894
	(0.00433)	(0.00349)	(0.00351)
Low income × 2nd quartile × Base	-0.0105*	-0.00941	-0.00168
	(0.00414)	(0.00507)	(0.00319)
Low income × 3rd quartile × Base	-0.0212**	-0.0280	-0.00457
	(0.00776)	(0.0140)	(0.00608)
Low income × Largest quartile × Base	-0.0295**	-0.0266	-0.00929
	(0.00985)	(0.0194)	(0.00638)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. In Panel A base category are students not receiving free or reduced lunches. In panel B, base category are students not receiving free or reduced lunches (High income) in the smallest district quartiles. Alternative groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, as well as a subject fixed effect, salaries (salary spending per number of staff members), and individual student controls. Clustered standard errors, by state, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Effect of Staffing Change on Math and Reading Achievement by parental education and enrollment quartile, 2003-2022.

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of staff by parental education			
High education	-0.00458	0.00186	-0.00537
(Base category)	(0.00809)	(0.00641)	(0.00550)
Base × Low education	0.00283	0.00140	0.00177
	(0.00320)	(0.00407)	(0.00266)
Panel B: Effect of staff by parental education and district size			
High education, smallest quartile districts	0.000856	-0.00437	0.000149
(Base category)	(0.00611)	(0.00396)	(0.00306)
Base × 2nd quartile	0.0193*	0.0194**	0.000879
	(0.00733)	(0.00683)	(0.00237)
Base × 3rd quartile	-0.00381	0.0319**	-0.00476
	(0.00659)	(0.0117)	(0.00585)
Base × Largest quartile	-0.0217	-0.0290	-0.0315
	(0.0145)	(0.0433)	(0.0180)
Low education × Base	0.00508	0.00163	0.00292
	(0.00361)	(0.00466)	(0.00287)
Low education × 2nd quartile × Base	-0.0142*	-0.0129*	0.0000560
	(0.00602)	(0.00559)	(0.00302)
Low education × 3rd quartile × Base	-0.00257	-0.0223	0.00373
	(0.00672)	(0.0152)	(0.00684)
Low education × Largest quartile × Base	-0.0107	-0.00353	-0.000590
	(0.00736)	(0.0177)	(0.00965)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. Low education is short for no parent who is college graduates, high education is at least one parent with college education. In Panel A, base category are students who have no parents with a college education, panel B base category are students who have no parents with a college education in the smallest district quartile. Alternative groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, salary controls, as well as a subject fixed effects and individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 5: Effect of Staffing Change on Math and Reading Achievement of white, Black and Asian students, by Enrollment Quartile, 2003-2022.

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of staff by race			
White (base category)	0.00123	0.00304	-0.00258
	(0.00528)	(0.00633)	(0.00409)
Black × Base	-0.00545	0.00515	-0.00564
	(0.00776)	(0.00991)	(0.00732)
Asian × Base	-0.000758	0.00237	0.00707
	(0.0131)	(0.0131)	(0.00967)
Panel B: Effect of staff by race and district size			
White, smallest quartile districts (Base category)	0.00486	0.000501	0.00151
	(0.00450)	(0.00335)	(0.00273)
Base × 2nd quartile	0.00741	0.00633	0.00205
	(0.00503)	(0.00510)	(0.00236)
Base × 3rd quartile	-0.00178	0.0160*	-0.00191
	(0.00533)	(0.00760)	(0.00484)
Base × Largest quartile	-0.00907	-0.00457	-0.0221*
	(0.00884)	(0.0278)	(0.00841)
Black × Base	-0.000839	-0.00290	0.00455
	(0.00626)	(0.00574)	(0.00431)
Black × 2nd quartile × Base	0.0125	0.0101	-0.00525
	(0.00804)	(0.0124)	(0.00696)
Black × 3rd quartile × Base	-0.00223	-0.00353	-0.0109
	(0.0111)	(0.0161)	(0.00650)
Black × Largest quartile × Base	-0.0214*	-0.0487	-0.0120
	(0.0101)	(0.0394)	(0.0133)
Asian × Base	0.0248*	0.00264	0.00800
	(0.0114)	(0.0108)	(0.00756)
Asian × 2nd quartile × Base	-0.00817	0.0226	-0.0113
	(0.0137)	(0.0172)	(0.0138)
Asian × 3rd quartile × Base	-0.0338	0.00497	0.00628
	(0.0215)	(0.0218)	(0.0100)
Asian × Largest quartile × Base	-0.00910	-0.0264	-0.00411
	(0.0260)	(0.0356)	(0.0253)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. In Panel A, base category are white students. In Panel B, Base category are white

Table 6: Effect of Staffing Change on Math and Reading Achievement of Hispanic and non-Hispanic Students, by Enrollment Quartile, 2003-2022.

