Combatting Poverty by Improving Children's Access to Health Services and Effective Teaching: An Evaluation of Three Social Policies

By

Walker Alexander Swain

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfilment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Leadership and Policy Studies

August 11, 2017

Nashville, Tennessee

Approved:

Matthew G. Springer, Ph.D.

Helen F. Ladd, Ph.D.

Mimi Engel, Ph.D.

Dale Ballou, Ph.D.

ACKNOWLEDGEMENTS

I feel fortunate to have had Matt Springer as an advisor, friend, collaborator, and mentor before, during, and after the completion of this dissertation. His support, feedback, and patience have been invaluable. I am grateful to my committee, Mimi, Dale, and Sunny, for their individual and collective wisdom, and for taking the time to challenge me to better develop this work. Thanks to Gary Henry and Ellen Goldring for their leadership, guidance, and mentorship, and to the rest of the continuously supportive, talented, and knowledgeable faculty in the department.

From the first days at Peabody to the last, my cohort, Ben, Chris, Dominique, Ngaire, and Richard, made the doctoral experience more enjoyable and intellectually rich than I ever could have hoped. To my dear friends and collaborators, Alvin, Norm, Jonathon, Jay, Luis, Sy, and so many others who inspire and make me happy on a daily basis from both near and far, you all know how much I appreciate you.

Finally, to my family, I would not be me without you, and certainly would not have completed this program without your support. Thanks to my mother, who taught and fought for me to read as a child, and who kept us safe, healthy, and attending diverse schools with good teachers. She also never stopped learning and growing herself. Thanks to the matriarch, Nana (Joan), who pushed me by instruction and example to think as deeply and creatively as I can. Thanks to my sisters, Caitlin, Erin, and Sunny, for always being there and always being wonderful. Dad, thank you for keeping me focused on what is right. Paul, thank you for being the model of passionate, dedicated scholarship. To my late stepmother, Ashley Osment, thank you for the confidence, advice, love, and example of how to live a full life in service to others, filled with joy and music. I am forever grateful.

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

CHAPTER 1

EARLY CHILDHOOD HEALTH INSURANCE EFFECTS ON HEALTH SERVICE UTILIZATION, ATTENDANCE, TEST SCORES, AND STUDENT HEALTH

Introduction

Research has long established that low-income children have higher rates of treatable health problems than their more affluent, White peers. Racially and economically disadvantaged students disproportionately suffer from vision impairment (Festinger & Duckman, 2004; Orfield et al., 2001), hearing problems (Egbuonu & Starfield, 1982), asthma (Forrest, Starfield, Riley, & Kang, 1997; Halfon & Newacheck, 1993), ADHD, and other learning disabilities (Stevens, Harman, & Kelleher, 2005).[1] Low-income and minority students also have less access to health services and health care coverage to pay for diagnoses and treatment (Currie, Decker, & Lin, 2008; Flores, Bauchner, Feinstein, & Nguyen, 1999). These types of problems can pose considerable barriers to students' academic success (Ding, Lehrer, Rosenquist, & Audrain-McGovern, 2009; Fletcher & Lehrer, 2009). However, although correlations between low-income and minority students' disparate health problems and their relative academic performance have been well-documented, empirical literature evaluating the effectiveness of health policy interventions in breaking the ties between poverty, health problems, and academic struggles remains sparse, especially during the first years of formal schooling.

[1] Low-income children experience vision impairment at twice the rate of their middle-class peers. Additionally, the National Institute of Health cites asthma as the leading cause of health-related school absences.

In 1997, President Clinton signed into law the largest single expansion of public health insurance for children since the introduction of Medicaid in 1965. The State Children's Health Insurance Program (SCHIP) provided federal dollars for states to expand health coverage to children under the age of 18 whose parental earnings were too high for them to qualify for Medicaid, but not high enough to afford private insurance. Most states chose eligibility requirements at 200% of the federal poverty level (FPL)—double the minimum Medicaid cutoff for children over the age of six. However, eligibility requirements still varied considerably across states, with Massachusetts setting the highest income threshold at 400% of the FPL. Indeed, the law gave states considerable flexibility in determining eligibility cutoffs, copayment amounts, and the types of treatments that could be covered, while simultaneously facilitating substantial, increased outreach and simplifications of enrollment processes, which drove major increases in coverage among children living in poverty (Aizer, 2007).

Other unique state program characteristics, such as presence-of-asset tests or required face-to-face interviews, have been linked to variations in insurance enrollment (Wolfe & Scrivner, 2005). Several studies have attributed the majority of the gains in increased children's health insurance coverage to the elevated take-up among families with incomes low enough to have qualified for public insurance even prior to the expansion. These studies also highlighted the importance of policy implementation differences (e.g., eliminating asset tests, simplifying application and renewal processes, or extending benefits to parents) in driving take-up among parents of low-income children (Bansak & Raphael 2007; Aizer, 2007). We have found that similar patterns hold for students in the ECLS-K sample, with especially exceptional reductions in uninsured children of parents with low-incomes and/or low education, and among students who identify as Hispanic.

Assessments of SCHIP's effects on rates of child health insurance coverage have consistently found that the proportion of uninsured children was significantly reduced as a result of

SCHIP/Medicaid expansion. Findings with respect to enrollment in private/employer-based plans were mixed, though the vast majority found no reduction in enrollment, with only limited evidence of crowd-out or reductions in private coverage among students at the highest incomes eligible for CHIP expansion (Howell & Kenney, 2012). The effect size of the reduction in uninsured children has ranged from a roughly 4-percentage-point drop to just over 10 percentage points. Notably, these findings are consistent with the overall reduction in uninsured children from 17% to 10% between 1987 and 2009, a period during which the uninsured rate among adults increased (Carmen, Proctor, & Smith, 2008).

In recent years, provisions of the Affordable Care Act (ACA) have expanded access to health care to millions of Americans, while requiring insurance providers to cover treatments and preventative measures for a broader range of health problems. It also extended federal funding of the SCHIP and modified elements of the SCHIP initiative to facilitate enrollment and improve coverage. In 2009, President Obama reauthorized the act with new incentives for states to simplify enrollment procedures.[2] Thirteen states expanded their eligibility cutoffs in 2010, and more than two million young people gained coverage (Health and Human Services, 2011). However, continued funding for the program, which is not an entitlement, hangs in the balance, plausibly due to the limited study of child insurance benefits.

In this study, we exploit the overlap between the administration of a nationally representative longitudinal education survey, the National Center for Education Statistics' Early Childhood Longitudinal Survey (ECLS-K), and the expansion of children's health coverage resulting from the initial implementation of the SCHIP, to examine the effects of a health policy intervention

[2] Marton (2007) examined the introduction of premiums into the SCHIP program in Kentucky, and found that premiums reduce the length of enrollment, with the impact concentrated in the first three months after the introduction of the premium.

on students' health service utilization, health status, attendance, and academic success. The ECLS-K is composed of a nationally representative sample of more than 20,000 kindergarteners in the 1998-1999 school year, who were surveyed in the fall of kindergarten and five more times before 2007. The survey also included annual standardized cognitive assessments and the collection of information on a range of school and social factors, including whether or not students had health insurance.

This study utilizes the rich longitudinal education and health data collected in ECLS-K to answer two related questions:

1. To what extend does children's health insurance coverage increase health service utilization or parent-rated student health for elementary school students?
2. To what extent does children's health insurance coverage improve elementary school students' academic achievement and attendance?

The sections that follow will describe the mechanisms by which expanded health care could improve student outcomes, provide a review of select literature, describe the data and analytical methods, summarize the findings, and discuss their implications.

Theoretical Framework and Hypotheses

There are two primary mechanisms by which expanded access to public health insurance might affect student health outcomes and influence student achievement. First, students who suffer from treatable health problems, such as visual or hearing impairments or asthma, could have illnesses and disabilities diagnosed and treated more effectively with increased access to health care. With their achievement-blocking symptoms ameliorated, they could potentially attend school more regularly and learn more throughout the academic year, as measured by standardized cognitive assessments. Second, decreased health care costs realized by newly insured families might increase the amount of resources families could put toward educational goods or other expenditures that might improve student achievement. We classify these as *income* and *substitution effects*.

Improved Student Health Status

Research has established a clear link between a child's economic well-being and rates of treatable health problems, like vision impairment (Festinger & Duckman, 2004; Orfield, et al., 2001), hearing problems (Egbuonu & Starfield, 1982), and asthma (Forrest et al., 1997; Halfton & Newacheck, 1993). These types of problems can pose considerable barriers to students' academic success. Students with health insurance should have greater access to health care, thus increasing the likelihood that these achievement-blocking problems are diagnosed and treated. Addressing issues like asthma or visual and hearing impairments should result in increased attendance and improve students' capacity to learn. However, the extent to which this health status effect will increase student achievement depends on the size of the academic obstacle posed by treatable health impairments, and the effectiveness of the health insurance as a driver of symptomatic relief.

Income and Substitution Effects

Individuals who lack health insurance experience higher costs when normal health problems arise. These higher health costs leave less income for families to spend on resources that support academic achievement. One study estimated that out-of-pocket healthcare expenditures for families that were uninsured for at least part of the year totaled over $30 billion in 2008 (Hadley, Holahan, Coughlin, & Miller, 2008). Gaining access to children's health insurance lowers their health-related realized costs, thereby increasing the family's expendable income, and decreasing the price of health care relative to educational services. Shaefer, Grogan, & Pollack (2011) conducted a study of financial benefits as a result of a family's transitions to SCHIP during a similar time window (1998-

2003), and found that families who enrolled in the program saved roughly the equivalent of a cash transfer of $1,500 in reduced out-of-pocket and healthcare premium expenses.

The extent to which this increased income could result in improved academic achievement depends on the families' relative valuation of education-related goods as well as their access to quality educational goods. For example, if a family placed a high value on the children's education, they might shift all of the money they previously spent on health costs to the purchase of tutoring, books, or other enrichment materials. If the educational goods and services the family purchased were of sufficient quality, student achievement would improve. Alternatively, a family that places a relatively low value on education, or that lacks information about the types of expenditures that might improve educational outcomes, would fail to realize any income-related improvements in students' academic outcomes.

Hypotheses

Given prior literature documenting greater health problems among uninsured children (Currie et al., 2008), and improved health indicators in the wake of insurance expansion (Currie & Gruber, 1996), we hypothesized that students who gain access to health insurance between the first and fifth grades will increase health service utilization, experience improved overall health, and, as a consequence, demonstrate elevated rates of school attendance and increased learning (as measured by standardized assessments). It is plausible that certain benefits, such as improved test scores, take longer to arise than the short window we observed in this study. For students to improve their academic performance as a result of improved health, they must have a treatable health problem, seek successful medical care, recover, perceive the improvement, and then realize the benefits through improved attendance and focus during school.

Select Literature on Impacts of Children's Health Insurance Coverage

Despite the long history of public policy interest in the links between health and learning, few studies have sought to evaluate the impact of health interventions on students' academic success. We reviewed several studies that have evaluated the association between poor health and students' struggles in school. We have also highlighted several seminal studies examining the impacts of healthcare expansions on student health and academic outcomes. Our study contributes to the literature by bridging the gap between a largely public health-oriented literature and a potentially vital education policy question: can health interventions improve student health and academic outcomes?

Student Health and Academic Outcomes

Empirical studies have documented a strong relationship between poor health and diminished academic performance, sometimes using rigorous research methods to support causal claims. Several innovative studies in health economics have utilized specific genetic differences between siblings in the same home to estimate significant causal effects of illnesses on academic achievement and student grade point averages (Ding et al., 2009; Fletcher & Lehrer, 2009). Other researchers have documented the effects of neonatal health on a range of academic-related outcomes, including IQ, educational attainment, and cognitive development (Black, Devereux, & Salvanes, 2007; Figlio, Guryan, Karbownik, & Roth, 2014; Oreopoulos, Stabile, Walld, & Roos, 2008; Royer, 2009).

Earlier, more conventional econometric analyses that controlled for a range of student and family characteristics have demonstrated that chronic illness is associated with a rate of absence from school that is more than twice as high as that of the average student (Fowler, Johnson, & Atkinson, 1985). These higher rates of absence associated with chronic illness are not generally related to lower achievement after controlling for socioeconomic status; however, for certain chronic illnesses like sickle cell disease, the negative impacts appear to be universal across income groups (Fowler et al., 1985).

Data from the North Carolina Child Health Assessment and Monitoring Program further demonstrates that poor oral health can also negatively impact children's school attendance and performance. Jackson and colleagues (2011) found that children with poor oral health were nearly three times more likely to miss school. Moreover, absences caused by poor oral health were associated with increased likelihood of poor school performance, whereas absences for routine dental care did not affect school performance.

Health Insurance and Health Outcomes

Research has found that expansions of public health insurance have resulted in increased access to care, show no signs of crowding out private insurance (Bansak & Raphael, 2007), and have improved health outcomes for low-income children (e.g., Currie & Gruber, 1996; Dafny & Gruber, 2005; Kenney, Dubay, Hill, Sommers, & Zuckerman, 2005). Other studies have highlighted the differential impacts of SCHIP on coverage across races and ethnicities. For example, Shone et al. (2003) found that, when controlling for socioeconomic status, children from minority groups were more likely to have lacked insurance before enrollment in SCHIP and to have had poorer general health prior to coverage. A RAND evaluation of SCHIP in California similarly found that the program decreased disparities in access to health care for language minorities, and was associated with improved health-related quality-of-life indicators, including physical, mental, and social well-being (Seid, Varni, Cummings, & Schonlau, 2006).

Currie et al. (2008) found that increased access to health care increased utilization of preventative care, and that early childhood eligibility for Medicaid improved health outcomes later in life. However, they found no significant effects on current health statuses for older children. Using ordered probit and linear probability models with a variety of state and time covariates, Currie and colleagues estimated the change in the relationship between income and health status within age groups, and found a significant decline in the income-to-health-status link for students age 4-17 in the years of Medicaid/SCHIP expansion. They also attempted to directly assess the effects of insurance expansion on student health statuses, and found a significant increase in the utilization of preventative care, though no changes in reported health statuses. Because they used data from the National Health Interview Survey (NHIS), which lacks information about the children's health insurance enrollment, they estimated the eligibility based on the family's income and the state's income eligibility cutoffs in the given year. ECLS-K allows us to improve on Currie et al.'s estimates both because of its longitudinal design (NHIS is cross-sectional), and because it asks parents about their children's health insurance enrollment as well as about the child's current health statuses over time.

However, a recent Urban Institute review of 38 rigorous evaluations of SCHIP's impacts on coverage, health service utilization, and health status concluded that, while studies tend to find an increase in utilization (i.e., a 6.3 percentage-point increase for having any medical visit, and a roughly 15 percentage-point increase in having had any dental visit), only one study (as of 2012) that accounted for endogeneity found even small improvements in perceived health statuses (Kenney, Haley, Pan, Lynch, & Buettgens, 2016). However, there are a number of studies that have found reductions in child mortality rates (e.g., Howell, Decker, Hogan, Yemane, & Foster, 2010), finding that a 10% increase in CHIP/Medicaid eligibility is associated with a 3% reduction in child mortality (Howell & Kenney, 2012). It is plausible that parents of children who gain health insurance do not perceive improvements in their child's health. However, one potential proxy for realized improvements in child health—increased school attendance—is presumably associated with better treatment of health problems that previously increased absenteeism.

Health Insurance and Academic Outcomes

Previous state-level analyses have shown some indication that state child health insurance expansions were associated with small but significant improvements in academic outcomes for at least some groups and measures. Levine and Schanzenbach (2009) took advantage of time- and state-level variations in early Medicaid and SCHIP expansions in the 1980s and 1990s to estimate the effects on state-level student performance on the National Assessment of Educational Progress (NAEP) exams. They found that a 50-percentage-point increase in access to health insurance at birth was associated with improved performance in reading (effect size of .09 SD), though not math.

Cohodes, Grossman, Kleiner, and Lovenheim (2016) conducted a study using a simulated instrumental variable strategy to estimate the effects of children's public insurance eligibility for Medicaid on long-run educational attainment in the 1980s and 1990s. The study, which used data from the American Communities Survey (ACS), found substantial improvements in high school completion for non-White students whose eligibility increased, and increased college completion for low-income White students whose eligibility increased. The benefits were primarily attributed to coverage expansions that occurred post-infancy. The largest educational benefits associated with expansions occurred during ages 4-8, and smaller but significant benefits occurred between the ages of 14 and 17. The variation in positive educational attainment suggests that, either benefits of insurance coverage take time to accumulate, or that there is some specific value to eligibility during the formative educational period of primary school.

This study builds on the extant literature in several important ways. First, it focuses on effects on both health and academic outcomes during elementary school for a nationally representative sample, which is a critically formative period in a child's development. Second, the inclusion of individual student-level data about parent-reported enrollment statuses allows for a more precise estimation of the effect on individual students than prior studies that relied on state or subgroup-level variation. Finally, the rich set of student-, family-, and school-level data in the ECLS-K allowed us to control statistically for a range of potential time-variant circumstances that might bias estimated health insurance effects, and that is typically absent from most longitudinal datasets.

Data, Measures, Sample

Data

This study utilizes data from the 1998-1999 cohort of the Early Childhood Longitudinal Study (ECLS-K). The survey includes a broad range of information about academic achievement (a cognitive exam of math and reading),[3] school features, and home characteristics (parental education, health information, income, welfare status, food security, etc.) for a nationally representative sample of children who were in kindergarten during the 1998-1999 school year. The survey features information gathered from children through researcher-administered exams, as well as surveys of parents, teachers, and school administrators. Base-year surveys were administered in the fall of 1998, and follow-ups were conducted in the spring of 1999 (kindergarten), fall of 1999 (first grade), spring of 2000 (first grade), spring of 2002 (third grade), spring of 2004 (fifth grade), and again in 2007 (eighth grade). For our primary analysis, we limit our focus to the first, third, and fifth grade years of the survey, where parents were consistently asked comparable questions about insurance coverage and data on attendance. Health ratings are also available. Appendix A contains survey items used to construct our dependent variables as well as key indicator variables.

[3] Tests on science and general knowledge are also administered to participants, but are omitted here, as they are not assessed in every year of the longitudinal survey.

Dependent Variables

This study is focused on the impact of expanded health insurance on student health and academic outcomes. We use three student health-related indicators, two of which measure health service utilization, and a third that measures a child's health. Our health service utilization measures come from questions that ask respondents how recently their child had been seen by a doctor (or dentist). Response options included, "never", "less than 6 months", "6 months to one year," "1 year to 2 years," and "more than 2 years." We code these health utilization measures so that "never" equals 1, "more than 2 years" equals 2, "1 year to 2 years" equals 3, "6 months to one year" equals 4, and "less than 6 months" equals 5. We treat respondents as missing if they answered "don't know" or "refused" to answer. These questions and response options were longitudinally consistent across all years of survey administration. For the primary analysis presented here, we converted these responses to binary indicators for recent visits, such that a child receives a 1 if they have seen a doctor within a year, and are otherwise coded zero.[4]

Our third health-related outcome enabled us to address the impact of health insurance expansion on a parent's self-report of their child's health. The measure of a child's health comes from a survey question that asks the parent: "Would you say your child's health is..." Response options include "excellent," "very good," "good," "fair," and "poor." We reverse coded these response options so that a larger value on the coefficient of interest means better health. We treated respondents as missing if they answered "don't know" or "refused" to answer. These questions and response options were longitudinally consistent across all years of survey administration. For the primary analysis presented here, we converted this response to a binary indicator of "good health,"

[4] We modeled health service utilization in a variety of ways, including treating the recency of visits as a continuous variable in an OLS framework, as well as an ordinal logistic regression modeling strategy. Because results were qualitatively similar, we present results from the binary linear probability model for ease of interpretability.

such that a child received a 1 if their health was rated as good or better, and were otherwise coded zero due to the potential nonlinearity of responses on the ratings scale.

To assess the impact of expanded health insurance on student academics, we focused on two outcomes: absences and test scores. We used the total number of student absences in a given school year as reported in the ECLS. Because data were not reported on this measure in eighth grade (2006-2007 school year), our analysis is restricted to a shorter panel for this outcome.

Our primary measures of student learning come from the battery of longitudinal cognitive assessments administered as part of the ECLS survey. These standardized assessments were designed to measure a child's knowledge and skills at a given point in time, as well as track their academic growth in different subject areas across time. Assessments were constructed using a series of items from previously administered large-scale exams for similar-aged children, including the National Assessment of Educational Progress (NAEP), the ECLS-K frameworks, and selected states' curriculum standards. The assessment, which covers math, reading, and science, contains overlapping content areas for longitudinal comparisons of student progress. Content area specialists reviewed individual items for appropriateness, relevance, and difficulty. Additionally, items were assessed for sensitivity to cultural differences among minority groups. ECLS-K reports criterion-referenced item response theory (IRT) scale scores to compare growth over time. For the purpose of this study, we focused on math and reading, as science was not assessed across all years. In all analyses, test scores are standardized within the year to have a mean of zero and a standard deviation of 1.

Indicator Variables

Our primary indicator variable of interest comes from the spring parent interviews, where parents were asked about their child's health insurance coverage during each wave of the longitudinal survey. The formatting of the question about health insurance changed twice, and in important ways. First, when the students were in kindergarten, the parents were asked only whether or not the student was covered by health insurance (including Medicaid). In subsequent years, the parents were given a list of insurance types, which included: “Private Plan,” “Medicaid,” “State CHIP program,” “Military Insurance,” or “No Insurance.” In the two most recent administrations of the survey, the “No Insurance” option was marked only if the interviewee responded “no” to all the other categories. To insure consistency, our primary analysis is restricted to the first through eighth grade years, where parents identified the type of health insurance the student had. Where parents responded “no” in the first or third grade, but also indicated that the student had a specific type of health insurance, we modified responses to reflect the decision rule applied in later survey administrations.

Time-Variant Controls

To account for changes in a student’s family situation that might coincide with the acquisition of health insurance coverage and bias-estimated effects on the outcomes of interest, we included an extensive set of time-variant control variables. We controlled for parental employment status, family size, total income, and two forms of public assistance: Supplemental Nutrition Assistance Program, and Temporary Assistance for Needy Families. We also controlled for school type (public or private), school size, and the percentage of minorities enrolled.

Sample

To construct a nationally representative sample of children entering kindergarten in the 1998-1999 school year, the ECLS-K utilized a multistage probability sample design. The first stage primary sampling units (PSUs) were geographic areas (counties or groups of counties). The second stage was schools within the PSUs, and the third and final stage was children within the schools. In the early stages of the survey (spring K and first grade), samples were freshened to add students in the same age cohort who were missed by the initial surveys. This study used the unique child identifier variable as the primary unit of analysis, and was only restricted based on the presence of primary dependent variables and information used to determine insurance eligibility. However, including only individuals who participated through the full survey cut the number of observations to 11,290 (nearly by half), and reduced restrictions to students who had complete test score data. For consistency, we used only test scores from tests administered in the spring term, excluding the estimates for the subsamples evaluated in the fall of kindergarten and the fall of first grade.[5]

Because our primary analytic strategy was a student-fixed effect model that relied on within-student variation in insurance coverage to estimate a local average treatment effect (LATE), it was instructive to examine how students whose health coverage changed over time differed from the students who were either consistently covered or consistently lacked insurance. Table 1 provides a subset of demographic descriptive statistics (parental educational attainment and race or ethnicity) comparing these "switchers" and "non-switchers." It also illustrates some of the differences between the analytic sample, which was restricted based on the consistent availability of key independent and dependent variables. Notably, the analytic sample, those who were located and responded to each of the questions of interest in this study, were slightly more likely to be white, with higher levels of parental education.

[5] Because the identification strategy focused on assessing effects for the subset of the population who experienced variation in coverage or insurance eligibility, the primary analytic models ignored ECLS-K sample weights designed to achieve representativeness of the national sample of kindergarten-age students in 1998. However, we did utilize these weights to determine population means, against which we compared the characteristics of our analytic sample, and the sample for which local average treatment effects were estimated.

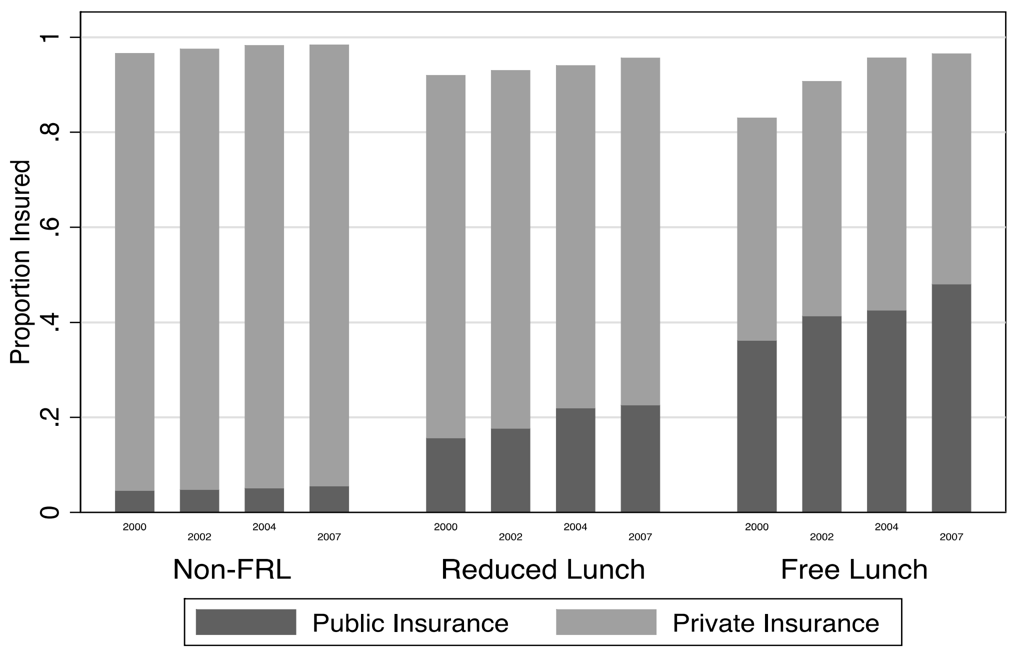
Table 1: Select Student-Level Summary Statistics By Sample and Insurance Status

	Full Sample	Analytic Sample	Unstable Insurance Coverage
Parental Education			
Non-HS Grad (%)	9.72	6.29	13.75
HS Grad (%)	23.80	22.39	35.87
Voc-Tech/Some College (%)	33.26	34.84	36.43
BA or more (%)	33.22	36.47	13.94
Race of Student			
Black (%)	15.17	10.85	11.52
Hispanic (%)	17.85	13.52	27.52
White (%)	55.33	64.54	50.37
Asian (%)	6.41	5.21	5.39
Other Race (%)	5.24	5.88	5.20
Below Federal Poverty Line			
Poverty (%)	19.14	14.61	26.39
N	16,020	7,740	540

Note. The Full Sample percentages show the characteristics of those students for whom basic demographic information was available in each of the 5 waves of the ECLS-K. The Analytic Sample (or complete case sample) describes those participants for whom all independent variables and outcome measures are present throughout the longitudinal panel. The final column highlights the substantive demographic differences between students whose reported health insurance changed (either gaining or losing) over the course of the panel. Sample sizes were rounded to the nearest ten due to use of restricted data.

Students whose health insurance coverage was unstable over the course of the panel were substantially more likely to have lower levels of parental education, and were much less likely to be White. To further highlight the socioeconomic differences in children's health coverage, both at baseline and over time, Figure 1 shows the proportion of students who are insured in each year broken down by students' eligibility for free and reduced-price lunch. Students in the lowest income group (free lunch) were much more likely to be uninsured at the outset of the survey, and were more reliant on public insurance coverage. The figure also shows that this subset of students experienced the greatest gains in insurance coverage (primarily in the form of public insurance), nearly matching the insured rate of non-FRL eligible students by the end of the study period.

