The Impact of School-Based Health Centers on Student Achievement

Claire Holleman
April 28th, 2021

Faculty Advisor: Professor Hyman

Submitted to the Department of Economics at Amherst College in partial fulfillment of the requirements for the degree of Bachelor of Arts with honors

Abstract

Children who grow up poor are more likely to have worse health outcomes and lower educational attainment than their wealthier peers (Case et al, 2002; Currie, 2009). School-based health centers (SBHCs) aim to break this cycle of poverty by increasing access to health care for low-income youth. An SBHC can be thought of as a pediatrician's office on a public-school campus; it employs advanced practitioners who are able to diagnose, prescribe, and manage care at a higher level than a school nurse. These services are provided at little to no cost to students through partnerships with local hospitals or community clinics. While there is a well-established link between health and education in the economics literature, there is a dearth of research specifically assessing the impact of SBHCs on education outcomes. In this study, I answer the research question: how does the opening of a school-based health center affect educational achievement, particularly among young children? I use test score, graduation, and attendance data from New York and California between 1997-2019 to answer this question. I use difference-in-differences and event study methodologies exploiting the plausibly exogenous timing of SBHC openings.

I find a small, positive, though not robust, effect on graduation rates, and see no significant impacts on attendance. In contrast, I find a robust, statistically significant 0.10 standard deviation decrease in test scores. This unexpected result is driven by decreases in test scores for more advantaged groups, including White and Asian students, students attending wealthier schools and students living in communities that have high-quality pre-existing clinical care. These results suggest that while SBHCs may be better than no care at all, they may be less effective than the existing health care services used by more resourced households.

JEL: H75, I14, I18, I24

Keywords: School-Based Health; Education; Inequality; Student Outcomes

Acknowledgments

Professor Hyman- there are no words to say just how thankful I am for your support this past year. Thank you for your remarkable insight, your thoughtful feedback, and for always hyping me up. Amherst is SO lucky to have you. I am your number one fan!!

Dana- my sweet, thoughtful friend. Far and away the best part of this thesis endeavor was getting to become your friend. I truly would not have made it without you and I'm forever grateful for the early morning facetimes, late night texts, and all the tiktoks in between. You are a superstar! Don't ever forget it!

Professor Reyes- thank you for all your support in the thesis seminar and for all the time spent in office hours answering my many questions as I was trying to get this project off the ground. I am forever grateful for you opening my eyes to the endless questions that applied economics can be used to answer.

Professor Theoharides, Professor Baisa, and Professor Ishii- thank you for encouraging me when I needed it most. In your own ways, you have each had a remarkable impact on my college experience, and without your kind words, I would never have embarked on this thesis adventure. I love economics and it's all thanks to you!

Mom and Dad- thank you for being so supportive and encouraging in everything I do. And Annie and Dougie- thank you for pretending to listen when I tell you what I'm working on. I love you all very much!

TABLE OF CONTENTS

I. Introduction **5**

II. Background **10**

III. Previous Literature **14**

IV. Data **17**

V. Methodology **27**

VI. Results **29**

VIa. Event Study 29

VIb. Main Results 30

VIc. Heterogeneity 33

VId. What is Driving Negative Test Scores? 39

VIe Robustness 45

VII. Conclusion **47**

References **51**

Data References **53**

I. Introduction

There is a well-established link in the literature between socioeconomic status, health, and education for children. Growing up in a low-income household results in worse health, particularly for children with chronic conditions, and this gap widens as children age (Case et al, 2002). Furthermore, poor health in childhood has been causally linked to lower educational attainment, worse future health outcomes, lower wages, and lower social status as adults (Case et al, 2005). The intergenerational transmission of health and socioeconomic status from parents to children creates a cycle of poverty that is difficult to break.

Additionally, access to quality health care is not equal across different socioeconomic groups. In 2010, close to 50 million people in the US did not have health insurance, with the largest segment of that population having an annual income of less than $25,000 (Maruthappu et al, 2013). Expanding public health insurance for low-income children and adolescents has been shown to have a major impact on child health outcomes. A Medicaid expansion between 1984 and 1992, which doubled the number of children eligible for insurance, led to increased utilization of primary care, lower fertility rates for young women, and a significant drop in child mortality (Currie et al, 1996). Beyond these health impacts, the same expansion of Medicaid increased rates of high school and college completion and led to higher wage attainment (Cohodes et al, 2016; Brown et al, 2020). This evidence suggests increasing access to health care for school-age children can improve both health and education outcomes. However, disparities in the health care system persist. School-based health centers (SBHCs) are another intervention that provides care to the most vulnerable child populations.

A school-based health center functions as a doctor's office on the site of a public-school campus. It goes above and beyond the role of a traditional school nurse by providing primary care, including annual physicals, immunizations, and reproductive counseling, helping manage chronic conditions like asthma and diabetes, providing mental health resources, and offering health education. The convenience of an SBHC is its biggest asset-students can go to the center during a lunch break so they are not being pulled out of class for appointments, and parents do not have to take valuable time off work to get their children to the doctor. Additionally, most of these SBHCs include provisions that require them to provide care to all students, regardless of their ability to pay (New York Department of Health, 2017). The reduced monetary, transportation, and opportunity costs of obtaining care at SBHCs make it an attractive model for delivering care to low-income students.

Currently, more than 6 million students receive care at 2,500 SBHCs across the country in both urban and rural communities. The centers are primarily located in low-income elementary, middle, and high schools. While the SBHC movement is growing in popularity, there is little evidence causally linking school-based care to improved health outcomes and academic achievement.

The current research is primarily comprised of case-study analyses in the public health and medical literature which provide useful context for the centers but limited evidence on the causal effect of its opening. To my knowledge, Lovenheim, Reback, and Wedenoja (2016) is the only existing study that examines the causal, longitudinal effect of SBHCs. They use a national sample to assess teenage pregnancy and high school dropout rates following the opening of an SBHC. While they find a significant reduction in teenage pregnancy, they find no effect on dropout rates.

My paper builds on the findings of Lovenheim et al (2016) by examining test scores and attendance rates in addition to graduation outcomes. Additionally, although about 40% of all SBHCs are in elementary schools, they focus only on high school students, and so I expand my research to include the SBHCs in elementary and middle schools that were omitted from their study. By increasing the number of outcomes for a wider range of students, this paper seeks to answer the question, how does the opening of a school-based health center affect educational achievement, particularly among young children?

To answer this question, I use school-level data between 1997 and 2019 from the California and New York Departments of Education to measure attendance rates, graduation rates, and standardized test scores. I get data on SBHC location and opening date from the New York Department of Health and the California School-Based Health Alliance. Additionally, I get data on my covariates from the National Center for Education Statistics' Common Core of Data, and data about health care services from the County Health Rankings and Roadmaps Project at the University of Wisconsin. I use difference-in-differences and event study methodologies that exploit the plausibly exogenous timing of SBHC openings in New York and California to estimate the effect on academic outcomes.

The main threat to validity is twofold; first, there may be trends in achievement prior to the opening of the SBHC. The second concern is that achievement outcomes may be trending differently between schools with and without SBHCs. I use the event study model to show that there are no pre-trends between the two groups and that there are no trends in the achievement data prior to the opening of the SBHC. I control for a host of time-varying variables to reduce the risk of omitted variables bias in my analysis. I also use state-by-year fixed effects to compare schools in a given state in a given year to help strengthen the parallel trends assumption by increasing the accuracy of the untreated comparison group.

I find that there are marginally significant improvements in graduation rates, however, the impacts are small, about a 1.48 percentage point increase, and attenuate in response to a variety of robustness checks. I find no effect on attendance outcomes; in fact, the results suggest a precise zero, since the coefficient is both economically and statistically insignificant. In stark contrast, however, I find a statistically significant 0.10 standard deviation decrease in standardized test scores following the opening of an SBHC in elementary and middle schools. This unexpected negative effect is relatively large and is robust to several different specifications and robustness checks.