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of staff by Hispanic status			
Non-Hispanic (base category)	-0.00544	0.000522	-0.00698
	(0.00819)	(0.00765)	(0.00578)
Hispanic × Base	0.0255*	0.0135	0.0140
	(0.0107)	(0.0194)	(0.0107)
Panel B: Effect of staff by Hispanic status and district size			
Non-Hispanic, smallest quartile districts	0.00633	-0.000978	0.000848
	(0.00417)	(0.00354)	(0.00277)
Base × 2nd quartile	0.00918*	0.00942	-0.00102
	(0.00377)	(0.00600)	(0.00261)
Base × 3rd quartile	-0.0145*	0.00329	-0.00577
	(0.00554)	(0.00581)	(0.00481)
Base × Largest district quartile	-0.0332**	-0.0245	-0.0284
	(0.0105)	(0.0483)	(0.0194)
Hispanic × Base	-0.0107*	-0.0105	0.0105*
	(0.00521)	(0.00591)	(0.00515)
Hispanic × 2nd quartile × Base	0.0140	0.0142	0.0170
	(0.0104)	(0.0185)	(0.00962)
Hispanic × 3rd quartile × Base	0.0552***	0.0651***	0.00706
	(0.00988)	(0.0146)	(0.0118)
Hispanic × Largest district quartile × Base	0.0485	-0.0250	-0.0246
	(0.0246)	(0.0568)	(0.0209)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. In Panel A, base category are non-Hispanic students. In Panel B, Base category are non-Hispanic students in the smallest district quartile by student size. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, salary controls (salary spending per number of staff members), as well as a subject fixed effects and individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 7: Effect of Staffing Change on Math and Reading Achievement, by Enrollment Quartile, 2003-2022, by Interaction between Growth and Size. (With Salary Controls.)

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of staff by district growth			
No growth	-0.00822	-0.000791	-0.00683
(Base category)	(0.00985)	(0.00671)	(0.00635)
Base × Growing	0.0125	0.00809	0.00547
	(0.00677)	(0.00428)	(0.00511)
Panel B: Effect of staff by district growth and district size			
No growth, smallest district	0.00928	-0.00202	0.00210
quartile (Base category)	(0.00547)	(0.00370)	(0.00275)
Base × 2nd quartile	0.00964	0.00882	0.000540
	(0.00542)	(0.00578)	(0.00318)
Base × 3rd quartile	-0.0139	0.0151*	-0.00266
	(0.00729)	(0.00724)	(0.00469)
Base × Largest	-0.0410***	-0.0702	-0.0467**
quartile	(0.00999)	(0.0543)	(0.0150)
Growing × Base	-0.0121*	0.00525	0.00203
	(0.00498)	(0.00441)	(0.00436)
Growing × 2nd quartile	0.00547	0.000803	0.0000403
× Base	(0.00625)	(0.00804)	(0.00844)
Growing × 3rd quartile	0.0207*	0.000188	-0.00772
× Base	(0.00879)	(0.00816)	(0.0102)
Growing × Largest	0.0475***	0.0732	0.0403*
quartile × Base	(0.0109)	(0.0486)	(0.0178)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. Growing districts are those that have grown by at leat 10% over the previous 10 years. Base category are students in the smallest district quartile by student size and no growth districts. Other groups are interacted with the base category. Models are weighted and include a growing indicator, growing indicator interacted with each quartile, district and state-by-year fixed effects, as well as a subject fixed effects, salaries (salary spending per number of staff members), and individual student controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 8: Effect of Salary Changes on Math and Reading Achievement, by Enrollment Quartile (controlling for staff numbers), 2003-2022.