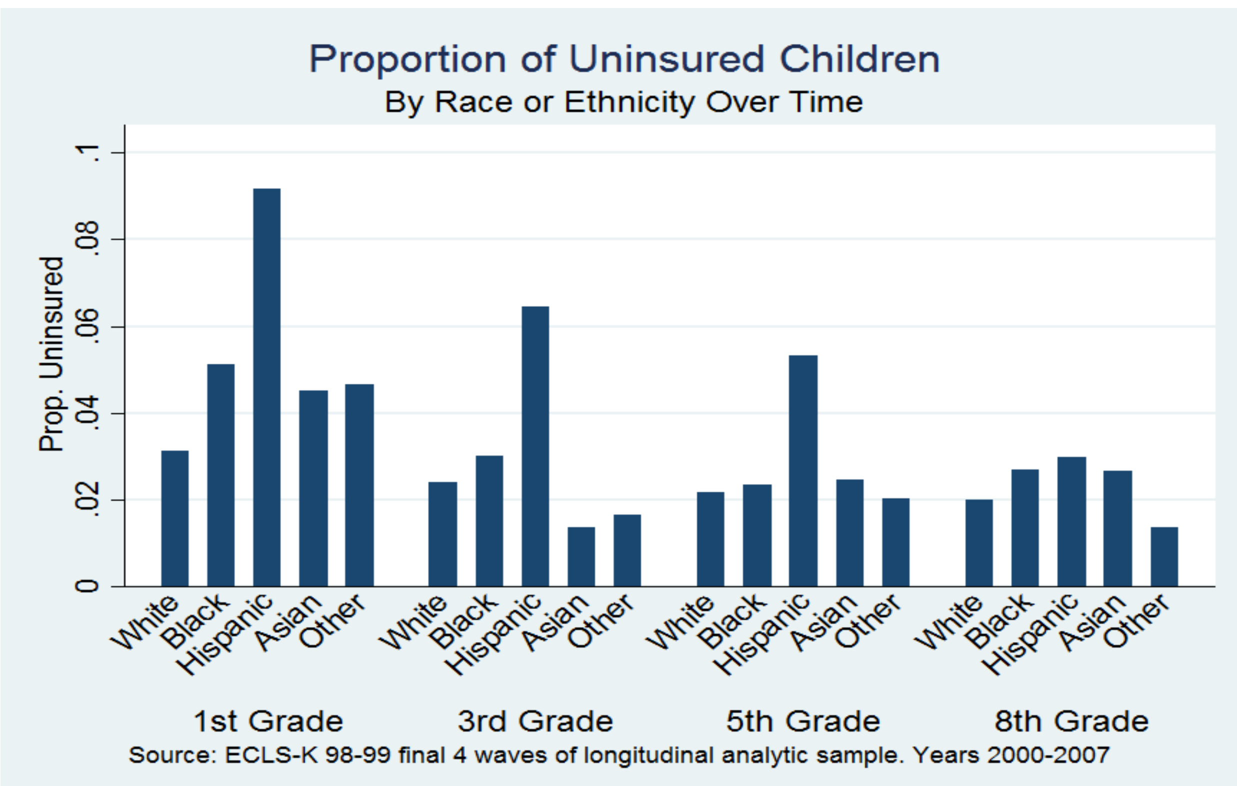
Figure 1: Insurance Coverage Gains by Free or Reduced-Price Lunch Status (2000-2007)



Source:ECLS-K 1998-1999

When we broke down coverage rates and types of insurance coverage by parental education level or race/ethnicity, the results followed a pattern similar to those depicted in Figure 2. We found that disadvantaged groups were more reliant on public insurance, and more likely to gain insurance over the course of the panel. Moreover, students from higher socio-economic status backgrounds were almost universally covered, while children of parents who did not complete high school were uninsured at rates above 20% in the 1998-1999 school year. However, following the full implementation of SCHIP, reports of children lacking insurance among those parents with the lowest levels of educational attainment dropped to approximately 5% by the final year of the survey (the 2006-2007 school year).

Figure 2: Reductions in Uninsured Children by Race or Ethnicity Over Time



Note. The figure above illustrates the changing proportion of students covered by any health insurance in the analytic sample over time.

Table 2: Descriptive Statistics by Baseline Insurance Status for Analytic Sample

	Uninsured (%)	Public (%)	Private (%)	N
Child Race or Ethnicity				
White	3.1	6.3	90.6	4,930
Black	5.9	31.0	63.1	820
Hispanic	8.5	17.1	74.4	980
Asian	4.6	9.2	86.2	390
Other	4.9	16.6	78.5	430
Parental Education				
Non-HS Grad	9.1	44.8	46.1	440
HS Grad	7.8	19.5	72.8	1,660
Voc-Tech, Some College	4.5	10.1	85.3	2,640
BA or More	1.2	1.8	97.0	2,810
Other Social Programs				
No SNAP	4.3	5.7	90.0	6,780
SNAP	4.4	58.3	37.3	770
No TANF	4.3	9.3	86.4	7,270
TANF	2.5	58.2	39.3	280
Mothers Employment				
35 HOURS OR MORE PER WEEK	3.9	8.7	87.4	3,680
LESS THAN 35 HOURS PER WEEK	3.6	7.8	88.6	1,780
LOOKING FOR WORK	7.7	43.5	48.8	170
NOT IN THE LABOR FORCE	5.3	16.0	78.7	1,910
Doctor Visit (Most Recent)				
NEVER	8.7	13.0	78.3	20
LESS THAN 6 MONTHS	3.5	13.5	83.0	3,700
6 MONTHS TO ONE YEAR	4.2	8.8	87.0	2,720
1 TO 2 YEARS	6.4	8.6	85.0	1,030
MORE THAN 2 YEARS	10.1	10.1	79.7	80
Dentist Visit (Most Recent)				
NEVER	9.9	15.1	75.0	250
LESS THAN 6 MONTHS	2.4	9.3	88.3	4,800
6 MONTHS TO ONE YEAR	6.1	13.2	80.7	1,980
1 TO 2 YEARS	11.2	16.7	72.1	420
MORE THAN 2 YEARS	13.0	23.0	64.0	100
Rating of Student Health				
EXCELLENT	3.8	7.5	88.8	4,180
VERY GOOD	4.8	13.2	82.1	2,370
GOOD	5.0	19.9	75.1	840
FAIR	6.4	30.0	63.6	140
POOR	5.0	20.0	75.0	20
Specific Public Insurance Program				
State CHIP	NA	57.6	42.4	270
Medicaid	NA	64.4	35.6	1,120
Military Health Insurance	NA	18.9	81.1	130

Total %	4.3	11.1	84.6	
N	322	837	6388	7,550

Notes: Descriptive statistics by insurance status in baseline year for analytic sample (Spring of first grade). All sample sizes are rounded to the nearest ten per NCES guidelines for restricted use data.

Table 2 presents baseline descriptive statistics for children falling into three major categories of health insurance coverage: no insurance, private/employer-based plans, or public insurance programs. Consistent with prior literature, Black and Hispanic students were much more likely to lack health insurance than their White peers (uninsured at roughly two and three times the rate of White students, respectively). Black and Hispanic students were also much more heavily reliant on public health insurance programs than White students (roughly 30% and 20% as compared to 6% for White students). Notably, even though they comprised a much smaller portion of the analytic sample, Native American students' insurance coverage patterns mirrored those of Black students, with nearly identical portions in each coverage category.

For our primary analytic sample, we also found that Black students were roughly 12% of the sample, and comprised an equal proportion of observations lacking insurance. Hispanic students accounted for roughly 17% of the sample, but nearly 36% of uninsured student observations. Hispanic students also experienced the largest reductions in lacking insurance over the course of the panel (Figure 2). Hispanic students had significantly lower baseline health ratings than non-Hispanic students in the sample (Black students' baseline health ratings were also poorer), even after controlling for income and parental education levels. Hispanic students also had higher rates of absence, and lower rates of dental care. Doctor usage was high from baseline.

Students who lacked health insurance in the first grade (the baseline year for our analysis) were more disadvantaged and lower performing by nearly every indicator. They missed 1.5 more school days (mean of 9 absences), were nearly 10 percentage points more likely to be chronically absent (31% vs. 40%) scored lower on both math (.29 SD) and reading (.23 SD) assessments, and had worse health ratings by .14 on a 5-point scaler (mean 1.77, where 1 is excellent). Uninsured students were 10 percentage points less likely to have had a recent doctor visit, and nearly 30 percentage points less likely to have recently seen a dentist. Average family incomes of students lacking health insurance were nearly half that of those who were covered (36,000/year vs. 71,000/year). Uninsured students were nearly twice as likely to be in poverty (31% vs. 17%), were half as likely to attend a private school (10 vs. 20%), and, on average, they attended significantly larger schools that were majority minority (51% vs. 36% for students with coverage). The percentage of uninsured students also dropped dramatically based on the level of maternal education, with rates near 10% among children of parents with less than a high school degree, and near zero for those with parents with a bachelor's degree or higher. Problematically, though unsurprisingly, children whose parents rated their child's health as less than good were also more likely to lack insurance.

Children who lacked health insurance were also much less likely to have seen a doctor or dentist recently, if at all. In fact, despite having poorer health, uninsured students were roughly 3 times more likely to have never seen a doctor or dentist, or to have seen neither in more than 2 years than students with public or private insurance. Notably, students whose parents reported reliance on public assistance programs (SNAP or TANF) had relatively low rates of being uninsured, but the highest rates of coverage by public insurance programs. This suggests that parents who are able to successfully navigate the system of attaining cash and in-kind welfare benefits do a better job than other low-income parents at finding insurance coverage for their children.

Analytic Method

To evaluate whether increased access to health insurance affects students' health and schooling outcomes, we employed a student-fixed effect model that exploits variation over time in a parent's reporting of the child's insurance coverage. We estimate an Ordinary Least Squares model that takes the following form:

$$(Y_{it} = \beta_0 + \beta_1 Covered_{it} + \boldsymbol{\omega_{it}}\beta_3 + \alpha_i + \gamma_t + \mu_{it} \ (1))$$

where Y_{it} represents the outcome of interest (test score, absences, doctor visits, dentist visits, or health rating) for student *i* in year *t*. $Covered_{it}$ is a binary variable indicating a parent's reporting of the child's insurance coverage that takes on a value of 1 if the parent reports that student *i* is covered by health insurance (including both private and public insurance) in year *t,* but was not covered. The $\boldsymbol{\omega_{it}}$ represents a vector of other resources that might influence students' insurance statuses, health, or academic outcomes, including income, family size, nutritional assistance receipt (SNAP), welfare support (TANF), mother's employment status, as well as controls for school type, size, and demographic composition. α_i and γ_t represent individual and year-fixed effects, respectively, accounting for characteristics of an individual that are constant over time and any trends within a year that are constant across individuals. Finally, μ_{ist} represents an individual error term.[6]

Results

Overall, we found that health insurance coverage resulted in increased utilization of medical and dental care, as well as marginally significant improvements in students' health ratings. We also found that insurance coverage increases student attendance, but there is no evidence that health insurance coverage gains improve in students' test scores in the short run. We discuss these findings below.

[6] To account for potential variation in the timing of effects, I also constructed alternative models that include lagged eligibility or coverage indicators, which allows for the estimation of longer-term benefits of improved health that may take time to set in. Lagged models mirror the ones described in the text, but include an indicator for whether the student was eligible during the previous survey administration. While this model allows for a more holistic examination of effects, it shortens the panel by one year, thus reducing power.

Effects of Health Insurance Coverage on Health Outcomes

(Doctor Visits, Dentist Visits, and Health Ratings)

Table 3 presents estimated effects of health coverage on health service utilization (doctor or dentist visits) and health status (parent rating of the child's health). Gaining health insurance was consistently followed by reports of more regular visits to doctors and dentists, but only a marginally significant increase, and only created a small change in a parent's perception of their child's health. The coefficients on coverage for recent doctor and dentist visits indicated that, all else equal, obtaining health insurance coverage is associated with a 6.4-percentage-point increase in the likelihood of having seen a doctor in the last calendar year, and only an 8.9-percentage-point increase in having recently visited a dentist. While these effects may seem modest, they represent roughly a doubling of the mean likelihood of recent health care visits for uninsured students. Effects on parents' perceptions of their children's health statuses were small and only marginally significant (a roughly 2-percentage-point increase in the likelihood of describing the child's health as good or better).

Table 3: Insurance Coverage Effects on Service Utilization and Health Status

	(1) Doctor Visit	(2) Dentist Visit	(3) Good Health
Covered	0.064***	0.089***	0.018+
	(0.022)	(0.022)	(0.009)
Observations	31,960	31,960	31,960
Student FE	YES	YES	YES
Year FE	YES	YES	YES

Notes: All models include controls for student family size, income, SNAP and TANF benefits, mother's employment status, school size, type, and demographics. Absences are not recorded in eighth grade survey year. Standard errors clustered at the student level (*** p<0.01, ** p<0.05, * p<0.1).

Effects of Health Insurance Coverage on Academic Outcomes (Test Scores and Attendance)

Table 4 presents the similarly mixed results from our primary models for the effects of students gaining health coverage on test scores and attendance. We find that gaining health insurance is associated with a significant increase in attendance—roughly two thirds of an additional instructional day in the year of insurance acquisition. However, test score benefits are statistically indistinguishable from zero. For comparison, attendance effects here are similar in magnitude to a recent study that found that severe carbon monoxide-level reductions drove an average per-student increase of 0.8 days per year (Currie, 2009).

Table 4: Effects of Gaining Insurance on Academic Performance and Attendance

	(1) Math Scale	(2) Read Scale	(3) Absences
Covered	-0.014	0.002	-0.644**
	(0.026)	(0.029)	(0.312)
Observations	23,829	23,829	23,829
Student FE	YES	YES	YES
Year FE	YES	YES	YES

Notes: All models include controls for student family size, income, SNAP and TANF benefits, mother's employment status, school size, type, and demographics. Absences not recorded in 8th grade survey year. Standard errors clustered at the student level (*** $p<0.01$, ** $p<0.05$, * $p<0.1$)

Robustness Checks

While our analytic strategy incorporates both student- and year-fixed effects with extensive controls, thus eliminating a broad set of alternative explanations for the association between health insurance and the outcomes we observed, it does not rule out the possibility that some unobserved time-variant parent or student characteristics drove both the change in insurance statuses and the outcome of interest. It is plausible that a child experiencing new health problems could motivate a parent to seek out public insurance options that had previously seemed less important for a healthy child. This scenario would be consistent with our findings that gaining coverage was associated with more doctor and dentist visits, but not necessarily improved health or academic performance, and would bias downward estimates of the benefits to coverage.

To test for the potential of pre-existing trends driving estimated effects, we conducted a placebo test, where we re-estimated our primary model, but replaced the coverage variable with an indicator that assumed the student gained insurance coverage in the year prior to actual coverage. The estimated value on this coefficient can be interpreted as pre-treatment trends or effects that precede the actual change in insurance status. While this model allows for a more holistic examination of effects, it shortens the panel by one year, thus reducing power. Results of the placebo test show no indication of positive trends in any of the outcomes of interest prior to gaining health insurance (see Table 5). In fact, while they are not statistically significant, each of the coefficients for the three areas where we found positive impacts (attendance, doctor visits, and dentist visits) indicated a slight downturn in the year prior to insurance enrollment.

Table 5: Placebo Test For Trends Predating Enrollment in Insurance Coverage

	(1) Math Scale	(2) Read Scale	(3) Absences	(4) Doctor Visit	(5) Dentist Visit	(6) Good Health
Placebo Gain	0.016	-0.031	-0.010	-0.007	-0.042	0.020
(year t-1)	(0.037)	(0.043)	(0.632)	(0.037)	(0.036)	(0.055)
Lose Insurance	0.032	-0.033	0.008	-0.008	-0.021	0.039
	(0.025)	(0.026)	(0.425)	(0.026)	(0.026)	(0.040)
Observations	27,130	26,910	21,820	27,450	27,480	27,510
Student FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: The coefficients on "Placebo Gain" represent the change in the specified outcomes associated with the year prior to a student's gaining health insurance coverage (year t-1). All models include controls for student family size, income, SNAP and TANF benefits, mother's employment status, school size, type, and demographics. Absences are not recorded in the eighth grade survey year. Standard errors clustered at the student level (*** $p<0.01$, ** $p<0.05$, * $p<0.1$)

While our primary analysis focuses on the effect of having health insurance coverage, setting aside potential differences in the types of coverage (private, employer-based vs. Medicaid, or specific state CHIP programs), results are robust to a focus only on the effects of public insurance enrollment, controlling for whether or not students also report having private coverage (Appendix B). Results are also robust to a variety of alternative sample restrictions, including the addition of eighth grade observations, all of which are available upon request.

Conclusions

One way of viewing the overall findings here is that, in the short run, expanded health coverage increased utilization of both health services (doctor and dentist visits) and schools (attendance), but not the outcomes associated with the services (measures of health and learning). An optimistic reading of these results could predict that, over time, the increased health treatment would result in improved health, and aggregate improvements in attendance would result in improved academic performance. Alternatively, it could be argued that the increased utilization of outputs (doctors, dentists, and schools) will not produce the improved outcomes until access to higher quality instruction or health care improves.

The marginal finding with respect to parental ratings of student health is perhaps unsurprising, as increased health care access could conceivably push perceptions of health status in competing directions. For example, if a child gets more regular doctor visits due to expanded health care coverage, and is subsequently diagnosed with a chronic illness such as asthma, the parent may view the child as less healthy because he or she is on medication and has a labeled disease. Conversely, a child who gains access to medical care that improves symptoms of illnesses that were previously undiagnosed or untreated may appear to be 'more healthy.' Consequently, the negating effects and loose construct of perceived health make it difficult to draw many conclusions from this seemingly key variable. Alternatively, the lack of effects here could be due to downward bias from a pre-existing trend, highlighted in the robustness check section of this study.

Our findings also suggest that experimental evaluations of health interventions, including major recent insurance expansions from the ACA, should look beyond health outcomes to quantify educational externalities. If we are to believe recent studies demonstrating the long-term economic benefits of early gains on test score measures of academic ability (e.g., Chetty, Friedman, & Rockoff, 2014; Hanushek, 2011), even a small boost in performance for such a large group of students could serve to substantially offset the costs of health interventions. Furthermore, evidence of the efficacy of health interventions in improving academic outcomes could inform efforts to address persistently vexing national education policy imperatives—thus elevating student achievement and narrowing persistent achievement gaps. Future research should also work to illuminate the differential benefits of health insurance expansions by race or ethnicity, particularly given the substantial gains among Hispanic students reported here.

APPENDIX

Appendix A: Coding and Other Details for Survey Instrument Items

Questions about Health Insurance Coverage

What kind of health insurance or health care coverage does {CHILD} have? By health insurance, I mean any kind of coverage that pays for health care expenses. Please do not include private plans that only provide extra cash while hospitalized. Does {he/she} have…

	YES	NO	REFUSE	DK
A private health insurance plan (from employer, workplace, or purchased directly through a state or local government program or community program)?	1	2	7	9
Medicaid (or name of state programs)?	1	2	7	9
CHIP (Children's Health Insurance Program) (or name of state program)?	1	2	7	9
Military health care/VA/CHAMPUS/TRICARE/CHAMP-VA?	1	2	7	9
Another government program (Indian Health Service, Medicare, State-sponsored health plan)?	1	2	7	9
No health insurance?	1	2	7	9

Note: There were minor changes to the prompt following the first-grade administration, though the core of each item remained intact. For example, the prompt for private health insurance plans did not include parentheticals. It read: "A private health insurance plan from employer, workplace, or purchased directly through a state or local government program or community program." Additionally, in the fifth- and eighth-grade administration, the "no health insurance" option was marked only if the respondent answered "no" to all other categories.

Question about Dentist Visits

How long has it been since {CHILD}'s last visit to a dentist or dental hygienist for dental care?

NEVER………………	1
LESS THAN 6 MONTHS………………	2
6 MONTHS TO YEAR………………	3
1 TO 2 YEARS………………	4
MORE THAN 2 YEARS………………	5
REFUSED………………	7
DON'T KNOW………………	9

Question about Doctor Visits

How long has it been since {CHILD}'s last visit to a clinic, health center, hospital, doctor's office, or other place for routine health care?

Probe: Routine health care may include check-ups or immunization appointments.

NEVER………………	1
LESS THAN 6 MONTHS………………	2
6 MONTHS TO YEAR………………	3
1 TO 2 YEARS………………	4
MORE THAN 2 YEARS………………	5
REFUSED………………	7
DON'T KNOW………………	9

Question about Student Health

Would you say {CHILD}'s health is…

EXCELLENT………………	1
VERY GOOD………………	2
GOOD………………	3
FAIR………………	4
POOR………………	5
REFUSED………………	7
DON'T KNOW………………	9

Appendix B: Effects of Child Health Insurance on Academic and Health Outcomes
(Public Insurance & Private Insurance Plans)

VARIABLES	(1) Math	(2) Read	(3) Absences	(4) Doctor	(5) Dentist	(6) Health
Public Ins.	-0.019	-0.010	-1.134*	-0.123***	-0.180***	0.016
	(0.024)	(0.028)	(0.583)	(0.038)	(0.038)	(0.034)
Private Ins.	-0.010	0.001	-0.748	-0.103***	-0.168***	0.024
	(0.023)	(0.026)	(0.491)	(0.035)	(0.036)	(0.031)
Observations	36,180	35,930	24,500	38,130	38,160	38,200
Student FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes: All models include controls for student family size, income, SNAP and TANF benefits, mother's employment status, school size, type, and demographics. Absences were not recorded in the eighth-grade survey year. Standard errors are clustered at the student level. Sample sizes vary based on availability of the data on the specified outcome (*** p<0.01, ** p<0.05, * p<0.1.).

REFERENCES

Ananat, E. O., Gassman-Pines, A., Francis, D. V., & Gibson-Davis, C. M. (2011). *Children left behind: The effects of statewide job loss on student achievement (National Bureau of Economic Research Working Paper No. 17104).* Retrieved from http://www.nber.org/papers/w17104

Angrist, J. D., Pathak, P. A., & Walters, C. R. (2011). Explaining charter school effectiveness. *American Economic Journal: Applied Economics 2013, 5*(4): 1-27. http://dx.doi.org/10.1257/app.5.4.1

Aizer, A. (2007). Public health insurance, program take-up, and child health. *The Review of Economics and Statistics*, *89*(3), 400-415.

Bansak, C., & Raphael, S. (2007). The effects of state policy design features on take-up and crowd out rates for the State Children's Health Insurance Program. *Journal of Policy Analysis and Management*, *26*(1), 149-175.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2007). From the cradle to the labor market?: The effect of birth weight on adult outcomes. *Quarterly Journal of Economics, 122*(1), 409-439.

Carmen, D. W., Proctor, B. D., & Smith, J. C. (2008). *Current population reports, P60-235, income, poverty, and health insurance coverage in the United States: 2007*. Washington, D.C.: U.S. Government Printing Office.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *The American Economic Review*, *104*(9), 2633-2679.

Cohodes, S. R., Grossman, D. S., Kleiner, S. A., & Lovenheim, M. F. (2016). The effect of child health insurance access on schooling: Evidence from public insurance expansions. *Journal of Human Resources*, *51*(3), 727-759.

Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economics Literature* 47(1), 87-122.

Currie, J., Decker, S., & Lin, W. (2008). Has public health insurance for older children reduced disparities in access to care and health outcomes? *Journal of Health Economics*, *27*(6), 1567-1581.

Currie, J., & Gruber, J. (1996). Health insurance eligibility, utilization of medical care, and child health. *Quarterly Journal of Economics, 111*(2), 431-466.

Dafny, L., & Gruber, J. (2005). Public insurance and child hospitalizations: Access and efficiency effects. *Journal of Public Economics*, *89*(1), 109-129.

Dastrup, S. R., & Betts, J. R. (2012). Elementary education outcomes and stress at home: Evidence from mortgage default in San Diego. Retrieved from http://www.terry.uga.edu/media/events/documents/SeminarFA14_Ross.pdf

Ding, W., Lehrer, S. F., Rosenquist, J. N., & Audrain-McGovern, J. (2009). The impact of poor health on academic performance: New evidence using genetic markers. *Journal of Health Economics, 28*(3), 578-597.

Dobbie, W., & Fryer, R. G. (2009). *Are high quality schools enough to close the achievement gap? Evidence from a social experiment in Harlem (National Bureau of Economic Research Working Paper No. 15473).* Retrieved from http://www.nber.org/papers/w15473

Egbuonu, L., & Starfield, B. (1982). Child health and social status. *Pediatrics*, 69(5), 550-557.

Festinger, T., & Duckman, R. (2000). Seeing and hearing: Vision and audiology status of foster children in New York City. *Journal of Behavioral Optometry*, *11*(3), 59-67.

Figlio, D. N., Guryan, J., Karbownik, K., & Roth, J. (2014). *The effects of poor neonatal health on children's cognitive development (National Bureau of Economic Research Working Paper No. 18846).* Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.104.12.3921

Fletcher, J. M., & Lehrer, S. F. (2009). *Using genetic lotteries within families to examine the causal impact of poor health on academic achievement (National Bureau of Economic Research Working Paper No. 15148).* Retrieved from http://www.nber.org/papers/w15148

Flores, G., Bauchner, H., Feinstein, A. R., & Nguyen, U. S. (1999). The impact of ethnicity, family income, and parental education on children's health and use of health services. *American Journal of Public Health, 89*(7), 1066-1071.

Forrest, C. B., Starfield, B., Riley, A. W., & Kang, M. (1997). The impact of asthma on the health status of adolescents. *Pediatrics, 99*(2), 1-7.

Fowler, M. G., Johnson, M. P., & Atkinson, S. S. (1985). School achievement and absence in children with chronic health conditions. *Journal of Pediatrics, 106*(4), 683-687.

Hadley, J., Holahan, J., Coughlin, T., & Miller, D. (2008). Covering the uninsured in 2008: Current costs, sources of payment, and incremental costs. *Health Affairs, 27*(5), 399-415.

Halfon, N., & Newacheck, P. W. (1993). Childhood asthma and poverty: Differential impacts and utilization of health services. *Pediatrics, 91*(1), 56-61.

Ham, J. C., & Shore-Sheppard, L. (2005). The effect of Medicaid expansions for low-income children on Medicaid participation and private insurance coverage: Evidence from the SIPP. *Journal of Public Economics*, *89*(1), 57-83.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review*, *30*(3), 466-479.

Hoxby, C., & Rockoff, J. (2004). *The impact of charter schools on student achievement* (Unpublished manuscript). Harvard University, Boston, Massachusetts.

Howell, E., Decker, S., Hogan, S., Yemane, A., & Foster, J. (2010). Declining child mortality and continuing racial disparities in the era of the Medicaid and SCHIP insurance coverage expansions. *American journal of public health*, *100*(12), 2500-2506.

Howell, E. M., & Kenney, G. M. (2012). The impact of the Medicaid/CHIP expansions on children: a synthesis of the evidence. *Medical Care Research and Review*, *69*(4), 372-396.

Jackson, S. L., Vann, W. F., Kotch, J. B., Pahel, B. T., & Lee, J. Y. (2011). Impact of poor oral health on children's school attendance and performance. *American Journal of Public Health, 101*(10), 1900 1906.

Kenney, G., Dubay, L., Hill, I., Sommers, A., & Zuckerman, S. (2005). *Congressionally mandated evaluation of the State Children's Health Insurance Program: Final report to Congress.* Washington, D.C.: Mathematica Policy Research, Inc.

