

School Personnel and Student Outcomes: The Role of Support Staff In North Carolina's Elementary and Middle Schools

Charles T. Clotfelter

Duke University, charles.clotfelter@duke.edu

Steven W. Hemelt

University of North Carolina at Chapel Hill, hemelt@email.unc.edu

Helen F. Ladd*

Duke University, hladd@duke.edu

February 2016

Abstract

This paper takes a new look at a perennial question in education: Do resources matter and, if so, which ones? In contrast to other studies, we focus directly on staffing levels rather than on funding levels or teacher quality. We do so in the context of variation in staffing levels over the period 2001-2013 in North Carolina's elementary and middle schools. Using variation in state allocations of staffing slots at the district level to generate plausibly causal estimates at the school level, we find that staffing in the form of regular teachers, teaching assistants, and health care providers have positive effects on student test scores and on some behavioral outcomes. Clearest and most consistent are the findings for teaching assistants. The evidence indicates that teaching assistants contribute to higher student test scores overall but especially so in schools serving disadvantaged students, and that they also lead to lower absentee rates in elementary schools. In addition, a larger number of regular teachers per student, which translates into smaller class sizes, generates effects similar in size to many other studies of class size effects. Although we find some positive effects of health providers on student test scores, data limitations keep us from making broad statements about that staffing category.

*Contact author; Adrienne Jones and Winnie Biwott provided excellent research assistance for this project. The authors thank the Center for the Analysis of Longitudinal Data in Education Research (CALDER) funded by the U.S. Department of Education (#R305C120008) for financial support.

This paper takes a new look at a perennial question in education: Do resources matter and, if so, which ones? Our perspective differs from that of most other studies in that we focus directly on staffing levels rather than on funding levels or teacher quality. We do so in the context of elementary and middle schools in North Carolina. The peculiarities of that state's approach to funding schools makes it possible for us to estimate plausibly causal estimates of the effects of teachers, teaching assistants, and health care providers on student outcomes.

The study builds most directly on two literatures related to school finance and policy. One is the extensive literature on whether money matters. In the context of much of that literature, especially the many early studies summarized by Hanushek (1986 and 1997), researchers explore whether spending on teachers in the form of higher salaries for experience, years of experience, master's degrees or National Board Certification, or on larger numbers of teachers in the form of smaller class sizes, pays off in the form of higher student outcomes, typically as measured by test scores. Based on Hanushek's conclusion of no clear and consistent effects of spending, for many policy makers, the standard mantra has been "money doesn't matter." That view has justified decisions either not to increase education spending or to implement accountability efforts designed to make schools use whatever resources they have, especially teachers, more productively. More recent work that makes use of funding changes induced by state school finance court cases challenges the Hanushek conclusion and indicates a more positive role for additional funding (Jackson et al, 2014). Moreover other recent studies confirm positive effects of school funding, especially in the lower grades and often for pupils in disadvantaged communities (e.g., Guryan 2001, Chadhary 2009, and Papke 2005).

A second smaller, but growing, body of literature focuses attention on expenditures for specific programs within schools, such as health centers or social workers, and often looks at a broader set of outcome measures, including graduation rates, teen pregnancy and grade retention (Carrell and Carrell 2006; Reback 2010a and Lovenheim et al. 2014).

A major challenge facing any study of school resources is reverse causation, namely that the amount of resources available to any specific school may well be driven by unmeasurable characteristics of the school. If schools with more able or motivated students have access to greater resources than those with less able or motivated students, for example, estimates of the impact of resources are likely to be upward biased. Conversely, if resources are allocated in a compensatory way, their estimated impacts may be downward biased. As we describe below, the North Carolina policy context allows us to overcome such endogeneity issues by using state-level changes in staffing policies over time at the district level as instruments for staffing at the school level.

Motivating our research are the significant reductions in state-funded staffing levels in North Carolina that took place during the Great Recession and that have continued to the present. The variation in staffing levels over the period 2001-2013 in three categories – regular teachers, teaching assistants and health care providers – is sufficiently large for us to estimate causal

effects of all three on school outcomes at the elementary and middle school level. In general we find stronger and more consistent effects of all three staffing categories in elementary schools than in middle schools. We find that all three staffing categories affect student achievement in elementary schools in predictable ways and in some cases affect other outcomes as well. Hence, cutbacks in resources of this type have clear adverse consequences for students.

Relevant literature on the key staffing categories

We highlight here the literature related to our three categories of primary interest: teachers, teaching assistants, and health providers. In addition, we refer briefly to other categories that we include as control variables in some of our models.

Teachers

As we explain further below, our measure of teachers is the number of regular teachers divided by the total number of students in each school. Although not specifically a measure of class size because we do not link teachers and students to particular classes, changes in this measure within a school over time are closely related to changes in class size, especially at the elementary school level. Hence the relevant literature for predicting the effects of such a change is the extensive class size literature.

At a theoretical level, class size often matters because of interactions among peers in the classroom. As spelled out in a well-known article by Edward Lazear (1999), education in the classroom is like a public good, but one in which negative externalities arise when the behavior of one or more students impede the learning of their classmates. This perspective predicts that lower class sizes may be most desirable in the lower grade levels and in schools serving disadvantaged children. The logic is that because younger children have shorter attention spans than older children the likelihood of disruption for any given class size is greater in the lower grades. An analogous argument applies for disadvantaged students who, because of their home situations, may come to class less ready to learn than their more advantaged counterparts.

The most well-known and credible evidence on class size comes from the Tennessee STAR class size reduction experiment in the 1980s. In this experiment, students in grades K-3 who were randomly assigned to small classes (13-17 students) were compared to those in regular classes (22-25 pupils) and also to those in regular classes with a teacher's assistant for children. Krueger's 1999 analysis of the experimental data, in which he corrects for non-random attrition, finds significant and sizable test score gains in the first year that pupils are in smaller classes and smaller but still positive gains in subsequent years as they progress through school. Notably the effects are larger for more disadvantaged students. In more recent studies that follow the experimental and control pupils into their 20's (Chetty et al., 2011; Dynarski et al., 2013) find that the pupils in the small classes were more likely to attend college and to have other positive outcomes.

Other evidence from quasi-experimental studies that use a variety of strategies to isolate the effects of class size generate more mixed results. One strategy pioneered by Hoxby (2000) relies on variation in class size driven by demographic changes that lead to differences in the size of student cohorts. Applying this method to elementary school pupils in Connecticut, she finds no effects of class size on student achievement. Using data for Texas, Rivkin et al. (2005) find small, but inconsistent, class size effects and Cho et al. (2012) find quite small effects for students in grades 3 and 5 in Minnesota. In contrast, some of the non U.S. studies find relatively large effects of class size reductions. A study using Israeli data by Angrist and Lavy (1999) makes use of rules relating to maximum class sizes and finds large effects of class size reductions on students in grades 4 and 5. Piketty (2004) and Bressoux et al. (2009) also find large class size effects in France, with larger effects for low-achieving students (see summary of evidence in Gibbons and McNally, 2013).

Emerging from this discussion is the prediction that reductions in the number of teachers per student – which in turn imply larger class sizes – are likely to have negative effects on student outcomes at the primary level. The prediction at the middle school level is far less clear. On the one hand middle schools students may have longer attention spans because they are older, and therefore be less likely to be disruptive, making larger class sizes more manageable. On the other hand, their vulnerability and hormonal changes may cause them to be more disruptive, which would call for smaller classes. In addition, the link between our measure of the total number of regular teachers in a school per student, and class size is less strong in middle schools. Changes in that measure at the middle school level need not lead to changes in the sizes of classes in math and English, which are the subjects in which they are tested. Instead it may show up in the form of changes in the number and variety of other course offerings in the school. Hence we expect to find weaker effects of teachers at the middle school level than at the elementary level.

Teaching assistants (TAs)

Nationally, teaching assistants account for close to 12 percent of the total elementary and secondary school labor force (*Occupational Employment Statistics*, 2014). The requirements to serve as a teaching assistant differ across states, and often, as is the case in North Carolina, also across districts and schools. Most districts in the nation historically required that TAs have at least a high school diploma. The 2002 federal No Child Left Behind Act (NCLB) raised the requirements for any TA working in a Title I school (a school with more than 35 percent of its students qualifying for federal Title I funding). Under that law, TAs in such schools are required to have either two years of higher education or an associate’s degree and to work with a “highly qualified teacher.” About 70 percent of North Carolina’s districts have implemented the higher standard for all its TAs, with the other 30 percent implementing it just for their Title 1 schools. North Carolina does not require TAs to have any form of state professional license.

TAs perform a variety of roles in classrooms. These include preparing classroom activities, working on instruction with individual and small groups, performing clerical tasks, managing student behavior, and helping to evaluate student work (Kerry, 2005, NCATA). The use of teaching assistants in any of these roles has the potential to free up time for teachers to focus on their main task of teaching and to facilitate better instructional differentiation within classrooms. The precise mix of activities engaged in by teaching assistants differs from classroom to classroom, depending on differences in teachers' ability to make good use of their assistants and in the skills of the teaching assistants themselves. The test-based accountability provisions of NCLB appear to have placed more pressure on teaching assistants to participate directly in instruction, for which they may or may not be qualified.

The research on teaching assistants is far more limited than that on teachers. One 2001 study has garnered attention because it uses data from the highly touted Tennessee STAR experiment referred to above. This study (Gerber et al, 2001) addressed three questions: 1) Does the presence of a full-time teacher aide in the classroom in grades K through 3 advance students' academic achievement? 2) If so, does the effect depend on how long the student attends class with an aide? 3) Do some functions of aids have greater impacts on student achievement than others? The findings were not positive. The authors found that teacher aides had little, if any, positive effects on students' academic achievement, with the one exception being students who were in classrooms with an aide for 2-3 years. Nor did the authors find that any specific type of TA activity had an effect on student achievement.

A thorough review of the literature as of 2009 provides a more mixed and nuanced picture (Alborz et al., 2009a and 2009b). This review was based on 35 high-quality studies, mainly from the U.S., England and Wales, most of which examined teaching assistant support in primary schools. Of the eight studies that examined actual or perceived impacts of TAs on literacy and language, seven "suggested that trained and supported teaching assistants working on a one-to-one basis or in a small group can help primary aged children with literacy and language problems to make significant gains in learning." (Alborz et al, p. 1). Only two studies looked at numeracy, yielding mixed results. Nine of the studies looked at the impact on teaching practices, such as the approaches that teachers took to organizing the classroom and facilitating learning. The impacts varied greatly. In some cases the presence of aides did not have much effect on teaching practices but in others, it actively facilitated student learning. Also there is some evidence that the presence of motivated support staff increases satisfaction, and reduced stress levels of teachers in mainstream classrooms (Alborz et al, technical report, p. 16). Overall, the authors conclude that, provided they are properly trained and supported, teaching assistants can have a positive impact on pupil progress.

None of this literature provides a particularly strong case for the efficacy of teaching assistants. The absence of recent studies and, in particular, studies that are grounded in the U.S. context of strong pressure on teachers to meet test-based accountability standards, represents a significant gap in knowledge. A primary aim of this paper is to start filling that gap by examining the

relationship between exogenous changes in the number of teaching assistants at the school level and changes in student achievement. The limited evidence we have suggests that the effects are likely to be larger in schools with larger proportions of disadvantaged students.

Health providers

In our empirical work, we define the category of health providers as those who are providing mental and allied health services, plus speech pathologists and audiologists. The three largest categories are the speech pathologists, psychologists, and social workers. We do not include guidance counselors in this category, a group that is categorized by North Carolina as providing school support services rather than health services. Nor do we include nurses, because local education agencies employ only about half of all school nurses in the state. Each of the categories we include requires professional qualifications. Psychologists require a degree in a field of school psychology and national certification by the national School Psychology Certification Board and social workers must have a degree in social work and be licensed by the NC Department of Public Instruction.