	Teachers	Paraprofessionals	School staff
Average effects	0.000443**	0.000166	0.0000774
of Salaries	(0.000165)	(0.000288)	(0.0000857)
Smallest quartile	-0.000758	-0.000648	-0.000331
(Base category)	(0.000576)	(0.000591)	(0.000195)
Base × 2nd quartile	0.000795	0.0973	0.000457*
	(0.000463)	(0.000743)	(0.000187)
Base × 3rd quartile	0.00130*	0.00106	0.000548*
	(0.000597)	(0.000668)	(0.000230)
Base × Largest	0.00189	0.00201	0.00135***
quartile	(0.00114)	(0.00125)	(0.000315)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of additional salary (in thousands), in separate regressions per staff category. Base category is the smallest district quartile by student size. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, as well as a subject fixed effect, staff numbers per 100 students and individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Additional Tables: robustness checks

Table A.1: Effect of Staffing Changes on Math and Reading Achievement, by Enrollment Quartile, 2003-2022.

	Teachers	Paraprofessionals	School staff
Average effects	-0.00204	0.00283	-0.00464
	(0.00826)	(0.00612)	(0.00572)
Personnel effects:			
Smallest quartile	0.00473	-0.00144	0.00178
(Base category)	(0.00439)	(0.00345)	(0.00266)
Base × 2nd quartile	0.0101*	0.0109*	0.000789
	(0.00429)	(0.00508)	(0.00262)
Base × 3rd quartile	-0.00600	0.0171**	-0.00467
	(0.00479)	(0.00616)	(0.00583)
Base × Largest	-0.0270*	-0.0322	-0.0316
quartile	(0.0122)	(0.0418)	(0.0168)
Control variables:			
Salary (average)	0.000470**	0.000573***	0.0000982
	(0.000147)	(0.000130)	(0.0000772)
Q2 size	-0.0610	-0.00299	0.00650
	(0.0337)	(0.0171)	(0.0179)
Q3 size	0.00877	-0.0299	0.0171
	(0.0336)	(0.0231)	(0.0215)
Largest q. size	0.148	0.0382	0.101
	(0.0885)	(0.0536)	(0.0522)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. Analogous to Table 2, but excluding growing district indicator as a control. Shows average effect (top) and effects by quartile (bottom). Bottom is base category is the smallest district quartile by student size. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, as well as a subject fixed effect and salary averages (thousands) individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.2: Effect of Staffing Change on Math and Reading Achievement by English language learner (ELL) categorization and by Enrollment Quartile, 2003-2022.

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of staff by English learner status			
Not English learner	-0.00312	0.000928	-0.00509 -0.00316
(Base category)	(0.00823)	(0.00601)	(0.00498)
Base × ELL	0.0258*	0.0319	0.00706
	(0.00377)	(0.0169)	(0.0155)
Panel B: Effect of staff by English learner status and district size			
Non-ELL, smallest quartile	0.00451	-0.00177	0.00153
(Base category)	(0.00461)	(0.00346)	(0.00263)
Base × 2nd quartile	0.0106*	0.0118*	0.0000650
	(0.00428)	(0.00513)	(0.00267)
Base × 3rd quartile	-0.00781	0.0151*	-0.00554
	(0.00465)	(0.00591)	(0.00496)
Base × Largest quartile	-0.0276*	-0.0329	-0.0307
	(0.0121)	(0.0441)	(0.0175)
ELL × Base	0.00495	0.000422	0.0123
	(0.0139)	(0.0206)	(0.0128)
ELL × 2nd quartile × Base	0.0306	-0.0227	0.0188
	(0.0166)	(0.0146)	(0.0102)
ELL × 3rd quartile × Base	0.0306	0.0257	0.00955
	(0.0166)	(0.0296)	(0.0164)
ELL × Largest quartile × Base	0.0195	0.0211	-0.0260
	(0.0203)	(0.0603)	(0.0163)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. Base category are students not classified as English learners in the smallest district quartiles. Alternative group is interacted with the base category. Models are weighted and include district and state-by-year fixed effects, as well as a subject fixed effect, salaries (salary spending per number of staff members), and individual student controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.3: Effect of District Administrative Staffing Change on Math and Reading Achievement, by Enrollment Quartile and controlling for Growth, 2003-2022.