Kenney, G. M., Haley, J., Pan, C., Lynch, V., & Buettgens, M. (2016). *Children's coverage climb continues: Uninsurance and Medicaid/CHIP eligibility and participation under the ACA*. Retrieved from http://www. urban. org/sites/default/files/alfresco/publication-pdfs/2000787-Childrens-Coverage-Climb-Continues-Uninsurance-and-Medicaid-CHIP-Eligibility-and-Participation-Under-the-ACA.

Levine, P. B., & Schanzenbach, D. W. (2009). The impact of children's public health insurance expansion on educational outcomes. *Forum for Health Economics and Policy, 12*(1), 1-26.

Marton, J. (2007). The impact of the introduction of premiums into a SCHIP program. *Journal of Policy Analysis and Management, 26*(2), 237-255.

Oreopoulos, P., Stabile, M., Walld, R., & Roos, L. L. (2008). Short-, medium-, and long-term consequences of poor infant health: An analysis using siblings and twins. *Journal of Human Resources, 43*(1), 88-138.

Orfield, A., Basa, F., & Yun, J. (2001). Vision problems of children in poverty in an urban school clinic: Their epidemic numbers, impact on learning, and approaches to remediation. *Journal of Optometric Vision Development, 32*(3), 114-141.

Peterson, P. E., & Hassel, B. C. (Eds.). (1998). *Learning from school choice.* Washington, D. C.: Brookings Institution Press.

Rosenbach, M., Ellwood, M., Irvin, C., Young, C., Conroy, W., Quinn, B., & Kell, M. (2003). *Implementation of the State Children's Health Insurance Program: Synthesis of state evaluations (background for the report to Congress).* Washington, D.C.: Mathematica Policy Research, Inc., for Centers for Medicare and Medicaid Services, Department of Health and Human Services.

Royer, H. (2009). Separated at girth: US twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics, 1*(1), 49-85.

Sebelius, K. (2010). Rising to the challenge: Tools for enrolling eligible children in health coverage. *Health Affairs, 29*(10), 1930-1932.

Seid, M., Varni, J. W., Cummings, L., & Schonlau, M. (2006). The impact of realized access to care on health-related quality of life: A two-year prospective cohort study of children in the California State Children's Health Insurance Program. *The Journal of Pediatrics*, *149*(3), 354-361.

Shaefer, L. H., Grogan, C. M., & Pollack, H. A. (2011). Transitions from private to public health coverage among children: Estimating effects on out-of-pocket medical costs and health insurance premium costs. *Health Services Research*, *46*(3), 840-858.

Shone, L. P., Dick, A. W., Brach, C., Kimminau, K. S., LaClair, B. J., Shenkman, E. A., . . . Bronstein, J. (2003). The role of race and ethnicity in the State Children's Health Insurance Program (SCHIP) in four states: Are there baseline disparities, and what do they mean for SCHIP? *Pediatrics, 112,* 521-532.

Swain, W. (2013). "F" is for foreclosed: The effects of the housing crash on North Carolina student achievement. *Sanford Journal of Public Policy, 4*, 3-23.

Wolfe, B., & Scrivner, S. (2005). The devil may be in the details: How characteristics of SCHIP programs affect take-up. *Journal of Policy Analysis and Management, 24*(3), 499-522.

Yeung, R., Gunton, B., Kalbacher, D., Seltzer, J., & Wesolowski, H. (2011). Can health insurance reduce school absenteeism? *Education and Urban Society, 43*(6), 696-721.

CHAPTER 2

SCHOOL-BASED BENEFITS OF SCHOOL-BASED HEALTH SERVICES: TEST SCORE AND ATTENDANCE EFFECTS ON NON-URBAN SBHCS

Introduction

Combined with increased federal infrastructure funding, recent expansions of health insurance coverage for children of low-income families have facilitated small but substantial increases in the availability of School-Based Health Centers (SBHCs) across the country. In addition to continued funding for the State Children's Health Insurance Program (SCHIP), the 2010 Affordable Care Act provided $200 million for the establishment and expansion of SBHCs through competitive grants to Local Education Agencies (LEAs) and health service provider partners. In Tennessee, approximately $6 million of new federal grant awards have funded space for an estimated 67 new or renovated school-based centers across the state. Like those that existed prior to the grant program, the majority of these new SBHCs in Tennessee are located in low-income, low-population density communities.

SBHCs typically provide a combination of primary, mental health, nutritional, and dental services, which, in many of the high-need areas they serve in Tennessee, could help mitigate substantial barriers to school attendance or academic success. While the primary goal of SBHCs is to improve children's health and quality of life, the well-documented associations between poor health and poor academic performance (e.g., Black et al., 2007; Ding, Lehrer, Rosenquist, & Audrain-McGovern, 2009; Figlio et al., 2014; Fletcher & Lehrer, 2009; Oreopoulos et al., 2008) suggest that well-targeted SBHCs might improve attendance and student learning. However, few empirical studies have evaluated whether health policy interventions can actually improve school-based outcomes, especially in poor rural areas, where medical care is often in short supply (Hartley, 2004; Valet, Perry, & Hartert, 2009).

The logic behind providing health services in schools to improve student health and academic performance is relatively simple. Low-income students suffer from higher rates of treatable health problems, and are substantially less likely to have access to primary care services than their middle-class peers (e.g., Orfield, Basa, & Yun, 2001; Festinger & Duckman, 2004; Egbuonu & Starfield, 1982; Forrest, Starfield, Riley, & Kang, 1997). While the concept of school-based health services dates back to as early as 1900, recent expansions of healthcare and an influx federal financial support have led to sharp upticks in the practice of school-based medicine during the last ten years (Keeton, et al., 2012). These centers generally represent a partnership between a school and an external healthcare provider that ultimately bills Medicaid for services rendered to covered students. This partnership model allows school systems to provide health benefits to students without directly employing or compensating medical personnel, and gives pediatric health service providers easier access to their target population.

There are at least three mechanisms by which expanded access to school-based health services could improve students' academic outcomes. First, low-income students who suffer from treatable health problems, such as visual problems (Orfield et al., 2001; Festinger & Duckman, 2004), hearing impairments (Egbuonu & Starfield, 1982), or asthma (Forrest et al., 1997; Halfon & Newacheck, 1993), could have illnesses and disabilities diagnosed and treated more effectively with increased access to healthcare. With their achievement-blocking symptoms ameliorated (Ding et al., 2009; Fletcher & Lehrer, 2009), they perform better on standardized assessments. Second, decreased healthcare costs realized by families who gain access to free medical care from SBHCs might increase the amount of resources families (Shaefer, Grogan, & Pollack, 2011) could put toward educational goods or other expenditures that improve student achievement. Finally, the placement of health professionals on school campuses could reduce the likelihood of unnecessary absences for doctor visits, especially for students with chronic illnesses who could receive care at school and return to class (e.g., Wade et al., 2008). None of these specific pathways are empirically tested here, though improved attendance—and especially reductions in excused absences—could serve as a strong proxy for improvement in child health. Instead, I focus here on describing the general services SBHCs provide, and deriving a clean estimate of their impacts on students' school-based outcomes.

I hypothesize that non-urban school-based health center openings will improve student attendance and academic performance because their services promote healthy school attendance for students who might otherwise stay home sick or attend school tired and unhealthy. The sample size of students in the more rural communities examined here are smaller than those used in prior work focused on urban SBHCs. However, utilization studies have found that students in rural areas tend to use SBHC services more frequently than their urban peers (Wade et al., 2008), indicating that the strength of treatment could be more substantial. Gains in academic performance and attendance should be larger among female students, who research has shown are more likely to actually utilize the SBHCs (Kerns et al., 2012), as well as students with low baseline attendance, which serves as a crude proxy for students having (potentially treatable) health problems (e.g., Forrest et al., 1997; Fowler, Johnson, & Atkinson, 1985; Halfton & Newacheck, 1993).

This study exploits 6 years of rich longitudinal student-level administrative records from Tennessee, which I pair with novel school-level data from the Tennessee Office of Coordinated School Health (TNCSH), which documents SBHC openings and current staffing levels. I also examine survey data from the School-Based Health Alliance to describe some of the variations in health service provision in non-Urban SBHCs across the state. I estimate effects of SBHC openings on student test scores and attendance, as well as additive effects over time from students' exposure thereto. Variation in student access to SBHCs results from SBHC openings, and student movement across schools, facilitating both a school-level difference-in-differences analysis, as well as a student-fixed effect modeling of SBHC impacts.

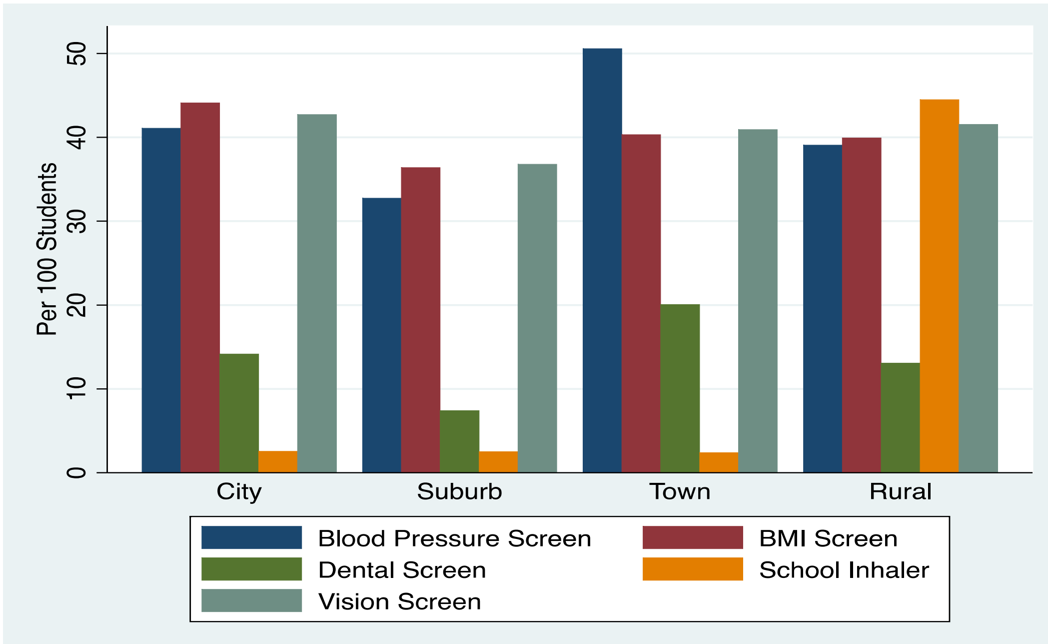
In the sections that follow, I describe the policy context of Tennessee's non-urban SBHCs, including a brief description of two rural districts for which I have more detailed information on specific services offered in their SBHCs. I then present a brief review of extant literature evaluating the impacts of SBHCs on student achievement and attendance, and outline the data, analytic strategy, and results of the two primary research questions of this study: (1) To what extent does increased access to school-based health centers improve students' test scores and absenteeism? (2) Do benefits from school-based health centers differ by student gender or baseline attendance rates (a proxy for student health)?

Policy Context of Tennessee's Non-Urban SBHCs

Tennessee presents a relatively unique opportunity to examine the impacts of SBHCs outside of major urban areas, which have been the primary focus of prior studies. Since the state's General Assembly passed the Coordinated School Health Improvement Act of 1999, the Commissioner of Education has been required (in consultation with the Tennessee Department of Health), to present an annual report on school health that outlines the efforts and needs of each LEA to the Governor and General Assembly. The Act also resulted in the establishment of a School Health Coordinator position for each district in the state; this individual is tasked with developing plans that "give priority to school health as a means to assist in meeting the education performance indicators," including attendance and test scores (T.C.A. § 49-5-415(a)(3)). In 2006, the Coordinated School Health Expansion and Physical Activity Law (T.C.A. § 49-6-1022) provided an additional $15 million in state funding to support coordinated school health activities across the state. This facilitated continued commitment to both school health staffing and data collection throughout the state. While the Act did not specifically emphasize the importance of rural areas, its focus on districts (the vast majority of which are non-urban) ensures disproportionate per-capita support of student health outside the major cities of Tennessee, and systematic centralized documentation of service provision in geographically dispersed rural and suburban districts.

Data from an annual survey of school-based health services conducted by the Tennessee Office of Coordinated School Health provides further insight into the school health infrastructure in the state, and in variations by district urbanicity. District-level responses to annual surveys provide rich health service data on over 100 metrics of school-based health service provisions, staffing, and utilization per year, showing a moderately robust school health infrastructure with some substantive variations across the state. Figure 3 visually depicts the distribution of select services based on district urbanicity, indicating that some health services, such as BMI or vision screenings, are consistently provided on a per-student basis, regardless of district urbanicity. On the other hand, some services, like nurse home visits, telemedicine, or access to an asthma inhaler at school, are considerably more common in more remote rural areas.

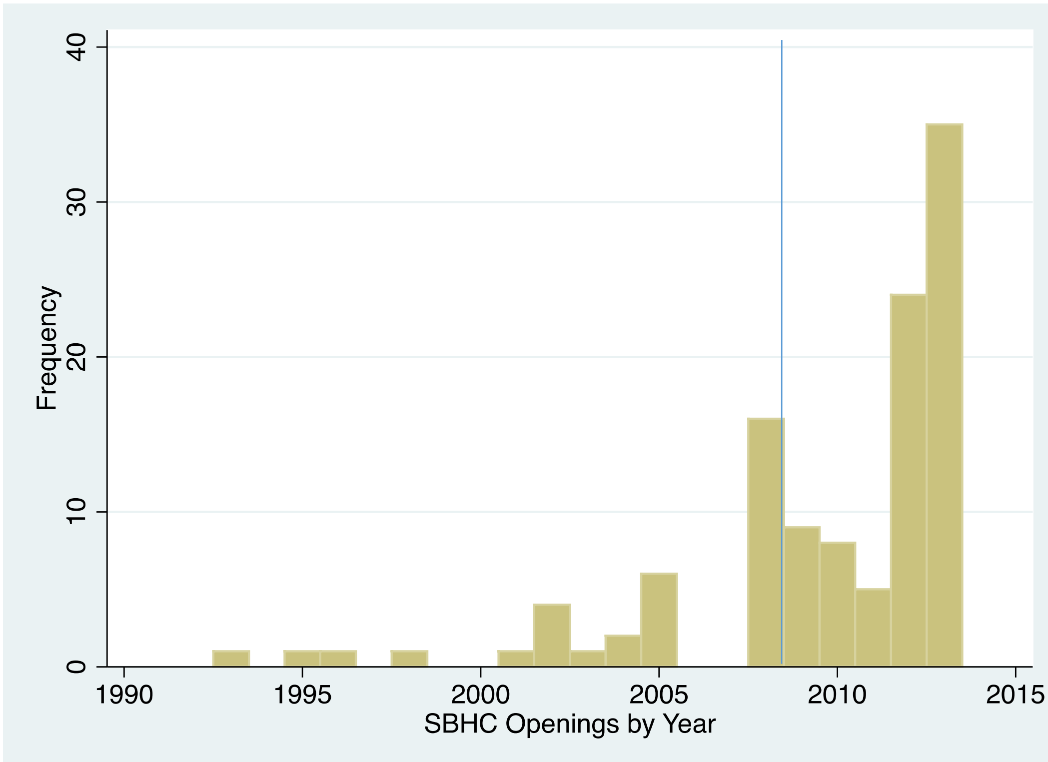
Figure 3: Differences in Health Services (Per 100 Students) by District Urbanicity



Source: Survey of District Coordinated School Health Officers (2013-14). Incomplete data from Memphis/Shelby County omitted from "city" totals

In addition to the general efforts of coordinated school health officers, student access to SBHCs has increased substantially in recent years across the state. While no health service data is collected explicitly at the school level, importantly, the 2013-14 school-based health services survey asked coordinators to document the opening dates, funding sources, and personnel for every active SBHC in the district, as well as the names of the schools where they operate. These dates allow for the construction of a longitudinal school-level SBHC dataset that I pair with school and student-level administrative records for the primary analyses in this study. Figure 4 illustrates the variation in SBHC openings in Tennessee over time, including a rapid uptick in the most recent years.

Figure 4: Frequency of Tennessee School-Based Health Clinic Openings by Year



Source: The above figure illustrates the reported opening dates of SBHCs in the analytic sample derived from the 2014 School-Based Health Services Survey, administered by the Office of Coordinated School Health. All of the SBHCs indicated in this figure were still active at the time of this study. The vertical line indicates the beginning of the panel. The 47 non-urban SBHCs that opened during this period constituted the treatment group in the analytic sample.

This study's primary analyses focus on the attendance and test score impacts of SBHC openings in the 47 non-urban Tennessee schools that opened centers between 2008 and 2014, regardless of whether they were supported by HRSA Grants.[7] However, in order to better understand what happens within rural SBHCs in the state of Tennessee, it is illustrative to examine more closely the services offered in two grant-winning districts that completed extensive surveys

[7] Notably, 23 of the 76 clinics in the school-level sample that opened after 2011 (the first year of HRSA Grants) report funding from the HRSA grant program described above, in spite of the lack of data on Nashville- and Memphis-based clinics. While not discussed in detail here, the SBHC Capital Grant awards drive considerable variation in access to SBHCs over time through new openings. Appendix A outlines the awards, districts, and partner organizations for the two waves of grants.

(SBHA, 2015) about the individual services and staffing in their SBHCS, beyond the general staffing information collected by the Office of Coordinated School Health. From July 2014 to May 2015, the School-Based Health Alliance conducted a national census of 2,315 SBHCs, identified through a mix of State Affiliate Rosters, State Government Office Rosters, School-Based Health Alliance membership lists, and daily compilations of online news articles about SBHC openings or closings since 1998. The extensive survey, funded through an agreement with the HRSA division of the Department of Health and Human Services, had a response rate of 82%, representing roughly 1,900 centers nationally.[8] Here, I focus on two non-urban eastern Tennessee districts, whose reporting was particularly thorough.

The two neighboring districts in the poor eastern region of the state both have more than 50% of their students enrolled in the federal free lunch program. These communities in the Appalachian Mountains have high rates of child poverty, low levels of insurance, few doctors per capita, and high rates of asthma. While the state of Tennessee did not expand Medicaid after the Affordable Care Act, thus leaving large portions of adults in the region uninsured, the state does have a relatively robust public SCHIP Program, upon which students in these low-income eastern counties heavily rely. Both of the case-study districts won Federal SBHCA capital grants, but the services they report providing differ substantially. McMinn County was awarded $402,006 in 2011, which it used to open nine new SBHCs, and Monroe County was awarded $200,000 in 2013, which was used toward renovating thirteen existing centers. In spite of the smaller federal investment, data from the School-Based Health Alliance's 2014 National Census of SBHCs indicate that the SBHCs of Monroe County offered more extensive services with longer hours to more community members,

[8] The School-Based Health Alliance (SBHA) made multiple attempts to confirm continued existence of centers with delayed responses to the survey. They requested that all surveys be completed by "the person who is most knowledgeable about the care provided in the health center, such as the SBHC administrator, nurse practitioner, or clinical director."

from a broader set of personnel than the centers in McMinn County. While I cannot rule out the potential that SBHCs in areas that failed to complete the survey provided more or less care than the centers in these two districts, in the context of prior qualitative studies and national survey data, they can be understood as representing a medium-high (Monroe) and a medium-low service (McMinn).

Data from the School-Based Health Alliance's 2014 National Census of SBHCs indicate that the centers in both Monroe and McMinn Counties operated inside school buildings with a typical external health provider partner, but varied with respect to the services they provided and who they treated. All centers in both districts allowed students who were enrolled in other schools to access the SBHC services, as could faculty and staff, but not out-of-school children. Only one SBHC in the two districts allowed for a student's siblings or family members, though eight of the thirteen centers in Monroe County provided services to the immediate family of faculty and staff. In both districts, the primary partner organization was a Federally Qualified Health Center (FQHC) that met all of the Health Center Program requirements. In Monroe County, SBHCs were all open five days a week for a total of 35 hours, while McMinn SBHCs operated only 2 days per week for a total of 8 hours. Only one center reported providing services before the school day began, though all the SBHCs in Monroe County reported having a prearranged source for after hours: on-call services. McMinn County SBHCs were classified as primary care only, while Monroe County also reported having behavioral health providers on staff. Monroe County SBHCs also reported providing a broad set of specific services on site, including a long list of vaccinations and screenings, some reproductive and relationship counseling, eye exams, and glasses, as well as referrals for services not provided on site, such as contraceptives (the distribution of which are reported as prohibited in all SBHCs outside of Memphis). While the center itself does not have oral health personnel, the students in the Monroe County centers are reportedly "transported by van to the county dental clinic" to receive dental screenings, education, guidance, and risk assessments. By contrast, McMinn County SBHCs offer a more restricted set of services, and rely more heavily on referrals.

While I do not have comparable data on health services provided or utilized in all the SBHCs that make up the "treatment" schools in the analytic sample, most centers submitted a list of the personnel types (e.g., nurse practitioners, social workers, and licensed counselors) currently staffing the SBHC in the 2014 school-based health services report. The characteristic staffing of Tennessee's rural SBHCs is consistent with a strong general emphasis on primary care, with many centers having multiple staff members with the ability to diagnose and prescribe treatments on site (e.g., doctors, nurse practitioners, and physician's assistants), only a few licensed counselors, and no documented dental affiliations. Appendix A depicts the staffing profiles of SBHCs in the analytic sample, with each bar representing an individual school's total count of documented SBHC staff, weighted by reported time at the school and sorted by category of credentials. Most SBHCs have at least a part-time nurse practitioner (NP) on staff, 45% report having a full-time NP, and roughly half have multiple full-time health professionals on staff.

Brief Review of SBHC Literature

Empirical studies have documented a strong relationship between poor health and diminished academic performance, sometimes using rigorous research methods to support causal claims. Several innovative studies in health economics have leveraged genetic differences between siblings living in the same home to estimate significant causal effects of illnesses on academic achievement and student grade point averages (Ding et al., 2009; Fletcher & Lehrer, 2009). Earlier, more conventional econometric analyses controlling for a range of covariates demonstrated that chronic illness is associated with a rate of absence from school that is more than double that of the average student (Fowler et al., 1985). These higher rates of absence associated with chronic illness are not generally related to lower achievement after controlling for socioeconomic status, though for certain illnesses like sickle cell disease, the negative impacts appear to be universal across income groups (Fowler et al., 1985). A 2012 study published in the *Journal of Pediatrics* found that students' dental health issues were associated with dramatically worse school performance and psychosocial wellbeing, even after controlling for a rich set of socioeconomic characteristics; this suggests that narrowing disparities in treatable health problems might narrow achievement gaps, even as other socioeconomic disadvantages persist (Guarnizo-Herreño & Wehby, 2012). However, evaluations of the capacity of school-based health interventions to reverse these negative effects (or any health interventions for that matter) are rare and mixed with respect to findings. For the purpose of concision, I will focus here on empirical evaluations that specifically examine school-based health centers' effects on academic outcomes.

Recent empirical analyses of SBHCs' effects on school-based outcomes are rare, are generally focused on high schools, and are frequently limited to a single large, urban school system. A 2004 review of links between SBHCs and academic outcomes conducted by Geierstanger, Amaral, Mansour, and Walters showed a mix of positive and null findings for impacts of provision and/or usage in early research that generally relied on less rigorous methods that were subject to extensive selection bias. More recently, Walker and colleagues' 2010 study of high school-based health clinics in Seattle, Washington, found benefits from SBHC utilization for both attendance and GPA. Analyses of differential effects based on the classifications of personnel whom students saw indicated that interactions with mental health professionals were associated with GPA increases, while care from medical professionals improved attendance. However, while the study does construct a propensity score-matched sample based on student characteristics, it is difficult to rule out the prospect of endogenous selection's driving perceived effects in what is ultimately a cross-sectional correlational study. Similarly, Kerns and colleagues' 2012 student-level study of high school

SBHCs in an unnamed urban school district found that moderate usage of SBHCs was associated with lower rates of dropout after controlling for students' propensity to utilize the SBHC. Importantly, the study also found that female students were much more likely to frequently utilize SBHCs, though similar concerns remain about the potential for reverse causality, where similarly situated students make better use of the SBHC because they are more committed to working towards graduation.[9]

Lovenheim, Reback, and Wedenoja (2016) conducted a high school district-level analysis and found that SBHC openings resulted in substantial reductions in teen births, especially where clinics provided access to contraception, though there were no significant effects on graduation rates. The study used rich longitudinal data on district-level SBHC openings, health service provision, and graduation rates to approximate a community-level causal estimate in a difference-in-differences framework. In their discussion, the authors note that, with respect to graduation rates, one plausible explanation for the null finding could be that high school is too late of a point for additional health services to impact students' trajectory, suggesting that earlier interventions like those examined here could be more effective for traditionally disadvantaged students.

By contrast, Rochmes' (2016) working paper using the National Longitudinal Study of Adolescent Health finds that school provision of preventative health services is associated with higher GPAs, greater attendance, and lower course failure rates. The author also counts increased likelihood of high school graduation as a benefit of higher levels of school-based health services, but the effects are not significant in models that account for other school-level characteristics. The study

[9] In fact, in an update on Kerns et al. (2012), published in response to criticism for inappropriately accounting for time-dependent survival bias, the authors of the original study re-estimate the models using the same data, and find no evidence of significant reductions in dropout rates.

acknowledges the potential for endogeneity in what is ultimately a cross-sectional analysis with rich controls at the student and school levels.

In regards to this study, the most similar analysis in design (e.g., a longitudinal analysis with student-level elementary and middle-grade test score data) and timing (e.g., it includes observations post-ACA implementation and SBHC Capital Grants) is Reback's and Cox's 2016 working paper, which uses data from New York City—possibly the sharpest contrast in context to the non-urban SBHCs of Tennessee. Like this study, Reback and Cox estimate the effects of SBHC openings on student test scores and attendance (as well as special education placements)[10] in a difference-in-differences framework. Because of the endogeneity threat posed by New York City Public Schools' extensive school choice system, the study employs a series of strategies to account for potential negative sorting of students who have become ill in schools that offer health services. While findings are sensitive to model specification, the authors' preferred model finds that student test scores improve in response to SBHC openings (effect size ~.03-.05), but find no evidence that SBHC openings impact attendance rates and Special Education placements (SPED). The study also finds substantial heterogeneity in SBHC impacts, such that girls benefit more from openings than boys, as do students with lower baseline test scores and attendance rates. Notably, this study's focus on non-urban centers, where schools are more geographically dispersed and attendance is tied to students' residence, radically reduces the threat of students' sorting into schools that open SBHCs, because parents' associated moving costs are so much higher than the cost of transferring one's child to another city school that could open an SBHC as nearby as a block away.

This study contributes to the research and policy communities' understanding of SBHCs in several important ways. First, to my knowledge, this is one of only two empirical studies of academic

[10] I also examined the relationship between openings and special education placements not reported here. The results were not significant at conventional levels.

impacts of SBHC openings in primary (K-8) schools, the other of which falls within the potentially unique policy context of New York City (Reback & Cox, 2016). Second, it is the first to specifically examine non-urban SBHCs, despite the fact that such SBHCs represent roughly half of all centers operating nationally (Lovenheim et al., 2016) and assess a potentially different theory of action in these areas, where physicians may be less likely to supply care.[11] Third, it is the first study to use longitudinal administrative data from a full state (excluding urban centers). Finally, this contemporary panel (data through 2014-15) is one of only two empirical studies to estimate impacts in the context of recent substantial investments in a healthcare landscape after ACA implementation.