The state provides clear guidelines for the services to be provided by each group. The main task for all of them is to reduce the health-related barriers to learning that individual children bring to the classroom, by attending to their mental and physical health needs. In addition, these professionals are charged with addressing broader health issues such as the control of communicable disease, and the creation of positive classroom and school environments. Social workers, in particular, are expected to assess and evaluate students to inform the design of appropriate interventions; provide crisis intervention services related to family violence, substance abuse, and behavioral disorders; and engage in advocacy that seeks to ensure all students have equal access to education and services to enhance their local academic progress.

Although the link between the health of children and student achievement is widely recognized, the research on how health professionals in schools affect student achievement is limited. A major challenge for researchers is the difficulty of ruling out the reverse causation that arises, for example, if such personnel are assigned to schools judged to be most in need. Carrell and Carrell (2007) overcome this problem by using cross-semester variation in the counselor-student ratios in schools across a large Florida School district in which the exogenous variation arises because of the relatively random availability of graduate students from the University of Florida's counseling program for internships and practicums. Reback (2010a) overcomes it by using Alabama's discrete cutoff for funding additional half-time appointments. Both studies find positive effects of school counselors on student behaviors such as reduced disciplinary or weapons-related incidents, but neither finds any effects on student achievement, perhaps because of the temporary nature of the changes.

In a far more ambitious study Reback (2010b) uses national data from various state sources along with national survey data on students to explore the effects of school-site mental health

services for children in third grade. His cross-state descriptive analysis provides evidence that students in states with more aggressive mental health counseling policies perform better both in terms of higher achievement and fewer behavioral problems than those in other states with weaker policies, even after controlling for many student background characteristics, including their test scores in kindergarten. A second, more causal difference-in-difference approach based on a different national data set on school staffing (the SASS) finds that the availability of counseling services generates positive effects in that it reduces the fraction of teachers reporting that students are misbehaving or that their instruction suffers due to student misbehavior. This study can say nothing directly about student achievement.

Other researchers have examined the extent to which school-based health centers affect a variety of student outcomes, but these studies typically focus on high schools, not the elementary and middle schools of interest in the current study. One recent working paper (Lovenheim et al, 2015), for example, uses national district-level data to analyze the effects of opening school health centers and finds reductions in teen births but no effects on graduation rates. Once again the existing evidence is limited, but it is consistent with the compelling logic that children who are experiencing mental health problems or who are dealing with stress that arises from parental job loss or their home environment will find it hard to learn.

Other staffing categories

Data are also available on other staffing categories that could potentially generate higher levels of achievement or other student outcomes. These categories include guidance counselors, school leaders, and noncertified academic staff.¹ Guidance counselors are trained to incorporate the state's Guidance Standard Course of Study within the school setting through large and small group activities and to work with individuals planning for their human growth and development. With the increase in student assessment required under federal and state laws, it appears that many guidance counselors in North Carolina were spending significant amounts of time coordinating standardized testing within the school. New guidelines from the state in 2014 now make it very clear that they should be spending at least 80 percent of their time providing direct services to students in both large and small group settings and that those activities do not include test coordination. Although the presence of guidance counselors in a school may well affect student achievement, the link is likely to be more tenuous than with the other three categories of professional staff.

Similar arguments apply to the category of school leaders. While additional leaders in a school may well lead to better behaved students, the link with achievement is likely to be tenuous. Although staff in the category of non-certified academic support, which includes tutors, interpreters, therapists, and non-certified instructors, could potentially have a more direct impact on student achievement, we have chosen to treat that category the same way we treat guidance

¹ We excluded from all the analysis an additional category of non-regular teachers on the ground that they were working with special populations.

counselors and school leaders to retain the focus on teachers, teaching assistants, and health personnel. All of our achievement models control for an aggregate of these three categories, but we explore their separate effects in the models explaining behavioral outcomes.

Background and data on school staffing in North Carolina

Our focus on staffing levels rather than funding reflects the particular system of school funding in North Carolina. Starting in 1933, and reinforced by policy actions in 1975, the state government assumed responsibility for funding a “sound basic education” for all students. It has done so primarily by using formulas to allocate positions (not dollars) to local districts, such as slots for teachers, school leaders, and support personnel, largely based on student populations. The allocation formulas are detailed and clear. With the noteworthy exception of teaching assistants, the state’s support works as follows. For each district in a particular year, the state provides funds to cover X number of teaching positions per student in grades 2-3 and Y teaching positions per student in grades 4-5. For the major staffing categories, districts have incentives to use all the slots funded in this way and to fill them with the highest quality personnel they can. Higher salaried teachers put no cost burden on the local school or district, because the state pays teachers based on the statewide salary schedule based on the qualifications of the teachers the district succeeds in hiring. The one exception to this funding approach is teaching assistants (TAs). In this case, the state provides a given dollar amount of funding per student to each district with the number of FTE state-funded TA positions a district can hire limited by the state appropriation.

Although the state is technically responsible for funding education, over time the local districts have shouldered a growing share of funding, now around 25 percent. The local funds are raised through taxes at the county level, because the elected county commissioners have the authority to levy taxes for education, not the elected school board members. In most cases, the counties are coterminous with the school districts but, largely as a historical anomaly, a few counties contain more than one school district.² Most of the local taxes for education are used for two main purposes – salary supplements to add to the amounts set by the state salary schedule for teachers, and funding for facilities. Small amounts are also used to hire additional staff and to pay for supplies. Any locally funded positions are likely to vary across counties in line with local preferences for education and their ability to pay. In recent years, some districts have used fiscal balances that the county accumulated from federal ARRA funds after the 2008 recession to offset cuts in positions funded by the state.

We focus on three main categories of instructional positions: Regular teachers, teaching assistants, and health providers. For the reasons explained above, each of them has the potential to affect student outcomes. The outcome variables (measured and analyzed at the school level)

² There are 100 counties in North Carolina and 115 school districts.

include normalized math and reading scores for all students in grades 3-5 in elementary schools and grades 6-8 in middle schools; percentages of students meeting state-defined proficiency levels in math and reading, and behavioral outcomes such as absence rates and suspensions.

We use school level data from 2001-2013 for 1,153 elementary schools with a total school sample size of 11,544 and 439 middle schools with a total school sample of 4,700 for all our test score analysis.³ Data availability considerations limit our analysis of behavioral outcomes to the shorter 2006-2013 period. The samples include all schools in each category other than those with atypical grade configurations, such as grades 5-8 or grades k-8, which might require unusual staffing patterns. For each category of staffing we have the number of slots broken down by whether the slots are funded by the state, local, or federal levels. The following three graphs depict the number of positions in each category per 100 students across our full sample of elementary schools. The patterns over time for staffing at the middle school level are similar (not shown) but the ratios for teaching assistants and health providers students in middle schools are lower than those for elementary schools (see averages in Table 1 below).

(Figures 1, 2, and 3)

Without exception, the state government funds the largest share of the positions in each category. The changes over time in state-funded staffing levels in all three categories are central to our analytical strategy. For teachers and teaching assistants the recent declines largely reflect a combination of recession-related pressures on the state budget and subsequent policy decisions by a Republican legislature after 2010 committed to reducing the size of government. The large drop in health care providers between 2007 and 2008 is more difficult to explain. While the school-level data we have indicates a drop in slots during that year for each of the subcategories that make up the health provider category, statewide budget data for all North Carolina public schools do not show comparable drops. Careful checking of our school-level data generated no obvious explanation, although we cannot rule out the possibility that many health workers were moved into specific program areas that we have excluded from our analysis. For that reason, the reader should be warned that we have less confidence in our findings for health providers than for the other categories. Importantly, however, we report models below to convince the reader that any possible problems with the health provider data do not interfere with the basic conclusions we draw about the impacts of teachers and teaching assistants.

The state allocates slots (or in the case of teaching assistants a pot of money) to districts not to schools. Districts may respond to changes in state-funded slots by changing levels of locally funded positions or by reallocating slots among schools.⁴ Based on our discussions with

³ The samples are somewhat smaller for the models in which we used lagged staffing variables because we lose the 2001 year.

⁴ While districts are not allowed to transfer teaching slots to other categories, they do have some flexibility to shift funding allocated for teaching assistants to teachers. Any reallocations of that type are already embedded in our data. That is, our data for state funded slots for teachers include any slots funded by a shift of funding from another state allocation.

district officials, it appears that districts use various district-specific formulas or policies for allocating positions across schools. For example they might allocate slots on a straight per-pupil basis, or they might use various allocation schemes that target slots across schools to give disadvantaged schools more resources. We do not know much about the methods they use, but we do have evidence that at least in some cases their allocations have changed significantly over time. We depict such shifts using allocations across elementary schools for one illustrative district: Durham County, for the years 2004 and 2012. The graph shows that with the decline in teacher slots, the county has flattened the distribution of teaching slots relative to a school's share of disadvantaged students (as measured by the percent eligible for free or reduced price lunch on the horizontal axis). In addition, the district appears to have also modified somewhat the allocation of teaching assistant positions.

(Figure 4 goes here)

These district level allocation decisions are relevant to our modeling challenge to the extent that changes in a school's slots over time are associated with changes in characteristics of a school correlated with student outcomes. As we explain below, we eliminate this correlation by using state funded (and also federally funded slots) at the district level as instruments for school level changes.

Statistical model

Our goal is to leverage changes over time in staff positions to determine how staffing patterns affect student outcomes at the school level. Letting s denote the school and t denote the year, we write the basic model as:

$$\text{Outcome}_{st} = \alpha + \beta_1 \mathbf{Staff}_{st} + \beta_2 \mathbf{X}_{st} + \beta_3 \mathbf{Z}_{st} + \epsilon_s + \delta_t + \mu_{st}. \quad (1)$$

Where \mathbf{Staff}_{st} is a vector of staffing categories school s in year t ;

\mathbf{X}_{st} is a vector of teacher quality characteristics in school s (e.g., average licensure test scores of the teachers);

\mathbf{Z}_{st} is a vector of time-varying student characteristics in school s (e.g percent on FRPL or special needs);

$\epsilon_s + \alpha_t$ are school and year fixed effects; and

μ_{st} is a randomly distributed error term.

Our key variables – the staffing measures – are all expressed as full time equivalent (FTE) positions at the school level. We have explicitly excluded variables that would control for differences across districts or schools in average teacher salaries. Such salary differences would reflect both the mix of teachers within a school – with more experienced teachers paid more – and the willingness of the local country to supplement the pay of teachers over and above the

statewide salary schedule. While to some extent, higher teacher salaries may reflect quality differences in teachers, they may also reflect cost-of-living differences, or differences in the salaries needed to attract a given quality teacher to a particular part of the state. Hence, instead of controlling for teacher salaries, we estimate models that include more direct measures of teacher quality at the school level, including their average licensure test scores, the proportion of teachers with more than three years of experience, and the share who are National Board Certified (\mathbf{X}_{st}). Research studies based on North Carolina data confirm that each of these measures of teacher quality are predictive of higher test scores (Clotfelter et al., 2006 and 2007). We have no similar measures for the quality of other school personnel; we simply assume that the quality of teachers also serves as a proxy for the quality of other staff in the school.

The standard statistical challenge in models of this type is the potential endogeneity that arises if staffing variables are correlated with the error term in the models. For example, districts may allocate fewer staff positions to schools with high test scores. Alternatively, schools with high test scores may receive more staffing because they are located in wealthy counties that are willing to raise local tax dollars to provide them with higher amounts of staffing. In the first case, simple cross-sectional estimates of how staffing affects student outcomes would be downward biased and in the second they would be upward biased. We address this challenge in three ways.