I find important heterogeneity in these results, which show that the negative effect in test scores seems to be driven by schools in areas with higher levels of existing primary care services and in schools that serve more privileged students. I see that schools which have fewer students who qualify for free and reduced lunch ("Not Poor" schools) have a 0.17 strongly significant standard deviation decrease in test scores following the opening of a center, compared to a much smaller, only marginally significant decrease in the "Poor" schools. Furthermore, I find White and Asian students have a significant decrease in their test scores, while their Black and Hispanic peers see no significant effect. I also see that SBHCs which open in counties with the best existing clinical services have a 0.24 standard deviation decrease in test scores, significant at the 99% confidence level. In comparison, I find that SBHCs opening in counties in the middle and bottom thirds of existing clinical care do not significantly decrease test scores at all, though there is not a significant increase either. Taken together, I find evidence that the decreases in test scores are coming primarily from more advantaged groups.

While SBHCs may be improving access to care for the most underserved populations, they may not be as effective as the existing health care services more resourced households are already using. The convenience of the SBHC is its biggest asset, so these higher-income parents may substitute away from their child's existing primary care provider to the SBHC because it is more accessible. In doing so, they may be unknowingly making their child worse-off. Conversely, students who had little or no prior access to care and who are the intended targets of the SBHC intervention are now able to receive basic check-ups, immunizations, and referrals. Although I do not see any positive spillovers onto education, I do not find these students are made worse off.

I also assess the differential impacts in New York and California and urban and non-urban schools. The patterns of small positive effects on graduation rates, no effects on attendance rates, and negative effects on test scores persist across these two different sample restrictions. I conduct two robustness checks, one restricting the treatment group to schools that open an SBHC between 1997-2019, and one using propensity score matching to create a comparison group that looks more similar to the SBHC treated schools. Both checks show that the null finding on attendance and significant decrease in test scores are robust, however, I find that the positive effect on graduation rates attenuates significantly when using propensity score matching.

Given these small or null effects on graduation rates, my results are consistent with Lovenheim et al (2016) and their null result on dropout rates. One reason for the slight discrepancy between my small positive result and their null result on graduation rates could be because of the school vs. district-level analysis, since students in a school may benefit more from the SBHC than students in the district at large. Overall, though, SBHCs do not seem to translate into major positive effects on education, and in some cases, they seem to make younger children worse off. Importantly, I do not see the negative effect on test scores translating into worse graduation rates, which is arguably the more important long-run outcome. Together our results suggest that SBHCs are not having far-reaching spillovers onto student academic achievement. However, the intended goal of SBHCs is to improve *health* outcomes. So, while the null findings on education are a bit disappointing, this evidence has no bearing on the effectiveness of SBHCs as a provider of health care for underserved populations.

In Section II I discuss the background and historical context of SBHCs. Section III explores the previous literature on child health more broadly, as well as the existing research on SBHCs. Section IV discusses my data collection and Section V addresses my empirical methodology. Section VI explores my results, and I conclude in Section VII.

II. Background

School-based health centers can be thought of as a pediatrician's office located on a public-school campus. All centers provide at least some sort of primary care, including annual physical exams and immunizations, and most also provide additional services, including behavioral and oral health. They are staffed by medical doctors or nurse practitioners and have registered nurses and administrative staff, just like any other health clinic does. There is often some confusion between the role of a school nurse and a school-based health center; the difference is that while the SBHC is certainly able to treat acute and chronic conditions, the primary focus is on preventative care.

The first school-based health center was founded in 1967 by Dr. Philip Porter in Cambridge, Massachusetts. His vision for an integrated and centralized model of care was inspired by children who were falling through the health care cracks- the existing services were fragmented and unimpactful (Love et al, 2019). Schools were chosen as the site of the center for convenience; no new construction was needed, and they were centrally located for children and their families.

The development of SBHCs was aided by the confluence of two other movements. In the 1960s, part of President Lyndon B. Johnson's "War on Poverty" campaign focused on increasing health care to poor families. Previously, these patients were often turned away from hospitals or asked to wait long hours in uncomfortable conditions to receive even the most basic care. The Office of Economic Opportunity began funding Federally Qualified Health Centers (FQHCs), which operated in underserved areas and provided a holistic approach to care (Bailey et al, 2015). These clinics relied on community input and required at least 51% of the board to be patients who use the clinic as their primary source of care; in other words, the board was (and still is) legally required to resemble its community in order to receive funding (Bureau of Primary Health Care, 2015). These community health centers embraced the vision of having medical services integrated with other social services in an accessible location, which became the fundamental basis for SBHCs when they were founded a decade later.

The second big, contemporary change was the foundation of the first nurse practitioner program in Denver, Colorado in 1965. Loretta Ford, a nurse working in rural

Colorado, saw the need for health care, particularly for children, in these underserved areas. She felt with specialized training, nurses could fill this gap in the health care system. Ford worked with pediatrician Henry Silver to create a program that enabled nurses to “become skilled in physical examination, diagnosis, and treatment” (Gustafson, 2005). With a new type of practitioner who was specifically trained to provide care in a school-based setting, and the increasing need for improved access to care for underserved populations, the school-based health center movement was poised to take-off.

Growth of SBHCs was steady for the first few decades; there were 200 centers open in 1990, and by 2000, more than 1,100 SBHCs were operating in 45 states. In 2010, the Affordable Care Act promised $50 million annually over the next four years for construction and infrastructure improvements for school-based health centers, leading to a major SBHC expansion. By 2017, there were more than 2,500 centers open in 48 states serving more than 6 million students. While SBHCs historically served low-income, urban high schools, today SBHCs are located in urban and rural settings and provide care to students of all ages. More than 40% of SBHCs serve elementary schools, 30% serve middle and high schools, and the remaining 30% provide services to students in all grades (Love et al, 2019).

The majority of school-based health centers operate through outside partnerships with health care organizations. Federally qualified health centers sponsor 51% of SBHCs, hospitals sponsor 21%, and the remaining 28% are sponsored by non-profits, local health departments, or the school-districts themselves. These partners fully fund and staff the centers; as such, the involvement of the school does not go much further than providing space for the SBHC to operate. Because of where the SBHCs are opened, the revenue comes primarily from Medicaid reimbursement, though they also receive private insurance revenue, state and government support, and rely heavily on in-kind donations. To the best of my knowledge, any student in a school who obtains written consent from their parents is eligible to receive care at their school-based health center. The literature finds that while females, non-whites, and students qualifying for free and reduced lunch tend to be overrepresented in the SBHC users relative to the rest of the student body, SBHC usage is by no means restricted just to these groups (Kerns, 2011). Many states have provisions that mandate the SBHC to serve any student, regardless of their ability to pay, though the percentage of uninsured students receiving care remains small (Love et al, 2019).

There is some evidence that SBHCs improve health care access for underserved populations. Recent studies have found that SBHCs are successfully providing immunizations to students, including one study showing a 3-percentage point increase in students completing the HPV vaccine cycle. (Oliver et al, 2019). Additionally, a study in Denver showed that although SBHC users were 36% less likely to be insured than peer community health clinic users, they were almost three times more likely to have received a flu shot and more than twice as likely to have received the Hepatitis B vaccine. Furthermore, the same study showed that SBHC users were two times less likely to have used emergency care in the past year (Allison et al, 2007). One common reason for emergency room visits for children are due to complications from asthma. A New York City study found a reduction in hospitalizations for children with asthma and a gain of three more school days each year for these same students (Webber et al, 2003). Though these studies are not causal, they provide important evidence that SBHCs seem to be correlated with positive primary care outcomes.