Data & Sample

Data Sources

This study utilized administrative data obtained from the Tennessee Department of Education (TDOE) and maintained by the Tennessee Education Research Alliance (TERA) at Vanderbilt University's Peabody College. It has been merged with data from the Tennessee Office of Coordinated School Health (TNCSH). Rich longitudinal data available for this study included student standardized test scores on the Tennessee Comprehensive Assessment Program (TCAP), and demographics from the 2008-9 to the 2013-14 school years. For the primary analysis, these data were combined with data from the Office of Coordinated School Health's (TNCSH) mandatory annual surveys of district school health coordinators on provided services, which include information on the placement and number of school-based health clinics, the employed personnel,

[11] The inclusion of data from outside school health service provision allows for a rough estimation of the degree to which clinics address gaps in access, as well as the estimation of effects with a better-defined counterfactual.

the availability of mental health or dental services, staff credentials, and the year each SBHC opened.[12]

School-level academic information came from multiple sources, including state school accountability reports, the National Center for Education Statistics' Common Core of Data, and by aggregating individual student- and teacher-level information at the school level. These school files contained typically-used information, such as level of schooling, school size, proficiency rates, as well as select student and teacher demographic information. I also incorporated information from annual U.S. census data to account for access to health services outside of schools, specifically the per-capita number of pharmacies, mental health, and physical health practitioners in the school's zip code.

Variables

The primary dependent variables in this study were annual student-level test scores on the states' NCLB-mandated TCAP exam, (math, reading, science, and social studies exams standardized within subject, grade, and year to have a mean of 0 and SD of 1) and attendance records (separate counts of excused and unexcused absences). The primary independent variable was a school-level indicator for whether or not the school where a child is enrolled has an SBHC open in a given year.[13] Other covariates included a set of school characteristics (total enrollment; percentages of students who identify as Hispanic, Black or non-native English speakers; and number of students participating in free or reduced-price lunch), community-based health service characteristics (the per-capita number of pharmacies, mental health practitioners, and non-mental health practitioners

[12] Future work will examine the viability of analyses using cross-sectional data that are also available at the district level on a variety of student health metrics, including obesity, oral health, and asthma.

[13] This SBHC indicator is used to calculate the years-of-exposure variable, which counts the number of years the student was enrolled in a school with an SBHC in an alternative analytic strategy described below.

within the school's zip code), and student characteristics (indicator variables for race/ethnicity, gender, native language, grade level, free or reduced-price lunch status, and an indicator for whether a student is new to a given school in a given year).

Analytic Sample Construction

I estimated the main effects of SBHC openings under two primary sample restrictions: one that limited the sample to students in schools that ultimately opened SBHCs during the course of the analytic sample, and one that matched schools that operated SBHCs to schools that never operated SBHCs, but were similar on observable characteristics at baseline. The "ever SBHC"[14] sample limitation accounted for the possibility that schools that opened SBHCs were simply too unique to compare, even against schools that were otherwise statistically identical. However, in addition to radically reducing sample size and statistical power for the school-level effect estimates, this sample restriction under-emphasized the effects of the most recent round of SBHC openings, as they had no comparison group in the final year, where all schools in the sample ultimately had operating SBHCs. To avoid the potential of the sample's being overly restrictive, maintaining a counterfactual through the entire panel, and increasing statistical power, the "matched sample" attempted to reduce the potential for bias from unobserved characteristics, while still capturing the primary variation of interest—changes in outcomes in treatment schools compared with those in similarly situated untreated schools.

Because the schools that opened non-urban SBHCs may have differed from other schools in the state in important ways that would affect the trajectories of students who attend them, I used a propensity score-matching technique to construct a sample of schools that did not participate in the program, but were equally likely to do so. Several studies have found that, in combination with a

[14] To reduce the potential downward bias introduced by inaccurate entries of SBHCs with early opening dates, but maximize the length of the panel, I dropped schools that have SBHCs with opening dates prior to the outset of the analytic sample (prior to 2009).

traditional difference-in-difference model, this method of matching on observables substantially reduces the susceptibility to bias from violations of the common trends assumption, discussed in greater detail below (e.g., O'Neill et al., 2016; Heckman, Ichimura, & Todd, 1997; Abadie, 2005). In the construction of the primary analytic sample, I used probit regression to estimate the likelihood of ever operating an SBHC, and the predicted school and community characteristics in the 2008-09 school year (the first year of full data availability). Specifically, the right-hand side of the equation includes the schools' average teacher test score value added (TVAAS Level); the average student math and reading scores on TCAP exams; and the percentage of students who are classified as native English speakers, Hispanic, Black, free or reduced-price lunch recipients, or those receiving special education services. I also accounted for several zip-code level community characteristics, including the poverty rate, the densities of pharmacies, mental health practitioners, and primary health practitioners, and the NCES urbanicity scale.

Among the 982 non-urban schools in Tennessee with complete data, indicators of poverty and rurality were significant predictors of opening an SBHC by 2014, such that schools that were more rural, had higher rates of poverty, and higher percentages of students participating in the federal free or reduced-priced lunch program were more likely to open SBHCs. All else being equal, the schools that opened SBHCs also had smaller proportions of students receiving special education services, and marginally significant larger proportions of Black and Hispanic students. Notably, though differences in community health service providers were not statistically significant in the multivariate context, which is likely due to their high correlation with poverty rates and urbanicity, simple t-test comparisons indicated that SBHCs were located in schools with lower per-capita mental and primary health practitioners.[15] Specific results of the propensity score estimation are included in Table 6.

Table 6: Probit Regression Predicting Schools' Opening an SBHC by 2014

Probit Regression: Predicted Outcome= Ever SBHC

	Coef.	Std. Err.	Z		P>z
Average TVAAS Level	-0.182	0.143	-1.27		0.203
% Native English	1.681	3.956	0.42		0.671
% Immigrant	12.838	7.806	1.64		0.100
% Black	9.278	4.824	1.92		0.054
% Hispanic	10.704	5.826	1.84		0.066
% FRPL	10.331	4.828	2.14		0.032
% Free Lunch	1.363	0.695	1.96		0.050
% SPED	-3.165	1.276	-2.48		0.013
Reading TCAP	0.563	0.516	1.09		0.275
Math TCAP	-0.871	0.412	-2.11		0.034
Urbanicity Scale	-0.063	0.026	-2.39		0.017
Zip Code Poverty Rate	-1.991	1.161	-1.72		0.086
Pharmacies	0.037	0.047	0.79		0.432
Mental Health Professionals	-0.092	0.188	-0.49		0.626
Non-Mental Health Professionals	0.008	0.010	0.78		0.434
Constant	-12.263	5.667	-2.16		0.030
Number of Schools =	982		Pseudo R2	=	0.0884

After matching on propensity scores in a nearest neighbor framework with a caliper of .001 and an allowance of five matches per treated school, baseline comparability on observables of SBHC schools and their untreated comparisons was substantially improved. The largest improvements in comparability came from the 411 schools in the untreated matched comparison group having more similar baseline test scores and attendance rates (treated schools were below average for non-urban

[15] SBHCs also tend to open in schools with lower test scores in all subject areas as well as lower attendance rates. Appendix B provides a more extensive set of univariate comparisons.

schools), as well as increases in the comparison groups schools' shares of White students and those qualifying for free lunch.[16] Table 7 presents select baseline school characteristics for the full set of non-urban Tennessee schools (Column 1), the matched comparison sample (Columns 2 and 3) and the "ever SBHC" sample (Column 4), which includes all the non-urban schools that opened SBHCs during the panel.

Table 7: Select Baseline Means Across Analytic Samples

		Matched Sample		
	Full Sample	Never SBHC	Ever SBHC	Ever SBHC Sample
Test Scores & Attendance				
Average TVAAS Level	2.94	2.92	3.01	2.98
Reading TCAP	0.10	0.02	0.01	-0.02
Math TCAP	0.08	-0.02	0.00	-0.04
Attendance Rate	95.18	95.01	94.25	93.72
Student Characteristics				
% Native English	0.94	0.95	0.97	0.96
% Immigrant	0.01	0.01	0.00	0.01
% Black	0.19	0.12	0.09	0.13
% Hispanic	0.04	0.04	0.03	0.03
% White	0.80	0.84	0.89	0.86
% Free Lunch	0.62	0.66	0.64	0.68
% SPED	0.16	0.17	0.16	0.16
Community Characteristics				
Urbanicity Scale	7.69	7.98	7.17	7.10
Zip Code Poverty Rate	0.17	0.17	0.17	0.18
Pharmacies	2.13	2.39	2.50	2.41
Mental Health Professionals	0.31	0.23	0.28	0.27
Non Mental Health Professionals	8.22	7.83	8.09	8.79
N Student Observations	587,389	167,104	19,711	23,653

Note: All samples exclude schools in urban districts in the state and omit schools where an SBHC is listed as open, but has no clear opening date.

[16] A visualization of sample restriction for common support is included in Appendix C.

Checks for Endogeneity Concerns

Two primary threats to a causal interpretation of the estimates in this study were the potential for endogenous student sorting in or out of SBHC schools, and the possibility that prior trends in the outcomes predated the SBHC openings. Table 8 displays the results from a series of regressions tests for compositional change in SBHC schools associated with the SBHC openings. Overall, there is limited evidence of any student sorting on the basis of observed characteristics. The only subgroups that have statistically significant reductions associated with the SBHC openings are students who qualify for free lunch (a 1.5 percentage-point decrease) and students who have low third-grade math scores (a marginally significant 1.8 percentage-point decrease). Otherwise, the composition of schools remained stable as SBHCs were introduced. While this finding of limited student sorting differs dramatically from the substantial negative sorting effects in New York City, it is not particularly surprising. Given the radical differential between the ease of such moves in a school-choice-rich, high-density urban center, and those in rural Tennessee, where changing schools could incur substantial costs because families would typically need to move their place of residence, it is far less plausible that short-run systematic sorting of students would result from SBHC openings.

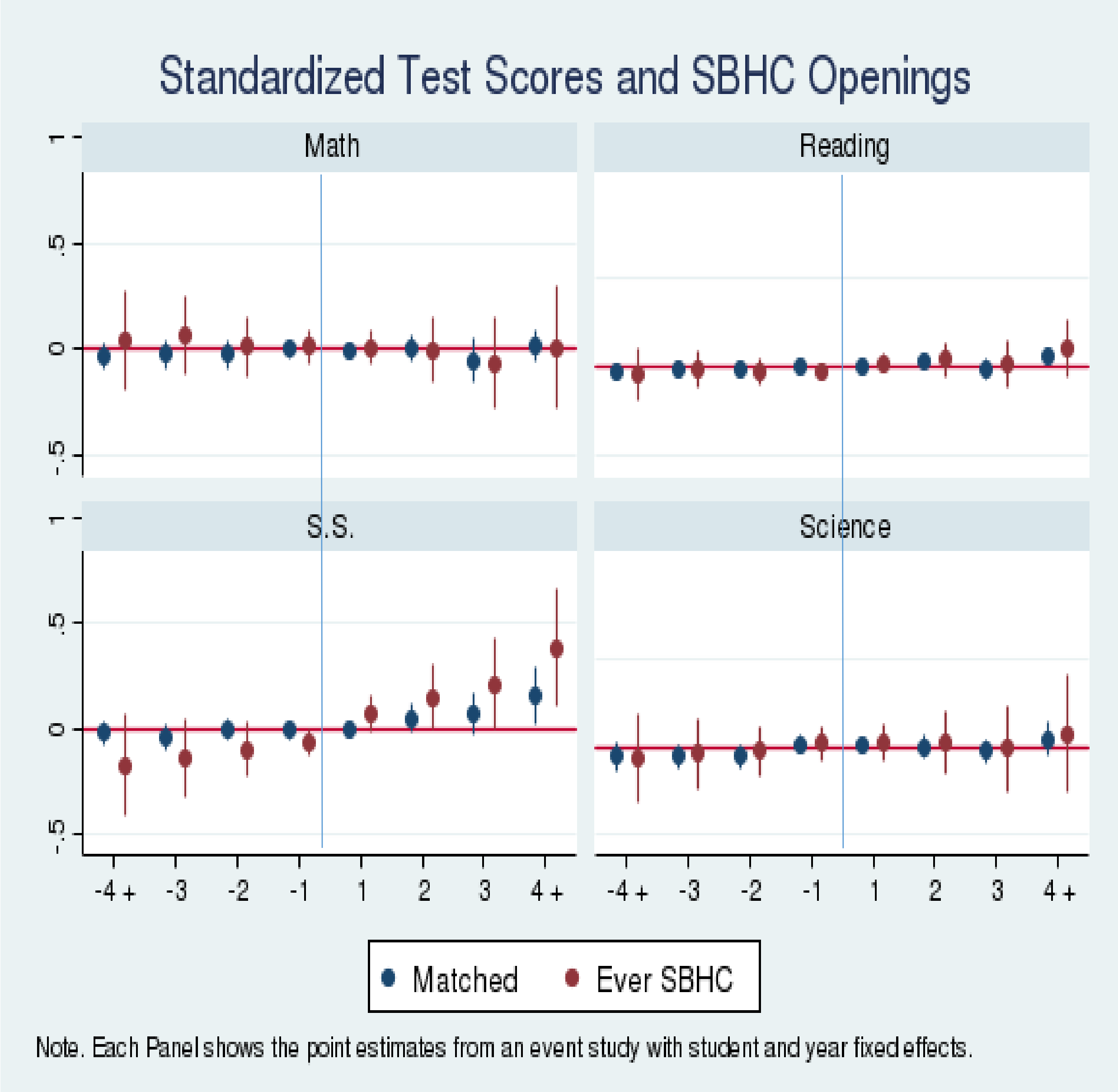
To assess whether the identification strategy satisfies the assumption that the openings of SBHCs were not related to pre-existing trends in student test scores, I present a visualization of event-study estimates for both the matched sample and most restrictive "Ever SBHC" sample. Figure 5 displays yearly estimated effects of SBHCs relative to the year before they opened (t=0 omitted for comparison), up to four years prior to their openings. The omitted reference year (zero on the x-axis) is the year prior to the stated SBHC opening date. The x-axes from left-to-right in each of the four panels plot the coefficients for schools adopting SBHCs four or more years prior to the actual opening, followed by three years prior, then two years prior, and so on. The model shows no clear indication of pre-existing trends, and none of the coefficients for the three years prior to openings is statistically significant at conventional levels.

Table 8: SBHC Openings and Student Composition

	% Free Lunch	% Black	% White	% Hispanic	% Prior SPED	Low 3rd Grade Math	Low 3rd Grade Reading	Low 3rd Grade Attendance
SBHC Open	-0.016**	-0.003	0.005	-0.004	0.002	-0.018+	-0.014	-0.012
	(0.006)	(0.003)	(0.004)	(0.003)	(0.005)	(0.010)	(0.010)	(0.020)
R2	0.11	0.32	0.29	0.08	0.01	0.08	0.08	0.06
N	971,301	971,036	971,036	967,676	828,370	971,308	971,308	971,308

Note: Each column displays the results from a separate regression analyzing the association between SBHC openings and changes in the proportion of specified student subgroups. All models control for school and year-fixed effects, as well as the full set of time-variant neighborhood level covariates; robust standard errors are clustered at the school level (+ p<0.1; * p<0.05; ** p<0.01).

Figure 5: Event Study Test for Pre-Existing Trends in Student Test Scores



Primary Analytic Strategy

To address the first stated research question, I estimate the effects of SBHC openings on student academic outcomes in a difference-in-differences and student-fixed effect framework. The primary analytic strategy provided a local average treatment effect for students who attended schools in which clinics opened in the 6-year window between 2008 and 2014. The main analytic model takes the following form:

$$(Y_{ist} = \beta_0 + \beta_1 SBHC_{st} + \boldsymbol{\lambda_{st}}\beta_2 + \boldsymbol{CBHS_{st}}\beta_3 + \beta_4 NewSchool_{it} + \beta_5 \gamma_{it} + \alpha_i + \tau_t + \epsilon_{ist} \quad (1))$$

Here, Y_{ist} represents the academic outcome of interest (e.g., attendance rate, standardized test scores on state exams, or disciplinary records) for student *i* in school *s* in year *t*. The $SBHC_{st}$ takes a 1 if there is an SBHC operating in school *s* in year *t*. Time-variant school characteristics (i.e., enrollment; percentages of students who identify as Hispanic, Black or non-native English speakers; and those participating in free or reduced-price lunch) are represented by the vector $\boldsymbol{\lambda_{st}}$, while all characteristics of the student *i* that are fixed over time are represented by the student-fixed effect, α_i. The $NewSchool_{it}$ is an indicator for whether a student is new to a given school in a given year, thus accounting for any shocks to student outcomes from school transitions; γ_{it} indicates the grade in which the student was enrolled in year *t*. The $\boldsymbol{CBHS_{st}}$ represents a vector of time-variant community-based health services, specifically the standardized per capita number of pharmacies, mental health practitioners, and non-mental health practitioners within the zip code where the school is located. I also include a series of year indicators τ_t to account for any phenomenon in a given year that might affect all schools in the state. The ϵ_{st} represents an error term clustered at the school level. The primary coefficient of interest $\boldsymbol{\beta_1}$is interpretable as the change in the outcome *Y* associated with the opening of the SBHC.

For comparison, I also present results from a traditional difference-in-differences model with school- and year-fixed effects, including a full set of student characteristics. Both models are estimated across three analytic samples. The first included all non-urban schools in Tennessee, the second relied on a propensity score matched sample of 411 schools with similar likelihood of ever opening an SBHC, and a third was limited to the 47 schools that opened an SBHC over the course of the analytic time window.

Given that the theory of action behind SBHCs' academic benefits requires student utilization and improved health, and/or decreased realized health costs for the family, it is plausible that benefits from attending a school with an SBHC take time to accrue within a student's academic career, such that additional years of exposure to care improve students' wellbeing and shift the student's academic trajectory upward. One way of modeling this type of additive function is to substitute a count of years of exposure to the SBHC for the simple binary indicator for whether the school has an SBHC open in a given year. In this model, for example, if a third-grade student is enrolled in a school that opens an SBHC during their fourth-grade year that remains open until they leave for a non-SBHC middle school in sixth grade, that student would be coded "0, 1, 2, 2, 2, 2," from third to eighth grade. However, if a student was enrolled in a school with a center that was already open by their third-grade year, and their middle school also had an SBHC, that student would be coded "1, 2, 3, 4, 5, 6" on the SBHC exposure variable. The model for the exposure effects takes the following form:

$$(Y_{ist} = \beta_0 + \beta_1 YearsSBHC_{it} + \boldsymbol{\lambda_{st}}\beta_2 + \boldsymbol{CBHS_{st}}\beta_3 + \beta_4 NewSchool_{it} + \beta_5 \gamma_{it} + \alpha_i + \tau_t + \epsilon_{ist} \quad (2))$$

To assess whether SBHCs are particularly beneficial to two student subgroups who research has shown are more likely to utilize SBHCs or benefit from their presence—girls and students with a history of absenteeism—I first re-estimate the main effects using the same model as that described above, restricted to each specified subsample of students. I then include the full sample of schools, but add an interaction term with the indicator variable for the grouping variable of interest. For example, the model below illustrates the strategy for estimating differences in effects for girls to assess whether differences in effect sizes for the subset are statistically significant:

$(Y_{st} = \beta_0 + \beta_1 SBHC_{ist} + \beta_2 SBHC_{ist} * \ Female_i + \boldsymbol{\lambda_{st}}\beta_3 + \boldsymbol{CBHS_{st}}\beta_4 + \beta_5 NewSchool_{it} + \beta_6 \gamma_{it} + \alpha_i + \ \tau_t + \epsilon_{ist}$ (3))

where all elements are identical to those described above, with the exception of the $SBHC_{ist} * \ Female_i$ indicator, which takes a 1 only when the student *i* is female and enrolled in a school *s* that has an SBHC open in year *t*. Because of the inclusion of the student-fixed effects, the main effect of being female is omitted. The β_2 is interpreted as the difference in effects of the SBHC opening for girls relative to boys, and β_1 becomes interpretable as the main effect of SBHC openings for the comparison category (here boys). To estimate the differential effects of the SBHC on students with low baseline attendance, I simply substituted an indicator for whether the student's third grade attendance rate was less than 94%, a threshold commonly applied in the SBHC literature (e.g., Reback & Cox, 2016). This roughly approximates the bottom quintile of student attendance rates. Thus, the coefficient on the low-attendance interaction term $SBHC_{ist} * \ Low\ 3rd\ Grade\ Attend_i$ is interpreted as the difference between SBHC effects for students with low third-grade attendance and those who had better than 94% attendance rates.

Results

I found consistent evidence that rural SBHC openings were associated with reduced absences, especially excused absences, which were likely to be related to avoided doctor visits. Findings with respect to test score benefits were generally positive, though less consistent across subject areas' sample restrictions and model specifications. However, in the most restrictive specification, I found that SBHC openings were associated with small but significant increases in scores on each subject area, with magnitudes (effect size .03 to .07 SD) comparable to findings from a recent study of SBHC openings in New York City (Reback & Cox, 2016). Furthermore, I found evidence of some additive benefits to continued exposure to SBHCs, such that additional years of enrollment in schools with an SBHC improved both test scores and attendance. Additionally, consistent with prior literature, a brief analysis of differential effects shows larger benefits for girls (whom studies consistently found were more likely to utilize SBHC services) as well as students with low baseline attendance rates (a crude proxy for children having health problems).[17]

Test Score Effects

Table 9 presents the estimated test score effects (annual math, reading, science, and social studies exams standardized within subject, grade, and year to have a mean of 0 and SD of 1) of SBHC openings from a series of alternative specifications and sample restrictions. The top panel shows results from a standard difference-in-differences model with school- and year-fixed effects, as well as a more restrictive student-fixed effect model, restricting the sample only to eliminate schools in major urban centers. The middle panel shows the results when the same-to-modeling strategies are applied to a matched sample of 411 schools with similar propensities to have ever opened an SBHC. Finally, in the bottom panel, I restrict the sample only to the 47 schools who ultimately open an SBHC over the course of the six-year panel, maximizing the comparability of the sample, but severely limiting its size, and focusing on the impacts of the earlier adopters in the difference-in-differences model.

In general, comparing the magnitude of effects across models and samples indicates that, as the comparison group grows more similar to the treatment group, the size of the apparent impact of SBHCs increases in the positive direction. This trend is indicative of a slight downward bias on

[17] Effects on special education placements are rarely statistically significant, but in places where they are, I find SBHC openings are associated with small reductions in the likelihood of a student's receiving disability services. This would be consistent with perhaps the least intuitive finding in the New York study, though the effect is near zero in all the school-level models, as well as in the most restrictive student-fixed effect model. Additionally, it is worth noting that SPED placements were typically lower in SBHC schools than comparison schools at baseline, suggesting the potential for omitted variable bias to drive this effect in the first two samples.

effect estimates from differences between SBHC schools and comparison schools on unobserved time-variant characteristics, such that something about the SBHC schools makes them less likely to improve in post-program-adoption years than other schools in the state. The increased magnitude of the effect estimates after matching on observed characteristics indicates that at least some of this downward bias could be attributable to the relatively stagnant outcomes for high-poverty schools, which compose a larger portion of the 'treated' schools than non-urban schools in general. Treatment schools also tend to have significantly lower average math scores than even comparison schools in the matched sample, which could explain the fact that math effects are only apparent when the sample is restricted to schools that ultimately open SBHCs.

In the school-level difference-in-differences model, when I characterize treatment as a simple indicator for whether or not a student's school has an SBHC open in a given year, effects on student test scores are modest, differ by subject area (there are some apparent increases in reading and science scores, but generally not math), and are sensitive to model specifications and sample selection. However, in the most restrictive model that is least subject to omitted variable bias, I find significant effects on student scores in all tested areas (effect size ranging from .03 to .07 SD).

Table 9: Effects of SBHC Openings on Student Test Scores

	Math		Reading		Science		S.S.	
	DID	FE	DID	FE	DID	FE	DID	FE
SBHC Openings and Students' Test Score (All Non-Urban Schools)								
SBHC Open	-0.022	-0.006	0.010	0.032**	0.026	0.014*	0.024	0.036**
	(0.025)	(0.007)	(0.015)	(0.006)	(0.027)	(0.006)	(0.027)	(0.006)
R2	0.17	0.82	0.18	0.85	0.20	0.83	0.21	0.83
N	2,454,732	2,454,732	2,473,021	2,473,021	2,467,740	2,467,740	2,465,867	2,465,867
SBHC Openings and Students' Test Score (Matched Sample)								
SBHC Open	-0.001	-0.012	0.029+	0.030**	0.058*	0.010	0.047	0.048**
	(0.026)	(0.008)	(0.016)	(0.007)	(0.028)	(0.008)	(0.029)	(0.008)
R2	0.12	0.85	0.13	0.87	0.16	0.85	0.16	0.85
N	925,988	925,988	931,052	931,052	929,480	929,480	927,937	927,937
SBHC Openings and Students' Test Score (Only if Ever SBHC)								
SBHC Open	0.007	0.072**	0.010	0.031**	-0.009	0.043**	0.038	0.060**
	(0.035)	(0.013)	(0.021)	(0.011)	(0.042)	(0.012)	(0.038)	(0.012)
R2	0.14	0.86	0.14	0.88	0.19	0.86	0.20	0.86
N	102,168	102,168	103,045	103,045	102,955	102,955	102,795	102,795

Note: Test scores are normed by grade and year to have a mean of zero and standard deviation of 1. All models control for the full set of time-variant neighborhood-, school-, and student-level covariates; robust standard errors are clustered at the school level for the DID analysis and at the student level for the fixed effect analysis for consistency with Reback & Cox (2016). The first panel only applies the restriction of omitting schools in major urban centers, the second is restricted to the matched sample of 411 schools with equivalent propensities to ever open an SBHC, and the bottom panel restricts the comparison group to only those schools who ultimately open an SBHC (+ p<0.1; * p<0.05; ** p<0.01).

Attendance Effects

In contrast to the relatively sensitive test-score effect estimates, the association between SBHC openings and reduced absences is markedly consistent across specifications and samples. The attendance effects are concentrated in the category of reduced excused absences, which conceivably serve as the best proxy for student health, given that confirmed illness is the primary accepted excuse for missing school. Table 10 mirrors Table 9 in structure, but focuses on the non-test score outcomes of student attendance (the number of excused and unexcused absences). In the preferred, most restrictive model, only the reduction in excused absences is statistically significant, with a magnitude of .55 fewer missed days. To put these attendance effects in context, early estimates from the New York City SBHCs were associated with 0.2 additional days of attendance (Reback & Cox, 2016), and Gottfried's, 2010 study of the causal link between attendance and achievement found that even small improvements in attendance can increase student's GPAs and test scores significantly for both elementary and middle school students.

Table 10: Effects of SBHC Openings on Attendance

	Excused Absences		Unexcused Absences	
	DID	FE	DID	FE
SBHC Openings and Students' Attendance, SPED placement (All Non-Urban Schools)				
SBHC Open	-0.805**	-0.446**	-0.480*	-0.114**
	(0.252)	(0.048)	(0.215)	(0.042)
R2	0.10	0.63	0.14	0.60
N	2,549,163	2,549,163	2,549,163	2,549,163
SBHC Openings and Students' Attendance, SPED placement (Matched Sample)				
SBHC Open	-0.700*	-0.460**	-0.296	-0.150**
	(0.271)	(0.058)	(0.227)	(0.050)
R2	0.08	0.69	0.14	0.67
N	962,511	962,511	962,511	962,511
SBHC Openings and Students' Attendance, SPED placement (Only if Ever SBHC)				
SBHC Open	-0.500	-0.550**	-0.479*	-0.032
	(0.420)	(0.088)	(0.226)	(0.067)
R2	0.09	0.71	0.16	0.70
N	106,651	106,651	106,651	106,651

Note: All models control for the full set of time-variant neighborhood-, school- and student-level covariates, and robust standard errors are clustered at the school level. The first panel only applies the restriction of omitting schools in major urban centers; the second is restricted to the matched sample of 411 schools with equivalent propensities to ever open an SBHC, and the bottom panel restricts the comparison group to only those schools who ultimately open an SBHC (+ $p<0.1$; * $p<0.05$; ** $p<0.01$).