Our main strategy for identifying the effects of staffing is to include school fixed effects in all models, which is feasible because we have a panel data set. This strategy means that we are estimating models that control statistically not just for measurable characteristics of each school but also unobservable characteristics that are time-invariant and that might be correlated with staffing levels. In effect, we are estimating the effects of within-school changes in staffing on within-school changes in outcomes. The staffing changes refer to all changes in a staffing category regardless of the funding source.

Although the use of school fixed effects helps because it breaks the link between the average characteristics of the schools and their staffing levels, it does not, rule out any time-varying correlation between school characteristics that could affect both achievement and the school's staffing levels. Because the school fixed effects control only for additive time-invariant characteristics of schools, we also need to account for the possibility that staffing in particular schools may be affected by time-varying characteristics of the schools, such as their proportions of low-income or minority students. That correlation would be present if district policymakers use school characteristics to allocate slots among the schools each year and those characteristics change differentially over time in individual schools. The most straightforward way to address this possibility is to include as explanatory variables a vector of school-level, time-varying characteristics such as the percent of low-income or African American students, both of which are likely to affect outcomes such as test scores and potentially to be correlated with staffing. Thus in all our models, we include a vector of time-varying school-level demographic variables (\mathbf{Z}_{st}).

While the combination of school fixed effects and time-varying covariates address many of the statistical problems related to identification, they do not completely rule out the potential for district-level decisions to bias the estimated effects on school staffing. This problem arises because of the power that district administrators possess to influence the staffing positions at each school. One way they can do so is by getting their local counties to appropriate funding for positions over and above the slots funded by the state or federal governments. They also have the authority to alter how they distribute the state-funded slots among schools, and in some cases to shift some slots among categories. Consider the possibility that local officials respond to a cutback in state funding for support staff either by redistributing the existing staff among schools or by changing or expanding their own funding but with a different allocation among the schools. To the extent that they pay attention to changes in student outcomes in making their allocation decisions (e.g. by providing more support staff to schools experiencing increases in absentee rates or declines in test scores), it would not be appropriate to view the change in staffing as exogenous to the individual schools. To avoid the possibility that our explanatory variables (changes in staffing) would be endogenous at the school level, we need to identify the effects of staffing by estimating models in which the variation in staffing reflect changes only at the state or federal level, which are exogenous to the district.

To that end we estimate 2SLS models in which we substitute predicted values for actual values of each of the staffing measures. This approach requires that we estimate first-stage regressions to predict levels of each of the staffing categories at the school level. Each first-stage equation includes as exogenous variables the state and federal allocations for all of the specific categories, aggregated to the district level (and normalized by the number of pupils), along with the other independent variables, including the school and year fixed effects. In this way, the variation in the staffing variables in the second stage of the model reflects only district-level variation not affected by local discretionary decisions.

We weight observations by the number of students in each school so that our estimated effects of staffing can be interpreted as average effects on students, not on schools. In addition, we cluster standard errors in our models at the school level to address the potential for student outcomes in a school to be correlated over time.

Data and Basic Results

Tables 1a and 1b provide descriptive statistics for all the variables in our test score analyses for elementary and middle schools, respectively. For elementary schools, the test scores of the students (normalized by school and year) refer only to students in grades 3, 4, and 5 because

those are the only students who are tested. At the middle school they include all the students who were tested in grades 6, 7 and 8. All the staffing variables, however, apply to the whole school.⁵

Consider first the data for the full sample of elementary schools in Table 1a. The mean teacher-to-student ratio is 0.052, that is, 5.2 teachers per 100 students. If each regular teacher had her own class, that would translate into an average class size of about 19 students, albeit most likely with smaller class sizes in the early grades and larger ones in the upper grades. The mean for teaching assistants is slightly more than half that at 2.8 per 100 students. Health providers are a much smaller category with only about 0.4 per 100 students, or about 1 provider per 250 students. The “other” combined category of guidance counselors, school leaders, and non-certified academic staff is somewhat larger with 0.7 per 100 students or about 1 per 140 students.

The student characteristics indicate that slightly less than half the students are eligible for free or reduced-price lunch (FRPL), 27 percent are black, 10 percent are Hispanic, and 10 percent have special needs. The teacher characteristics indicate that teachers have slightly above average licensure scores (i.e., Praxis scores), about 10 percent are national board (NBPTS) certified, and about 78 percent have more than three years of experience. The next two columns of Table 1 provide comparable information for schools with above average proportions of low-income students (indicated by high-FRPL) and those with below average proportions (indicated by low-FRPL).⁶ The staffing variables are generally comparable to the full sample, although the high FRPL schools have slightly higher ratios of teachers and teaching assistants, but lower quality teachers as measured by their lower average Praxis scores and lower percentages of Board Certified and experienced teachers. The final two columns show comparable information for minority students (defined as black, Hispanic and students classified in other non-white categories) and white students. The staffing variables are almost identical across these two groups and are similar to those for the full sample.

The middle school summary statistics in Table 1b are comparable, although the far lower ratios for teaching assistants and health providers are worthy of note. These lower ratios justify our greater attention to elementary than to middle schools in the following discussion.

Basic test score models

In this section, we provide results for three versions of the basic model that differ only in how they treat the potentially problematic health provider variable. All the results in Tables 2a (for elementary schools) and 2b (for middle schools) are based on the 2SLS models that we described above, all include the full set of control variables for time-varying school characteristics and

⁵ The average test scores for the full samples are not precisely zero because both samples exclude tested students in schools that offer non-standard sets of grades.

⁶ We group schools into these categories based on the average proportion of FRPL students in a school over the time period of our panel. We consider schools with average share of FRPL students above X as “high-FRPL” schools and those below Y as “low-FRPL” schools.

teacher quality as well as school and year fixed effects, and all are weighted by the number of students in a school. In addition, errors are clustered at the school level.

We present results separately for elementary and middle schools, using two specifications for each. In one specification (Panel A), the staffing variables are all in the current year and in the other they are lagged one year (Panel B) The lagged specification accounts for the possibility that when staffing levels change either by rising, and perhaps, even more so when they are falling, schools may take time to adjust. In the short run they may improvise by moving staff around and it is not until the following year that the effects of the changes are fully felt. We present both, but find that the lagged specifications generate the stronger and more consistent results.

Full results for the initial models in both unlagged and lagged form are reported in the appendix. As can be seen there, the various time-varying control variables describing student characteristics enter with predictable signs. For example, higher proportions of students eligible for free and reduced-price lunch or of minority students predict lower average test scores, and higher proportions of girls predicts higher test scores. Among the teacher qualifications variables, teacher experience plays the largest and more consistent role: an increase in the proportion of teachers with more than 3 years of experience predicts higher test scores.

The results for our initial specifications are presented in the first two columns of Tables 2a and 2b. Consider first the results for elementary schools in Table 2a. Across both panels, all 12 coefficients of the major staffing categories are positive and all but one (teachers per student in math) are statistically significant at least at the 10 percent level. Although the “other” category enters with some unexpected negative coefficients, the coefficients are not statistically significant. By controlling for this other category we avoid any bias that may arise from correlations between that category and any of our three categories of major interest.⁷

To determine the effect of one additional staff member in each category we need to multiply the reported coefficients by 0.01. For teachers an increase of that magnitude would mean going from 5.2 teachers per 100 students to 6.2 teachers per 100 students, which would be equivalent to a reduction in average class size from 19 to 16 students (bearing in mind, however, that some classrooms may initially be larger than others). Our estimates for regular teachers suggest that such a change would have an impact of about 0.010 to 0.018 standard deviations in reading and 0.009 to 0.020 in math, with the larger estimates coming from the lagged results in Panel B.

We find that teaching assistants have impacts that are smaller than those for teachers by about 20 percent in reading and by as much as 47 percent in math. The estimates indicate that one additional teaching assistant would increase reading scores from 0.008 to 0.013 standard

⁷ We have run some models in which we split the “other” category into its components of guidance counselors, school leaders and non-certified academic support but find some implausibly large negative coefficients that we believe reflect some unusual correlations in the data. Later in the paper, we present some disaggregated results for the non-test score outcomes based on a shorter sample.

deviations in reading and from 0.007 to 0.011 in math, with the larger effects appearing once again in the lagged models. In the concluding section of this paper we compare these magnitudes to those reported elsewhere in the literature and to the relative costs of hiring teachers and teaching assistants.

Although the much larger coefficients for health providers (4.0 to 5.4 or reading and 5.6 to 5.9 in math) may appear to suggest that the health personnel are the most productive, with one additional staff member per student leading to a 0.04 to 0.06 standard deviation predicted change in student achievement, that conclusion would be misleading. The reason is that such a change would be far outside the range of our data, and would only make sense if the relationship remained linear throughout. Hence, for some purposes, it may be useful to consider the relative effects of a common percentage change, say 10 percent, in each of the categories, starting from the average. That would require us to multiply coefficients by 0.005, 0.0028, and 0.0004 for teachers, teaching assistants and health personnel, respectively. The results of doing so are shown in square brackets under the standard errors for each estimate. For example, based on the lagged results in panel B, the effect of a 10 percent increase in teachers (in standard deviation units) would be 0.009 in reading and 0.010 in math, of a 10 percent increase in teaching assistants 0.004 in reading and 0.003 in math, and for health personnel, about 0.002 in both subjects. We note that a 10 percent increase in teachers from the average would reduce the class size from about 19 to about 17 students and that such a change would have more than twice as much impact on student test scores as a comparable percentage increase in teaching assistants and about five times that of a comparable increase in health providers.

Given our concerns about the health provider variables, we use columns 3-6 to show comparable results for two variants of the initial model. In the first variant, we simply eliminate the health provider variable completely and in the second we split the variable into two variables, one for the period 2001-2007 and the other for the period, 2008 to 2013. Eliminating the health provider variables leads to larger coefficients for all the teacher and teaching assistant variables, whether in current or lagged form. The second variant, in contrast, generates coefficients for those two staffing categories that are quite close in magnitude to those for the basic model, especially for the lagged staffing variables. We note in addition that the estimated coefficients for the health variable in the lagged version for the first period (about 7.5 for reading and 5.9 for math) are very comparable to our initial estimates and the coefficients (one of which is extremely large and negative) for the second period are not significant. This pattern for the health providers makes sense in that we observe more variation in health providers during the first period than during the more recent period. We put little stock in the coefficients for the later health provider variable given our concerns about the validity of our data for that period.

From this exercise we conclude that it is important to include health providers in the model and that the best way to do so is to include two separate variables. Hence we have labeled the version with the two health-provider variables as our preferred model and report results for that one in

most of the subsequent tables. In all cases, however, we find comparable results for teachers and teaching assistants if we estimate models using the simpler initial model.

The results for middle schools are reported in Table 2b. The most consistent finding is that teaching assistants have statistically significant positive effects in reading in all of the specifications, and also in math in the lagged versions of the model. To translate these estimates into the effects of 10 percent change, we need to multiply each by 0.0008 (which is 0.10 times the average number of TAs per student which is only 0.008 as shown in Table 1b). Once again, we show these percentage effects in square brackets under the standard errors.

In contrast to the results for teaching assistants, the coefficients on both the teacher variables and the health personnel variables at the middle school level are far more sensitive to the specification of the equation. Moreover, the statistically significant negative coefficients for some of the “Other” category estimates suggests that something is going on at the middle school level that we have not fully sorted out. Given this inconsistency in our subsequent analyses of the middle school patterns, we pay most attention to the category of teaching assistants.