III. Literature Review

Much of the economics literature regarding child health has focused on prenatal interventions. The “Fetal Origins Hypothesis” refers to the causal impact that events during pregnancy have on long-term outcomes, specifically physical health, disability, earnings, and education. Increased maternal stress, contraction of the 1918 Spanish Flu, and fasting for Ramadan during pregnancy have all been shown to have negative impacts on a baby’s outcomes (Almond et al, 2006; Almond et al, 2011; Aizer et al, 2016). The prenatal period is favored by economists because sudden shocks offer a clear natural experiment that disentangles the event from other factors that could have a confounding effect (Almond and Currie, 2011). This work highlights the role that health can have on long-term outcomes.

Additionally, there has been substantial economics work studying interventions in the early childhood period, particularly with regards to education. Recent literature examining the causal effect of Head Start suggests that the large-scale, publicly provided intervention leads to increased high school and college completion among its participants (Bailey et al, 2020). Furthermore, new research shows that this effect has intergenerational benefits, and leads to a decrease in teen pregnancy and an increase in educational attainment in the second generation (Barr and Gibbs, 2017). Head Start, in addition to providing free preschool, also has a substantial health component, including medical and oral health services. This idea of a comprehensive social service provider is replicated in the SBHC model.

Despite the abundant research on child health in the early stages, there has not been as much work on older children and adolescents, and such work is almost non-existent in the school-based health setting. Nearly all of the existing literature on school-based health centers comes from cross-sectional case-study analyses in the public health and medical fields. The evidence shows that opening a school-based health center results in positive health outcomes, including fewer emergency room visits and hospitalizations, higher rates of immunization, more primary care visits, and an overall higher health-related quality of life (Guo et al, 2005; Wade et al, 2008; Allison et al, 2007). The education outcomes are slightly more mixed; SBHCs have been associated with improvements in attendance rates, grade point average, dropout rates, and graduation in some studies, while others find no effect at all (Geierstanger et al, 2004; Walker et al, 2010). Kerns (2011) finds a decrease in dropout rates following the opening of an SBHC for Black and Hispanic students and students who qualify for free and reduced lunch but finds no impact on dropout rates for White students or students who do not qualify for free and reduced lunch. In fact, she finds a statistically significant increase in dropout rates for White students following the SBHC opening. These studies typically compare students who use the SBHC in the school to those who do not, or they compare a small number of schools that have an SBHC to a few similar schools that do not. Unfortunately, due to the non-random nature of the studies, the causal impact of SBHCs on health and education is still in question.

Lovenheim, Reback, and Wedenoja (2016) is the only longitudinal, large-scale, arguably causal evaluation of the impacts of school-based health centers. In order to analyze teenage pregnancy outcomes and high school dropout rates, the authors use nationwide census survey data on SBHCs. These data come from the National Alliance on School-Based Health Care (NASBHC), which is an organization that collects SBHC census data every three years; Lovenheim et al (2016) have census data from 1998, 2001, 2004, 2007, and 2011. The data include information on the zip code of the center, the services provided, which populations are eligible for care, and the total hours an SBHC is open. From the

NASBHC data, they match each center to its respective county (for birth outcomes) or district (for dropout outcomes). Their data on teenage pregnancy come from the CDC's National Vital Statistics and contain information on race, ethnicity, and age of mothers based on birth records. The dropout rates are calculated for 10^{th}, 11^{th}, and 12^{th} graders from the NCES Common Core of Data.

Lovenheim et al (2016) use difference-in-differences and event study methodologies at the district level, exploiting the plausibly random timing of the openings of centers across the country to assess the impact of the SBHC. Additionally, they assess the impact of the "intensity" of the SBHC treatment, which is based on the total hours of operation across SBHCs in a county in a given week. They find there is a significant decrease in teen pregnancy, about 5% for girls aged 15-18, following the opening of the center. They find important heterogeneity in these results; while Black and Hispanic girls each see an 8% decrease in pregnancy, the results for White girls are universally small and insignificant. There is no evidence, however, that SBHCs are affecting dropout rates. The results on dropout rates are precise zeroes, and Lovenheim et al (2016) rule out the possibility of any effect on dropout greater than 1%. They hypothesize this is because the intervention is coming too late in adolescence to significantly impact behaviors that would spill over onto graduation outcomes.

This paper will build upon the Lovenheim et al (2016) findings by increasing the scope of education outcome measures for a broader group of students. While they only look at high schoolers, I also look at outcomes for elementary and middle school students. Younger children are at a critical age when most childhood chronic conditions, such as asthma and diabetes are diagnosed, and they may be more malleable than their high school counterparts in forming healthy habits, all of which could potentially have an impact on their educational achievement. By including elementary and middle schools in the study, I am able to estimate the effect of an SBHC at an earlier stage in the life cycle. Furthermore, by looking at attendance rates and test scores in addition to graduation rates, I may be able to see a more moderate increase in achievement outcomes than the graduation rate alone is able to provide. The attendance rates in particular may show a more direct effect on health, especially if students are no longer missing class due to illness or other health related concerns. Furthermore, Lovenheim et al (2016) uses a district level analysis, while I am assessing the impact at the school level. This is important because of the accessibility component of the SBHC, where having an SBHC down the hall may make more of a difference than an SBHC that opens across town. Lastly, I am able to explore heterogenous treatment effects depending on the existing health services in the county in which the SBHC is located.

IV. Data

My data are from the New York Department of Education, the New York Department of Health, the California Department of Education, the California School-Based Health Alliance, the NCES Common Core of Data, and the University of Wisconsin Population Health Institute.

School-Based Health Data

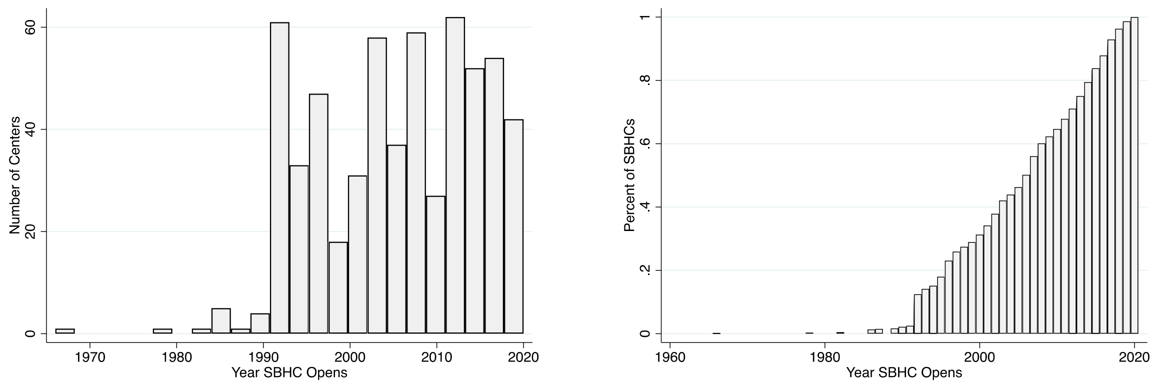
Information and data about school-based health centers are collected at the state level. The California School-Based Health Alliance is a private, non-profit organization that advocates for increased health services in schools. They provided me with a comprehensive list of active school-based health centers in the state which contained data on year opened, address, sponsor, and services provided. I used the center's address to match it with the school it served using the National Center for Education Statistics (NCES) Search for Public Schools. Of the 294 centers in California, I matched 237 to public schools. The remaining centers were mobile units or off-site centers, which I drop from my sample.

The New York Department of Health provided me with the New York SBHC data, which included the sponsor, the name of the SBHC site, and the effective date of opening. I was able to find the addresses of each health center using different hospital and community health clinic websites and matched them to their school using the NCES database. Many schools throughout New York, both in rural towns where there is one centrally located campus containing an elementary, middle and high school, and in cities, where one building may serve many schools, have a school-based health center that shares resources across the campus. So, although there are just over 200 unique SBHCs in New York, they were matched to 499 different schools.