Average effects	0.0102		
	(0.0163)		
Personnel effects by quartile:			
Smallest district quartile	0.0145	-0.00144	0.00178
(Base category)	(0.0168)	(0.0168)	(0.00266)
Base × Q2 size	-0.000729	-0.000729	0.000800
	(0.0291)	(0.0291)	(0.00261)
Base × Q3 size	-0.0130	-0.0130	-0.00466
	(0.0265)	(0.0265)	(0.00479)
Base × Largest district quartile	-0.0559	-0.0321	-0.0316
	(0.0643)	(0.0422)	(0.0168)
Personnel effects by income:			
High income		0.00672	
(Base category)		(0.0164)	
Base × Low income		0.0140	
		(0.0200)	
Personnel effects by parental education:			
High education			0.00694
(Base category)			(0.0155)
Base × Low education			0.0159
			(0.0120)
Observations	2,106,953	2106388	1,937,741

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. For reference, shows the coefficients for control variables. Base category is the smallest district quartile by student size. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, salary controls, as well as a subject fixed effects and individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.4: Effect of Salary difference on Math and Reading Achievement by household income status (free or reduced lunch eligibility) and by Enrollment quartile, 2003-2022.

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of salary by household income			
High income	0.000635*	0.000393	0.000323**
(Base category)	(0.000295)	(0.000403)	(0.000116)
Base × Low income	-0.000347	-0.000422	-0.000432**
	(0.000441)	(0.000568)	(0.000127)
Panel B: Effect of salary by household income and district size			
High income , smallest quartile	-0.000925	-0.000906	-0.000231
(Base category)	(0.000476)	(0.000476)	(0.000195)
Base × 2nd quartile	0.000801	-2.61e-05	0.000320
	(0.000403)	(0.000718)	(0.000203)
Base × 3rd quartile	0.00157**	0.00134*	0.000655*
	(0.000580)	(0.000640)	(0.000285)
Base × Largest quartile	0.00210	0.00257	0.00200***
	(0.00161)	(0.00190)	(0.000536)
Low income × Base	-0.000192	-0.000127	-0.000234
	(0.000507)	(0.000545)	(0.000246)
Low income × 2nd quartile × Base	0.000644	0.000926*	0.000323
	(0.000408)	(0.000417)	(0.000213)
Low income × 3rd quartile × Base	-0.0212**	0.000336	-0.000114
	(0.000371)	(0.000445)	(0.000272)
Low income × Largest quartile × Base	-0.0000292	-0.000462	-0.00126
	(0.000991)	(0.00139)	(0.000668)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of additional salary, in separate regressions per staff category, controlling for staffing levels. In Panel A base category are students not receiving free or reduced lunches. In panel B, base category are students not receiving free or reduced lunches (High household income) in the smallest district quartiles. Alternative groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, as well as a subject fixed effect, salaries (salary spending per number of staff members), and individual student controls. Clustered standard errors, by state, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Effect of Additional Salary on Math and Reading Achievement by parental education and enrollment quartile, 2003-2022.

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of salary by parental education level			
High education	0.00103***	0.000940**	0.000245
(Base category)	(0.000254)	(0.000317)	(0.000131)
Base × Low education	-0.00103***	-0.00131***	-0.000392**
	(0.000265)	(0.000361)	(0.000119)
Panel B: Effect of salary by income and district size			
High education, smallest quartile districts	-0.000191	0.000284	-0.000288
(Base category)	(0.000679)	(0.000639)	(0.000258)
Base × 2nd quartile	0.000340	-0.000670	0.000427
	(0.000719)	(0.00104)	(0.000245)
Base × 3rd quartile	0.000983	0.000388	0.000641*
	(0.000731)	(0.000790)	(0.000315)
Base × Largest quartile	0.00253	0.00303*	0.00182***
	(0.00129)	(0.00147)	(0.000467)
Low education × Base	-0.00144***	-0.00198***	-0.000277
	(0.000250)	(0.000350)	(0.000178)
Low education × 2nd quartile × Base	0.00116*	0.00170*	0.000106
	(0.000508)	(0.000652)	(0.000212)
Low education × 3rd quartile × Base	0.00106**	0.00165**	0.0000318
	(0.000332)	(0.000488)	(0.00684)
Low education × Largest quartile × Base	-0.0000734	-0.000355	-0.000590
	(0.000545)	(0.000758)	(0.000289)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of additional salary, in separate regressions per staff category. Low education is short for no parents who are college graduates, high education is at least one parent with college education. In Panel A, base category are students who have no parents with a college education, panel B base category are students who have no parents with a college education in the smallest district quartile. Alternative groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, salary controls, as well as a subject fixed effects and individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Effect of Salary on Math and Reading Achievement of white, Black and Asian students, by Enrollment Quartile, 2003-2022.