Before moving to patterns by subsample and subgroup, we perform a falsification test for our preferred model at both the elementary and middle school levels (see table 3). This test substitutes the values of the staffing variables in the following year (year $t+1$) for the values in either the current or lagged year. If our models are correct, none of the staffing variables should have positive effects on student test scores. Looking first at the results for elementary schools, we find no evidence that any of the staffing variables (other than the problematic health variable in the later period) has a positive effect. Although we do not have a good explanation for the negative and statistically significant estimates for regular teachers, the absence of positive coefficients provides support for our approach. At the middle school level, the falsification test is somewhat less successful in that one of the key variables, teaching assistants for math, enters the equation with a positive and significant coefficient, a finding that requires that we be careful not to oversell conclusions about teaching assistants in math at that level.

Subsample and subgroup patterns

The literature we reviewed earlier suggested that smaller class sizes and teaching assistants may be more effective in schools serving higher proportions of disadvantaged students than in other schools. We look for differences by disadvantage first by comparing effects for schools serving high and low proportions of students eligible for free and reduced-price lunch, where schools are categorized into above average or below average proportions based on their percentages of such students averaged over the full period. In addition, we explore whether there are any differences by the race or ethnicity of the individual student by estimating separate models for two groups of students: minority students (i.e., black, Hispanic, and students of all other races or ethnicities) who are typically viewed as disadvantaged and white students, regardless of the mix of students

in particular schools.⁸ Table 4a reports the patterns from the basic model for elementary schools and Table 4b for middle schools. Once again, the staffing variables are unlagged in Panel A and lagged in Panel B.

For both elementary and middle schools, the most consistent findings emerge for teaching assistants, especially when the staffing variables are in lagged form (see panel B in both tables). Strikingly, for that specification, all the estimated effects for teaching assistants serving disadvantaged groups – whether that be the students in the high-FRPL schools or minority students in any school – are positive and statistically significant. Stated differently, at the current levels of teaching assistants in both elementary and middle schools, the evidence shows that additional teaching assistants for disadvantaged student groups contribute to higher test scores.

Further, the evidence suggests that teaching assistants in elementary schools also contribute to higher test scores for the more advantaged groups in reading, but generally not for math. Interestingly though, in every case, the estimated effects for the advantaged groups are smaller than the comparable effects for the less advantaged groups. Although large standard errors in many cases make us reluctant to conclude with confidence that the differences are statistically significant, the consistency of the general patterns suggests that TA's are more productive for the disadvantaged students than for their more advantaged counterparts.

The patterns across subgroups for the other two staffing variables are less consistent. Indeed, for teachers, many of the estimated effects in the unlagged version are not statistically significant. Based on the results in the lagged version (in panel B), it is hard to make the case that more teachers per student (and hence smaller classes) are more productive for the disadvantaged than for the advantaged groups. While there is some weak evidence for that for test scores in reading, the reverse is true for math. Indeed, contrary to our expectations, many of the estimated effects for math are larger for the low-FRPL than for the high-FRPL sample and are larger for the white group than for the minority group. Once again, though, large standard errors make it hard to pin down differences with confidence.

With respect to health providers at the elementary level, the only statistically significant coefficients emerge for reading (and once again more clearly and consistently for the lagged version in panel B), but no clear patterns emerge across the groups. At the middle school level, none of the health provider estimates are statistically significant

In sum, the clearest patterns by subsample and subgroup emerge for teaching assistants. For this category of staffing, the estimated effects are consistently (though not always statistically significantly) larger for teaching assistants serving disadvantaged students.

⁸ In the equations by racial group we did not include as control variables the time-varying racial mix of students in each school.

Academic proficiency results

We report one final set of test score results (see Table 5). In this case, the outcome measure is the proportion of tested students in each school who score at or above subject- and grade-specific proficiency levels. The averages for these composite proficiency rates are 74.2 percent in reading and 79.2 percent in math in elementary schools, and 73.9 and 75.3 percent, respectively in middle schools. In contrast to the other test score measures that incorporate all student test scores, this one focuses attention on students at the margin of success. One might expect that additional teaching assistants could play a key role in helping students close to passing to perform well enough on the test to reach the proficiency level. The reason is that teaching assistants – assuming they have sufficient training – can work with small groups of students who need additional help to get over the hurdle as the teacher continues to work with the bulk of the class or they can free up time for the teacher to work more intensively with such students.

Consistent with this prediction we find positive and statistically significant effects for all eight of the coefficients for teaching assistants in the table. The coefficients, which range from 0.41 to 0.77 in elementary schools, imply that one additional teaching assistant would increase the proportion of students passing by 0.4 to 0.8 percentage points. The coefficients are far larger at the middle school level although 1 more teaching assistant at that level is a far larger percentage change than in elementary schools. Nonetheless as shown by the figures in square brackets, a 10 percent increase in teaching assistants at the middle school level, starting from the small base of teaching assistants at that level, is predicted to have a larger impact on proficiency rates than a comparable percentage change from a larger base at the elementary level. These statistically significant findings simply reinforce our conclusion from other tables that teaching assistants matter for student achievement.

The clearest results for teachers – and hence for the effects of class size – emerge for the lagged specification at the elementary level (Panel B). The coefficients imply that one additional teacher per student raises the proportions passing by 0.9 percentage points in reading and 1.6 percentage points in math. This finding is consistent with the conclusion that smaller class sizes can generate higher passing rates at the elementary level. Not surprisingly given the range of classes that students in middle schools take, none of the teacher coefficients are significant at that level of schooling.

In terms of health providers, we find essentially no effects on average proficiency rates. This finding makes sense in that it would be far more difficult for health providers to target services to students at the margin of passing than would be the case for teaching assistants.

Impacts on student behavioral outcomes

In addition to affecting test scores, staffing levels may affect other student outcomes of policy interest. We are able to examine effects on three behavioral outcomes -- absences, tardies, and within-school suspensions – but only for the shorter period of 2006 to 2013.⁹ Prior research has shown that such measures may serve as reasonable proxies for what some people refer to as socio-emotional or “non-cognitive” skills that contribute to student flourishing. For example, a low absence or tardy rate may signify that a student is motivated to show up when expected. Moreover, attendance as early as grade six has been shown to predict the likelihood that a student will persevere and graduate from high school (Allensworth et al., 2007 and Balfanz et al., 2007). Finally, student suspensions are often indicative of student misbehaviors that may interfere with the learning of not only the misbehaving student but also the learning of others through peer effects.

One potential limitation of the data available for all three outcomes is that it is reported by the schools. That raises the possibility that different schools may have somewhat different reporting policies which could, potentially, bias estimates of the effects of staffing levels. The inclusion of school fixed effects helps somewhat. Nonetheless it could still be that the more staff there are in a school over time, the more likely they are to report higher levels of one or more of these outcomes regardless of the true levels. Given our prediction that higher staffing would reduce the true incidence of all three of these behaviors, this reporting bias would work against detecting effects of these staffing categories on our set of behavioral outcomes.

Table 6 shows the results. The absence rate is measured as the total number of absences divided by the total number of students in the school. The average rate in elementary schools is 4.75 days and in middle schools 5.86 days. Because of the non-normal distributions of the tardy rate and the in-school suspension rate, we treat the dependent variable in each case as a zero-one variable which takes on the value one if the school’s rate is above the 75th percentile rate for the schools in the sample (with the sample for the tardy rate being far smaller than that for the other two variables because of missing data).¹⁰ Thus, we are predicting the probability that staffing of different types affects the probability of a high tardy or of a high suspension rate.

As was the case for the test score results in prior tables, the clearest pattern emerges for the teaching assistants. In particular we find that teaching assistants reduce both absentee and tardy rates at the elementary level, and possibly also at the middle school level. Thus, we conclude that teaching assistants contribute in positive ways to student behavior, as well as to student test

⁹ We also have data on out-of-school suspensions but find they are generally not affected by the levels of staffing so we do not discuss them here.

¹⁰ The 75th percentile cut offs at the elementary level are 2.7 times for tardies and 0.32 days for suspensions and for middle schools 2.8 times and 2.4 days. We have experimented with different cut points but they do not affect the basic patterns shown in Table 6.

scores, especially in elementary schools. Indeed, the reduction in absentee and tardy rates could be one of the mechanisms through which TAs contribute to higher student test scores.

The pattern for regular teachers is harder to explain.¹¹ The presence of more regular teachers per student reduces the tardy rate and the in-school suspension rate as we would have predicted. Surprisingly, however, in elementary schools more teachers appear to increase the rate of student absenteeism, a finding for which we have no good explanation. With the health providers, the results jump around and most are not statistically significant.

We estimated separate coefficients for each of the three components of the “other category” with the expectation that one or more of these staffing categories might have an effect on student behaviors. The most interesting finding is the negative and statistically significant coefficient on school leaders in the school suspension model for elementary schools. This finding is indicative of their ability to reduce within-school suspension rates which would be consistent with the fact that assistant principals are often charged with enforcing school discipline policies. Finally, although we had expected the school guidance counselors to play a similar role through their efforts to promote a positive school climate, the estimated negative effects of school counselors on the probability of high in-school suspensions at both levels of schooling are not statistically significant.

Conclusions

It is difficult to determine how staffing levels within schools affect student outcomes such as test scores or rates of absenteeism given that school staffing levels are typically not exogenous to the school. The resulting challenge for empirical research is to rule out the possibility of reverse causation, namely that student outcomes may affect staffing levels which in turn would lead to biased estimates of staffing levels. The estimates would be upward biased if staffing levels were typically higher in schools serving more advantaged and higher-performing students and downward biased if staffing levels were higher in schools serving more disadvantaged, and lower-performing students.

One of the contributions of this study is our use of a new strategy for ruling out such bias-inducing reverse causation. In particular, the specific manner in which the state of North Carolina funds its schools allows us to instrument for changes in staffing at the school level with state-determined changes in slot allocations to districts. This two-stage approach combined with school fixed effects and time-varying demographic and teacher quality variables at the school level provides plausibly causal estimates of staffing levels, such as teachers, teaching assistants, and health providers on student outcomes.

¹¹ We note that none of the teacher quality variables such as years of experience enter any of these models with statistically significant effects. That contrasts with their role in the test score models.

In terms of substantive findings, the most important finding is the remarkably strong and consistent evidence on the role of teaching assistants, a staffing category that has been growing over time, but that has been woefully understudied. Positive effects of teaching assistants emerge regardless of the outcome measure and across most of the specifications that we present. Moreover, the evidence is consistent with the reasonable hypothesis that teaching assistants are more productive in terms of academic achievement in disadvantaged schools and for disadvantaged students than in other schools or for more advantaged groups.

Further findings regarding our teacher variable provide new support for the conclusion that class size matters in elementary school. Unfortunately, due to data limitations, our positive findings that relate health staff to measures of academic achievement and behavior should be interpreted as a first step and one that highlights the need for additional future research on the effects of a broader range of health care workers, including nurses, in schools.

We can put our estimated effects for the teacher variables into perspective by comparing the implied class size effects to other estimates of class size in the literature (see Gibbons and McNally, 2013). As we noted earlier, the most well-known study of class size is the Tennessee class size experiment, which reduced class sizes from about 22-25 students down to 13-17 students in grades K-3. A 2006 review by Diane Schanzenbach of the many studies of that experiment concluded that the test score impact was about 0.15 standard deviations. Given the average reduction was 8 students (from about 23 students per teacher down to about 15 students per teacher), the teacher-to-student ratio would have risen from 0.043 to about 0.066, or about 0.023 units in our metric. Applying a change of this magnitude to the coefficients in our basic model for elementary schools based on the lagged staffing variables, we find the following:

Coefficients		Estimated impact of a 0.023 change	
Reading	Math	Reading	Math
1.90	1.97	0.044 SD	0.045 SD

Although it appears that our estimated effects are quite small, a few points are worth noting. First, in comparing our estimates to the effects of the STAR class reductions, one should bear in mind that those reductions applied only to the early grades for which one might expect larger effects. Second, other studies using different methods such as Hoxby (2000) found no effects at all and another study using the same method she used (Cho et al., 2012) from Minnesota found effects that are comparable in size to our estimates (i.e., effects of 0.04 and 0.05 standard deviations for class size reductions of 10 students). Using the California class size reduction, Jepson and Rivkin (2009) find effects of 0.06 in reading and 0.10 in math. In light of this literature, we conclude that our estimates appear to be reasonable, albeit substantially smaller than effects that emerged from the Tennessee STAR context.