Figure 1a shows the opening date of each of the school-based health centers, while Figure 1b shows the cumulative total of the number of centers. Importantly, more than 70% of all centers in New York and California opened between 1997-2019, the period for which I have education outcome data.

Figure 1: SBHC Opening Dates for New York and California

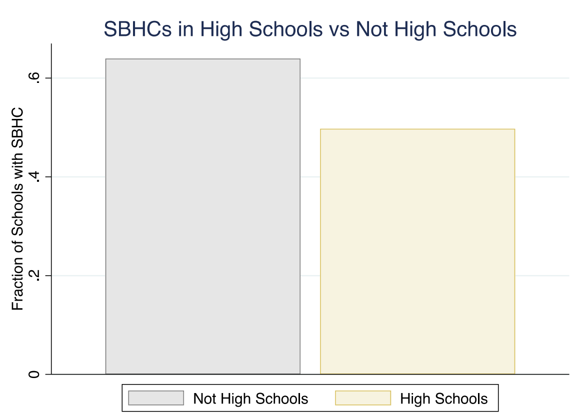
(a) SBHC Opening Years (b) Cumulative Total SBHCs Open



Note: Each bar in Figure 1a represents a two-year interval. In Figure 1b, each bar is a single year.

Figure 2 shows that the location of school-based health centers is fairly evenly split between high schools and elementary or middle schools in both states, suggesting there is a substantial treatment group for younger children when looking at test scores, and a treated group for high schoolers when looking at graduation rates.

Figure 2: Fraction of SBHC Schools in High Schools vs Non-High Schools



Note: The total appears greater than 1 because there are a few schools in my sample which are some sort of combination school (K-12, 6-12, etc) and appear in both categories. These will contribute to both the test score and graduation rate outcomes and therefore are counted in each sample.

Education Outcome Data
New York

The education outcomes for New York include test scores for grades 3-8, graduation rates, and attendance rates and come from the New York Department of Education Report Card Database. Importantly, test scores and attendance rates are the outcomes of interest for elementary and middle schools, while graduation and attendance rates will be used to measure the effect in high schools.

The test score data are comprised of school-level observations from 2000 to 2019. For data between 2006 and 2019, the raw mean scores for Math and ELA are given by subgroup for a given school in a given year. The subgroups include different racial and ethnic categories- African American, Hispanic, Native American, Asian or Pacific Islander, White- as well as gender, English proficiency, migrant status, and socioeconomic status. Subgroup data were not reported for categories with less than five students to protect student identity.

To create a single composite test score for a school in a given year, I standardize the scores within a subject-grade to a mean of zero and a standard deviation of one. I take an average of these standardized scores first across subjects and then across grades. The data ultimately consist of school-year observations for the twenty-year period between 2000-2019. At this point, I did a final standardization of the test scores for all schools across the whole period.

The New York attendance data are measured beginning in 2005 for all schools in the state. The attendance data are interesting because attendance is arguably a more direct measure of the impact of an SBHC on health. Attendance rates may be more responsive to improved health benefits than other outcome measures, particularly in the short-term.

Beginning in 2005, the 4-year graduation rate was calculated by dividing the number of graduates by the cohort enrollment from four years prior. So, the number of graduates in 2007 was divided by the 9th grade cohort enrollment from the 2003-2004 school year. The cohort enrollment includes any student who started in that school in ninth grade and remained for all four years as well as students who transferred in, but it excludes students who transferred out of the school. The data were also reported by subgroup cohort enrollment and subgroup graduates, and so the subgroup graduation rates were calculated in the same way as the sample at large. Before 2005, the cohort enrollment data were not reported. Using 9th grade enrollment data as a proxy for the cohort, the graduation rate was calculated as above, by dividing the number of graduates by 9th grade enrollment from four years prior.

Although the cohort enrollment data are able to keep track of students when they transfer between schools, the 9th grade enrollment proxy is not. Thus, before 2005 the graduation rates are somewhat depressed because any student that transfers schools appears in the data as a dropout. Additionally, there are some schools that increase enrollment in successive grades over time, meaning that they have graduation rates that appear to be greater than 100%. Schools with calculated graduation rates of greater than 100% were recorded as missing. Schools that had a 9th grade cohort enrollment of less than ten in any year were dropped to reduce measurement error.

California

The test score data in California come from the California Standards Test (CST) from 2003-2013. In 2014, California switched from CST to STAR, and subsequently there is no data for this transition year. From 2015-2019, the test data come from the California STAR test. In each year, the data in my sample come from the math and language arts tests for grades 3-8. I create a composite score for a given school in a given year by standardizing each subject-grade test score so the mean is zero and the standard deviation is one. The data were reported in different racial, ethnic, and socioeconomic subgroups. Data were not reported for subgroups that contained less than 10 students in order to maintain confidentiality. I average the standard score across math and language arts tests, and the subsequently take the average of all grades in the school. The result was a single composite score for a given school in a given year between 2003-2019. I finish with a final standardization such that the mean test score in California across the whole period is zero and the standard deviation is one.

I calculate the graduation rates using the number of graduates at the school level divided by 9th grade enrollment from four years prior. Observations from continuation schools, virtual schools and charter schools were dropped from the data set, as they often had growing enrollment from ninth grade to twelfth grade indicating an influx of transfer students. As such, for these schools the calculated rate was often 200% or 300%, which is not an accurate representation of the graduation rate but more likely an indication of the large number of transfer students into the school.

Additionally, schools that ever have a ninth-grade enrollment of less than 10 students were dropped from the dataset to minimize volatility. These schools had very noisy enrollment and graduation values which led to many observations with infeasible graduation rates. Observations from the remaining sample that had graduation rates greater than 100% were recategorized as missing.

This graduation rate estimation does leave some room for error. It does not take into account students transferring between schools, nor students who take more than four years to graduate. Notably, though, while the measure is not perfect, it should not bias my analysis as long as the SBHC is not having a significant impact on the number of students transferring into and out of the school, which seems like a fair assumption given the many important factors aside from an SBHC that a family considers when deciding where to attend school.

Existing Health Services Data

I use data from the University of Wisconsin's County Health Rankings Project to determine potential heterogeneity in my results based on existing health services in the area. Their clinical care index is comprised of seven measures of health: rate of primary care providers per 100,000 people, rate of mental health providers per 100,000 people, rate of dentists per 100,000 people, percentage of people under 65 without health insurance, number of preventable hospitalizations per 100,000 Medicare enrollees, percentage of female Medicare enrollees who received mammography screenings, and percentage of fee-for-service Medicare enrollees who received an annual flu vaccination. Each county receives a ranking within its state for each of these individual measures. I take an average of the rankings across these seven measures to divide the counties within California and New York into thirds, such that the top third has the best existing clinical care services.

Covariate Data

Covariates for both New York and California were collected at the school and district level. At the school level, I collect data from the respective state departments of education on the race and ethnicity of the student population. I also have data on the fraction of students who are English language learners and students who qualify for free and reduced lunch. I divide the number of students in a category by the total school enrollment to create a variable equal to the fraction of the school that belongs to a certain subgroup. The variables were recategorized as missing if they had values greater than one or less than zero. I then multiply by 100 to see the percent of students in a school belonging in each category.

At the district level, I collect data from the NCES's Common Core of Data on the race and ethnicity of all students in the district, the district student-teacher ratio, and the district student-guidance counselor ratio. The student-teacher and student-guidance counselor ratios were calculated by dividing the total number of students in the district by the total number of full-time teachers and guidance counselors in the district, respectively. I censor extreme values so that any ratio below the 1^{st} percentile is equal to the 1^{st} percentile, and any value greater than the 99^{th} percentile is equal to the 99^{th} percentile. Prior to 2010, the NCES did not report subgroups of students belonging to each racial and ethnic cohort in the district. For those years, I create district-level averages using the school level data weighted by enrollment to create district race and ethnicity covariates between 1997 and 2010.