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of salary by race			
White (base category)	0.000281	0.0287	0.0850
	(0.177)	(0.000291)	(0.0848)
Black × Base	0.000685	0.000747	-0.000149
	(0.000576)	(0.000702)	(0.000227)
Asian × Base	0.00149*	0.00215*	0.0243
	(0.000721)	(0.000885)	(0.000275)
Panel B: Effect of salary by race and district size			
White, smallest quartile districts (Base category)	-0.000430	-0.000724	-0.000282
	(0.000539)	(0.000535)	(0.000181)
Base × 2nd quartile	0.000534	0.000233	0.000388*
	(0.000393)	(0.000664)	(0.000181)
Base × 3rd quartile	0.000772	0.000975	0.000462*
	(0.000589)	(0.000642)	(0.000209)
Base × Largest quartile	0.00136	0.00162	0.00150**
	(0.00110)	(0.00117)	(0.000453)
Black × Base	0.00141	0.00167*	0.000278
	(0.000894)	(0.000829)	(0.000607)
Black × 2nd quartile × Base	-0.000365	-0.000185	-0.00000945
	(0.000949)	(0.000795)	(0.000411)
Black × 3rd quartile × Base	-0.000574	-0.000801	-0.000332
	(0.000927)	(0.00102)	(0.000531)
Black × Largest quartile × Base	-0.000722	-0.000642	-0.00124
	(0.00120)	(0.00144)	(0.000849)
Asian × Base	0.00226***	0.00340***	0.00123
	(0.000575)	(0.000931)	(0.000614)
Asian × 2nd quartile × Base	-0.00147	-0.00235*	-0.000878
	(0.000872)	(0.00103)	(0.000932)
Asian × 3rd quartile × Base	-0.000171	-0.000969	-0.000931
	(0.00134)	(0.00202)	(0.000556)
Asian × Largest quartile × Base	-0.00157	-0.00212	-0.00269
	(0.00154)	(0.00238)	(0.00179)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of additional salary, in separate regressions per staff category. In Panel A, base category are white students. In Panel B, Base category are white students in the smallest district quartile by student size. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, salary controls (salary spending per number of staff members), as well as a subject fixed effects and individual controls. Clustered

Table A.7: Effect of Salary on Math and Reading Achievement of Hispanic and non-Hispanic Students, by Enrollment Quartile, 2003-2022.

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of salary by Hispanic status			
Non-Hispanic (base category)	0.000972**	0.00101*	0.000162
	(0.000283)	(0.00765)	(0.000109)
Hispanic × Base	-0.00281**	-0.00375**	-0.000429*
	(0.000965)	(0.00127)	(0.000209)
Panel B: Effect of staff by Hispanic status and district size			
Non-Hispanic, smallest quartile districts	-0.000861	-0.000467	-0.000308
	(0.000564)	(0.000554)	(0.000197)
Base × Q2 size	0.000917	0.000552	0.000532**
	(0.000471)	(0.000656)	(0.000181)
Base × Q3 size	0.00193**	0.00166*	0.000611*
	(0.000685)	(0.000716)	(0.000279)
Base × Largest district quartile	0.00384***	0.00428***	0.00150***
	(0.000796)	(0.000890)	(0.000394)
Hispanic × Base	-0.00114	-0.00184	-0.000332
	(0.000922)	(0.00118)	(0.000262)
Hispanic × Q2 size × Base	-0.000458	-0.000348	-0.000266
	(0.00107)	(0.00127)	(0.000324)
Hispanic × Q3 size × Base	-0.00185	-0.00212	-0.000168
	(0.00127)	(0.00161)	(0.000427)
Hispanic × Largest district quartile × Base	-0.00428**	-0.00499**	-0.000512
	(0.00137)	(0.00178)	(0.00116)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. In Panel A, base category are non-Hispanic students. In Panel B, Base category are non-Hispanic students in the smallest district quartile by student size. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, salary controls (salary spending per number of staff members), as well as a subject fixed effects and individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Effect of Staffing Change on Math and Reading Achievement of pre- and post-2013, 2003-2022.