We have essentially no studies with which to compare the magnitudes of our estimated achievement effects of teaching assistants. The best we can do is to compare the productivity of teaching assistants to that of teachers. By comparing the estimated effects in the lagged form of our preferred model for elementary schools (see Table 2, panel B), we conclude that a teaching assistant is about 86 percent as productive as a regular teacher in reading and about 54 percent as productive in math. That would imply that teachers and teaching assistant would be equally good investments if the salary of a typical TA were somewhere between 53 and 86 percent of the salary of a regular teachers, provided that the goal is to raise student achievement.

We use these comparisons primarily to indicate that our preferred estimates of the effects of staffing categories are reasonable and to emphasize our conclusion that staffing levels matter. We remind the reader, however, that student achievement as measured by test scores in math and reading is not the only goal of schooling, and that teachers, teaching assistants, and health providers may also contribute to student outcomes in ways that we have not measured here. We have provided some hints of these other outcomes in our analysis of behavioral outcomes, but more research is needed to present a full picture of how school staff contribute to valued outcomes for students.

References

- Alborz, A., Pearson, D. Farrell, P. and Howes, A. 2009a. "The Impact of Adults Support Staff on Pupils and Mainstream Schools, A Systematic Review of the Evidence" A research brief from the EPPI-Center, Social Science Research Unit, Institute of Education. University of London. DCSF-EPPI—09-01.
- Alborz, A., Pearson, D. Farrell, P. and Howes, A. 2009b. *The Impact of Adults Support Staff on Pupils and Mainstream Schools. Technical Report*. Research Evidence in Education Library. London: EPPI-Center, Social Science Research Unit, Institute of Education, University of London. Report no. 1702T. <http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=w423SGddbGM%3d&tabid=2438&mid=4540>
- Allensworth, E.M. and J.Q. Easton. 2007. *What matters for staying on track and graduating in Chicago Public Schools*. Chicago, IL: Consortium on Chicago School Research at the University of Chicago.
- Angrist, J. and V. Lavy (1999), Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement, *Quarterly Journal of Economics*, 114 (2) 533-75.
- Balfanz, Robert, Lisa Herzog, and Douglas J. Mac Iver. 2007. Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist* 42 (4): 223-235.
- Bressoux, P., F. Kramarz and C. Prost (2009), Teachers' Training, Class Size and Students' Outcomes: Learning from Administrative Forecasting Mistakes, *Economic Journal*, 119: 540-61.
- Carrell S. and Carrell S. (2006). "Do lower student to counselor ratios reduce school disciplinary problems?" *Contributions to Economic Analysis & Policy*. 5(1).
- Chaudhary, L. (2009), Education Inputs, Student Performance and School Finance Reform in Michigan, *Economics of Education Review*, 28: 90-98.
- Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. Whitemore Schanzenbach, and D. Yagan (2011), How Does Your Kindergarten Classroom Effect Your Earnings? Evidence from Project STAR, *Quarterly Journal of Economics*, 126(4): 1593-660.
- Cho, H., P. Glewwe, and M. Whitley (2012), Do Reductions in Class Size Raise Student's Test Scores? Evidence from Population Variation in Minnesota's Elementary Schools, *Economics of Education Review*, 31: 77-95.
- Clotfelter, C.T., H F. Ladd and J L. Vigdor. 2007. "Teacher credentials and student achievement: Longitudinal analysis with student fixed effects" *Economics of Education Review*, Dec.

- Clotfelter, C.T., H F. Ladd and J L.Vigdor. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness (with Charles Clotfelter and Jacob Vigdor). *Journal of Human Resources*. vol. 41, number 4, Fall 2006, pp. 778-820. (Also available as NBER Working Paper 11936, January 2006; <<http://www.nber.org/papers/w11936>>).
- Dynarski, S., Hyman, J., & Schanzenbach, D. W. (2013). Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management*, 32(3), 392-417.
- Gerber, S. B., Finn, J. D., Achilles, C. M., & Boyd-Zaharias, J. (2001). Teacher aids and students' academic achievement. *Educational Evaluation and Policy Analysis*, 23(2), 123-143.
- Gibbons, S. B., & McNally, S. (2013). The effects of resources across school phases: A summary of recent evidence. Center for Economic Performance. DEP Discussion Paper, No. 1226. June.
- Guryan, J. (2001), Does Money Matter? Regression Discontinuity Estimates from Education Finance Reform in Massachusetts, NBER Working Paper 8269, May.
- Hanushek, E. A. (1986), The Economics of Schooling: Production and Efficiency in Public Schools, *Journal of Economic Literature*, 24(3): 1141-77.
- Hoxby, C. (2000), The Effects of Class Size on Student Achievement: New Evidence from Population Variation, *Quarterly Journal of Economics*, 115(4): 1239-85.
- Jepsen, C. and S. Rivkin (2009). Class Size Reductions and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size, *Journal of Human Resources*, 44(1), 223-50.
- Krueger, A. (1999), Experimental Estimates of Education Production Functions, *Quarterly Journal of Economics*, 114(2): 497-532
- Kerry, T. (2005). Towards a typology for conceptualizing the roles of teaching. *Educational Review*, 57(3), 20-30.
- Krueger, A., (1999) "Experimental Estimates of Education Production Functions." *Quarterly Journal of Economics*, vol. 114 , pp. 497–532.
- Lazear, E.P. (2001 Educational Production. *Quarterly Journal of Economics* 116 (3): 777-803.
- Papke, L. E. (2005), The Effects of Spending on Test Pass Rates: Evidence from Michigan, *Journal of Public Economics*, 8: 821-39.
- Lovenheim, M.F., Reback, R &Wedenoja,L. (2014). "How does access to Health Care Affect Helath and Education? Evidence from school-based health center openings?" Working paper, Cornell University, Ithaca, NY.

Piketty, T. and M. Valdenaire (2006), L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français: Estimations à partir du panel primaire 1997 et du panel secondaire 1995, Ministère de l'Éducation Nationale, Les Dossiers n173.

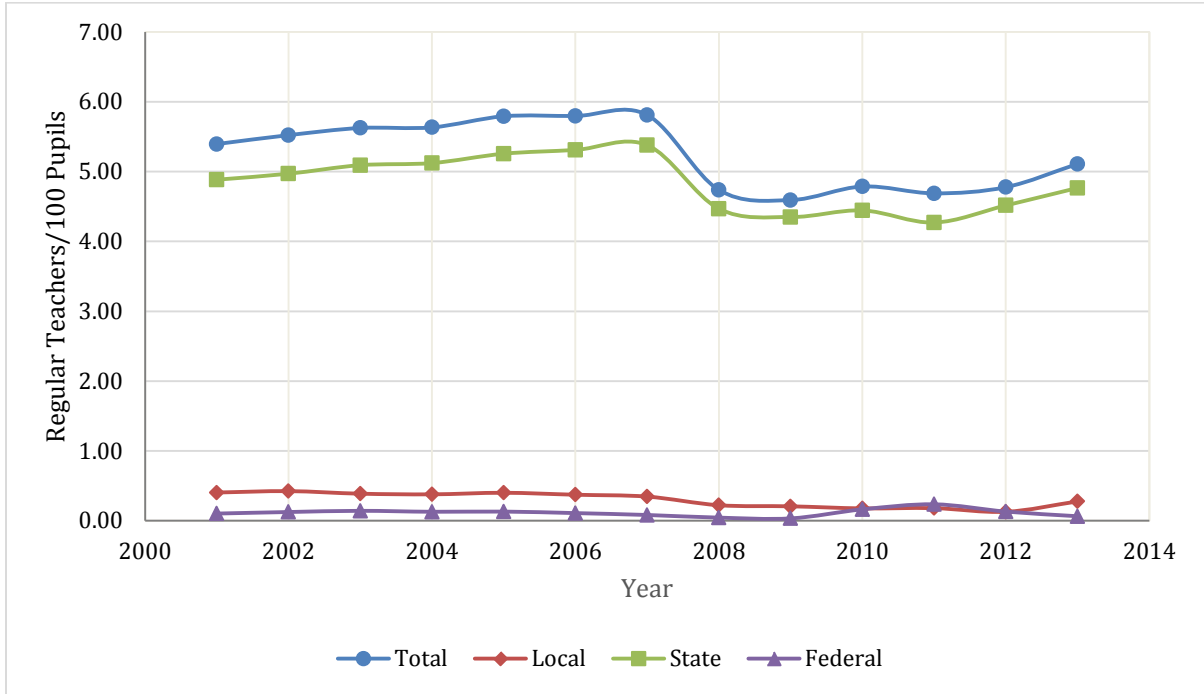
Reback, R. 2010. Schools' Mental Health Services and Young Children's Emotions, Behaviors and Learning. *Journal of Policy Analysis and Management*, 29(42001) : 698-725.

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

Rivkin, S., E. Hanushek and J. Kain (2005), Teachers, Schools, and Academic Achievement, *Econometrica*, 73 (2) 417-58.

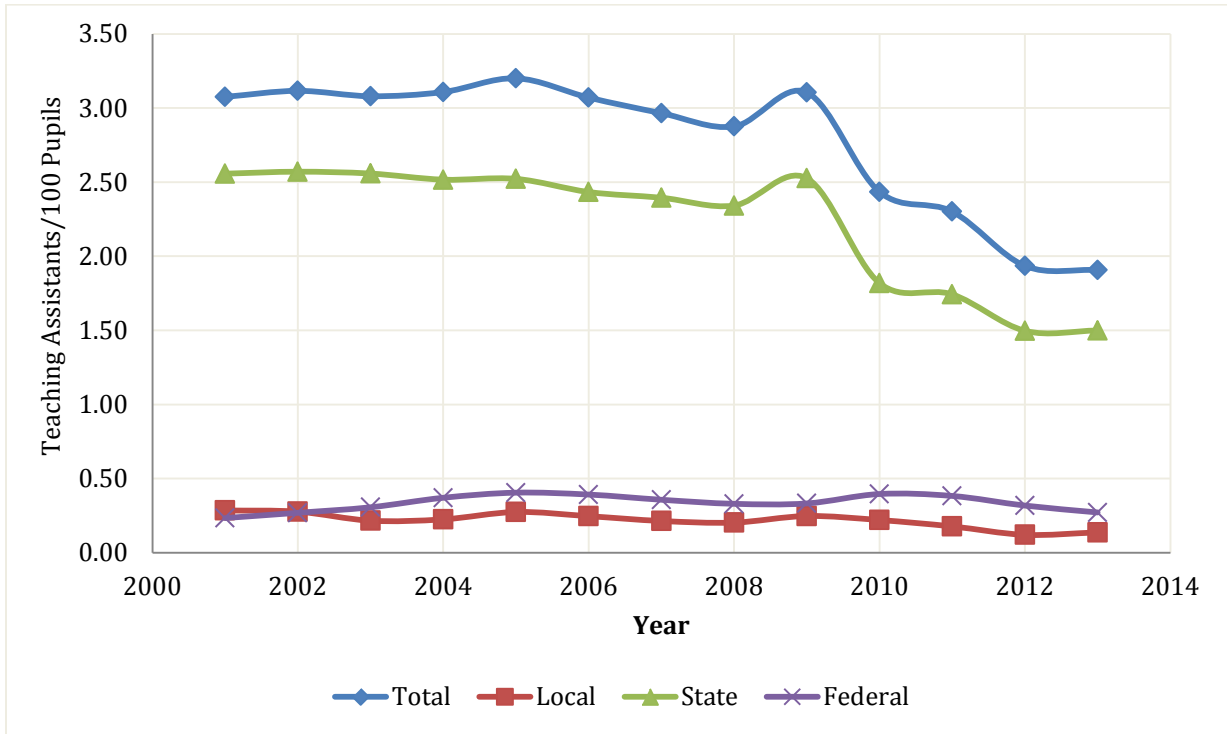
Schanzenbach, D. (2006) What have Researchers Learned from Project STAR? Bookings Paper on Education Policy: 205-28.