Additive Effects of SBHC Exposure

I also examine the additive impact of a student's being enrolled in a school with an SBHC, testing the extent to which additional years of exposure to care improve student attendance and shift the student's achievement trajectory upward. Results from the student-fixed effect model, where I substitute a count of years of exposure to the SBHC for the simple binary indicator for whether the school has an SBHC open in a given year (described above), indicate that benefits of SBHC exposure indeed accrue over time.

Table 11: SBHC Exposure and Student Outcomes

	Math	Read	Sci.	S.S.	Excused	Unexcused
Exposure to SBHC (Years)	0.055*	0.072**	0.009	0.168**	-0.193**	-0.101+
	(0.010)	(0.009)	(0.010)	(0.010)	(0.075)	(0.055)
R2	0.86	0.88	0.86	0.86	0.71	0.70
N	102,168	103,045	102,955	102,795	106,651	106,651

Note: Coefficients represent the effects of an additional year of exposure to SBHCs in a student-fixed effect framework, with the full set of time-variant student and school-level covariates. Robust standard errors are clustered at the student level. Sample is restricted to the matched comparison schools (* p<0.05; ** p<0.01).

Table 11 displays results from the exposure model, and indicates a substantial additive effect of access to SBHCs over time. Specifically, each additional year of access to an SBHC is associated with significant improvements in test scores in three of the four tested subject areas, as well as reduced absences—especially for excused absences. To put the magnitude of these effects in context, consider that the average years of exposure to an SBHC for students who are ever enrolled in a school where one operates is 2.2 by the end of the panel. The coefficients in the exposure model indicate that this average level of exposure would produce an increase in reading scores of roughly .16 SD, and a .12 SD bump in math scores. In the fourth year of exposure to SBHC services, the average student would be expected to have 1.18 fewer absences than a similarly situated student with zero exposure.

Differential Benefits by Gender and Low Baseline Attendance Rates

Following prior SBHC literature, I also briefly examine differential effects of the centers by the gender of students and baseline attendance rates, which serve as a crude proxy for student health. More than many conventional educational interventions, we might expect the opening of health centers in schools to have substantially different benefits for girls than boys, especially during adolescent years that comprise an important portion of the analytic sample. Prior studies of SBHC utilization have also found that girls comprise a much larger portion of SBHC visits than boys (Kerns, et al., 2012). Similarly, it is well established that students with chronic illnesses (Forrest et al., 1997; Halfton & Newacheck, 1993) or dental problems (Guarnizo-Herreño & Wehby, 2012) miss more school than their healthier peers. Thus, we can use early attendance problems as crude proxy for student health concerns that might identify students as especially in need of the type of services provided in SBHCs.

Table 12 displays both the interaction effect and overall effects for the subset of students identified as female. While there are some notable exceptions, the interaction on female and SBHC openings is generally positive, such that female students appear to benefit more from SBHC openings than their male classmates. However, the estimates for the female sample alone are not qualitatively different from those for all students presented above.

Table 13 mirrors Table 12, but instead focuses on students who had low attendance rates in the third grade (or, less than 94% attendance rate in the first year students appear in the analytic sample for state testing). Again, the interactions are not entirely consistent, though, fittingly, this group appears to experience significantly larger reductions in absences (nearly two full days fewer excused absences), suggesting that the presence of the SBHCs helped stabilize the ability to attend school (a crude proxy for health) for this group that started off struggling to do so. I cannot rule out that these apparent differential benefits are driven by a regression to the mean or ceiling effect for students whose attendance before SBHC openings was too high to improve. However, these findings could also be driven by greater utilization, and increased likelihood of returning to class after SBHC visits by students with chronic medical conditions (Wade et al., 2008).

Table 12: Differential Effects by Gender

Differential Effects by Gender (Matched Sample, with Student-Fixed Effects)						
	(Math)	(Read)	(Sci.)	(S.S.)	(Excused)	(Unexcused)
SBHC Open	-0.054*	0.013	-0.019+	0.067**	-0.411**	-0.112
	(0.012)	(0.010)	(0.011)	(0.011)	(0.078)	(0.071)
SBHC Open * Female	0.085*	0.036**	0.058**	-0.038*	-0.099	-0.079
	(0.016)	(0.014)	(0.015)	(0.015)	(0.113)	(0.097)
R2	0.85	0.87	0.85	0.85	0.69	0.67
N	925,976	931,040	929,468	927,925	962,498	962,498
SBHC Effects for Girls Only (Matched Sample, with Student-Fixed Effects)						
SBHC Open	-0.014	0.032**	0.022*	0.045**	-0.511**	-0.136*
	(0.011)	(0.009)	(0.010)	(0.010)	(0.085)	(0.069)
R2	0.85	0.87	0.85	0.85	0.69	0.66
N	453,743	456,411	455,690	454,956	467,579	467,579

Note: Coefficients represent the effects of SBHC openings in a student-fixed effect framework, with the full set of time variant student- and school-level covariates. Robust standard errors (in parentheses) are clustered at the student level. Sample is restricted to the matched comparison schools (* p<0.05; ** p<0.01).

Table 13: Differential Effects by Baseline Attendance (<94% Third Grade Attendance Rate)

Differential Effects by Baseline Attendance (Matched Sample, with Student-Fixed Effects)						
	(Math)	(Read)	(Sci.)	(S.S.)	(Excused)	(Unexcused)
SBHC Open	0.000	0.030**	0.024**	0.045**	0.006	0.010
	(0.009)	(0.008)	(0.009)	(0.009)	(0.057)	(0.048)
SBHC Open * Low 3rd	-0.046+	-0.001	-0.055**	0.010	-1.766**	-0.608**
Grade Attend	(0.019)	(0.016)	(0.018)	(0.017)	(0.152)	(0.137)
R2	0.85	0.87	0.85	0.85	0.69	0.67
N	925,988	931,052	929,480	927,937	962,511	962,511
Only Students with Low Baseline Attendance (Matched Sample, with Student-Fixed Effects)						
SBHC Open	-0.027	0.035*	-0.007	0.061**	-0.567**	0.021
	(0.017)	(0.015)	(0.015)	(0.015)	(0.144)	(0.130)
R2	0.82	0.85	0.83	0.83	0.64	0.62
N	156,255	157,013	156,649	156,235	166,217	166,217

Note: Coefficients represent the effects of SBHC Openings in a student-fixed effect framework, with the full set of time-variant student- and school-level covariates. Robust standard errors (in parentheses) are clustered at the student level. Sample is restricted to the matched comparison schools (* p<0.05; ** p<0.01).

Conclusions

The positive findings presented here regarding the efficacy of rural SBHCs provide important insights about a potential tool for combatting poverty, promoting child health, and enhancing learning in parts of the country that are often overlooked in contemporary education policy research and popular discourse. While it is difficult to conclusively identify a causal relationship outside of an experimental setting, the findings from this student-level longitudinal analysis paint a largely consistent picture of students benefiting from the expanded access to primary health care services, and the affirming robustness checks substantially limit the set of plausible alternative explanations for the positive associations. In the years following SBHC openings, students' excused absences reduced significantly, and test scores improved by .03 to .07 SD. Academic benefits of the SBHCs accrued annually with increased student exposure, and were largest among the student subgroups (girls and students with low baseline attendance) that were most likely to visit the centers (based on prior literature).

One of the largest debates in current education policy centers on the question of whether schools alone can close the gaps in achievement between low-income students and their middle-class peers, or whether a broader range of social policy interventions are necessary to combat the many facets of poverty that pose obstacles to academic success. The schools-alone advocates tend to argue for greater use of financial incentives and accountability for teachers and administrators, privatization of school management, and other market-oriented reforms that seek to improve school quality through parent choice. The research body examining these types of market-oriented reforms is rich, with mixed and sometimes contradictory findings (e.g., Hanushek et al., 2007; Springer et al., 2010; Springer, Swain, & Rodriguez, 2016; Angrist, Pathak, & Walters, 2011; Ballou, 2001; Dobbie & Fryer, 2009; Hoxby & Rockoff, 2004; Peterson, 1998; Bifulco & Ladd, 2007). Alternatively, those who favor a "broader, bolder" approach to education policy generally highlight disparities in access to stable or integrated housing, employment, and extracurricular learning opportunities as drivers of racial and income achievement gaps (e.g., Dastrup & Betts, 2012; Bradbury, Burke, & Triest, 2013; Schwartz, 2010; Ananat, Gassman-Pines, Francis, & Gibson-Davis, 2017; Condron, 2009). Another factor frequently cited as a success barrier for students in poverty is their disproportionate experience of treatable health problems and lack of access to care (e.g. Rothstein, 2004; Ladd, 2012). It is plausible that, even in cases where school-based reforms appear to be successful, such as evaluations of high-performing charter schools, the non-negligible number of students who fail to respond to the relatively robust intervention still struggle, due to rigorous instruction or treatable health problems that limit the effectiveness of longer school days.

While this study finds evidence of significant academic benefits from SBHC openings, it is worth emphasizing that a failure to find academic returns to health service interventions should not be viewed as a condemnation of the policy of providing health services to low-income students in schools. Well-established estimated effects on health outcomes and the utilization of preventative care measures support their worth as a public health policy. However, the findings presented in this study also suggest that experimental evaluations of health interventions should look beyond health outcomes to quantify educational externalities. If research continues to confirm that health policies are causally linked to even small increases in student achievement, the economic returns associated with the improved academic performance could significantly shift cost-benefit analyses of interventions targeting student health. It would also give those interested in educational equity another tool to address persistently vexing national education policy imperatives, thus elevating student achievement and closing achievement gaps.

Finally, both the construct of the SBHC and the measurement of its effects in this study represent a muted estimate of the fully realized policy's potential effects. While the SBHCs in rural Tennessee tend to be focused on providing primary care to students enrolled in the schools

(SBHA, 2015), prior literature suggests that students in these underserved populations might benefit even more from SBHCs that cover broader sets of health problems—especially mental health (e.g. Reback, 2010), and dental health (Guarnizo-Herreño & Wehby, 2012)—and that share access to these vital services with parents (Claessens, Engel, & Curran, 2015) and siblings (Heissel, 2017), whose own health problems can place significant strains on students' ability to learn. Furthermore, because of the state's longstanding commitment to the principle of coordinated school health, it is plausible that the estimated impacts of SBHC openings are smaller in Tennessee schools than they would be in schools that were devoid of health services prior to their introduction. Like other recent empirical analyses of SBHCs, results here represent the effects of the centers on all students who technically gained access to school-based health services due to a center's opening in their school—whether they needed them, took advantage of them, or not. Thus, they could be interpreted as an intent-to-treat average treatment effect, attenuated by the inclusion of numerous students who may have rarely, if ever, visited the SBHC. In other words, better longitudinal data on student health and utilization, paired with increasingly rich educational data systems, could yield substantially larger effect sizes for treated students, even more so than promising estimates presented here. Additionally, more comprehensive implementation of the SBHC policy concept that covers a broader set of health problems and community members could substantially lessen the drag poverty imposes on students' learning, development, and prospects for success later in life.

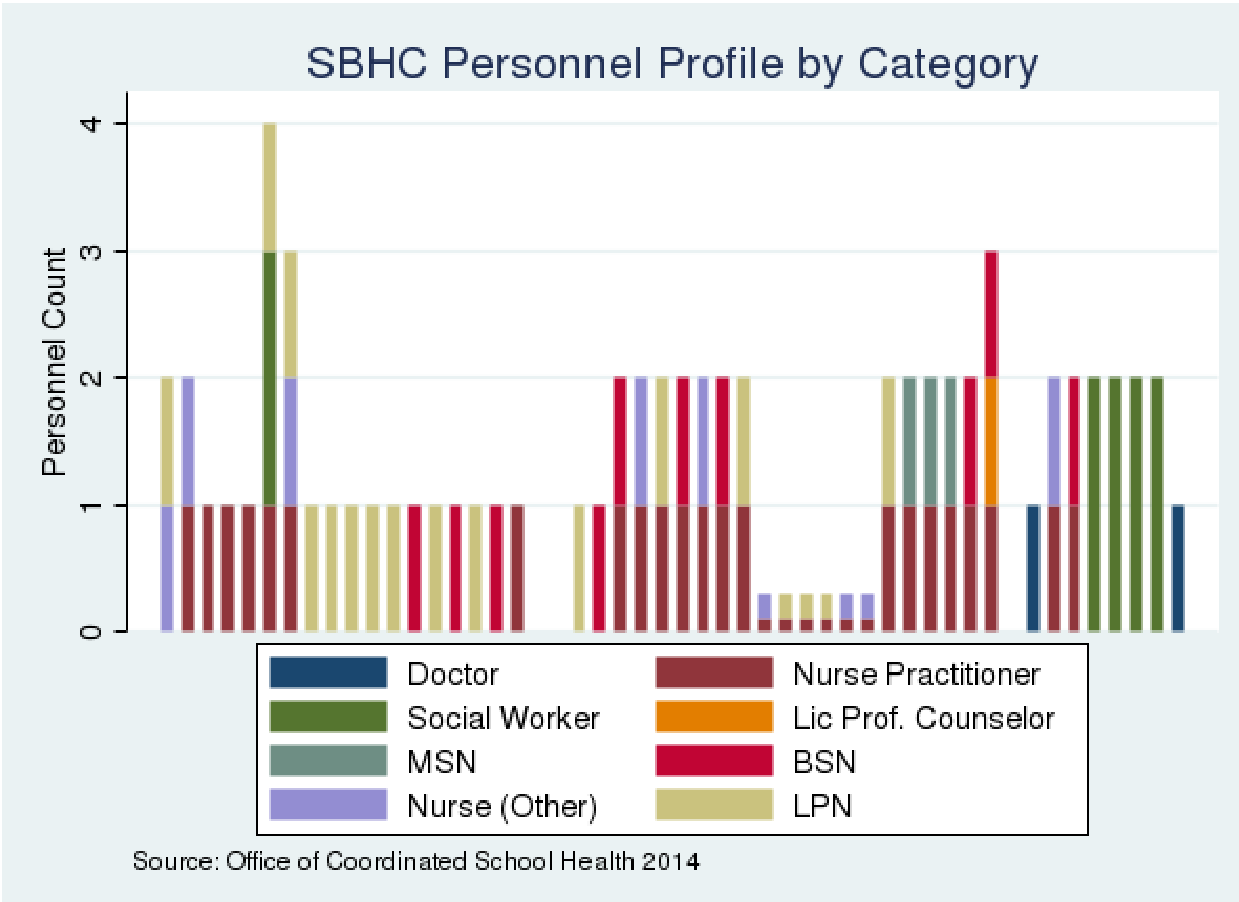
APPENDIX

Appendix A: SBHC Capital Grant Winners in Tennessee

Awarded 2011		
Named Applicant	**Location**	**Grant Amount**
Cheatham County School District, Inc.	Ashland City	$499,926
McMinn County	Athens	$402,065
Humboldt Family Resource Center	Humboldt	$393,034
Community Health of East Tennessee, Inc.	Jacksboro	$202,422
Le Bonheur Community Health and Well-Being	Memphis	$492,500
Bedford County Schools	Shelbyville	$497,936
Franklin County School District	Winchester	$499,997
Total		$2,987,880
Awarded 2012		
Named Applicant	**Location**	**Grant Amount**
Cherokee Health Systems	Knoxville	$443,465
Knox County Schools	Knoxville	$212,190
Matthew Walker Comprehensive	Nashville	$299,500
United Neighborhood Health Services	Nashville	$500,000
Memphis City Schools	Memphis	$499,167
Monroe County Department of Education	Madisonville	$200,000
The Medicine and Education Group	Lebanon	$500,000
Total		$2,654,322

Source: Health Resources and Services Administration

Appendix B: SBHC Staffing Profile for Analytic Sample



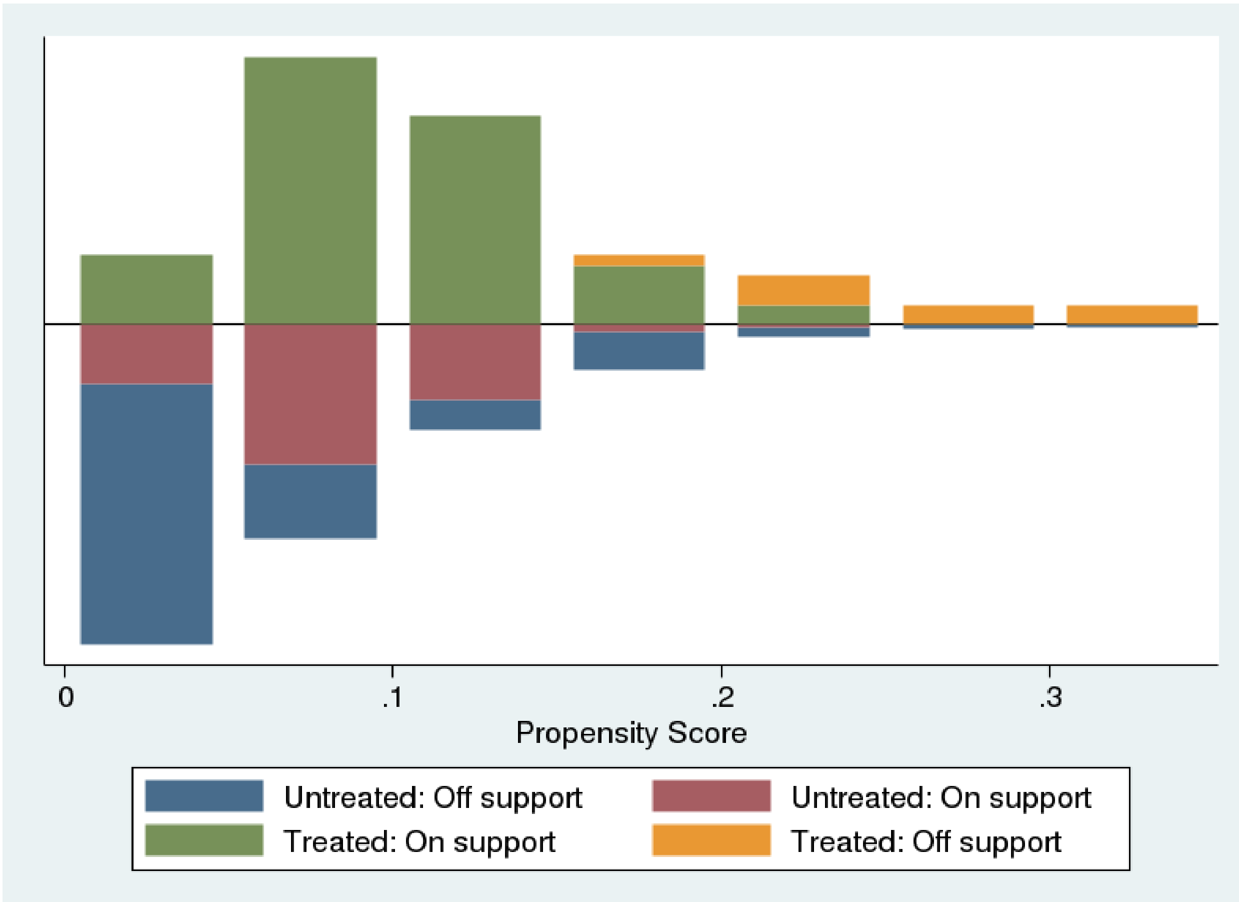
Note: The bottom four categories are all listed categories of nursing credentials, where MSN stands for Master of Science in Nursing, BSN stands for Bachelor of Science in Nursing, LPN stands for Licensed Practical Nurse, and Nurse (other) captures unspecified nurses. Several SBHCs also have administrative staff or uncategorized support personnel not depicted here.

Appendix C: Baseline Comparison of Schools that Ever Had SBHCs in Their Sample to Schools in Sample that Never Had SBHCs

	Ever SBHC	Never SBHC	Mean Diff	Std. Error
Mean School Characteristics				
Female (%)	48.22	48.41	-0.19	0.31
White (%)	85.21	77.38	7.83**	2.34
Black (%)	9.72	16.53	-6.81**	2.06
Hispanic (%)	4.57	5.9	-1.33+	0.69
Free lunch (%)	61.95	54.65	7.30**	1.99
Free or Reduced (%)	70.96	61.12	9.84**	1.99
Average Daily Membership (#)	479.41	490.11	-10.7	22.71
Community-Based Health Service Statistics (Standardized by School Zip Code)				
Pharmacies	0.1	0.03	0.08	0.14
Mental Health Practitioners	-0.14	-0.02	-0.12*	0.05
Physical Health Practitioners	-0.11	0.02	-0.13*	0.06
K-8 Attendance, and Test Scores (Standardized by School)				
Attendance	-0.64	0.04	-0.68**	0.1
Math	-0.21	0.01	-0.23*	0.1
Reading	-0.23	0.01	-0.25*	0.1
Science	-0.24	0.01	-0.25*	0.11
Social Studies	-0.21	0.01	-0.22+	0.12
Number of school	*91*	*1039*	*Total:*	*1130*

Note: School characteristics above are presented as percentages or whole numbers for the baseline year of the panel analysis, and come from the CCD and the TN Department of Education. The community health service statistics represent the standardized (mean 0, SD 1) per capita number of pharmacies as well as mental and physical health practitioners in the school's zip code, according to the U.S. Census Bureau. K-8 attendance and standardized test scores are also adjusted to have a mean of 0 and standard deviation of 1 in order to be comparable across years. Subject scores reflect the school's growth measure on the state exam used for accountability. The statistical significance of above differences is estimated using regression t-tests, with heteroscedasticity-robust standard errors. The Never SBHC Group omits schools from the large urban districts.

Appendix D: Graph of Common Support for Matched Sample



Note: This figure provides a visualization of the proportions of schools falling into the specified categories post matching on propensity scores, in a nearest neighbor framework with a caliper of .001 and an allowance of five matches per treated school.

REFERENCES

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, *72*(1), 1-19.

Ananat, E. O., Gassman-Pines, A., Francis, D. V., & Gibson-Davis, C. M. (2017). Linking job loss, inequality, mental health, and education. *Science*, *356*(6343), 1127-1128.

Angrist, J. D., Pathak, P. A., & Walters, C. R. (2011). *Explaining charter school effectiveness* (*National Bureau of Economic Research Working Paper Series*, *No. 17332).* Retrieved from http://www.nber.org/papers/w17332

Ballou, D. (2001). Pay for performance in public and private schools. *Economics of Education Review*, *20*(1), 51-61.

Bansak, C., & Raphael, S. (2007). The effects of state policy design features on take-up and crowd-out rates for the State Children's Health Insurance Program. *Journal of the Association for Public Policy Analysis and Management*, *26*(1), 149-175. Retrieved from http://doi.wiley.com/10.1002/pam.20231

Bifulco, R., & Ladd, H. F. (2007). School choice, racial segregation, and test-score gaps: Evidence from North Carolina's charter school program. *Journal of Policy Analysis and Management*, *26*(1), 31-56.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2007). From the cradle to the labor market?: The effect of birth weight on adult outcomes. *Quarterly Journal of Economics, 122*(1), 409-439.

Bradbury, K., Burke, M. A., & Triest, R. K. (2013). *Do foreclosures affect Boston Public School student academic performance? (Public Policy Brief No. 13-5).* Retrieved from http://www.bostonfed.org/economic/ppdp/index.htm.

Claessens, A., Engel, M., & Curran, F. C. (2015). The effects of maternal depression on child outcomes during the first years of formal schooling. *Early Childhood Research Quarterly, 32*(3), 80-93.

Cohodes, S., Grossman, D., Kleiner, S., & Lovenheim, M. F. (2016). The effect of child health insurance access on schooling: Evidence from public insurance expansions. *Journal of Human Resources, 51*(3), 727-759. doi:10.3368/jhr.51.3.1014-6688R1

Condron, D. J. (2009). Social class, school and non-school environments, and black/white inequalities in children's learning. *American Sociological Review*, *74*(5), 685-708.

Currie, J. 2009. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economics Literature, 47*(1), 87-122.

Currie, J., Decker, S., & Lin, W. (2008). Has public health insurance for older children reduced disparities in access to care and health outcomes? *Journal of Health Economics*, *27*(6), 1567-1581. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18707781

Currie, J., & Gruber, J. (1996). Health insurance eligibility, utilization of medical care, and child health. *Quarterly Journal of Economics*, *111*(2), 431-466. doi:10.2307/2946684

Dafny, L., & Gruber, J. (2005). Public insurance and child hospitalizations: Access and efficiency effects. *Journal of Public Economics*, *89*(1), 109-129. doi:10.1016/j.jpubeco.2003.05.004

Dastrup, S. R., & Betts, J. R. (2012). *Elementary education outcomes and stress at home: Evidence from mortgage default in San Diego*. Retrieved from http://econ.ucsd.edu/~sdastrup/pdfs/DastrupJobMarketPaper.pdf

Ding, W., Lehrer, S. F., Rosenquist, J. N., & Audrain-McGovern, J. (2009). The impact of poor health on academic performance: New evidence using genetic markers. *Journal of Health Economics*, *28*(3), 578-597.

Dobbie, W., & Fryer, R. G. (2009). *Are high quality schools enough to close the achievement gap?: Evidence from a social experiment in Harlem.* Retrieved from http://www.nber.org/papers/w15473

Egbuonu, L., & Starfield, B. (1982). Child health and social status. *Pediatrics*, *69*(5), 550-557.

Ethier, K. A., Dittus, P. J., DeRosa, C. J., Chung, E. Q., Martinez, E., Kerndt, P. R. (2011). School-based health center access, reproductive health care, and contraceptive use among sexually experienced high school students. *Journal of Adolescent Health, 48*(6), 562-565.

Festinger, T., & Duckman, R. H. (2004). Vision Status Among Foster Children in NYC: A Research Note. *Social work in health care*, *38*(4), 77-81.

Figlio, D. N., Guryan, J., Karbownik, K., & Roth, J. (2014). *The effects of poor neonatal health on children's cognitive development (National Bureau of Economic Research Working Paper No. 18846).* Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.104.12.3921

Fletcher, J. M., & Lehrer, S. F. (2009). *Using genetic lotteries within families to examine the causal impact of poor health on academic achievement (National Bureau of Economic Research Working Paper No. 15148).* Retrieved from http://www.nber.org/papers/w15148

Forrest, C. B., Starfield, B., Riley, A. W., & Kang, M. (1997). The impact of asthma on the health status of adolescents. *Pediatrics*, *99*(2), E1.

Fowler, M. G., Johnson, M. P., & Atkinson, S. S. (1985). School achievement and absence in children with chronic health conditions. *Journal of Pediatrics, 106*(4), 683-687.

Geierstanger, S. P., Amaral, G., Mansour, M., & Walters, S. R. (2004). School-based health centers and academic performance: Research, challenges, and recommendations. *Journal of School Health, 74*(9), 347-352.

Gottfried, M. A. (2010). Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach. *American Educational Research Journal*, *47*(2), 434-465.

Guarnizo-Herreño, C. C., & Wehby, G. L. Children's dental health, school performance, and psychosocial well-being. *The Journal of Pediatrics, 161*(6), 1153-59.

Halfon, N., & Newacheck, P. W. (1993). Childhood asthma and poverty: Differential impacts and utilization of health services. *Pediatrics, 91*(1), 56-61.

Hanushek, E. A., Kain, J. F., Rivkin, S. G., & Branch, G. F. (2007). Charter school quality and parental decision making with school choice. *Journal of Public Economics*, *91*(5), 823-848.

Hartley, D. (2004). Rural health disparities, population health, and rural culture. *American Journal of Public Health*, *94*(10), 1675-1678.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, *64*(4), 605-654.

Heissel, J. (2017). *Spillover effects within families: Evidence from teenage motherhood and sibling academic performance*. Retrieved from http://www.aefpweb.org.