Enrollment data were collected at the school level. Schools that ever have enrollments of less than 10 students were dropped from the data to prevent noise. Schools were categorized as urban, suburban, town, or rural based on the 2008 NCES designation for their district.

Table 1 shows the sample mean for the covariates used in the regressions and Table 2 shows the sample means for the education outcomes. The tables show that schools which receive an SBHC have different student body compositions and education outcomes than the untreated comparison group. Table 1 shows that schools with SBHCs tend to have higher fractions of minority students, particularly Black and Hispanic students, and substantially fewer white students. They also have higher percentages of students who qualify for free and reduced lunch and more students who are English language learners. The difference in enrollment size between the treated and comparison schools is a reflection of the fact that high schools, which tend to have much larger student bodies, are overrepresented in the SBHC sample.

Table 1- School Characteristic Summary Statistics

			SBHC Schools	
	All Schools (1)	Never SBHC (2)	Before SBHC Opens (3)	After SBHC Opens (4)
School Characteristics				
Enrollment	713	699	1210	1014
White (%)	39.68	40.59	22.83	18.65
Black (%)	11.74	11.23	23.54	23.20
Hispanic (%)	37.77	37.25	45.46	49.66
Asian (%)	7.61	7.65	7.15	6.64
Native American (%)	1.04	1.05	0.90	0.80
Multiracial (%)	2.72	2.79	0.96	1.01
Free and Reduced Lunch (%)	52.86	52.18	64.11	68.46
English Language Learner (%)	16.58	16.46	17.82	19.36
Number of Observations	223,874	77,790	1,589	4,102
Number of Schools	15,990	15,407	333	574

Notes: Data are collected at the school-year level and includes observations from 1997-2019.

Table 2 shows that schools with SBHCs tend to have significantly lower graduation rates than the comparison group of schools. Standardized test scores are on average 0.67 standard deviations below the whole sample mean for SBHC schools. Attendance rates follow similar trends, with treated schools having lower attendance than the control group, although here the difference is less pronounced

Table 2- Education Outcome Summary Statistics

	All Schools (1)	Never SBHC (2)	SBHC Schools: Before SBHC Opens (3)	SBHC Schools: After SBHC Opens (4)
Graduation Rate (%)	71.94	72.55	62.80	67.31
	(21.02)	(21.01)	(20.34)	(19.85)
Number of Observations	48,181	43,932	1,530	2,719
Number of Schools	3,594	3,326	164	259
Standardized Test Scores	0.00	0.02	-0.67	-0.67
	(1.00)	(0.99)	(0.83)	(0.91)
Number of Observations	184,088	178,428	1,573	4,087
Number of Schools	13,548	13,079	198	373
Attendance (%)	93.30	93.58	88.17	90.48
	(7.06)	(6.93)	(10.02)	(6.75)
Number of Observations	57,375	53,010	1,029	3,336
Number of Schools	4,792	4,415	155	372

Notes: Data are collected at the school-year level. Test scores include data from 2000-2019 in NY and 2003-2019 in CA. In CA, there are no test scores from 2014 because no state-wide test was given that year. Graduation rates include data from 1997-2019 for both states. Attendance includes only New York schools from 2005-2019.

The difference in sample means is not necessarily surprising. Students that attend lower quality schools are more likely to be low-income or identify as a racial or ethnic minority (Darling-Hammond, 1998). It may also be true that these students are the ones who are most in need of increased access to health care. The fact that I see SBHCs opening in this kind of school suggests the intervention is occurring where it is most needed.

V. Methodology

Difference-in-Differences

To answer the question of whether a school-based health center has an impact on student achievement, I begin with the simple difference-in-differences equation:

$$Y_{idst}=\beta_0 + \beta_1 SBHC_{idst} + X_{idt} + \alpha_i + \gamma_{st} + \varepsilon_{idst} \qquad \text{(Equation 1)}$$

where Y_{idst} indicates the outcome variable of interest in school i in district d in state s in year t. The variable $SBHC_{idst}$ is a dummy for whether a school-based health center is open in a particular school in a given year. β_1 is the main coefficient of interest and describes the effect of a school getting treated with a school-based health center. X_{idt} is a vector of school and district demographic characteristics, and controls for race, ethnicity, and socioeconomic status of the student body in a given year. α_i are the school fixed effects, which control for school-specific characteristics that are unchanging over time, while γ_{st} controls for state-by-year fixed effects, which control for unobservable exogenous shocks that may have affected all schools in a state in a given year. The state-by-year fixed effects also ensure I compare schools within a given state in a given year. The error term is described by ε_{idst}, and the standard errors are clustered at the school level.

The key identifying assumptions necessary for β_1 to represent the causal effect of SBHCs are, first, that the location and timing of the opening of an SBHC is uncorrelated with trends in student achievement outcomes, and second, that the school-based health center is the only change occurring when a school does get treated with an SBHC. The school-fixed effects address the concern that SBHCs are generally opening in more disadvantaged schools, however I am still making the assumption that the SBHC schools and the non-SBHC schools are trending similarly over time, prior to the intervention. I use the event study described below to show this assumption is plausibly valid. Additionally, I am assuming the decision to open a school-based health center is not made in tandem with other decisions, such as lowering the student-teacher ratio, introducing new meal programs, or lengthening the school day, which could have a confounding effect on the treatment. Starting an SBHC is a complex process that requires collaboration between many key stakeholders over the course of several years. Given this time-intensive development, it is unlikely the opening would be correlated with any of these other changes.

Event Study

To supplement the difference-in-differences model, I use an event study specification, which is helpful for two reasons. First, the event study can identify any existing pre-trends in the data prior to the opening of the SBHC. Additionally, the model can show whether there is a time dependent effect of the school-based health center, perhaps having a stronger impact the longer it has been open. The event study model is as follows:

$$Y_{idst} = \sum_{j=-6}^{8} \theta_j T_{j,idst} + X_{idt} + \alpha_i + \gamma_{st} + \varepsilon_{idst}. \qquad \text{(Equation 2)}$$

where Y_{idst} describes the outcome variable, and T_j is a series of lead and lag indicator variables for when school *i* opened an SBHC. The period of interest ranges from six years prior to the health center opening to eight years after. T_j is equal to one for the given year of interest and zero if else. Because many schools have observations during my sample period that are more than six years before or eight years after their SBHC opens, $T_{-6,idst}$ equals one for all observations that are 6 or more years before an SBHC opens, and $T_{8,idst}$ equals one for all observations that are 8 or more years after the health center has opened. X_{idt} is the vector of school and district characteristics, α_i are the school fixed effects, γ_{st} are the state-by-year fixed effects, and ε_{idst} is the error term.

The coefficient of interest is θ_j which describes the SBHC's effect on the outcome variable relative to the omitted year, which in this case is the year before the health center opens, $T_{-1,idst}$. For j=-2 through -6, θ_j identifies any pre-trends in the data prior to the SBHC opening and should be small and statistically insignificant for us to believe the identifying assumptions hold. From year j=0 through 8, θ_j will show the difference-in-differences effect of the opening of the SBHC in each year. This separates the "after" period evaluated by the difference-in-differences analysis into unique, year-by-year effects, allowing for a dynamic treatment effect.