Figure 1. Regular Teachers Per 100 Pupils: NC Elementary Schools, 2001-2013



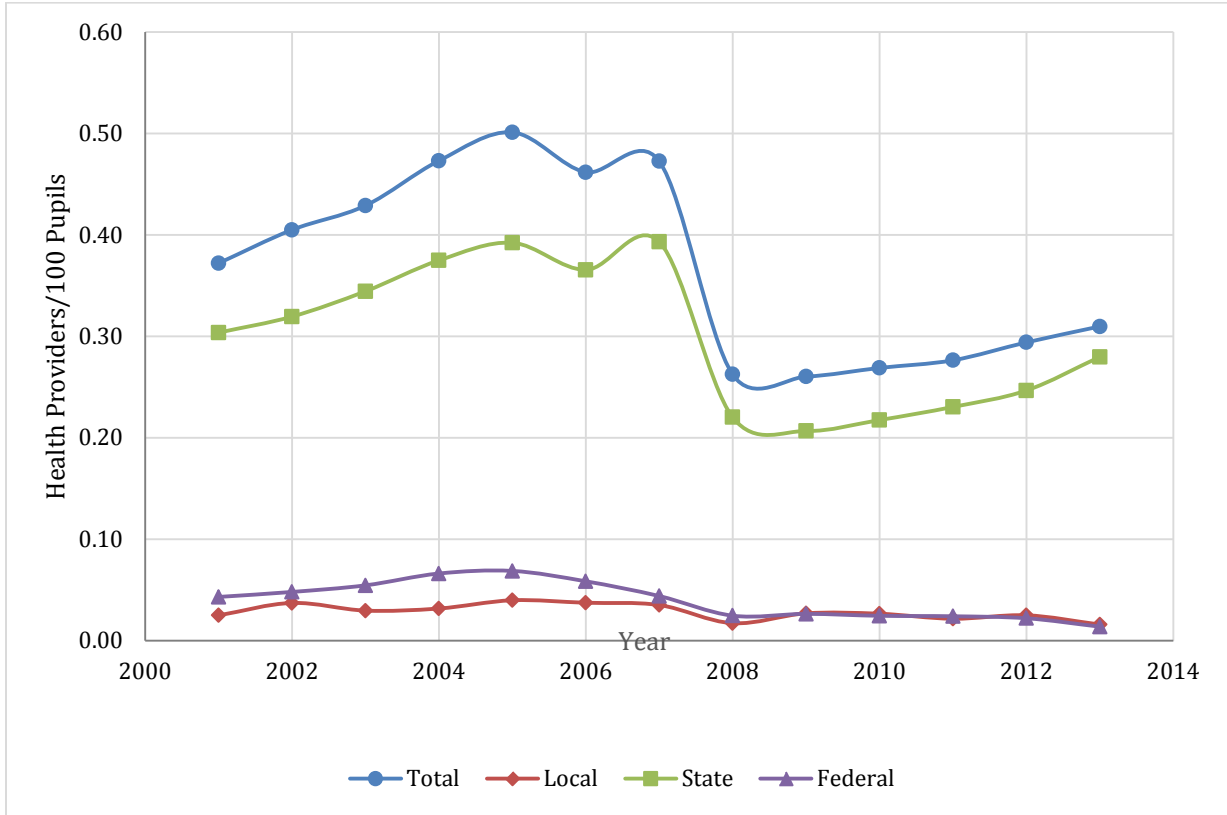
Source: Authors' calculations for all elementary schools in sample.

Figure 2. Teaching Assistants Per 100 Pupils: NC Elementary Schools, 2001-2013



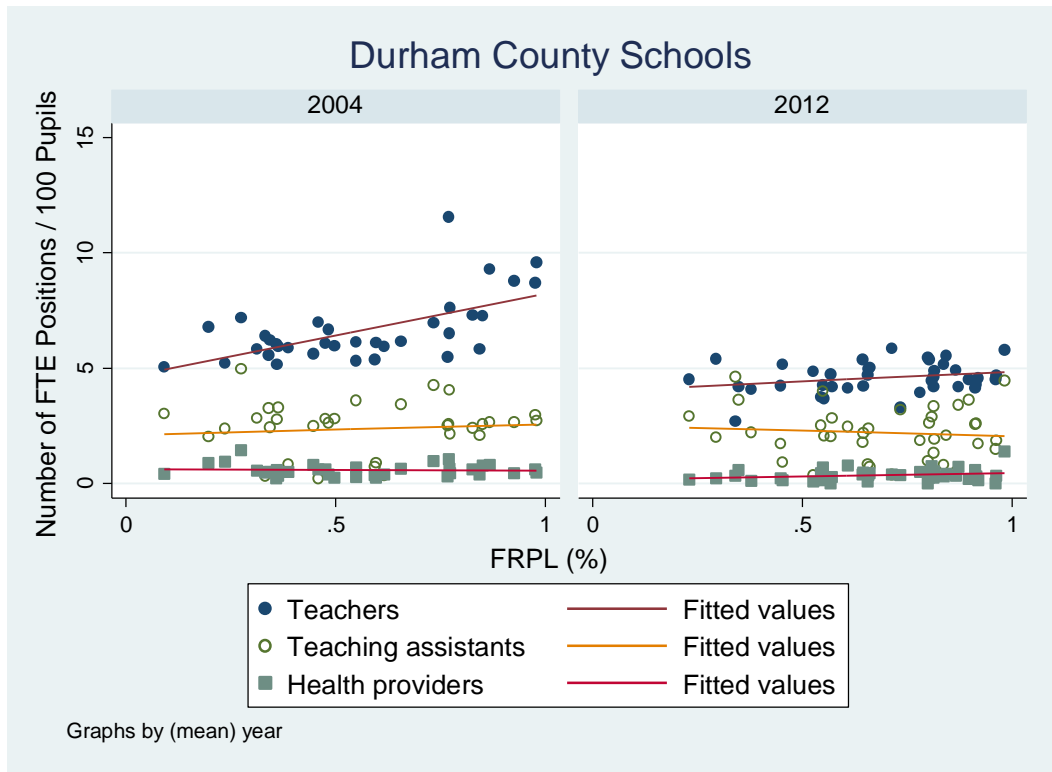
Source: Authors' calculations for all elementary schools in sample.

Figure 3. Health Providers Per 100 Pupils: NC Elementary Schools, 2001-2013



Source: Authors' calculations for all elementary schools in sample.

Figure 4. Example of Within-District Distribution of Staffing Positions



Notes: Graph only includes elementary schools; FRPL = free or reduced-price lunch.

Table 1a. Summary Statistics: Elementary Schools, 2001-2013

	Full Sample	High FRPL	Low FRPL	Minority Students	White Students
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
No. of schools	1,163	583	580	1,163	1,160
No. of observations	12,867	6,368	6,499	12,822	12,530
Test Scores					
Reading (SD)	0.020 (0.358)	-0.258 (0.265)	0.230 (0.264)	-0.256 (0.308)	0.256 (0.351)
Math (SD)	0.030 (0.380)	-0.250 (0.279)	0.235 (0.306)	-0.242 (0.333)	0.257 (0.378)
Staffing Categories					
Regular teachers/student	0.052 (0.009)	0.053 (0.011)	0.052 (0.008)	0.052 (0.009)	0.053 (0.009)
Teaching assistants/student	0.028 (0.014)	0.030 (0.016)	0.026 (0.013)	0.028 (0.014)	0.028 (0.014)
Health providers/student	0.004 (0.003)	0.004 (0.003)	0.004 (0.003)	0.004 (0.003)	0.004 (0.003)
Other staff/student	0.007 (0.004)	0.008 (0.005)	0.006 (0.003)	0.007 (0.004)	0.007 (0.004)
Regular teachers/student (lagged)	0.053 (0.009)	0.053 (0.011)	0.052 (0.008)	0.053 (0.009)	0.053 (0.009)
Teaching assistants/student (lagged)	0.029 (0.014)	0.031 (0.015)	0.027 (0.013)	0.029 (0.014)	0.029 (0.014)
Health providers/student (lagged)	0.004 (0.003)	0.004 (0.003)	0.004 (0.003)	0.004 (0.003)	0.004 (0.003)
Other staff/student (lagged)	0.007 (0.004)	0.008 (0.005)	0.006 (0.003)	0.007 (0.004)	0.007 (0.004)
Student Characteristics					
FRPL (%)	0.479 (0.233)	0.681 (0.164)	0.331 (0.150)	0.480 (0.233)	0.478 (0.231)
Black (%)	0.268 (0.235)	0.408 (0.258)	0.164 (0.148)	N.A.	N.A.
Hispanic (%)	0.101 (0.105)	0.144 (0.134)	0.070 (0.059)	N.A.	N.A.
White (%)	0.546 (0.282)	0.370 (0.268)	0.675 (0.215)	N.A.	N.A.
Other (%)	0.066 (0.065)	0.066 (0.079)	0.066 (0.052)	N.A.	N.A.
Female (%)	0.491 (0.041)	0.492 (0.042)	0.491 (0.040)	0.492 (0.034)	0.492 (0.034)
Special needs (%)	0.096 (0.053)	0.098 (0.054)	0.095 (0.053)	0.096 (0.053)	0.097 (0.054)
Teacher Characteristics					
Average Praxis score	0.091 (0.356)	-0.041 (0.382)	0.189 (0.301)	0.091 (0.356)	0.093 (0.353)
NBPTS certified (%)	0.099 (0.129)	0.073 (0.116)	0.118 (0.134)	0.099 (0.128)	0.098 (0.128)
>3 years of experience (%)	0.780 (0.180)	0.750 (0.194)	0.803 (0.166)	0.781 (0.180)	0.781 (0.180)
Number of Students					
In grades 3-5	290 (104)	248 (90)	321 (104)	290 (104)	289 (104)
In whole school	603 (207)	529 (181)	658 (207)	603 (206)	602 (205)

Notes: SD = Standard deviation, FRPL = Free or reduced-price lunch. Lagged staffing variables are based on a somewhat smaller sample because we lose the first year.