Hoxby, C., & Rockoff, J. (2004). The impact of charter schools on student achievement (vol. 2005). Retrieved from http://www.rand.org/content/dam/rand/www/external/labor/seminars/adp/pdfs/2005hoxby.pdf

Kenney, G., Dubay, L., Hill, I., Sommers, A., & Zuckerman, S. (2005). *Congressionally mandated evaluation of the state children's health insurance program: Final report to Congress.* Washington, D.C.: Mathematica Policy Research, Inc.

Kerns, S. E. U., Pullmann, M. D., Walker, S. C., Lyon, A. R., Cosgrove, T. J., & Bruns, E. J. (2012). School-based health center use and high school dropout rates—reply. *Archives of Pediatrics and Adolescent Medicine Journal*, *166*(7), 675-677. doi:10.1001/archpediatrics.2012.553

Ladd, H. F. (2012). Education and poverty: Confronting the evidence. *Journal of Policy Analysis and Management*, *31*(2), 203-227.

Levine, P. B., & Schanzenbach, D. W. (2009). *The impact of children's public health insurance expansions on educational outcomes (National Bureau of Economic Research Working Paper 14671).* Retrieved from http://www.nber.org/papers/w14671

Lovenheim, M. F. (2011). The effect of liquid housing wealth on college enrollment. *Journal of Labor Economics, 29*(4), 741-771.

Lovenheim, M. F., Reback, R., & Wedenoja, L. (2016). *How does access to health care affect teen fertility and high school dropout rates? Evidence from school-based health centers* (No. w22030). Cambridge, MA: National Bureau of Economic Research.

O'Neill, S., Kreif, N., Grieve, R., Sutton, M., & Sekhon, J. S. (2016). Estimating causal effects:

considering three alternatives to difference-in-differences estimation. *Health Services & Outcomes Research Methodology*, *16*, 1-21. http://doi.org/10.1007/s10742-016-0146-8

Oreopoulos, P., Stabile, M., Walld, R., & Roos, L. L. (2008). Short-, medium-, and long-term consequences of poor infant health: An analysis using siblings and twins. *Journal of Human Resources, 43*(1), 88-138.

Orfield, A., Basa, F., & Yun, J. (2001). Vision problems of children in poverty in an urban school clinic: Their epidemic numbers, impact on learning, and approaches to remediation. *Journal of Optometric Vision Development, 32*(3), 114-141.

Peterson, P. E. (1998). School choice: A report card. In P. E. Peterson & B. C. Hassel (Eds.), *Learning from school choice* (pp. 3-32). Washington, D.C.: Brookings Institution.

Reback, R. (2010). Schools' mental health services and young children's emotions, behavior, and learning. *Journal of Policy Analysis and Management, 29*(4), 698-725. doi:10.1002/pam.20528

Reback, R., & Cox, T. (2016). Primary health care and children's academic achievement (Working Paper). Retrieved from http://www. csd.wustl.edu/publications/documents/wp13-15.pdf

Rochmes, J. (2016). *School-based healthcare and academic performance: Implications of physical health services for educational outcomes and inequality (CEPA Working Paper No. 15-07).* Retrieved from Stanford Center for Education Policy Analysis Web site: http://www.cepa.stanford.edu/wp15-07

Rosenbach, M., Ellwood, M., Irvin, C., Young, C., Conroy, W., Quinn, B., & Kell, M. (2003). *Implementation of the State Children's Health Insurance Program: Synthesis of state evaluations: Background for the Report to Congress*. Retrieved from http://eric.ed.gov/?id=ED475969

Rothstein, R. (2004). *Class and schools*. New York, NY: Teachers College, Columbia University.

School-Based Health Alliance (SBHA). (2015). *Census of school-based health centers for 2013-14*. Retrieved from http://www.sbh4all.org/school-health-care/national-census-of-school-based-health-centers/

Schwartz, H. (2010). *Housing policy is school policy: Economically integrative housing promotes academic success in Montgomery County, Maryland*. New York, NY: Century Foundation.

Shaefer, L. H., Grogan, C. M., & Pollack, H. A. (2011). Transitions from private to public health coverage among children: Estimating effects on out-of-pocket medical costs and health insurance premium costs. *Health Services Research*, *46*(3), 840-858.

Sebelius, K. (2010). Rising to the challenge: Tools for enrolling eligible children in health coverage. *Health Affairs*, *29*(10), 1930-1932.

Seid, M., Varni, J. W., Cummings, L., & Schonlau, M. (2006). The impact of realized access to care on health-related quality of life: A two-year prospective cohort study of children in the California State Children's Health Insurance Program. *The Journal of Pediatrics*, *149*(3), 354-361.

Shone, L. P., Dick, A. W., Brach, C., Kimminau, K. S., LaClair, B. J., Shenkman, E. A., . . . Bronstein, J. (2003). The role of race and ethnicity in the State Children's Health Insurance Program (SCHIP) in four states: Are there baseline disparities, and what do they mean for SCHIP? *Pediatrics*, *112*(Supplement E1), e521-e532.

Springer, M. G., Ballou, D., Hamilton, L., Le, V. N., Lockwood, J. R., McCaffrey, D. F., . . . Stecher, B. M. (2011). Teacher pay for performance: Experimental evidence from the Project on Incentives in Teaching (POINT). Retrieved from https://eric.ed.gov/?id=ED518378

Springer, M. G., Swain, W. A., & Rodriguez, L. A. (2016). Effective teacher retention bonuses: Evidence from Tennessee. *Educational Evaluation and Policy Analysis*, *38*(2), 199-221.

Valet, R. S., Perry, T. T., & Hartert, T. V. (2009). Rural health disparities in asthma care and outcomes. *Journal of Allergy and Clinical Immunology*, *123*(6), 1220-1225.

Wade, T. J., Mansour, M. E., Guo, J. J., Huentelman, T., Line, K., & Keller, K. N. (2008). Access and utilization patterns of school-based health centers at urban and rural elementary and middle schools. *Public Health Reports*, *123*(6), 739-750.

Walker, S. C., Kerns, S. E., Lyon, A. R., Bruns, E. J., & Cosgrove, T. (2010). Impact of school-based health center use on academic outcomes. *Journal of Adolescent Health, 46*(3), 251-257.

Webber, M. P., Hoxie, A. M., Odlum, M., Oruwariye, T., Lo, H., and Appel, D. (2005). Impact of asthma interventions in two elementary school-based health centers in the Bronx, New York City. *Pediatric Pulmonology, 40(6)*, 487-493.

CHAPTER 3

SELECTIVE TEACHER RETENTION BONUSES (SRBs) IN HIGH-POVERTY SCHOOLS: EVIDENCE FROM THE TENNESSEE PRIORITY SCHOOL RETENTION BONUS PROGRAM

Introduction

High-quality teachers are one of the most important school-based components in the production of student achievement; however, value-added studies of teacher effectiveness consistently find large variations in teacher classroom performance (Aaronson, Barrow, & Sander, 2007; Hanushek, Kain, O'Brien, & Rivkin 2005; Kane, Rockoff, & Staiger, 2008; Rockoff, 2004; Sanders & Rivers, 1996). While the majority of test score gaps along the lines of income and race are attributable to outside-of-school factors and are evident very early in life (e.g., Fryer & Levitt, 2004; Heckman, 2006), the distribution of effective teachers represents a potentially important equity-promoting lever that is susceptible to policy intervention. Top-performing teachers, defined as teachers whose value-added estimates are at or above the 95th percentile, produce as much as three times the achievement growth in students when compared to low-performing teachers (Hanushek, 2011). Simulations further indicate that the systematic reassignment of highly-effective teachers to low performing schools could radically reduce across-school achievement gaps (Chetty, Friedman, & Rockoff, 2014).

Research has equally established that minority and low-income students—particularly those in schools with high concentrations of poverty or racial minorities—are more likely to be staffed by teachers graduating from less competitive colleges, teachers instructing out-of-field, and novice teachers (Clotfelter, Ladd, & Vigdor 2011; Iaterola & Steifel, 2003; Lankford, Loeb, & Wyckoff, 2002; Podgursky, Monroe, & Watson, 2004; Podgursky & Springer, 2011). For these reasons, a strong body of research has sought to better understand what makes highly qualified or effective teachers decide to stay or leave a school, or exit the profession altogether, with particular emphasis on schools' racial composition, disparities in working conditions, the distribution of economically disadvantaged students, and compensation (Boyd, Lankford, Loeb, & Wyckoff, 2008; Clotfelter et al., 2011; Feng, 2010; Feng, Figlio, & Sass, 2010; Scafidi, Sjoquist, & Stinebrickner, 2007; Steele, Pepper, Springer, & Lockwood, 2015).

In recent years, a growing number of states and districts have implemented various forms of financial incentives to shift the performance distribution of teachers within and across schools. Selective retention bonuses (SRBs) are a frequently considered strategy to curtail attrition of targeted groups such as highly-rated teachers, or teachers of hard-to-staff subject areas like math and science. SRBs differ from conventional pay-for-performance plans by systematically targeting the composition (increasing the average level of teacher effectiveness) and continuity (reducing the amount of unwanted turnover) of a school's faculty. Several studies have found relatively promising results when estimating the effects of financial incentives' capacity to promote retention at high-need, low-performing schools (e.g., Clotfelter, Glennie, Ladd, & Vigdor, 2008; Glazerman et al., 2013; Springer, Swain, & Rodriguez, 2016). However, extant literature tells us relatively little about the impact of this class of incentive pay policies on their ultimate goal: improving student academic outcomes.[18]

[18] A recent study by Adnot, Dee, Katz, and Wyckoff (2016) attempts to assess the effects of selective teacher turnover on student achievement in the context of the DC IMPACT reforms. The study finds (perhaps not surprisingly) that, when exiting teachers are replaced by teachers with higher demonstrated effectiveness, students perform better the following year. The theory of action in this paper is that the opposite should hold: if teachers who are retained in response to a policy intervention have greater measured effectiveness than their likely replacements, students' test scores should improve.

This study bridges a gap in the literature by estimating the effect of a SRB policy for high-rated teachers on low-performing schools' ability to elevate student test scores. The theory of action behind SRBs is simple: SRBs result in greater numbers of highly effective teachers at participating schools, who subsequently drive more measurable student learning than the teachers who would otherwise fill their positions. To examine whether this conceptualization holds true, this study uses rich longitudinal administrative data from the State of Tennessee, where a retention bonus program for highly effective teachers in Priority Schools (high poverty schools identified as having the lowest test scores in the state) was recently implemented, and was found to stem unwanted attrition.
The estimated effects of Tennessee's SRB program on teacher retention were relatively modest in magnitude, increasing the likelihood of a highly-effective teacher returning by roughly 20%. High-performing teachers represented less than 25% of the faculty in Priority Schools. However, the radical differences between the retained Level 5 teachers' estimated effectiveness and that of their likely replacements results in the equivalent of a relatively profound intervention. As reported in Springer et al. (2016), the 321 teachers[19] who accepted bonuses and remained in their schools had overall teacher effectiveness ratings more than a full standard deviation above the state average, and the average of teachers hired by Priority Schools was rated roughly two-thirds of a standard deviation below the state average. Thus, for every teacher retained as a result of the bonus, students taught by that teacher experienced an increase in teacher effectiveness of 1.7 standard deviations.

Furthermore, there are several other pathways through which SRBs might improve school-level performance. Recent studies have found or predict significant positive effects on student achievement through a number of factors, including positive peer effects on colleagues who work

[19] TDOE records indicated that 361 bonuses were distributed, though 40 bonuses appeared to be the result of administrative errors, where teachers were either not in a Priority School, did not remain at a Priority School, or did not satisfy the performance requirement. Issues concerning program implementation are detailed at length in Springer, Swain, and Rodriguez (2016).

with highly effective teachers (Jackson & Bruegmann, 2009; Sun, Loeb, & Grissom, 2017); a reduction of teacher churn within and across schools (Redding & Henry, 2017; Ronfeldt, Loeb, & Wyckoff, 2013); and effective pay-for-performance programs, where teachers face the threat of losing a pre-paid bonus if their estimated effectiveness falls below average (Fryer, Jr., Levitt, List, & Sadoff, 2012).[20] For these reasons, I hypothesize that schools who distribute retention bonuses for highly effective teachers will elevate student test scores by increasing access to effective instruction, faculty stability, and positive peer effects on colleagues.

In the sections that follow, I first describe the specific retention bonus policy context in Tennessee, then review background literature on selective retention bonuses. Next, I describe the data sources, construction of the analytic sample, and the primary strategy for estimating the impacts of the bonus program. Finally, I present results for both student achievement and changes in teacher composition by subject area, followed by a discussion of policy implications of the findings with respect to the following research question: To what extent do students' test scores at low-performing schools benefit from offering retention bonuses to highly effective teachers?

Tennessee's Highly Effective Teacher Bonus Program

The distribution and mobility patterns of highly effective teachers in the Tennessee public school system, as defined by a composite of observation ratings and value-added measures of teacher effectiveness, works to the detriment of students in low-test-score schools with large concentrations of economically disadvantaged and non-White students. In the year prior to the TN SRB program (the 2011-12 school year), approximately 17% of teachers left their schools, and the

[20] The retention bonus program was offered in conjunction with a recruitment bonus, with similar eligibility requirements for participating schools. While the uptake of this program was extremely low (fewer than 20 teachers), it should bias estimates of retention bonus effects upwards, assuming that these teachers were still highly effective in this new setting.

attrition rate for the state's most-effective teachers was around 7%. However, in the predominantly Black, low-income, bottom 5% of schools in the state, the attrition rate for highly effective educators increases to 23% (more than three times greater than the statewide attrition rate of highly-effective teachers).

In the spring of 2013, in an effort to combat these high rates of teacher turnover among highly-effective teachers in chronically low-performing schools, the Tennessee Department of Education (TDOE) and the Tennessee Governor's Office announced a teacher retention bonus program for priority schools. Under the program, all priority schools were eligible to participate by applying for the funding to offer $5,000 retention bonuses to any Level 5 teacher who was teaching in a Priority School.[21] On average, for teachers in Tennessee Priority Schools, the $5000 bonus constituted approximately a 10% salary increase, or the equivalent of a teacher with a master's degree moving from 10 to 15 years of experience on a district salary schedule. Statewide, roughly one-third of teachers earned a Level 5 rating in the 2012-13 school year. In total, 473 teachers—roughly 18% of those working in Priority Schools—were eligible for the Highly Effective Teacher Retention Bonus due to their Level 5 rating, and 321 were retained and paid the extra $5,000.

Level 5 teachers at Priority Schools who accepted retention bonuses were required to complete the 2013-14 school year at a Priority School and maintain at least a Level 4 effectiveness rating in order to keep the bonus. For the purposes of this program, a teacher is defined as a classroom teacher with assigned students and associated evaluation scores. It excludes principals, school counselors, and school services personnel. Itinerant teachers can receive a pro-rated amount

[21] TDOE and the Governor's office also implemented a teacher signing bonus program. To help attract the most effective teachers to Priority Schools, a signing bonus of $7,000 was offered to every new Level 5 teacher who transferred from a non-Priority School into a priority school during the 2013-14 school year. Only 59 teachers received the signing bonus. Due to a small sample size, I do not specifically evaluate this aspect of the program.

of the retention bonus based on the number of days per week that he or she is actually working in a Priority School.

Priority Schools

In 2012, the TDOE secured waivers from certain portions of the federal No Child Left Behind (NCLB) Law. The waiver allowed Tennessee to replace NCLB's Adequate Yearly Progress proficiency targets with a system that focuses on "ensuring growth for all students every year and closing achievement gaps by ensuring faster growth for those students who are furthest behind" (TDOE, 2012). Additionally, the state identifies individual schools based on these relative performance measures, ranging from high-performing "reward" schools to low-performing "priority" schools.

Tennessee identified 82 Priority Schools based on their three-year composite proficiency rates for all enrolled students. The bottom 5% of schools in the state were assigned Priority Status. The composite proficiency rate used to determine schools' eligibility for Priority School status is based on the percentage of students rated as proficient or advanced in math, reading/language arts, and science in grades 3-8. The state uses up to three years of data for determining Priority School status.

Teacher Evaluation

In January 2010, the Tennessee General Assembly passed Senate Bill 5, also known as the First to the Top Act, which reformed dozens of aspects of state education policy. As part of the federal Race to the Top Competition, the ambitious reforms helped Tennessee win a $501 million award to implement and institutionalize innovative policy changes statewide. One of the most contentious provisions of the new law required that all school personnel be evaluated annually, and personnel decisions be based, in part, on those evaluations.

In the year prior to the first round of Priority School designations, the Tennessee State Board of Education approved a series of teacher evaluation models as districts began implementation of the state's First to The Top Act requirement that all school personnel be evaluated annually, and personnel decisions be based (in part) on those evaluations. The evaluation models all follow the requirements set forth by Tennessee's Teacher Effectiveness Advisory Committee, and are adopted by the State Board of Education, as described in detail in Springer et al. (2016). In short, the annual evaluations differentiate teacher performance, based on an overall level of effectiveness score (often referred to as *overall teacher evaluation rating* or *teacher rating* for short), groups teachers into five discrete effectiveness categories (Level 1: "Significantly Below Expectation"; Level 2: "Below Expectation"; Level 3: "At Expectation"; Level 4: "Above Expectation"; and Level 5: "Significantly Above Expectation").

The overall teacher evaluation rating is calculated using individual and school-level student growth scores and achievement data as well as teacher observations for teachers in tested and untested subjects and grades. For tested teachers, state law specifies that 50% of the evaluation score be based on student achievement data. Of this 50%, 35% is comprised of value-added student achievement data as calculated using the Tennessee Value Added Assessment System (TVAAS), and 15% is based on alternative measures of student achievement approved by the State Board of Education and selected through joint agreement by the educator and evaluator. The remaining 50% of an evaluation must be based on qualitative measures, including teacher observations, student perception surveys, personal conferences, and reviews of prior evaluations and work. For untested teachers, 40% of the evaluation is comprised of student achievement data: 25% is based on school- or district-wide student growth as measured by TVAAS, and 15% is based on additional approved measures of student achievement. The remaining 60% of the overall evaluation scores are determined through qualitative measures that are similarly used for tested teachers.

Tested subject area teachers are designated as Level 5 if they score a Level-5 rating on a three-year composite TVAAS, or if they achieve Level 5 status on overall evaluation. Teachers of non-tested subject areas must achieve Level 5 on overall evaluation. In Tennessee, roughly 7% of teachers with the highest teacher effectiveness ratings exit their schools each year. However, in the schools designated as Priority Schools (the bottom 5% on school-level student test scores, whose students are roughly 90% Black and 90% eligible for free or reduced-price lunch), attrition rates for highly effective educators is 23% (more than three times the statewide average).

Literature on Selective Retention Bonuses

Several recent studies have evaluated the effectiveness of offering SRBs to attract or retain desirable teachers in hard-to-staff, disadvantaged, and low-performing schools, with mixed results. The nature, size, and context of the evaluated bonuses vary considerably, as do the methods used to assess their impacts. One of the difficulties SRB studies have faced is the fact that policymakers often introduce retention bonuses in the context of a broader set of reforms (i.e., Balch & Springer, 2014; Dee & Wyckoff, 2013; Hough, 2012; Sun, Penner, & Loeb, 2017).[22] For example, in their evaluation of a pilot supplemental funding program for a group of educationally disadvantaged schools in North Carolina, Henry and colleagues (2010) found that approximately half of the money went towards salary bonuses that gave the schools a comparative advantage in hiring and retaining

[22] Another common form of recruitment incentive is the use of scholarship programs for teachers that condition receipt of payment on teachers serving in disadvantaged schools for a specified period of time (Johnson, 2005). Steele, Murnane, and Willett (2010) evaluated California's Governor's Teaching Fellowship (GTF) Program, which offered $20,000 conditional scholarships ($5,000 per year over 4 years) to attract and retain academically talented, newly licensed teachers to low-performing schools. The program had a significant effect on teacher recruitment, but did not differentially affect teacher retention among GTF recipients and non-recipients.

teachers. Teacher turnover decreased significantly at the schools with the supplement—in spite of having the most disadvantaged students in the state—while turnover rates increased at non-supplement schools. However, while the RD design of the study allowed the authors to attribute the increased retention to the supplemental funding, they were unable to distinguish the effects of salary bonuses from other expenditures, such as expanded professional development, reduced class size, or upgraded materials and equipment—all of which might have made teachers more likely to stay.[23]

Dee and Wyckoff's 2013 analysis of salary bonuses in IMPACT, a high-stakes teacher evaluation system implemented in Washington, D.C. and designed to improve teacher quality and student achievement, faced similar challenges. Using a RD design, the authors compared teachers near the IMPACT score threshold that separated "Effective" from "Highly Effective" teachers. Like the evaluation program in Tennessee, the D.C. system utilized a mix of observation and value-added metrics to generate a continuous composite score with sharp cut points to group teachers into consequential categories of effectiveness. Teachers qualified for a large one-time bonus (up to $25,000) after being rated "Highly Effective" for one year, and a sizable and permanent base salary increase (as large as $27,000 per year) upon achieving "Highly Effective" status in a second consecutive year. While the IMPACT incentive had positive effects on teacher performance, impacts on the retention of effective teachers were not statistically significant. In a follow-up study, the compositional effects of selective forced attrition at the low end of the IMPACT teacher rating system found that, when exiting teachers were replaced by teachers with higher demonstrated effectiveness, students perform better the following year (effect size: .08 SD), in spite of the increase

[23] In a similar vein, Hough (2012) assesses the effect of a salary increase on teacher retention in the San Francisco Unified School District as part of the Quality Teacher and Education Act of 2008 (QTEA). QTEA introduced an overall salary increase ranging between $500 and $6,300 based on placement on the salary schedule, a $2,000 bonus for teaching in a hard-to-staff school, and a $2,500 retention bonus after the 4th year of the program followed by $3,000 after the 8th year. Hough finds that the QTEA salary increase did not affect the retention of targeted teachers, though overall teacher retention rates increased following the implementation of the program, to which the author attributes the null finding.

in faculty churn (Adnot, et al., 2016). However, the theory of action in this paper is that the opposite should hold. If teachers who are retained in response to a policy intervention have greater measured effectiveness than their likely replacements, students' test scores should improve in those subjects.

More directly focused on SRB policies, Clotfelter and colleagues (2008) evaluated the impact of a $1,800 teacher retention bonus offered in North Carolina between 2001 and 2004 to certified math, science, and special education teachers in a set of low-performing and/or high-poverty secondary schools. The authors found modest but significant effects on teacher turnover. The difference-in-difference-in-difference analytical strategy indicated that the bonuses reduced turnover rates of eligible teachers in eligible schools by 17%, or five percentage points. Survey results also indicated widespread misunderstandings about the nature of the retention incentive offered, and skepticism among teachers and administrators that the size of the bonus would be sufficient. The North Carolina bonus program differed from the Tennessee retention bonuses both in its smaller magnitude ($1,800 vs. $5,000) and in the fact that it was not tied to any measure of teacher quality, but rather specified credentials (rewarding teachers who were trained in math, science, and special education).

Finally, though selective transfers or signing bonuses, which elicit the transfers of desired teachers to high-need schools, differ in important ways from those seeking to retain the existing talent pool, it is worth describing one large-scale experiment that financially incentivized both the transfer and (for a time) retention of high value-added instructors in high-poverty schools. Glazerman and colleagues (2013) evaluated a substantial monetary incentive offered through the Talent Transfer Initiative (TTI) across 10 school districts in seven states, which was designed to recruit high-performing teachers to low-performing schools. Using a randomized controlled trial design, teachers who demonstrated a sufficient level of value-added effectiveness (roughly the top 20% for their subject and grade) were eligible for a $20,000 bonus—paid in installments over a two-year period—if they transferred into and remained in schools that had low average test scores. Results showed that the transfer incentive had substantial positive impact on teacher-retention rates during the payout period; retention rates were significantly higher for high-performing teachers compared to their counterparts: 93% versus 70%. However, not surprisingly, the difference was no longer statistically significant after the payments stopped. With respect to student achievement, the study also found significant impacts of the induced placement of high value-added instructors in elementary schools, but not middle schools, where the context of the new school could presumably play a larger role. This is an important distinction between transfer incentives and SRB incentives, where the bonus recipients have already demonstrated their context-specific effectiveness.

The current study provides an important follow-up to a 2016 evaluation of the teacher retention effects of the same SRB program in Tennessee. Springer et al. (2016) used a RD design that exploited the sharp cutoff in teachers' eligibility to receive a SRB worth $5,000. The researchers found that teachers of tested subject areas who were eligible for the bonus were roughly 20% more likely to remain in their priority schools than those who were rated just below the eligibility threshold. The study found no effects on teachers of untested subject areas, who face greater accountability pressure because a larger proportion of their overall rating is tied to school-level performance. Teachers who accepted bonuses had overall teacher effectiveness ratings of more than a full standard deviation above the state average, and roughly 1.7 standard deviations greater than the average teacher hired to replace them. While the retention effects are arguably modest, the substantial difference in expected outcomes for students taught by each retained teacher informs the hypothesized improvement in student outcomes tested in this study.

Data

This study utilizes administrative data obtained from the Tennessee Department of Education (TDOE) and maintained by the Tennessee Education Research Alliance (TERA) at Vanderbilt University's Peabody College. The data available for this study includes student standardized test scores on the Tennessee Comprehensive Assessment Program (TCAP), and student and teacher demographics from 2007-8 to the 2013-14 school years. School-level information comes from multiple sources, including state school accountability reports, The National Center for Education Statistics' Common Core of Data, and by aggregating individual student- and teacher-level information at the school level. These school files contain basic information, such as average daily membership, proficiency rates, as well as select student and teacher demographic information. TDOE also provided our research team at TERA (formerly known as the Tennessee Consortium) with details on the design and implementation of the teacher retention bonus program. The teacher retention bonus program file contains the teacher's name, the school's name, and the local education agency for all teachers who received a retention bonus. School files contain typically-used information, such as level of schooling, a measure of total enrollment, average daily membership (ADM), and proficiency rates, as well as select student and teacher demographic information.

I also incorporate teacher-level data from the Tennessee Value-Added Assessment System (TVAAS) and Tennessee's online teacher evaluation platform, CODE. The TVAAS data file, created by SAS Institute in Cary, North Carolina, contains value-added estimates for teachers in grades 4 through 8 in math, reading/language arts, science, and social studies, and end-of-course reporting for high school educators in English I, II, III, Algebra I and II, Biology I, and U.S. History. Teacher effect estimates are calculated for specific subject, grade, and year pairings as well as for composites across subjects, grades, and years. All scores are expressed in state normal curve equivalents, using the 2008-09 school year as the reference year. My analysis uses subject-specific 3-year-averaged TVAAS estimates for math and reading/language arts teachers, where teachers are given a numeric rating on a 5-point scale.[24]

The primary depended variables of interest in this study are student-level test scores on the state's Tennessee Comprehensive Assessment Program's (TCAP) annual accountability exams, and the subject-specific school-level average TVAAS scores described above. Students' 3rd to 8th grade reading and math scores are standardized within grade level, subject, and year to have a mean of zero and a standard deviation of one, such that they represent a student's performance relative to students' taking the same exam in any given year. They are also comparable over time.

Additionally, I include a series of control variables for student characteristics, including indicators of student gender, race, or ethnicity; whether the student is identified as an immigrant; whether s/he participates in the free or reduced-price lunch program (FRPL); and whether the student has a special education placement.[25] In addition to the school-fixed effect that accounts for any invariant school-level characteristics, all analyses include a set of time-variant school-level controls—specifically, the percentage of students in a school who are identified as Black, Hispanic, or participate in FRPL, as well as the schools' average daily membership (ADM)—to account for changes in the composition of the size of the student population.