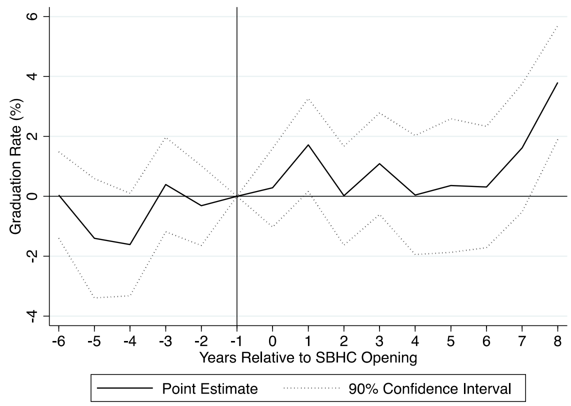
VI. Results

VIa. Event Study

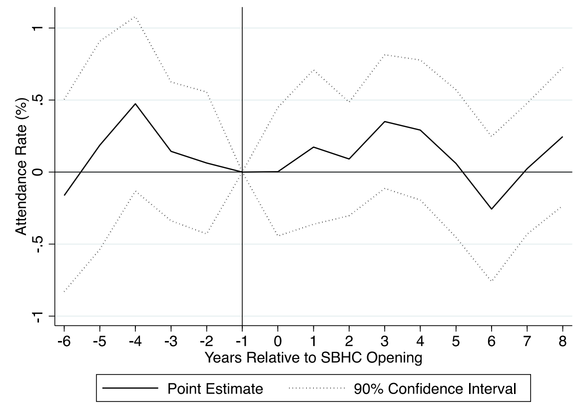
I begin with the event study analysis estimated using Equation 2 and shown in Figure 3. The point estimates are indicated by the dark line, and the 90% confidence intervals by the dotted lines. I see a positive effect on graduation rates, particularly in the later years after a center has been opened, and I find no effect on the attendance rate. The standardized test data show a significant decrease in test scores following the opening of an SBHC. Explanations for this unexpected result are discussed in the following sections. Across my outcomes, I see no evidence of pre-trends, suggesting that the SBHC opening is uncorrelated with prior trends in student achievement. To increase statistical precision and provide a single parameter summarizing the SBHC impact, I show difference-in-differences estimations in the next section.

Figure 3: Effects of Opening an SBHC on Academic Achievement

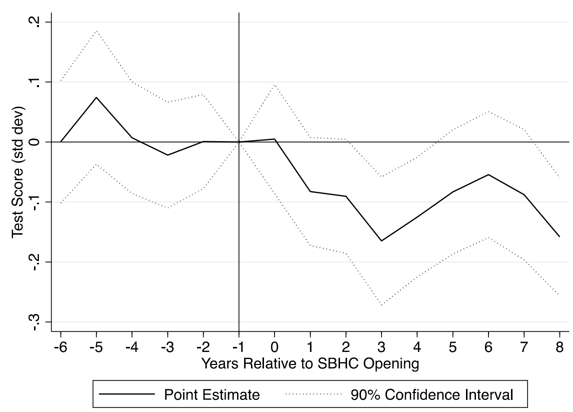
(a) Graduation Rates



(b) Attendance Rates



(c) Standardized Test Scores



Notes: The event study is estimated using Equation 2. The solid line is the point estimate, and the dotted line is the 90% confidence interval.

VIb. Main Results

Table 3 shows the results from Equation 1 on my main outcome variables: graduation rates, attendance rates, and standardized test scores. I estimate each regression twice: once with only school and state-by-year fixed effects (shown in Columns 1,3, and 5) and once with the fixed effects and controls (Columns 2,4, and 6). The controls include race and ethnicity covariates at the school and district level, percent of students qualifying for free and reduced lunch, percent of low English proficiency students, district student-teacher ratio, district student-guidance counselor ratio, and school enrollment. My preferred specification includes the covariates. The results in Table 3 show that opening a school-based health center has no effect on attendance. The point estimate is very small, showing just a 0.07 percentage point increase, and statistically insignificant, meaning I can rule out with confidence any economically significant effects on attendance. Graduation rates increase by 1.48 percentage points, significant at the 90% confidence level. However, this result is not robust and attenuates with different specifications, heterogeneity, and robustness checks.

Table 3: Difference-in-Differences SBHC Opening, All Outcomes

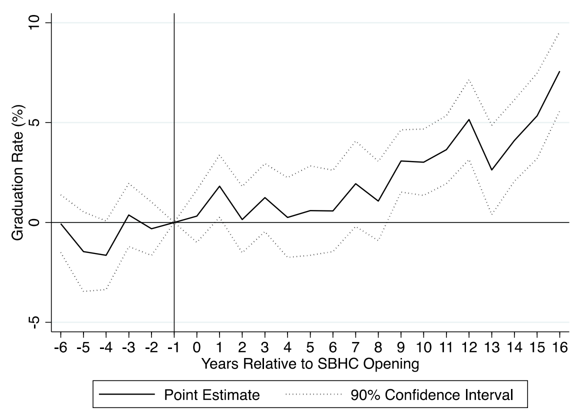
	Attendance Rate		Graduation Rate		Test Scores	
	(1)	(2)	(3)	(4)	(5)	(6)
SBHC Opening	0.20	0.07	2.53***	1.48*	-0.08*	-0.10**
	(0.23)	(0.22)	(0.76)	(0.77)	(0.04)	(0.04)
Observations	57,114	57,114	47,908	47,908	183,612	183,612
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-by-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes

Notes: Each column presents results from a separate regression. The dependent variable is in the top row. Attendance rates are calculated for New York only from 2005-2019. Standard errors are clustered at the school level. * significant at 10%, ** significant at 5%, *** significant at 1%

The increase in graduation rates seems to be driven almost entirely from increases eight or more years after the school-based health centers open. Given Figure 3a, this suggests it takes a full two cohorts of high school students for the SBHC to begin having an effect. Figure 4 shows the event study picture for graduation rates extended to 16 years after an

SBHC opens. The figure shows an increasing effect beginning about 7 to 9 years after an SBHC opens.[1] One potential hypothesis for this increasing trend could be explained by the population the SBHCs are serving. Many schools, particularly in California and upstate New York, offer SBHC services to any student in the district. Some schools also offer services to family members or younger siblings. Perhaps if younger children in the area began receiving care at the SBHC when it first opened, this large increase in graduation rates is a result of the cohort having been exposed to care at a young age rather than just in their high school years.

Figure 4: Effects of Opening an SBHC on Graduation for 16 Years After Opening



Notes: Graduation rates from 1997-2019. Point estimates in black, 90% confidence interval in dotted line.

The final academic outcome shown in Table 3 looks at test scores. Column 6 shows that following the opening of an SBHC in elementary schools, test scores significantly decrease by 0.10 standard deviations. These test score and non-test score outcomes are somewhat at odds with each other. SBHCs seem to be having a neutral to positive effect on non-test score outcomes and a negative effect on test scores. Most importantly, the

[1] Figure 4 only includes SBHCs opened in high schools because only high schools have non-missing graduation rates.

magnitude of these effects is not necessarily comparable. The effects on the graduation and attendance are small, relatively imprecise, and attenuate throughout the study, while the decrease in test scores is strongly statistically significant, persistent, and relatively large.

Given this caveat, however, there are a few reasons why I might expect this pattern to emerge. One important consideration is in the different services that elementary and high school SBHCs provide. While elementary school SBHCs focus primarily on giving annual primary care physicals, treating acute incidents, and managing chronic conditions, high school SBHCs have a much larger focus on providing contraceptives, family planning, and mental health services. Perhaps these types of services have larger effects on educational attainment and could be driving the positive effect that is absent from the test-score outcomes. In an extreme example, we might think that being able to meet with a mental health provider each week affects education more than getting a flu shot at an annual check-up each September. However, this hypothesis fails to explain why test scores are significantly smaller following the opening of a center. The following sections delve deeper into explanations for this surprising result.