Table 1b. Summary Statistics: Middle Schools, 2001-2013

	Full Sample	High FRPL	Low FRPL	Minority Students	White Students
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
No. of schools	439	217	222	439	439
No. of observations	4,700	2,219	2,481	4,695	4,633
Test scores					
Reading (SD)	0.013 (0.320)	-0.244 (0.237)	0.187 (0.243)	-0.290 (0.266)	0.258 (0.302)
Math (SD)	0.014 (0.363)	-0.261 (0.252)	0.200 (0.303)	-0.267 (0.305)	0.241 (0.353)
Staffing Categories					
Regular teachers/student	0.052 (0.008)	0.053 (0.009)	0.051 (0.007)	0.052 (0.008)	0.052 (0.008)
Teaching assistants/student	0.007 (0.005)	0.008 (0.005)	0.007 (0.004)	0.007 (0.005)	0.007 (0.003)
Health providers/student	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)
Other staff/student	0.007 (0.003)	0.008 (0.004)	0.007 (0.003)	0.007 (0.003)	0.007 (0.003)
Regular teachers/student (lagged)	0.052 (0.008)	0.053 (0.009)	0.052 (0.007)	0.052 (0.008)	0.052 (0.008)
Teaching assistants/student (lagged)	0.008 (0.005)	0.009 (0.005)	0.007 (0.004)	0.008 (0.005)	0.008 (0.005)
Health providers/student (lagged)	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)
Other staff/student (lagged)	0.007 (0.003)	0.008 (0.004)	0.007 (0.003)	0.007 (0.003)	0.007 (0.003)
Student Characteristics					
FRPL (%)	0.451 (0.198)	0.615 (0.153)	0.339 (0.139)	0.451 (0.198)	0.451 (0.198)
Black (%)	0.276 (0.214)	0.412 (0.230)	0.184 (0.143)	N.A.	N.A.
Hispanic (%)	0.085 (0.077)	0.108 (0.101)	0.069 (0.050)	N.A.	N.A.
White (%)	0.563 (0.255)	0.410 (0.243)	0.667 (0.205)	N.A.	N.A.
Other (%)	0.058 (0.062)	0.061 (0.082)	0.057 (0.042)	N.A.	N.A.
Female (%)	0.493 (0.030)	0.493 (0.028)	0.493 (0.031)	0.493 (0.024)	0.493 (0.024)
Special needs (%)	0.088 (0.046)	0.091 (0.046)	0.086 (0.046)	0.088 (0.046)	0.088 (0.046)
Teacher Characteristics					
Average Praxis score	0.066 (0.328)	-0.100 (0.347)	0.179 (0.259)	0.066 (0.328)	0.064 (0.328)
NBPTS certified (%)	0.087 (0.103)	0.057 (0.087)	0.107 (0.108)	0.087 (0.103)	0.085 (0.101)
>3 years of experience (%)	0.789 (0.149)	0.748 (0.163)	0.817 (0.132)	0.789 (0.149)	0.789 (0.149)
Number of Students					
In grades 6-8	782 (267)	673 (230)	856 (265)	782 (267)	778 (265)
In whole school	813 (270)	708 (237)	883 (268)	812 (270)	808 (268)

Notes: SD = Standard deviation, FRPL = Free or reduced-price lunch. Lagged staffing variables are based on a smaller sample since we lose the first year.

Table 2a. Basic Results for Staffing Variables: Elementary Schools, Full Sample

Panel A. Current Staffing Variables

Independent variables	Initial Model		No Health		Preferred Model	
	Reading (1)	Math (2)	Reading (3)	Math (4)	Reading (5)	Math (6)
Regular teachers	1.038+ (0.554) [0.005]	0.869 (0.741) [0.005]	1.425** (0.510) [0.007]	1.405* (0.688) [0.007]	1.173+ (0.639) [0.006]	1.375 (1.046) [0.007]
Teaching assistants	0.796** (0.274) [0.002]	0.699+ (0.405) [0.002]	0.873** (0.272) [0.002]	0.811* (0.404) [0.002]	0.718* (0.312) [0.002]	0.405 (0.475) [0.001]
Health	4.041** (1.419) [0.002]	5.567** (1.890) [0.002]				
Health (early)					2.869 (2.593) [0.001]	1.181 (3.918) [0.000]
Health (late)					19.182 (29.736) [0.008]	62.222 (44.470) [0.025]
Other staff	0.306 (0.957) [0.000]	-0.603 (1.408) [0.000]	0.408 (0.936) [0.000]	-0.479 (1.367) [0.000]	0.343 (0.991) [0.000]	-0.464 (1.593) [0.000]
No. of observations	12,839	12,839	12,839	12,839	12839	12839
No. of unique schools	1,163	1,163	1,163	1,163	1,163	1,163

Panel B. Lagged Staffing Variables

Regular teachers	1.776** (0.608) [0.009]	1.972* (0.782) [0.010]	2.348** (0.552) [0.012]	2.580** (0.720) [0.014]	1.903* (0.755) [0.010]	1.972* (0.787) [0.010]
Teaching assistants	1.365** (0.307) [0.004]	1.055* (0.433) [0.003]	1.496** (0.302) [0.004]	1.199** (0.432) [0.003]	1.650** (0.398) [0.005]	1.056* (0.481) [0.003]
Health	5.417** (1.378) [0.002]	5.923** (1.952) [0.002]				
Health (early)					7.525** (2.106) [0.003]	5.930* (2.597) [0.002]
Health (late)					-42.434 (31.020) [-0.020]	5.750 (37.600) [0.0023]
Other staff	-1.407 (1.062) [-0.001]	-2.054 (1.542) [-0.001]	-1.117 (1.040) [-0.001]	-1.744 (1.505) [-0.001]	-1.568 (1.258) [-0.001]	-2.055 (1.538) [-0.001]
No. of observations	11,554	11,554	11,554	11,554	11554	11554
No. of unique schools	1,153	1,153	1,153	1,153	1,153	1,153

Notes: Each column comes from a separate regression that includes all the control variables listed in Table 1, plus year and school fixed effects. The staffing variables are expressed per student. Each school level observation is weighted by the number of students in the school, and standard errors are clustered at the school level. The full equations for 1 and 2 are in the appendix. **p<0.01, *p<0.05, +p<0.1

Table 2b. Basic Results for Staffing Variables: Middle Schools, Full Sample

Panel A. Current Staffing Variables

Independent variables	Initial Model		No Health		Preferred Model	
	Reading (1)	Math (2)	Reading (3)	Math (4)	Reading (5)	Math (6)
Regular teachers	0.821 (0.732) [0.004]	1.497 (0.931) [0.008]	0.999 (0.733) [0.005]	1.527 (0.931) [0.008]	2.726 (1.900) [0.014]	3.383 (2.355) [0.018]
Teaching assistants	3.769** (1.075) [0.003]	8.570** (1.525) [0.006]	3.739** (1.061) [0.003]	8.648** (1.536) [0.006]	4.403** (1.320) [0.003]	9.198** (1.751) [0.006]
Health	4.649* (2.114) [0.001]	-1.901 (2.688) [0.000]				
Health (early)					17.4757 (11.025) [0.003]	10.7953 (13.737) [0.002]
Health (late)					-60.056 (54.114) [-0.012]	-65.951 (67.262) [-0.013]
Other staff	-5.468** (1.586) [-0.003]	-4.419* (1.884) [-0.003]	-5.194** (1.550) [-0.004]	-4.420* (1.869) [-0.003]	-4.1374* (2.011) [-0.003]	-3.1019 (2.386) [-0.002]
No. of observations	4,691	4,691	4,691	4,691	4691	4691
No. of unique schools	439	439	439	439	439	439

Panel B. Lagged Staffing Variables

Regular teachers	2.342** (0.798) [0.012]	2.686** (1.039) [0.014]	2.565** (0.795) [0.013]	2.818** (1.044) [0.015]	9.067 (6.760) [0.047]	10.255 (7.721) [[0.053]
Teaching assistants	6.025** (1.198) [0.005]	9.947** (1.643) [0.008]	6.075** (1.186) [0.005]	10.076** (1.653) [0.008]	9.096* (4.600) [0.007]	13.403* (5.229) [0.011]
Health	3.654+ (2.186) [0.001]	-1.050 (2.749) [0.000]				
Health (early)					50.039 (40.057) [0.010]	51.153 (46.073) [0.010]
Health (late)					-229.339 (202.204) [-0.050]	-263.269 (233.322) [-0.053]
Other staff	-4.744** (1.535) [-0.003]	-4.919* (1.968) [-0.003]	-4.411** (1.491) [-0.003]	-4.838* (1.955) [-0.003]	0.343 (5.289) [0.000]	0.806 (6.168) [0.001]
No. of observations	4,181	4,181	4,181	4,181	4,181	4,181
No. of unique schools	430	430	430	430	430	430

Notes: Each column comes from a separate regression that includes all the control variables listed in Table 1, plus year and school fixed effects. The staffing variables are expressed per student. Each school level observation is weighted by the number of students in the school, and standard errors are clustered at the school level. The full equations for 1 and 2 are in the appendix. **p<0.01, *p<0.05, +p<0.1

Table 3. Falsification Test for Preferred Model: Elementary and Middle Schools

Independent variables	Elementary		Middle	
	Reading (1)	Math (2)	Reading (3)	Math (4)
Regular teachers (t+1)	-2.270* (0.909)	-3.204* (1.508)	-1.520 (0.997)	-2.354+ (1.253)
Teaching assistants (t+1)	-0.155 (0.444)	-0.970 (0.652)	0.959 (1.178)	5.847** (1.726)
Health (early) (t+1)	3.813 (2.702)	4.496 (4.221)	3.644 (6.821)	-7.934 (9.141)
Health (late) (t+1)	50.158 (37.572)	101.359+ (58.430)	-1.478 (31.399)	20.613 (41.529)
Other staff (t+1)	0.321 (1.248)	-0.841 (1.991)	-5.904** (2.283)	-5.851* (2.918)
No. of observations	10,635	10,635	3,843	3,843
No. of unique schools	1,153	1,153	427	427

Notes: Each column comes from a separate regression that includes all the control variables listed in Table 1, plus year and school fixed effects. The staffing variables are expressed per student, and all for year t+1. Each school-level observation is weighted by the number of students in the school, and standard errors are clustered at the school level. **p<0.01, *p<0.05, +p<0.1

Table 4a. Results for Subsamples and Subgroups: Elementary Schools

Panel A. Current Staffing Variables

Independent variables	High FRPL		Low FRPL		Minority Students		White Students	
	Reading (1)	Math (2)	Reading (3)	Math (4)	Reading (5)	Math (6)	Reading (7)	Math (8)
Regular teachers	0.950 (0.929) [0.005]	-0.174 (1.182) [-0.001]	2.039 (1.383) [0.010]	4.375+ (2.500) [0.023]	0.876 (0.760) [0.005]	0.040 (1.165) [0.000]	1.195 (0.761) [0.006]	1.823 (1.317) [0.010]
Teaching assistants	1.252+ (0.729) [0.004]	1.278 (0.957) [0.004]	0.835* (0.373) [0.002]	0.655 (0.654) [0.002]	1.828* (0.859) [0.005]	2.627* (1.105) [0.007]	0.522+ (0.292) [0.001]	-0.009 (0.451) [0.000]
Health (early)	5.105* (2.435) [0.002]	5.245 (3.204) [0.002]	4.134 (3.854) [0.002]	-0.489 (6.486) [0.000]	7.240 (6.264) [0.003]	11.621 (8.473) [0.005]	2.544 (2.004) [0.001]	0.430 (3.211) [0.000]
No. of observations	6,358	6,358	6,481	6,481	12,795	12,795	12,501	12,501
No. of unique schools	583	583	580	580	1,163	1,163	1,160	1,160

Panel B. Lagged Staffing Variables

Regular teachers	2.663* (1.346) [0.014]	1.578 (1.521) [0.008]	1.963+ (1.117) [0.010]	3.550* (1.759) [0.018]	2.731* (1.129) [0.014]	1.890+ (1.027) [0.010]	2.169** (0.763) [0.011]	2.613* (1.014) [0.014]
Teaching assistants	2.511* (0.992) [0.008]	2.271* (1.116) [0.007]	1.175** (0.383) [0.003]	0.597 (0.657) [0.002]	3.239** (1.111) [0.009]	2.420* (1.147) [0.007]	0.924** (0.322) [0.003]	0.384 (0.426) [0.001]
Health (early)	5.111* (2.398) [0.002]	4.411 (2.922) [0.002]	7.684** (2.976) [0.003]	5.268 (4.698) [0.002]	8.057* (3.951) [0.003]	7.020+ (4.093) [0.003]	5.625** (1.713) [0.002]	3.372 (2.439) [0.001]
No. of observations	5,717	5,717	5,837	5,837	11,519	11,519	11,232	11,232
No. of unique schools	575	575	578	578	1,151	1,151	1,149	1,149