Analytic Sample Construction

Priority Schools targeted by the retention bonus program differ dramatically from other schools in the state in important ways that could have influenced both the patterns in their teachers' labor market decisions as well as student outcomes. Following Abadie (2005) and Heckman,

[24] For more information on TVAAS, see http://www.tn.gov/education/TVAAS.shtml.
[25] Notably, students who were administered modified assessments are omitted here for comparability concerns.

Ichimura, and Todd (1997), I use a multivariate matching strategy to balance treatment and control groups on a broad set of pre-treatment characteristics. Several studies have found that this method of matching on observables, in combination with a traditional difference-in-difference model, substantially reduces the susceptibility to bias from violations of the common trends assumption, which I discuss in greater detail below (e.g., Abadie, 2005; Heckman et al., 1997; O'Neill, Kreif, Grieve, Sutton, & Sekhon, 2016). To construct the primary analytic sample, I use probit regression to estimate the likelihood of a school's participation in the SRB program, as predicted by teacher and student characteristics in the 2012-13 school year (the year prior to bonus distribution). School characteristics in the model include the average teacher age; value-added test scores; teachers' overall evaluation ratings; the proportions of teachers who are women and/or have one or fewer years of experience; average daily membership (ADM); and the percentage of students who are Hispanic, Black, are on free or reduced-price lunch, and who score as proficient or advanced on state exams. Not surprisingly, given the definition of Priority School status and eligibility rules for distributing bonuses, the results displayed in Table 14 indicate that the strongest predictors of participation in the SRB program are student demographics, student test scores, and teacher effectiveness ratings. Specifically, schools with more Black and low-income students with low proficiency rates but high teacher test score value-added ratings are more likely to have participated.

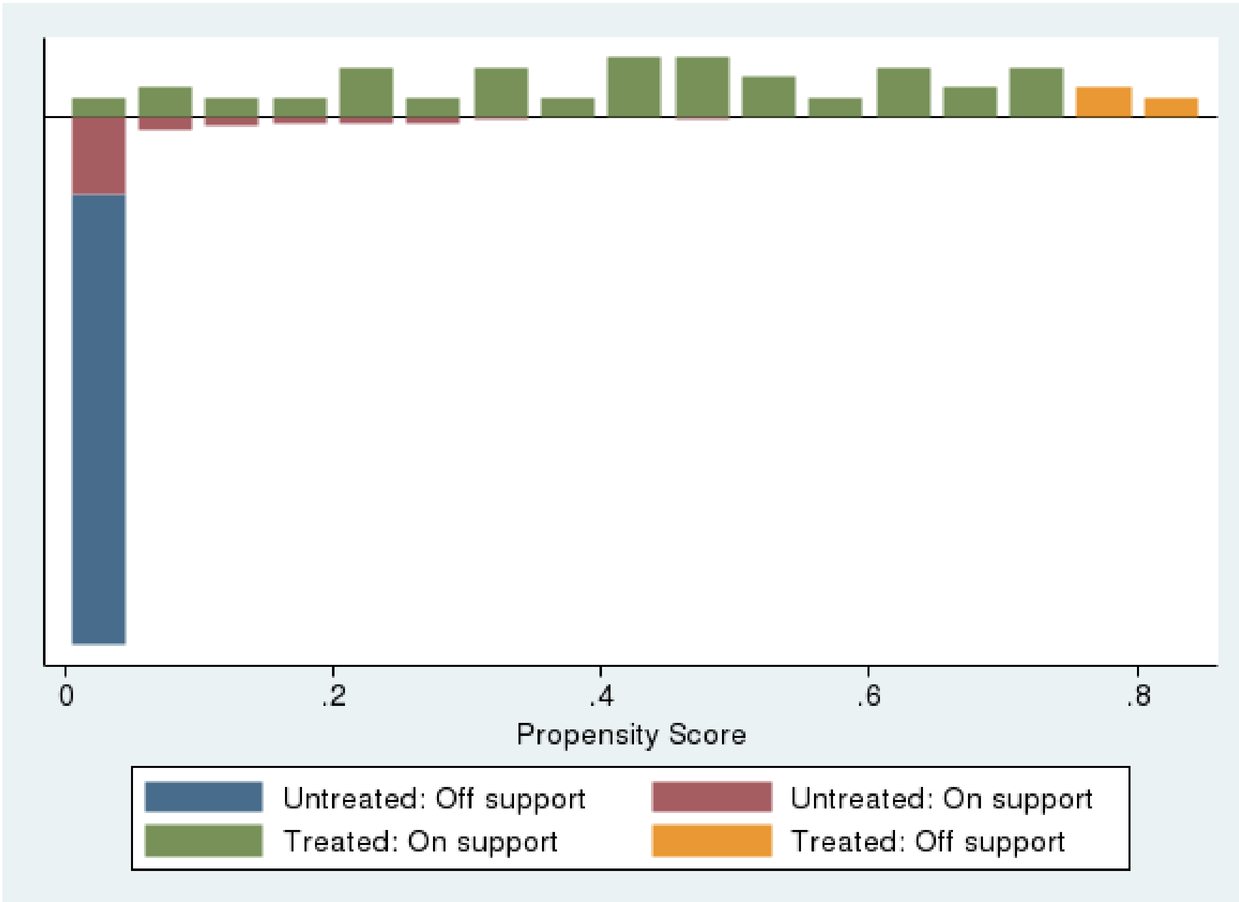
Table 14: Results from Probit Regression Predicting Bonus Program Participation in The 2012-13 School Year

	Coef.	Std. Err.	P-Value
% Novice	1.167	1.272	0.359
% Female Teachers	-1.073	0.866	0.215
Avg. Age Teachers	-0.001	0.036	0.971
Avg. Overall Effectiveness	-0.074	0.254	0.770
% level 5 Teachers	1.140	0.781	0.144
Avg. TVAAS Level	0.624	0.192	0.001
Urbanicity	0.001	0.065	0.982
ADM	<0.000	<0.000	0.418
% Black Students	2.152	1.047	0.040
% Hispanic Students	-0.028	1.748	0.987
% FRL	4.071	2.312	0.078
% Proficient/Advanced	-0.056	0.014	<0.000
Constant	-5.832	2.816	0.038
Pseudo R2	0.580	N	1602

After estimating the likelihood of participation, I use a kernel matching technique as described in Imbens (2004) with a bandwidth of .06, as suggested by Heckman et al. (1997) in a linear probit framework to identify a set of SRB-eligible schools and non-participant schools that are observably similar. This procedure resulted in nine of the 63 SRB-eligible schools' elimination from the analytic sample, which is primarily driven by abnormally low proficiency rates. A total of 349 out of 1,543 non-participant schools were matched to one or more of the 63 treated schools for comparison purposes. Figure 6 plots the distribution of TN schools included or omitted from the treatment and comparison group samples.[26]

[26] Alternative bandwidths, as well as different underlying matching strategies, such as nearest-neighbor matching with varying caliper size, were also tested for robustness to sample selections that were more or less inclusive, and results were qualitatively similar.

Figure 6: Graph of Common Support for Matched Sample of Schools Based on Propensity to Participate in the Bonus Program



Note: A total nine of the 63 participant schools (treated) were considered "Off Support" due to high propensity scores that lack a comparable school, and 349 of 1,543 non-participant schools (Untreated) were matched using a kernel match with a .06 bandwidth in a linear probit framework.

Table 15 reports descriptive statistics and mean comparisons for both student and teacher demographics for the state as a whole (Column 1); the 63 Priority Schools that distributed at least one retention bonus (Column 2); and the 26 Priority Schools that did not (Column 3). Priority Schools differ dramatically from the rest of the state on academic performance by design. They are also distinctive outliers demographically. In the 2011-12 school year, the average public school in the state was roughly 70% White, and less than 60% of students qualified for free or reduced-price lunch. In contrast, the average Priority School was less than 3% White, more than 90% Black, and nearly 90% of students qualified for free or reduced-price lunch.

When comparing participating and non-participating Priority Schools, few differences are substantively or statistically significant. Student demographics are near identical, with only slightly lower proficiency rates in non-participating Priority Schools (27% vs. 32%). However, Non-participant Priority Schools had slightly more male teachers (29% vs. 21%), lower average test score value-added levels (2.2 vs. 3.2 on a 5-point scale), and notably fewer teachers who surpass the Level 5 rating (5% vs. 31%), which rendered most non-participating Priority Schools functionally ineligible to participate in the SRB program.

The final two columns of Table 15 present select descriptive statistics for the schools that represent the SRB treatment, and matched comparison schools in my primary analytic sample. Like Priority Schools in general and bonus participants specifically, the matched comparison sample of schools have lower test scores, with nearly 30% fewer students surpassing proficiency thresholds on state exams. The match comparison sample schools are also less rural, and have much higher proportions of free or reduced-price lunch students, non-White students, and teachers with fewer than two years of experience. At the same time, comparison schools differ slightly on average from SRB-participant schools, as they have lower percentages of Black students, higher percentages of

Hispanic students, are more likely to be located in smaller metro areas, and have less extreme proficiency rates (44% vs. 32%).[27]

[27] It is also instructive to assess the comparability of the participating schools to the non-participant schools, which constitute an important component of my primary comparison group due to their potential selection out of treatment. A more inclusive table showing the consistent comparability of these groups is included in Appendix A.

Table 15: Select Summary Statistics on Alternative Comparison Groups for Analytic Sample

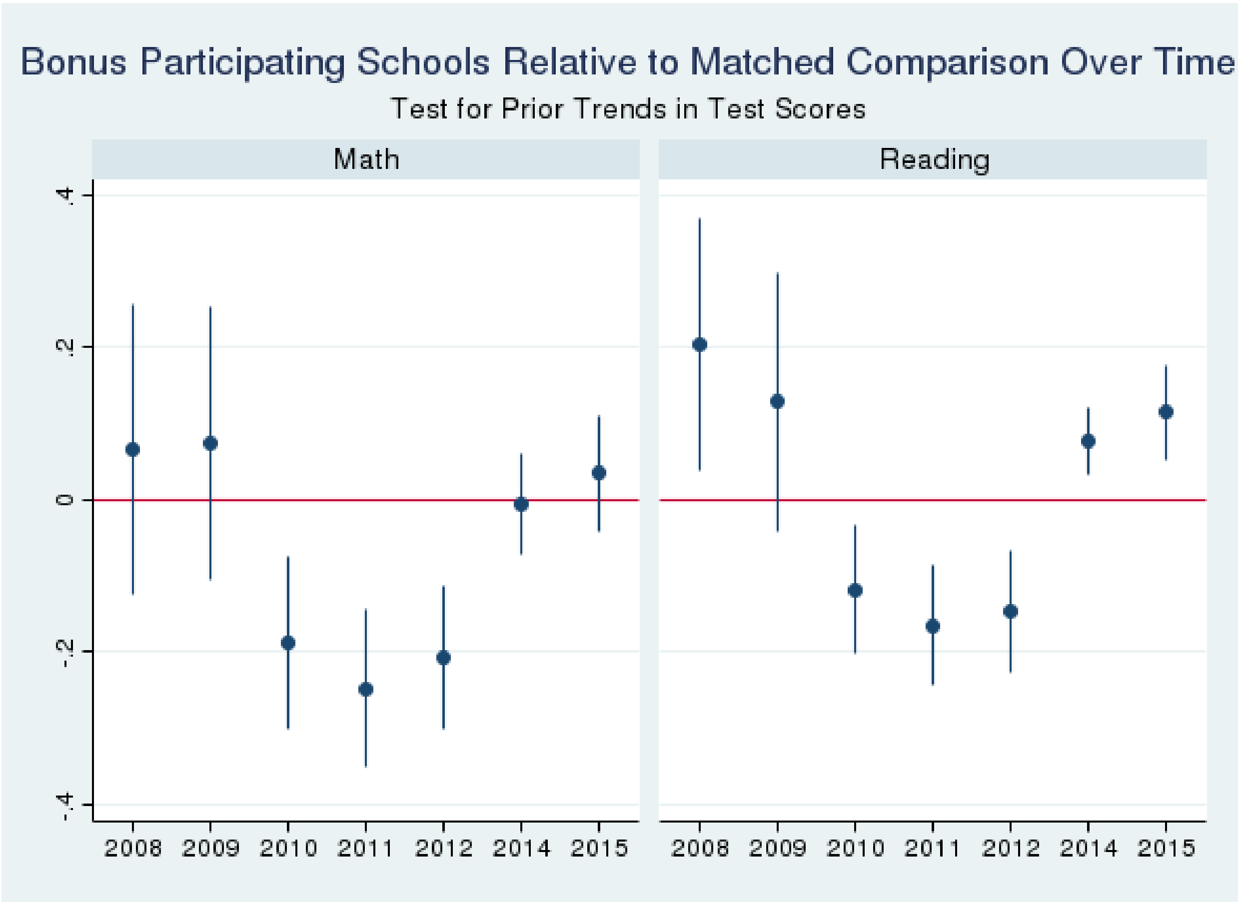
	All TN Schools		Bonus Participants		Eligible Non-Participants		Matched Comparison Schools		Bonus Participants (after Sample Restriction)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Teacher Characteristics										
% Novice	0.14	0.12	0.18	0.11	0.23	0.15	0.18	0.14	0.18	0.11
% Female Teachers	0.81	0.17	0.79	0.13	0.71	0.16	0.80	0.14	0.80	0.13
Avg. Age Teachers	43.04	4.53	42.47	3.65	42.85	3.58	42.31	4.26	42.96	3.58
Avg. Overall Effectiveness	3.76	0.71	3.60	0.82	2.72	0.51	3.44	0.84	3.53	0.83
% Level 5 Teachers	0.31	0.24	0.31	0.23	0.05	0.08	0.25	0.23	0.29	0.22
Avg. TVAAS Level	3.37	0.70	3.24	0.70	2.25	0.75	3.17	0.76	3.17	0.67
Urbanicity rating	6.66	3.86	1.35	1.18	1.48	1.12	3.50	3.52	1.41	1.27
Student Characteristics										
ADM	568.65	355.65	460.39	233.87	431.95	164.92	577.72	338.92	484.98	236.91
% Black Students	0.24	0.32	0.92	0.13	0.92	0.09	0.61	0.31	0.92	0.13
% Hispanic Students	0.07	0.09	0.05	0.08	0.03	0.05	0.12	0.14	0.05	0.09
% FRL	0.64	0.22	0.94	0.05	0.93	0.06	0.86	0.10	0.94	0.05
%Proficient/Advanced	58.91	14.96	32.26	8.72	26.95	4.77	44.32	11.29	32.57	9.08
	N	1,860	N	63	N	25	N	349	N	54

Test for Pre-Existing Trends

While I was able to establish a reasonably comparable set of SRB-participant and comparison group schools for the primary analytic sample, the fact that SRB-eligible schools, by definition, included the lowest performing schools in the state according to a school accountability metric, made it difficult to construct a comparison sample of schools that were identically equivalent. However, in the difference-in-difference framework, the comparison groups are not required to be equivalent at baseline; instead, they must satisfy the parallel trends assumption (e.g., Abadie, 2005; Heckman et al., 1997). A violation of the parallel trends assumption for the comparison group poses the greatest threat to an unbiased estimation of the SRB impact. Priority Schools are uniquely low-performing, and were subsequently subject to overlapping interventions. Thus, it can be difficult to rule out the prospect that coincident pressures or shocks are influential in the changes I attribute to the SRB policy.

To test for the presence of these secular trends, I first ran a conventional event history analysis, where the treatment indicator was interacted with each of the years in the panel, omitting as a reference category only the year prior to the implementation of the SRB policy (here, the 2012-13 school year). While some significant differences are observed in prior years, they generally favor the matched comparison group of schools in periods preceding the SRB intervention (see Appendix B). Visually, Figure 7 shows that, if anything, the difference in trends for students' performance levels in SRB schools were negative in direction prior to the policy intervention, which appears to have substantially redirected their trajectory.

Figure 7: Tests for Prior Trends Associated with Bonus Participation



Note: The coefficients plotted above represent the results from a standard test for pre-existing trends, wherein the treated school indicator (Bonus Program Participants) is interacted with indicators for each year in the panel except the one directly preceding intervention, which is omitted to serve as reference. (e.g., Autor, 2003).

Test for Student Sorting

Another potential threat to the interpretation of the achievement effects as causal impacts of retained teachers on student performance is a concurrent compositional shift in the student population of the participating schools who comprise our treatment group. While it is plausible that the retention bonus itself might drive achievement-positive changes in the student composition (e.g., greater continuity of the effective faculty could disproportionately attract and retain students of more highly engaged parents), and could be construed as an indirect policy effect, this is not the primary mechanism through which SRBs are designed to bolster achievement. More important, in the context of the school-choice-friendly, large urban districts where Priority Schools are concentrated, changes in the composition of schools participating in the SRB program could, by happenstance or through some characteristic correlated with participation, experience shifts in the student population that bias estimates of the program's effects.

To assess the extent of this threat, I estimated an identical difference-in-differences model to those used to derive treatment effects, but substituted measures of student composition as the outcome of interest. Table 16 displays estimates of the association between bonus participation and student composition of schools, with limited evidence of significant student sorting. Demographically, the participating schools are largely consistent over the treatment period, with changes in the proportion of students who are Black, Hispanic, or qualifying for free or reduce-priced lunch indistinguishable from zero. While the coefficient on the change in the proportion of White students is statistically significant, it is questionable whether the magnitude is substantively significant, when you consider that the 2-percentage point increase only changes the participating schools from being 3% White to 5% White on average. Slightly more concerning is the reduction in average daily membership (ADM) associated with SRB participation. All models statistically control for these time-variant school characteristics, but it is possible that their association with program participation is indicative of some unobserved characteristic of the participating schools that would cause them to violate the common trends assumption underlying the causal interpretation of the difference-in-differences strategy. The reductions in student totals could also facilitate greater attrition of novice teachers who would generally be the first to be reassigned to other schools in response to under-enrollment.

Table 16: Bonus Participation and Student Composition Change

	Black	Hispanic	White	FRPL	ADM	SPED	Low Math	Low Reading
Bonus	-0.007	-0.009	0.022**	0.020	-91.869**	0.001	-0.014	-0.007
	(0.006)	(0.006)	(0.003)	(0.018)	(23.756)	(0.005)	(0.016)	(0.015)

Note: Robust Standard Errors clustered at the school level (N=403 Schools + p<0.1; * p<0.05; ** p<0.01).

Primary Analytic Strategy

As discussed above, I exploit the timing of policy implementation and the school-level threshold in eligibility for schools to offer SRBs in order to estimate the effects on subsequent student achievement in a difference-in-differences (DiD) framework. Shifts in outcomes for students and teachers in participating schools, relative to matched comparison schools in the post-treatment period, are attributed to the SRB intervention, after controlling for an extensive set of time-variant characteristics, and any characteristics of the school that are fixed over time. The primary analytic model takes the following form:

$$(y_{ist} = \beta_0 + \beta_1 bonus_{st-1} + \beta_2 priority_{st} + \beta_3 ASD_{st} + \beta_4 iZone_{st} + \beta_5 y_{ist-1} + \boldsymbol{\lambda_{ist}}\beta_6 + \alpha_s + \gamma_t + \epsilon_{ist})$$

where, y_{ist} represents the standardized test score in math or reading of student *i* in school *s*, in year *t*. $Bonus_{st}$ is an indicator for whether school *s* distributed any bonuses in the spring of the preceding academic year *t*, and can be understood as the conventional second difference, approximating the SRB treatment effect. $Priority_{st}$ is an indicator for whether school *s* was designated a Priority School. The Priority School designation preceded the announcement of the bonus program by 1 year, so this indicator captures pre-treatment shocks experienced by participant schools. y_{ist-1} represents the lagged value for the outcome of interest and makes the treatment effect interpretable as an impact on student test score growth. I also include indicators for whether a school is incorporated into the Achievement School district ASD_{st} or the state's Innovation Zone $iZone_{st}$ to account for potentially overlapping treatments.[28]Here, α_s represents a school-fixed effect, γ_t is a year-fixed effect, $\boldsymbol{\lambda_{ist}}$ is a vector of student characteristics (gender race/ethnicity, immigrant status, free or reduced-price lunch status, special education placement, and grade level) and time-variant school characteristics (ADM, racial composition, %FRPL). ϵ_{ist} represents an individual error term clustered at the school level. The coefficient of interest is β_1, which represents the estimated treatment effect of the SRB program.

DiD methods comparing pre- and post- intervention outcomes for a treatment and control group identify causal effect estimates, where the parallel trends assumption holds (e.g., Ashenfelter 1978; Bertrand et al., 2004; Jones, Rice, & Rosa-Dias, 2011). However, to better account for the potential of endogenous selection into treatment at either the student or school level, I estimate a series of alternative model specifications. The first is a conventional difference-in-differences model (DiD) described above, that should produce the most efficient estimate of treatment effects if the common trend assumption holds. I then estimate an identical model incorporating a lagged dependent variable (LDV) to account for potential bias on unobserved student characteristics (O'Neill et al., 2016). Finally, I estimate a model that incorporates a student-fixed effect to control for any student characteristics that are constant over time. This most restrictive model estimates treatment effects based on within-student variation in student outcomes for students who have test scores in bonus schools before and after the selective retention bonus program was implemented.

Ultimately, the primary identification strategy here produces a treatment on the treated (TOT) estimate of participation in the SRB program by comparing the change in student test score

[28] Over the course of the last three years of the panel, 23 Priority Schools were incorporated into the privately managed Achievement School District, and 26 were designated as iZone schools, operating as semiautonomous district-within-district schools. Both sets of schools were given greater latitude to promote differential attrition of their teacher labor force (Zimmer, Henry, and Kho, 2017). In the final year of the panel, eight Priority Schools were closed, and are omitted from two-year estimates of treatment effects.

gains among SRB-eligible schools to those of similar, nearly eligible schools that did not participate in the SRB program. The distinction between the TOT estimate here and an intent to treat (ITT) model that assesses the policy effect of simply offering principals the opportunity to grant bonuses but ignores the minimal non-compliance or opt-outs, is functionally nominal; nearly all schools that were actually eligible (being both designated Priority Schools and having at least one teacher rated Level 5 in the year the bonus program was announced) distributed at least one retention bonus.

In the following sections, I describe the primary results of these analyses, including alternative specifications and robustness checks to add confidence to a causal interpretation of the findings. I then discuss implications for public policy.

Results

After accounting for a range of potential confounding factors, I find evidence that participation in the SRB program drove improvements in student test scores. Students' test scores in reading improved significantly more (effect size ~.10 sd) in schools participating in the SRB program than in otherwise similar non-participant schools in the years following implementation. Reading achievement effects also appear to persist at least one year after the teacher retention incentives are removed. Impacts on math scores were smaller in magnitude (effect size ~.05 sd) and only marginally significant ($p<.10$) in more restrictive model specifications. I find suggestive evidence that the differential effects by subject area may be driven by smaller effects of the retention bonus on effective math teachers' decisions to remain at Priority Schools.

Effects on Student Test Scores

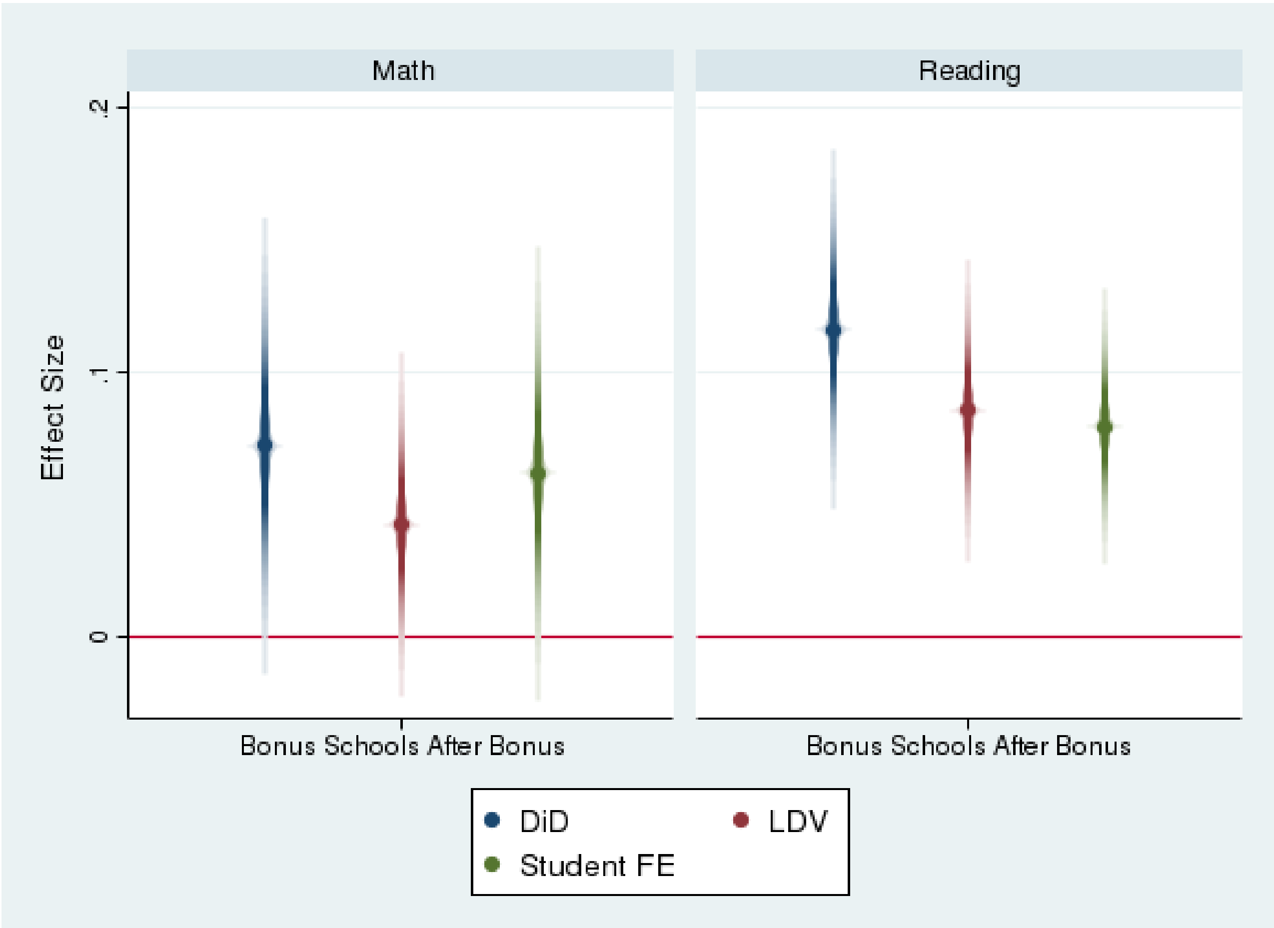
Table 17 shows the primary results from the difference-in-difference model comparing changes in math and reading scores among schools participating in the SRB program to those of students in a matched sample of schools that did not participate. Controlling for student characteristics and time-variant school characteristics, reading scores of students in SRB schools improved by roughly 0.10 of a standard deviation in the years following program implementation. Students' math scores were also higher (0.07 SD) on average in SRB schools in the years following bonus distribution, though effect estimates are smaller in magnitude (0.04-0.06 SD) and only marginally significant in more restrictive models that account for students' prior test scores or student-fixed effects. I also display results visually in Figure 8 below.