VIc. Heterogeneity

I first break down my sample to look at outcomes in each state to determine whether one or the other is driving the trends in my data. Table 4 shows the results on academic outcomes separately for California and New York. I see that while I lose statistical power due to a smaller sample size and the estimates are no longer statistically significant, the patterns are fairly consistent with the overall trends in Table 3. Graduation rates are small but positive, and there is a negative effect on test scores hovering around 0.10 standard deviations.[2]

Table 4: Difference-in-Differences SBHC Opening, By State

	New York		California	
	Grad Rate	Test Score	Grad Rate	Test Score
	(1)	(2)	(5)	(6)
SBHC Opening	0.60	-0.08	1.05	-0.07
	(1.14)	(0.05)	(1.01)	(0.06)
Observations	20,569	68,274	27,323	115,338
School Fixed Effects	Yes	Yes	Yes	Yes
State-by-Year Fixed Effects	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Notes: Each column presents results from a separate regression. The dependent variable is in the top row. Standard errors are clustered at the school level. * significant at 10%, ** significant at 5%, *** significant at 1%.

The trends in my data may also be driven by differences in the effect of a school-based health center depending on where it is located or the students it serves. Table 5 breaks down my sample into “Poor” vs “Not Poor” schools depending on the fraction of students who qualify for free and reduced lunch in the school. Schools that have a fraction of FRL students greater than the median are considered “Poor”. Table 6 compares the results in urban vs non-urban schools.

Table 5 shows that in “Not Poor” schools, attendance and graduation rates are actually slightly positive and marginally statistically significant, though the effects, particularly for attendance, are quite small. Columns 5 and 6 in Table 5, however, uncover some previously masked heterogeneity. In “Not Poor” elementary and middle schools, shown in Column 6, I find a 0.17 standard deviation decrease in test scores, significant at the 99%

[2] I only have attendance data for New York which is why attendance outcomes are not shown in Table 4.

confidence level, relative to a 0.09 standard deviation marginally significant decrease for the poor schools shown in Column 5. The unexpected, negative effect on test scores seems to be coming primarily from these “Not Poor” schools. In other words, places with fewer students who qualify for free and reduced lunch are seeing greater decreases in test scores following the opening of an SBHC.

Table 5: Difference-in-Differences, Poor vs Not Poor Schools

	Attendance Rate		Graduation Rate		Test Scores	
	Poor	Not Poor	Poor	Not Poor	Poor	Not Poor
	(1)	(2)	(5)	(4)	(5)	(6)
SBHC Opening	-0.15	0.60*	0.45	1.61*	-0.09*	-0.17***
	(0.29)	(-0.32)	(1.28)	(0.89)	(0.05)	(0.06)
Observations	22,613	34,501	15,164	32,744	98,403	85,209
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-by-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column presents results from a separate regression. The dependent variable is in the top row. Attendance rates are calculated in New York only from 2005-2019. Standard errors are clustered at the school level. * significant at 10%, ** significant at 5%, *** significant at 1%.

Similar to the main results, the opposite trends between test-score and non-test score outcomes are present and particularly strong in the “Not Poor” schools. One possible explanation for the large negative effect on test scores, which is discussed in much greater detail below, is that children in “Not Poor” schools may already have good health care, for which a lower quality SBHC is a poor substitute. If students ultimately begin receiving lower quality care following the opening of an SBHC, then it may not be entirely surprising that the test scores go down.

It is a bit more difficult to explain, however, the opposite, positive trend which occurs in attendance and graduation rates in the "Not Poor" schools. One hypothesis, which is beyond the scope of my paper to test empirically, is that in wealthier schools there is potentially less of a stigma surrounding the use of mental health resources and contraceptives. Bharadwaj et al (2008) finds that there is more under-reporting of mental illness and greater stigma around seeking mental health care for minorities and low socioeconomic status individuals in Australia. Furthermore, Dereuddre et al (2016) finds that higher socioeconomic status is associated with greater contraceptive use in Europe. If wealthier students are more motivated to ask for help or feel less stigmatized in seeking out care, then perhaps they are making greater use of the SBHCs. This is one potential explanation for why I find a greater effect on graduation and attendance in "Not Poor" schools relative to schools that have more students who qualify for free and reduced lunch.

It is a fair question to wonder why the "quality" argument I present for test scores does not seem to also be true for these non-test score outcomes. If elementary school SBHCs tend to focus more on traditional aspects of primary care, while high schools are emphasizing mental health, then it is important to consider what existing services are available outside the SBHC. Perhaps students are less likely to already be seeing an outside mental health provider than a primary care practitioner, so the quality of SBHCs matters less because some care does seem to be an improvement over no care at all. Cronin et al (2020) found that in 2011, that while just 3% of Americans saw a therapist in the past year, this number is closer to 55% for primary care visits by adolescents (Rand et al, 2018). Connecting students to resources they might not already have access to could be one reason that I see this differing effect between test score and non-test score outcomes.

Table 6 shows the results from the urban and non-urban schools. Here, the results are much more similar across groups. Results for attendance rates and graduation rates are not statistically significant, and the point estimates are fairly consistent, though imprecise, across the divisions. The test scores, while showing statistical significance, also do not seem to vary greatly between these urban and non-urban schools.

Table 6: Difference-in-Differences, Urban vs Not Urban Schools

	Attendance Rate		Graduation Rate		Test Scores	
	Urban (3)	Not Urban (4)	Urban (3)	Not Urban (4)	Urban (3)	Not Urban (4)
SBHC Opening	0.11	0.07	1.39	1.13	-0.10*	-0.12**
	(0.30)	(0.19)	(1.06)	(1.08)	(0.05)	(0.06)
Observations	21,804	35,310	16885	31023	78,362	105,187
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State-by-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column presents results from a separate regression. The dependent variable is in the top row. Attendance rates are calculated for New York only from 2005-2019. Standard errors are clustered at the school level. * significant 10%, ** significant at 5%, *** significant at 1%.

Importantly, there could also be a differential impact of SBHCs on different students. Given the existing disparities in the health care system, certain students could potentially derive more benefit than others from the increased accessibility and convenience of an SBHC. Tables 7 and 8 explore this heterogeneity at the student level. Table 7 shows that across racial and ethnic subgroups, graduation rate estimates remain positive, small, and statistically insignificant, making it difficult to compare the effects between different populations.

Table 7: Graduation Rate Difference-in-Differences, By Subgroups

	White	Black	Hispanic	Asian
	(1)	(2)	(3)	(4)
SBHC Opening	0.77	1.05	0.63	1.81
	(1.51)	(1.32)	(0.98)	(1.35)
Observations	29,475	25,021	29,798	20,783
School Fixed Effects	Yes	Yes	Yes	Yes
State-By-Year Fixed Effects	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Notes: Each column presents results from a separate regression. The dependent variable is in the top row. * significant at 10%, ** significant at 5%, and *** significant at 1%

Table 8, however, which assesses the impact on test scores, tells a different story. For Hispanic and Black students, although the results still have a negative sign on the point estimates, they are smaller, and the effects are not statistically significant. However, Column 1 shows a 0.16 statistically significant standard deviation decrease in test scores for White students, and Column 4 shows 0.22 marginally significant standard deviation decrease for Asian students. The results at the student level suggest that the largest drivers of the negative trends in test scores seem to be coming from more historically advantaged groups.[3] This is consistent with the results showing that the wealthier schools have larger decreases in test scores. The negative test score results are discussed further in the following section.

[3] Unfortunately, I do not have attendance data broken down into subgroups at the student level, and so I am unable to assess these same impacts on the attendance rate outcome.

Table 8: Test Score Difference-in-Differences, By Subgroup

	White	Black	Hispanic	Asian
	(1)	(2)	(3)	(4)
SBHC Opening	-0.16**	-0.08	-0.00	-0.22*
	(0.07)	(0.06)	(0.05)	(0.12)
Observations	109,116	56,345	130,207	48,807
School Fixed Effects	Yes	Yes	Yes	Yes
State-by-Year Fixed Effects	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes

Notes: Each column presents results from a separate regression. The dependent variable is in the top row. Test scores are shown in standard deviations. * significant at 10%, ** significant at 5%, and *** significant at 1%

VId. What is Driving Negative Test Scores?