Notes: Each column comes from a separate regression that includes all the control variables listed in Table 1, plus year and school fixed effects. The staffing variables are expressed per student. Each school level observation is weighted by the number of students in the school, and standard errors are clustered at the school level. **p<0.01, *p<0.05, +p<0.1

Table 4b. Results for Subsamples and Subgroups: Middle Schools

Panel A. Current Staffing Variables

Independent variables	High FRPL		Low FRPL		Minority Students		White Students	
	Reading (1)	Math (2)	Reading (3)	Math (4)	Reading (5)	Math (6)	Reading (7)	Math (8)
Regular teachers	1.106 (0.943) [0.006]	2.648* (1.095) [0.014]	0.984 (3.063) [0.005]	2.953 (4.626) [0.015]	-0.388 (1.157) [-0.002]	0.648 (1.119) [0.003]	3.142 (1.961) [0.016]	3.648 (2.925) [0.019]
Teaching assistants	5.914** (1.648) [0.005]	10.932** (2.018) [0.009]	2.286 (1.477) [0.002]	7.760** (2.617) [0.005]	4.803** (1.698) [0.003]	9.117** (1.897) [0.006]	3.325* (1.359) [0.002]	7.634** (1.958) [0.005]
Health (early)	10.299 (9.681) [0.002]	-3.051 (12.215) [-0.001]	8.019 (5.260) [0.002]	3.070 (7.797) [0.001]	19.468 (16.473) [0.004]	-0.373 (18.476) [0.000]	11.270 (7.085) [0.002]	9.916 (10.219) [0.002]
No. of observations	2,213	2,213	2,478	2,478	4,686	4,686	4,624	4,624
No. of unique schools	217	217	222	222	439	439	439	439

Panel B. Lagged Staffing Variables

Regular teachers	1.544 (1.194) [0.008]	2.792* (1.363) [0.015]	2.895 (2.900) [0.015]	2.366 (4.433) [0.012]	1.634 (2.927) [0.008]	2.487 (2.623) [0.013]	7.593+ (4.086) [0.039]	9.206 (5.907) [0.048]
Teaching assistants	7.107** (1.997) [0.006]	11.768** (2.258) [0.011]	5.453** (1.743) [0.004]	9.038** (2.753) [0.006]	9.851+ (5.265) [0.008]	12.979** (4.749) [0.010]	6.980* (2.956) [0.006]	11.198** (4.016) [0.009]
Health (early)	16.733 (10.884) [0.003]	13.167 (12.511) [0.003]	5.107 (5.270) [0.001]	-0.872 (7.947) [0.000]	55.623 (43.613) [0.011]	37.273 (41.496) [0.007]	20.740 (13.927) [0.004]	26.149 (20.143) [0.005]
No. of observations	1,963	1,963	2,218	2,218	4,176	4,176	4,115	4,115
No. of unique schools	212	212	218	218	430	430	430	430

Notes: Each column comes from a separate regression that includes all the control variables listed in Table 1, plus year and school fixed effects. The staffing variables are expressed per student. Each school level observation is weighted by the number of students in the school, and standard errors are clustered at the school level. **p<0.01, *p<0.05, +p<0.1

Table 5. Academic Proficiency Results: Elementary and Middle Schools

Panel A. Current Staffing Variables

Independent variables	Elementary		Middle	
	Reading: Share Proficient (1)	Math: Share Proficient (2)	Reading: Share Proficient (3)	Math: Share Proficient (4)
Regular teachers	0.086 (0.340) [0.000]	0.230 (0.425) [0.001]	0.313 (0.799) [0.002]	-0.109 (0.781) [-0.001]
Teaching assistants	0.410* -0.167 [0.001]	0.541** (0.190) [0.002]	1.888** (0.583) [0.001]	2.863** (0.620) [0.005]
Health (early)	-2.375+ (1.310) [0.000]	-2.677+ (1.581) [0.001]	3.318 (4.823) [0.001]	-3.263 (4.841) [-0.001]
Health (late)	19.355 (15.332) [0.004]	33.683+ (18.724) [0.007]	-21.592 (23.377) [-0.004]	8.551 (23.187) [0.002]
Other staff	0.026 (0.479) [0.000]	0.450 (0.634)	-1.973* (0.847) [-0.001]	-1.972* (0.966) [-0.001]
No. of observations	12,833	12,833	4,689	4,689
No. of unique schools	1,163	1,163	439	439

Panel B. Lagged Staffing Variables

Regular teachers	0.866** (0.321) [0.005]	1.597** (0.333) [0.008]	2.396 (2.000) [0.012]	3.800 (2.489) [0.020]
Teaching assistants	0.769** (0.174) [0.002]	0.606** (0.186) [0.002]	3.477* (1.354) [0.003]	4.747** (1.657) [0.004]
Health (early)	-0.700 (0.760) [0.000]	-0.554 (0.849) [0.000]	12.443 (11.865) [0.002]	15.852 (14.764) [0.003]
Health (late)	-5.504 (11.587) [0.002]	10.056 (13.222) [0.004]	-69.628 (59.472) [-0.014]	-78.841 (74.295) [-0.016]
Other staff	-0.548 (0.568) [0.000]	-0.714 (0.643) [0.000]	-0.863 (1.599) [-0.001]	-1.045 (2.039) [-0.001]
No. of observations	11,544	11,544	4,176	4,176
No. of unique schools	1,153	1,153	430	430

Notes: Each column comes from a separate regression that includes all the control variables listed in Table 1, plus year and school fixed effects. The staffing variables are expressed per student. Each school level observation is weighted by the number of students in the school, and standard errors are clustered at the school level. "Share Proficient" represents the proportion of students who performed at or above grade level proficiency on end-of-grade (EOG) reading or math tests. **p<0.01, *p<0.05, +p<0.1

Table 6. Results for Behavioral Outcomes: Elementary and Middle Schools

	Elementary Schools			Middle Schools		
	Absentee Rate	Tardy Rate	In-School Suspension Rate	Absentee Rate	Tardy Rate	In-School Suspension Rate
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
Regular teachers	15.623** (4.040)	-12.749* (5.882)	-6.546** (2.368)	13.055 (13.846)	-7.481 (9.174)	10.337+ (6.226)
Teaching assistants	-6.053* (2.979)	-6.451* (2.556)	-0.777 (1.296)	-27.614+ (14.313)	-10.204 (10.801)	-3.547 (8.180)
Health (early)	32.993 (41.122)	-15.999* (7.517)	20.885 (30.517)	112.042 (110.208)	-12.188 (17.173)	173.388** (57.397)
Health (late)	137.066 (115.335)	74.504 (84.279)	88.755 (56.006)	-142.612 (249.280)	-102.653 (74.168)	-259.395* (106.701)
Counselors	34.329 (66.063)	24.691 (38.504)	-42.506 (40.066)	46.172 (59.534)	-11.884 (24.231)	-41.040 (29.297)
Schools leaders	40.531 (31.684)	32.581 (23.362)	-48.318* (20.383)	67.506 (42.776)	70.054** (25.460)	25.777 (28.637)
Non-certified academic support	21.889 (15.877)	-15.976 (10.688)	2.640 (6.632)	71.870 (69.429)	-39.418* (16.092)	22.045 (21.612)
No. of observations	8,015	3,622	6,938	2,964	1,452	2,809
No. of unique schools	1,125	1,048	1,067	424	391	415

Notes: Each column comes from a separate regression that includes all the control variables listed in Table 1, plus year and school fixed effects. The staffing variables are expressed per student. Each school level observation is weighted by the number of students in the school, and standard errors are clustered at the school level. **p<0.01, *p<0.05, +p<0.1

Appendix Table A1. Initial Model

Independent variables	Elementary				Middle			
	Current Staffing		Lagged Staffing		Current Staffing		Lagged Staffing	
	Reading (1)	Math (2)	Reading (3)	Math (4)	Reading (5)	Math (6)	Reading (7)	Math (8)
Teachers	1.038+ (0.554)	0.869 (0.741)	1.776** (0.608)	1.972* (0.782)	0.821 (0.732)	1.497 (0.931)	2.342** (0.798)	2.686** (1.039)
Teaching assistants	0.796** (0.274)	0.699+ (0.405)	1.365** (0.307)	1.055* (0.433)	3.769** (1.075)	8.570** (1.525)	6.025** (1.198)	9.947** (1.643)
Health	4.041** (1.419)	5.567** (1.890)	5.417** (1.378)	5.923** (1.952)	4.649* (2.114)	-1.901 (2.688)	3.654+ (2.186)	-1.050 (2.749)
Other staff	0.306 (0.957)	-0.603 (1.408)	-1.407 (1.062)	-2.054 (1.542)	-5.468** (1.586)	-4.419* (1.884)	-4.744** (1.535)	-4.919* (1.968)
% FRPL	-0.347** (0.031)	-0.320** (0.034)	-0.340** (0.031)	-0.314** (0.034)	-0.280** (0.048)	-0.300** (0.048)	-0.281** (0.046)	-0.309** (0.048)
% Black	-0.422** (0.036)	-0.489** (0.043)	-0.372** (0.035)	-0.429** (0.042)	-0.489** (0.073)	-0.494** (0.073)	-0.442** (0.066)	-0.457** (0.073)
% Hispanic	-0.544** (0.051)	-0.242** (0.067)	-0.536** (0.054)	-0.217** (0.070)	-0.652** (0.101)	-0.368** (0.127)	-0.648** (0.104)	-0.362** (0.133)
% Other	0.258** (0.064)	0.449** (0.083)	0.281** (0.066)	0.460** (0.084)	0.624** (0.139)	0.848** (0.163)	0.578** (0.135)	0.815** (0.172)
% Female	0.318** (0.039)	0.238** (0.048)	0.307** (0.040)	0.214** (0.049)	0.356** (0.065)	0.236** (0.083)	0.338** (0.064)	0.217** (0.081)
% Special Needs	-0.430** (0.047)	-0.373** (0.060)	-0.403** (0.047)	-0.351** (0.060)	-0.893** (0.096)	-0.951** (0.105)	-0.811** (0.094)	-0.870** (0.109)
Average Praxis score	0.012 (0.007)	0.022* (0.009)	0.009 (0.007)	0.022* (0.009)	0.017 (0.011)	0.025* (0.013)	0.010 (0.011)	0.019 (0.013)
% NBPTS certified	0.016 (0.017)	0.026 (0.021)	0.013 (0.018)	0.018 (0.022)	0.001 (0.026)	0.008 (0.033)	0.007 (0.027)	0.019 (0.033)
% with years of expr>3	0.089** (0.011)	0.087** (0.014)	0.080** (0.011)	0.077** (0.014)	0.077** (0.020)	0.096** (0.023)	0.066** (0.022)	0.085** (0.025)
No. of observations	12,839	12,839	11,554	11,554	4,691	4,691	4,181	4,181
No. of unique schools	1,163	1,163	1,153	1,153	439	439	430	430

Notes: Each column comes from a separate regression that includes all the control variables listed in Table 1, plus year and school fixed effects. The staffing variables are expressed per student. Each school level observation is weighted by the number of students in the school, and standard errors are clustered at the school level. The full equations for 1 and 2 are in the appendix. **p<0.01, *p<0.05, +p<0.1