Table 17: Bonus Participation and Student's Math & Reading Test Score

	Diff-in-Diff		LDV		Student-Fixed	
	Math	Reading	Math	Reading	Math	Reading
Bonus Schools After Bonus	0.072*	0.116**	0.042+	0.085**	0.062+	0.079**
	(0.033)	(0.026)	(0.025)	(0.022)	(0.033)	(0.020)
R2	0.15	0.17	0.53	0.61	0.80	0.85
N	733,107	741,186	565,551	571,726	733,107	741,186
Year FE	✓	✓	✓	✓	✓	✓
School FE	✓	✓	✓	✓	✓	✓
Student FE					✓	✓

Notes: The first two columns present results from the standard difference-in-differences model with school- and year-fixed effects. The second two mirror the first, but add a lagged version of the dependent variable, making the outcome interpretable as a gain score. The final two columns present the results from a student-fixed effect model, accounting for any student characteristics that are constant over time. All models contain the full set of time-variant school-level controls (% Black, % Hispanic, % FRPL, ADM), student demographics (gender, race, immigrant status, FRPL, grade-level) and year-fixed effects. The sample is restricted to the 403 schools in the matched sample, for years 2009-2015. Robust standard errors clustered at the school level are in parentheses (+ p<0.1; * p<0.05; ** p<0.01; + p<0.1; * p<0.05; ** p<0.01).

Figure 8: Visualization of Achievement Effect Estimates by Modeling Strategy



Notes: This figure plots coefficients and their confidence intervals from six separate regression models predicting two test-score outcomes using alternative modeling strategies. The first coefficient plot in each panel displays the effect estimate from the conventional Difference-in-Differences (DiD) model. The second shows the effect estimate from a similar model that incorporates a lagged dependent variable (LDV), and the final coefficient plot comes from a student-fixed effect model.

Effects on Teacher Composition

While prior work established that SRBs increased the likelihood of retention of highly effective teachers of tested subject areas by roughly 20 percentage points (Springer et al., 2016), the causal estimates were based on teachers near the threshold for eligibility, and no distinctions were made based on the subjects teachers taught. Here, mirroring the analytic strategy applied to the student test scores, I test whether participation in the bonus program was associated with increases in the composition of teacher effectiveness as measured by the time-lagged subject-specific test score value-added estimates. I also estimate the effect of the bonus offer on the share of teachers who had a Level 4 or 5 rating on their TVAAS score, which most closely approximates the target of the bonus eligibility requirements.[29]

Table 18 displays the results from the difference-in-differences estimation of bonus program impacts on schools' average teacher effectiveness in math and reading, as well as the average years of experience and education levels of teachers. Mirroring the primary results for student achievement, bonus program participation was associated with a substantial increase (roughly 0.6 higher average on a 5-point scale) in the average reading test score value-added of the schools' faculty, and a smaller marginally significant increase in the average level of math value-added. I find no significant effects on the levels of education or experience of teachers, neither of which were directly targeted by the SRB program. These findings are consistent with the theory of action underlying the mechanism by which the SRB program would influence student achievement. They are also consistent with findings from prior teacher labor market literature documenting the particular challenges of recruiting and retaining math instructors, whose credentials often open more lucrative employment opportunities outside of teaching (e.g., Rumberger, 1987).

[29] During the year that the bonuses were offered, roughly 75% of teachers who had a Level 4 or 5 TVAAS in reading earned an overall rating of Level 5 on their composite evaluations. The same was true for 60% of reading teachers.

Table 18: Bonus Participation and Teacher Composition Change

	Average TVAAS Level		% L4 or L5		Non-VAM	
	Math	Reading	Math	Reading	Experience	Education level
Bonus School After Bonus	0.205	0.606**	0.071+	0.169**	-0.767	0.398
	(0.130)	(0.095)	(0.038)	(0.032)	(0.521)	(0.594)
R2	0.26	0.25	0.21	0.17	0.20	0.11
N	15,514	17,614	15,514	17,614	80,530	85,125
Year FE	✔	✔	✔	✔	✔	✔
School FE	✔	✔	✔	✔	✔	✔

Notes: Estimates from panel 2009-2010 through 2014-15 school years. All models contain the full set of time-variant school- level controls (% Black, % Hispanic, % FRPL, ADM), student demographics (gender, race, immigrant status, FRPL, grade), and year-fixed effects. The sample was restricted to the 403 schools in the matched sample, for years 2009-2015. Robust standard errors clustered at the school level are in parentheses (+ p<0.1; * p<0.05; ** p<0.01).

Because there are two years of outcome data following the implementation of the SRB program, the second of which provided no additional incentives for teachers to stay, we can separate the effects of the program to test for persistence of benefits beyond the initial influx of funds. Here, I am also concerned about the potential for a mass exit in the second year, once teachers are informed that the bonus program will not be continued. Similar to the combined difference-in-differences results, where post-treatments are pooled (as displayed in Table 17), the apparent effects of the bonus are stronger for reading in both the first and second years following the bonus distribution. However, the benefits of the bonus on math teacher retention subside substantially after the state removed funding for the program, and math achievement effects are inconsistently significant when the effects are separated by year. On the other hand, the persistence of effects on reading teacher retention and reading achievement could be indicative of some stickiness in bonus effects, where a one-year reduction in turnover of effective reading teachers shifts the culture of the school to one of greater stability, reducing the subsequent attrition of teachers who are successful in the improved environment.

Table 19: Persistence of Bonus Effects

Bonus Participation Effects (Differential by Year) Test Scores & Teachers				
	Student Test Scores		High VAM Teacher	
	Math	Reading	Math	Reading
Bonus Schools After Bonus	0.022	0.085**	0.095*	0.155**
	(0.032)	(0.023)	(0.040)	(0.042)
Bonus Schools Year 2	0.065+	0.086**	0.012	0.197**
	(0.034)	(0.025)	(0.064)	(0.042)
R2	0.53	0.61	0.21	0.17
N	565,551	571,726	15,514	17,614

Note: All models contain the full set of time-variant school-level controls (% Black, % Hispanic, % FRPL, ADM), student demographics (gender, race, immigrant status, FRPL, grade), and year-fixed effects. Sample restricted to the 403 schools in the matched sample, for years 2009-2015. Robust standard errors clustered at the school level are in parentheses (+ $p<0.1$; * $p<0.05$; ** $p<0.01$).

Another way to examine the effect of participation in the SRB program is to examine differences by the extent of take-up. Substituting a series of indicators for the number of bonuses administered in each participating school for the simple treatment indicator (as displayed in Table 20), I found limited evidence of a "dosage" effect, such that the number of bonuses does not substantively differ from simple participation. However, because the number of bonuses distributed is also a function of the number of "highly effective" teachers in a school during the pre-treatment year (which could independently affect the average turnover), it is difficult to characterize these relatively flat dosage effects as decreasing returns to investment. Furthermore, as teacher exit decisions are not always entirely independent, the retention effects of the bonus are not necessarily restricted to the teacher who received a bonus. That is, if ineligible teachers in these low-performing schools are encouraged to stay based on the retention of the schools' most effective instructors, the bonuses could have a multiplicative impact on teacher composition and stability, which could be particularly strong in schools where prior turnover rates were exceptionally high.

Table 20: Bonus Participation Effects by Number of Bonuses (Teachers & Test Scores)

	Student Test Scores		High VAM Teacher	
	Math	Reading	Math	Reading
1 Bonus	0.018	0.085**	0.161*	0.111*
	(0.044)	(0.032)	(0.064)	(0.051)
2-5 Bonuses	0.047	0.134**	-0.007	0.175**
	(0.044)	(0.020)	(0.069)	(0.057)
6-10 Bonuses	0.019	0.058	0.152+	0.189*
	(0.053)	(0.057)	(0.080)	(0.085)
10 or more Bonuses	0.069*	0.048	0.043	0.199**
	(0.030)	(0.032)	(0.047)	(0.052)
R2	0.53	0.61	0.21	0.17
N	565,551	571,726	15,514	17,614

Notes: All models contain the full set of time-variant school-level controls (% Black, % Hispanic, % FRPL, ADM), student demographics (gender, race, immigrant status, FRPL, grade level), and year-fixed effects. Sample restricted to the 403 schools in the matched sample, for years 2009-2015. Robust standard errors clustered at the school level are in parentheses (+ p<0.1; * p<0.05; ** p<0.01).

Conclusions

As long as there are schools with high concentrations of racially and economically disadvantaged students, we must have a policy imperative to keep effective teachers from leaving these challenging—albeit rewarding—working environments. While the results presented here are modest in magnitude (effect size of roughly 0.05 to 0.10 SD), they highlight the potential of targeted incentives as a tool for minimizing the harm caused by racially and economically isolated schools. Additionally, the findings support the hypothesis that even modest increases in the retention of effective teachers in low-performing schools, where turnover is high and replacement teachers tend to be less effective or less experienced on average, can measurably improve student learning.

Proponents of teacher evaluation and tenure reform often argue that if we could identify the least effective teachers in the profession and somehow replace them with teachers of average effectiveness, the gains in student achievement and other long-run student outcomes would be substantial (e.g., Chetty et al., 2011; Chetty et al., 2014b; Hanushek, 2011). However, in many of the lowest performing schools, which also tend to have the highest concentrations of poor and non-White students as well as the highest rates of teacher turnover, the larger challenge may be identifying and retaining the most effective teachers, who are typically replaced by teachers whose measured effectiveness is well *below average*. Findings presented here indicate that Tennessee's $2.1 million pilot program offering SRBs to the highest-rated teachers in the lowest-rated schools succeeded, not only in increasing the share of high-value-added teachers, but in substantially elevating student performance in subsequent years.

Optimistically, retention bonuses tied to estimates of teacher effectiveness could serve as a tool for policymakers to improve the quality of the teachers instructing disadvantaged students without implementing layoffs or other punitive measures. Because teachers across the effectiveness spectrum often leave high-poverty, high-minority schools of their own volition, and are generally replaced by less-experienced, less-effective teachers, bonuses that retain teachers at the higher end of the effectiveness distribution can have substantial impacts on the quality of a school's faculty. In contrast to policies that would target teachers with poor evaluations or low value-added estimates for dismissal (which introduce churn and instability), SRBs can mitigate unwanted turnover and have the potential to strengthen leadership and institutional knowledge among schools' faculty, while avoiding the financial and social burdens associated with turnover.

Moving forward, policymakers implementing similar programs could benefit from additional steps to ensure principals and teachers in eligible schools are aware of bonuses and are supported throughout the implementation process. Also, given the year-to-year volatility test score-based, value-added measures of effectiveness, and the value employees place on predictability in their compensation, policymakers should consider creating opportunities for permanent or longer-term increases to base salaries for teachers with consistently high ratings. Finally, policymakers at all levels should make a concerted effort to address the broad range of factors that reduce the desirability of teaching in low-performing, high-poverty schools, including the concentrated poverty itself. Even marginal improvements on any of these fronts could increase the likelihood of this type of performance-based retention bonus's serving as both an incentive to stay at a hard-to-staff school and a reward for laudable work in a critical, challenging setting.

As is true for any policy that relies on observations and test-score-based, value-added estimates to differentiate teachers, the benefits of retention bonuses are only as strong as the measures of effectiveness are accurate. If, for example, the designation of "highly effective," based on the composite evaluation, is functionally random—or even falls more frequently on less desirable teachers—then the policy would not have the desired effects on the teaching pool, and could have discouraging effects on effective teachers who failed to receive the designation and monetary reward. However, the negative consequences of such mis-categorizations in the context of retention bonuses are seemingly less severe than those for teacher quality policies that rely on terminations. Alternatively, assuming that the effects of the bonus on teacher exit decisions were not entirely a function of its selectivity, a broader definition of Highly Effective Teachers that makes more teachers eligible (e.g., teachers who earn a Level 4 or 5) could facilitate a larger school-level ecological impact through improvements in faculty stability at the top end of the effectiveness distribution, and reduced reliance on teachers who are unfamiliar with the school or the profession. While more expensive up front, a policy that allowed for bonus pay to become a permanent salary supplement, conditional on continued service in high-need schools, could increase the perceived desirability of retention enough to be justified as cost-effective. Further research is necessary to estimate the need to tie these high-need-school retention bonuses or salary bumps to consistent high-rated instruction (e.g., experimentation comparing conditional and unconditional retention bonus offers in high-need schools).

The high rates of teacher turnover in low-performing, high-poverty schools are concerning on multiple fronts. It is not hard to imagine how difficult it is to create a supportive learning culture in a school where sizable portions of the teaching force are new to the building—if not the profession—each year. Even more problematic are the exits of teachers who have proven themselves highly effective in these challenging settings. The loss of these exemplary instructors not only lowers the average efficacy of the schools' teaching force, but also limits the potential for leadership development, mentoring, and support for the teachers who remain. The systematic exit of teachers who are, based on a mix of metrics, thriving in a challenging but vital setting handicaps the capacity of the school to develop a culture of learning, consistency, and support for students. As these schools struggle to develop a stable environment, and as school choice options expand, parents who have the means to do so avoid enrolling their students in these schools, thus exacerbating the underlying isolation of disadvantaged students.

With respect to the generalizability of the findings, there are a number of notable factors that are potentially unique to this specific sample, place, and time. Numerous urban schools across the country have extremely high concentrations of low-income, racial minority students, with relatively low teacher salaries, and disparate access to highly effective teachers (e.g., Isenberg et al., 2016; Goldhaber, Lavery, & Theobold, 2015; Clotfelter et al., 2011). However, as noted above, the schools here who were eligible to participate were all designated as Priority Schools in the state's post-Race to the Top accountability program, and were thus disproportionately subjected to overlapping policy interventions. While I control statistically for the two largest overlapping school-level policies (incorporation into the privately managed Achievement School District or iZone), it is difficult to rule out the potential that overlapping interventions, unobserved in the data, could have facilitated or magnified the achievement boosting impacts of the teacher retention bonuses. That said, the Priority designation, which predates the bonus intervention by nearly two school years, could have alternatively weakened the impacts of a retention bonus, depending on whether the effective teachers targeted by the bonus view the official labeling of their school's low ranking as a stigma or a challenging inspiration.

It may be the case that this bump in achievement is a result of student sorting directly following the implementation of the bonus program, specifically in the form of differential reductions in the average total enrollment at SRB-participant campuses. However, the consistent results from the models, which include the lagged dependent variable and the student-fixed effect model, minimize the number of alternative explanations for the bump in performance in participating schools. The coherence of the differential effects by subject areas, the effects on the composition of teachers by subject, and the limited evidence of any factor determining participation outside of the policy eligibility rules themselves strengthen the plausibility of a causal interpretation. Furthermore, the prior trends analysis indicates that, if anything, the schools that participated in the bonus program were substantively underperforming the comparison schools prior to the intervention.

In line with several studies before it, the findings presented here indicate that financial incentives can marginally shift teachers' decisions to persist in the challenging work environments of high-accountability, high-poverty, racially isolated schools, and promote higher levels of learning than would have occurred had they left. However, for many teachers, additional pay alone is inadequate to overcome pressures to leave, and only affects the underlying learning and working conditions to the extent that retained teachers improve the leadership culture in the building. Moving forward, research should examine the roles of non-pecuniary incentives, and the interactions between conditions and simple salary improvements. Ultimately, policies that improve working conditions and better integrate student populations across schools (thus minimizing the concentration of economic disadvantage) would likely have larger, more sustainable effects on the stability and equitable distribution of effective instruction. However, given the current state of hyper-segregated schooling, and the disparate distribution of resources associated therewith, financial incentives to attract and retain effective teachers and leaders in "hard-to-staff" disadvantaged schools like the bonus program evaluated here could both mitigate the damages of isolation and make the prospect of integration more attractive to both parents and policymakers.

APPENDIX

Appendix A: Summary Statistics on Students and Teachers by Participation Status of Priority Schools

	(1)	(2)	(3)	(4)
	All Priority Schools	Distributed Bonuses	Did Not Participate	Difference (2) - (3)
Student Characteristics				
Percent Female	48.22%	48.51%	47.49%	1.03%
Percent White	2.60%	1.98%	4.16%	-2.18%*
Percent Black	93.23%	93.52%	92.50%	1.02%
Percent Asian	0.25%	0.23%	0.29%	-0.06%
Percent Hispanic	3.76%	4.10%	2.90%	1.20%
Percent Other	0.16%	0.16%	0.14%	0.02%
Percent FRPL	88.85%	90.04%	85.83%	4.21%*
Percent Special Education	16.96%	16.51%	18.09%	-1.57%
Percent ELL	2.26%	2.40%	1.92%	0.48%
Teacher Characteristics				
Percent Female	77.21%	79.41%	72.28%	7.13%**
Percent White	31.09%	30.14%	33.21%	-3.07%
Percent Non-White	68.91%	69.86%	66.79%	3.07%
Percent with Bachelors	33.67%	34.56%	31.66%	2.90%
Percent with Masters	36.46%	36.07%	37.32%	-1.25%
Percent with > Masters	29.87%	29.36%	31.03%	-1.67%
Average Years of Experience	11.70	11.73	11.66	0.07
Average Salary	52414.40	52811.91	51523.97	1287.95
N	82	56	26	

Note: Data presented above come from the 2011-12 school year, in which bonuses were first distributed. Mean differences were assessed using regression based t-tests, with statistical significance designated as follows: * significant at the 10% level; ** 5% level; *** 1% level.

Appendix B: Test for Prior Trends in Student Achievement

	Gains		Levels	
	Math	Reading	Math	Reading
Bonus 2008	-0.002	0.072+	0.066	0.205*
	(0.058)	(0.038)	(0.097)	(0.084)
Bonus 2009	0.005	-0.002	0.075	0.129
	(0.048)	(0.035)	(0.091)	(0.086)
Bonus 2010	-0.205**	-0.156**	-0.188**	-0.118**
	(0.052)	(0.048)	(0.057)	(0.043)
Bonus 2011	-0.127**	-0.047+	-0.248**	-0.165**
	(0.043)	(0.025)	(0.053)	(0.040)
Bonus 2012	-0.063	-0.005	-0.207**	-0.146**
	(0.039)	(0.030)	(0.048)	(0.041)
Bonus 2014	-0.021	0.068**	-0.007	0.076**
	(0.035)	(0.024)	(0.033)	(0.022)
Bonus 2015	0.022	0.069**	0.035	0.115**
	(0.039)	(0.025)	(0.038)	(0.032)
R2	0.53	0.61	0.15	0.17
N	565,551	571,726	733,107	741,186

(+ $p<0.1$; * $p<0.05$; ** $p<0.01$)

Note: The coefficients plotted above represent the results from a standard test for pre-existing trends, wherein the treated school indicator (Bonus Program Participants) was interacted with indicators for each year in the panel except the one directly preceding intervention, which is omitted to serve as reference.

REFERENCES

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, *25*(1), 95-135.

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, *72*(1), 1-19.

Adnot, M., Dee, T., Katz, V., & Wyckoff, J. (2016). *Teacher turnover, teacher quality, and student achievement in DCPS (CEPA Working Paper No. 16-03).* Retrieved from Stanford Center for Education Policy Analysis Web site: http://www.cepa.stanford.edu/wp16-03

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *The Review of Economics and Statistics*, 47-57.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics*, *21*(1), 1-42.

Balch, R., & Springer, M. G. (2015). Performance pay, test scores, and student learning objectives. *Economics of Education Review*, *44*, 114-125.

Ballou, D. (2001). Pay for performance in public and private schools. *Economics of Education Review*, *20*(1), 51-61.

Ballou, D., & Podgursky, M. J. (1997). *Teacher pay and teacher quality*. Kalamazoo, MI: W. E. Upjohn Institute.

Bertrand, M., Mullainathan, S., & Shafir, E. (2004). A behavioral-economics view of poverty. *The American Economic Review*, *94*(2), 419-423.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2008). *Who leaves?: Teacher attrition and student achievement (National Bureau of Economic Research Working Paper Series, No. 14022).* Retrieved from http://www.nber.org/papers/w14022

Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation & Policy Analysis*, *31*(4), 416-440. doi:10.3102/0162373709353129

Boyd, D., Lankford, H., Loeb, S., Rockoff, J., & Wyckoff, J. (2008). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools. *Journal of Policy Analysis and Management*, *27*(4), 793-818.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2008). The impact of assessment and accountability on teacher recruitment and retention: Are there unintended consequences? *Public Finance Review*, *36*(1), 88–111. doi:10.1177/1091142106293446

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics*, *126*(4), 1593-1660.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *The American Economic Review*, *104*(9), 2593-2632.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014b). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *The American Economic Review*, *104*(9), 2633-2679.

Clotfelter, C., Glennie, E., Ladd, H., & Vigdor, J. (2008). Would higher salaries keep teachers in high-poverty schools?: Evidence from a policy intervention in North Carolina. *Journal of Public Economics*, *92*(5), 1352-1370.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2011). Teacher mobility, school segregation, and pay-based policies to level the playing field. *Education Finance and Policy*, *6*(3), 399-438.

Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., & Diaz, R. A. (2004). Do school accountability systems make it more difficult for low-performing schools to attract and retain high-quality teachers? *Journal of Policy Analysis and Management*, *23*(2), 251-271. doi:10.1002/pam.20003

Dee, T. S., & Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, *34*(2), 267-297.

Feng, L. (2010). Hire today, gone tomorrow: New teacher classroom assignments and teacher mobility. *Education Finance and Policy*, *5*(3), 1-39. doi:10.1162/EDFP_a_00002

Feng, L., Figlio, D. N., & Sass, T. (2010). *School accountability and teacher mobility (National Bureau of Economic Research Working Paper Series. Working Paper 47, CALDER)*. Retrieved from http://www.nber.org/papers/w16070

Figlio, D. N., & Kenny, L. W. (2007). Individual teacher incentives and student performance. *Journal of Public Economics*, *91*(5), 901-914.

Fryer, Jr., R. G., & Levitt, S. D. (2004). Understanding the Black-White test score gap in the first two years of school. *Review of Economics and Statistics*, *86*(2), 447-464.

Fryer, Jr., R. G., Levitt, S. D., List, J., & Sadoff, S. (2012). *Enhancing the efficacy of teacher incentives through loss aversion: A field experiment (No. w18237)*. Retrieved from http://www.nber.org/papers/w18237

Glazerman, S., Protik, A., Teh, B., Bruch, J., Max, J. (2013). *Transfer incentives for high-performing teachers: Final results from a multisite experiment (NCEE 2014-4003)*. Washington, D.C.: National Center for

Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven playing field?: Assessing the teacher quality gap between advantaged and disadvantaged students. *Educational Researcher*, *44*(5), 293-307.

Goodman, S. F., & Turner, L. J. (2013). The design of teacher incentive pay and educational outcomes: Evidence from the New York City bonus program. *Journal of Labor Economics*, *31*(2), 409-420.

Greenberg, D., & McCall, J. (1974). Teacher mobility and allocation. *The Journal of Human Resources*, *9*(4), 480-502.

Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). *The market for teacher quality* (No. w11154). National Bureau of Economic Research.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review*, *30*(3), 466-479.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*(5782), 1900-1902.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, *64*(4), 605-654.

Henry, G. T., Bastian, K. C., & Fortner, C. K. (2011). Stayers and leavers: Early-career teacher effectiveness and attrition. *Educational Researcher*, *40*(6), 271-280. doi:10.3102/0013189X11419042

Henry, G. T., Fortner, C. K., & Thompson, C. L. (2010). Targeted funding for educationally disadvantaged students. *Educational Evaluation & Policy Analysis*, *32*(2), 183-204. doi:10.3102/0162373710370620

Hock, H., & Isenberg, E. (2017). Methods for accounting for co-teaching in value-added models. *Statistics and Public Policy, 4*(1), 1-11. http://dx.doi.org/10.1080/2330443X.2016.1265473

Hough, H. J. (2012). Salary incentives and teacher quality: *The effect of a district-level salary increase on teacher retention*. Retrieved from https://cepa.stanford.edu/content/research-brief-the-effect-of-a-district-level-salary-increase-on-teacher-retention

Iatarola, P., & Stiefel, L. (2003). Intradistrict equity of public education resources and performance. *Economics of Education Review*, *22*(1), 69-78.

Isenberg, E., Max, J., Gleason, P., Johnson, M., Deutsch, J., & Hansen, M. (2016). *Do low income students have equal access to effective teachers?: Evidence from 26 districts (NCEE 2017-4008).* Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*, *1*(4), 85-108.

Johnson, S. M., Berg, J. H., & Donaldson, M. L. (2005). *Who stays in teaching and why?: A review of the literature on teacher retention*. Retrieved from http://hub.mspnet.org/index.cfm/12908

Jones, A. M., Rice, N., & Rosa-Dias, P. (2011). Long-term effects of school quality on health and lifestyle: Evidence from comprehensive schooling reforms in England. *Journal of Human Capital*, *5*(3), 342-376.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness?: Evidence from New York City. *Economics of Education Review*, *27*(6), 615-631.

Kukla-Acevedo, S. (2009). Leavers, movers, and stayers: The role of workplace conditions in teacher mobility decisions. *The Journal of Educational Research*, *102*(6), 443-452.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation & Policy Analysis*, *24*(1), 37-62. doi:10.3102/01623737024001037

Loeb, S., Darling-Hammond, L., & Luczak, J. (2005). How teaching conditions predict teacher turnover in California schools. *Peabody Journal of Education*, *80*(3), 44-70. doi:10.1207/s15327930pje8003_4

Murnane, R. J. (1981). Teacher mobility revisited. *The Journal of Human Resources*, *16*(1), 3-19.

O'Neill, S., Kreif, N., Grieve, R., Sutton, M., & Sekhon, J. S. (2016). Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health Services & Outcomes Research Methodology*, *16*, 1-21. http://doi.org/10.1007/s10742-016-0146-8

Podgursky, M. J., & Springer, M. G. (2007). Teacher performance pay: A review. *Journal of Policy Analysis and Management*, *26*(4), 909-949.

Podgursky, M., & Springer, M. (2011). Teacher compensation systems in the United States K-12 public school system. *National Tax Journal*, *64*(1), 165.

Redding, C. H. (2017). *The determinants and consequences of within-year teacher turnover* (Unpublished doctoral dissertation). Vanderbilt University, Nashville, TN.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, *94*(2), 247-252. doi:10.1257/0002828041302244

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*(1), 4-36.

Rumberger, R. W. (1987). The impact of salary differentials on teacher shortages and turnover: The case of mathematics and science teachers. *Economics of Education Review*, *6*(4), 389-399.

Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Retrieved from http://www.mdk12.org/practices/ensure/tva/tva_2.html

Scafidi, B., Sjoquist, D. L., & Stinebrickner, T. R. (2007). Race, poverty, and teacher mobility. *Economics of Education Review*, *26*(2), 145-159. doi:10.1016/j.econedurev.2005.08.006

Springer, M. G., Hamilton, L., McCaffrey, D. F., Ballou, D., Le, V. N., Pepper, M., . . . & Stecher, B. M. (2010). *Teacher pay for performance: Experimental evidence from the project on incentives in teaching*. Nashville, TN: National Center on Performance Incentives at Vanderbilt University.

Springer, M. G., Swain, W. A., & Rodriguez, L. A. (2016). Effective teacher retention bonuses: Evidence from Tennessee. *Educational Evaluation and Policy Analysis*, *38*(2), 199-221.

Steele, J. L., Murnane, R. J., & Willett, J. B. (2010). Do financial incentives help low-performing schools attract and keep academically talented teachers? Evidence from California. *Journal of Policy Analysis and Management*, *29*(3), 451-478.

Steele, J. L., Pepper, M. J., Springer, M. G., & Lockwood, J. R. (2015). The distribution and mobility of effective teachers: Evidence from a large, urban school district. *Economics of Education Review*, *48*, 86-101.

Sun, M., Loeb, S., & Grissom, J. A. (2017). Building teacher teams: Evidence of positive spillovers from more effective colleagues. *Educational Evaluation and Policy Analysis*. doi:0162373716665698

Sun, M., Penner, E. K., & Loeb, S. (2016). Resource- and approach-driven multidimensional change: Three-year effects of school improvement grants. *American Educational Research Journal*. doi:0002831217695790

Zimmer, R., Henry, G. T., & Kho, A. (2016). The effects of school turnaround in Tennessee's achievement school district and innovation zones. *Educational Evaluation and Policy Analysis*. doi:0162373717705729