	Teachers	Paraprofessionals	School staff
Panel A: Average effect of staff by period			
Pre-2013 (base category)	-0.00775	0.00336	-0.0110
	(0.0106)	(0.00901)	(0.00794)
Post-2013 × Base	0.00683	-0.00934	0.00964
	(0.0107)	(0.00741)	(0.00574)
Panel B: Effect of staff by period and district size			
Pre-2013, smallest quartile districts (Base category)	0.00102	0.000501	-0.00212
	(0.00476)	(0.00409)	(0.00284)
Base × 2nd quartile	0.0178*	0.0189**	0.00366
	(0.00685)	(0.00602)	(0.00479)
Base × 3rd quartile	-0.00543	0.0133	-0.00457
	(0.00655)	(0.00882)	(0.00542)
Base × Largest quartile	-0.0334**	-0.0486	-0.0345
	(0.0107)	(0.0495)	(0.0192)
Post-2013 × Base	0.00287	-0.00232	0.00524
	(0.00441)	(0.00488)	(0.00323)
Post-2013 × 2nd quartile × Base	-0.00660	-0.00960	-0.00363
	(0.00697)	(0.00835)	(0.00521)
Post-2013 × 3rd quartile × Base	-0.00245	0.00107	-0.00367
	(0.0108)	(0.0116)	(0.00850)
Post-2013 × Largest quartile × Base	0.0152	-0.0143	0.0158
	(0.0134)	(0.0353)	(0.0151)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category. In Panel A, base category are pre-2013 observations. In Panel B, Base category are pre-2013 observations in the smallest district quartile by student size. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, salary controls (salary spending per number of staff members), as well as a subject fixed effects and individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.9: Effect of Staffing Changes on Math and Reading Achievement, by Region, 2003-2022.

	Teachers	Paraprofessionals	School staff
Midwest	0.00299	-0.00593	-0.000140
(Base category)	(0.00339)	(0.00864)	(0.00251)
Base × Northeast	0.00305	0.00107	-0.00458
	(0.00651)	(0.0112)	(0.00580)
Base × South	-0.00726	-0.00177	-0.0100
	(0.0119)	(0.0206)	(0.00892)
Base × Mountain	0.00554	0.00989	0.0205
	(0.00936)	(0.0134)	(0.0135)
Base × Pacific	-0.0452***	0.00887	(0.00552)
	(0.0122)	(0.0103)	(0.00949)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category using US Census region definitions and dividing "West" region into its two divisions (Pacific and Mountain). Midwest region is base category. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, as well as a subject fixed effect and salary averages (thousands) individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.10: Effect of Staffing Changes on Math and Reading Achievement, by contemporaneous collective bargaining for teachers, 2003-2022.

	Teachers	Paraprofessionals	School staff
Non-duty to bargain state	-0.00608	0.00843	-0.0115
(Base category)	(0.0127)	(0.0146)	(0.00976)
Base × Duty to bargain	0.00896	-0.00853	0.0123
state	(0.0131)	(0.0154)	(0.0102)
Observations	2,520,504	2,412,961	1,848,283

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category using an indicator of whether collective bargaining for teachers was active in the state at the time (based on Lovenheim and Willén (2019), updated with the 2011 weakening of collective bargaining in five states (García and Han, 2021)) . Democratic states are base category. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, as well as a subject fixed effect and salary averages (thousands) individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.11: Effect of Staffing Changes on Math and Reading Achievement, by Democratic and Republican state in 2000 presidential elections, 2003-2022.

	Teachers	Paraprofessionals	School staff
Democratic state	-0.00131	-0.00102	0.0000500
(Base category)	(0.00475)	(0.00643)	(0.00395)
Base × Republican	-0.00101	0.00738	-0.00660
state	(0.0120)	(0.0120)	(0.00873)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category using 2000 presidential election partisanship. Democratic states are base category. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, as well as a subject fixed effect and salary averages (thousands) individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.12: Summary statistics of selected variables

Summary of threshold levels for density quartiles and share of observations in NAEP analytical sample, 2003–2022, by District Enrollment

Variable	**Min. density**	**All Obs.**	**Smallest size Quartile**	**2nd size Quartile**	**3rd size Quartile**	**Largest size Quartile**
Lowest teacher dens. qtile.	0	0.238 (0.426)	0.138 (0.345)	0.241 (0.428)	0.351 (0.477)	0.200 (0.400)
2nd teacher dens. qtile.	5.5	0.178 (0.382)	0.166 (0.372)	0.202 (0.402)	0.213 (0.410)	0.145 (0.352)
3rd teacher dens. qtile.	6.3	0.295 (0.456)	0.275 (0.447)	0.352 (0.478)	0.320 (0.466)	0.255 (0.436)
Highest teacher dens. qtile.	7.5	0.289 (0.453)	0.420 (0.494)	0.205 (0.403)	0.115 (0.319)	0.401 (0.490)
Lowest paraprofessional dens. qtile.	0	0.179 (0.384)	0.206 (0.404)	0.179 (0.383)	0.231 (0.421)	0.133 (0.339)
2nd paraprofessional dens. qtile.	0.77	0.226 (0.418)	0.157 (0.364)	0.214 (0.410)	0.289 (0.454)	0.217 (0.412)
3rd paraprofessional dens. qtile.	1.25	0.228 (0.419)	0.186 (0.389)	0.250 (0.433)	0.263 (0.440)	0.208 (0.406)
Highest paraprofessional dens. qtile.	1.79	0.367 (0.482)	0.451 (0.498)	0.357 (0.479)	0.216 (0.412)	0.442 (0.497)
Lowest school staff dens. qtile.	0	0.121 (0.326)	0.149 (0.356)	0.148 (0.355)	0.136 (0.342)	0.084 (0.277)
2nd school staff dens. qtile.	1.83	0.101 (0.301)	0.104 (0.305)	0.127 (0.333)	0.123 (0.328)	0.070 (0.255)
3rd school staff dens. qtile.	2.50	0.193 (0.395)	0.182 (0.386)	0.226 (0.418)	0.202 (0.402)	0.174 (0.379)
Highest school staff dens. qtile.	3.55	0.585 (0.493)	0.564 (0.496)	0.499 (0.500)	0.539 (0.498)	0.673 (0.469)

*M*ean and standard deviations (in parentheses) for all observations and observations in each size quartile. Observations weighted by NAEP-provided analytical weights. Excludes 1,603 smallest 10% of districts (less than 83 students). Share students by quartile displayed are the share of students in each district quartile in 2003 in the NCES Common Core of Data 2003-22, which differs from the share in each of those quartiles in the NAEP analytical sample 2003-2022.

Table A.13: Effect of Staffing Changes on Math and Reading Achievement, by staff density in 2003, 2003-2022.

	Teachers	Paraprofessionals	School staff
Lowest staff density quartile	0.00336	0.0128	0.00467
(Base category)	(0.0101)	(0.00741)	(0.00397)
Base × 2nd staff	0.0375	-0.0136	-0.00332
density quartile	(0.0200)	(0.00791)	(0.00414)
Base × 3rd staff	0.00148	-0.0121	-0.00897
density quartile	(0.0119)	(0.00859)	(0.00494)
Base × Highest staff	-0.0262*	-0.0215*	-0.0183*
density quartile	(0.0125)	(0.00992)	(0.00829)
Observations	2,425,231	2,314,088	1,778,569

Notes: Shows effects of an additional staff member per 100 students, in separate regressions per staff category by quartile of staff total per 100 students in the district in 2003 (defined separately by regression). The first density quartile is base category. Other groups are interacted with the base category. Models are weighted and include district and state-by-year fixed effects, as well as a subject fixed effect and salary averages (thousands) individual controls. Clustered standard errors, by state, in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.