The negative effects on test scores are at first glance inconsistent with the existing literature showing the strong correlation between improved health and improved educational attainment over the course of the life cycle. Furthermore, within the scope of my study, the negative test score results contradict the small but positive effect on graduation rates. I empirically test three different theories to show why this might be the case.

School-based health centers are primarily targeted at the students who will benefit the most from the intervention. Given the link in the literature between socioeconomic status, health, and education, there is likely overlap between students who are in need of access to care and students who are suffering academically. After an SBHC opens in a school, that school could potentially attract a different type of student; one who needs the health center and who also has lower test scores. As a result, I am concerned that the student body composition may change after an SBHC opens which could be driving the negative effect on test score outcomes.

Table 9 shows the effect of opening a center on the percent of students who qualify for free and reduced lunch, as well as different racial and ethnic groups. The effects for each subgroup are quite small, showing percentage point increases on a 0 to 100 scale. Column 1 shows a very small increase in White students, Column 2 shows a very small decrease in Black students, and Column 5 shows a small decrease in students who qualify for free and reduced lunch. The estimates for Hispanic and Asian students in Columns 3 and 4 are precise zeroes. Because these point estimates are so small and if anything in the opposite direction that I initially hypothesize, I do not expect that changing student body composition is driving the negative effect on test scores.

Table 9: Change in Student Body Composition After SBHC Opens

	White (1)	Black (2)	Hispanic (3)	Asian (4)	Free and Reduced Lunch (5)
SBHC Opening	0.19**	-0.25**	-0.00	-0.02	-1.58*
	(0.08)	(0.12)	(0.09)	(0.10)	(0.86)
Observations	223,361	223,361	223,361	223,361	223,361
School Fixed Effects	Yes	Yes	Yes	Yes	Yes
State-By-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes

Note: Each column presents results from a separate regression. The dependent variable is in the top row. The student body composition measured on a scale of 0 to 100. Column 1 should be interpreted as a 0.19% increase in the number of white students in a school after an SBHC opens.

I hypothesized also that there may be differential effects on ELA and Math scores, given that each subject tests different skills. My initial test score results are a composite score which combines Math and ELA scores, but to test whether there might be a differential impact, I separate the data. The results are shown in Table 10. I find that although the results are negative on both tests, the math results are more negative and statistically significant at the 95% confidence level. This suggests that the negative effect on test scores is coming in part from worse outcomes in mathematics.

Table 10: Difference-in-Differences, Math vs ELA Test Scores

	Math	ELA
	(1)	(2)
SBHC Opening	-0.10**	-0.03
	(0.04)	(0.04)
Observations	179,591	180,526
School Fixed Effects	Yes	Yes
State-by-Year Fixed Effects	Yes	Yes
Covariates	Yes	Yes

Notes: The test scores should be interpreted in terms of standard deviations. * is significant at 10%, ** is significant at 5%, and *** is significant at 1%

The existing health services in a community may also influence the effectiveness of an SBHC. If students' health care needs are already being met, it is possible the SBHC will not have as much of an effect on student academic outcomes. Furthermore, if students' existing health care providers are of higher quality than the SBHCs, it is possible student academic achievement could even suffer as a result. Tables 11, 12 and 13 use clinical care rankings on rates of primary care physicians, mental health providers, and dentists by county, as well as numbers on the uninsured population, preventable hospitalizations, rates of flu vaccination, and mammography screenings to assess the differential impacts of SBHC depending on the existing services in the county. Table 11 shows that SBHCs do seem to have slightly more of an impact on attendance in places with less existing clinical care. Although the results are statistically insignificant, the negative coefficient in the upper third and positive coefficients for the middle and lower thirds of clinical care status are suggestive of this pattern.

Table 11: Attendance Rates, Clinical Care Index

	Upper Third	Middle Third	Lower Third
	(1)	(2)	(3)
SBHC Opening	-0.26	0.13	0.09
	(0.30)	(0.30)	(0.34)
Observations	24,896	11,256	18,007
School Fixed Effects	Yes	Yes	Yes
State by Year Fixed Effects	Yes	Yes	Yes
Covariates	Yes	Yes	Yes

Notes: Each column shows a separate regression. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 12 finds that rates of graduation do not vary widely between different exposures to existing health services; the results are positive and statistically insignificant within each third of clinical care.

Table 12: Graduation Rates, Clinical Care Index

	Upper Third	Middle Third	Lower Third
	(1)	(2)	(3)
SBHC Opening	1.01	1.81	0.80
	(1.19)	(1.33)	(1.28)
Observations	16,912	10,909	16,947
School Fixed Effects	Yes	Yes	Yes
State by Year Fixed Effects	Yes	Yes	Yes
Covariates	Yes	Yes	Yes

Notes: Each column shows a separate regression. * significant at 10%, ** significant at 5%, *** significant at 1%

Table 13 shows the effect of SBHCs on standardized test scores and finds important heterogeneity. The negative effect on test scores seems to be coming entirely from the upper third of clinical care rankings. Column 1 shows a 0.24 standard deviation reduction in test scores statistically significant at the 99% confidence level. Essentially, schools with the best existing health care services in the surrounding area are seeing the biggest decline in academic achievement for elementary and middle school students. Columns 2 and 3 show much smaller point estimates that are not statistically significant. These findings are consistent with the heterogeneity at the student and school level in showing that the negative test scores seem to be driven predominantly by more privileged groups. While the absence of a negative is certainly not a positive, the data do seem to suggest that the most vulnerable student populations are not being made worse off by the opening of an SBHC.

Table 13: Standardized Test Scores, Clinical Care Index

	Upper Third (1)	Middle Third (2)	Lower Third (3)
SBHC Opening	-0.24***	0.03	-0.06
	(0.07)	(0.09)	(0.05)
Observations	66,613	36,831	69,884
School Fixed Effects	Yes	Yes	Yes
State by Year Fixed Effects	Yes	Yes	Yes
Covariates	Yes	Yes	Yes

Notes: Standardized test scores are shown in standard deviations. * significant at 10% level** significant at 5% confidence level *** significant at 1% confidence level

One possible explanation is that SBHCs are simply less effective or lower quality than the existing health providers that the more advantaged children are already seeing. When the SBHC opens, families may switch to the new center because of its convenience and students are no longer receiving as high-quality care. On the flip side, in more underserved communities, the SBHCs does not seem to be having a negative effect. So, in other words, SBHCs are not as effective as existing health services in places with high levels of clinical care, but they do seem to be at least of equal quality to the care more underserved populations are currently using.

For example, consider a situation with Billy and Tommy who are similar in every way except that Billy's family has good health insurance and so he goes to his pediatrician at a nearby hospital once a year for his annual physical. For Tommy's family, however, the nearest doctor's office is 40 minutes away. As he has gotten older his parents have not been able to find the time to get off work to take him to his appointments, and he's gone a few years now without seeing a pediatrician. When an SBHC opens in their respective schools, both families give consent for their children to be seen there alone. The SBHC has relieved parents of the burden of coordinating work obligations, transportation, and childcare in order to attend appointments with their children. Relative to his initial condition, Tommy is now slightly better-off. He gets his necessary immunizations, receives a normal physical exam, and interacts with the health care system in a way he has not done in a while. On the other hand, Billy's family perhaps unknowingly switched to a worse alternative. Billy gets the same shots and examinations that he did before, but he is now left without his parents to retain the information, ask the right questions, and make the necessary changes that his doctor recommends. Perhaps suggestions like "eat more vegetables" or "less screen time" that would have been implemented by his parents now never make it home. Placing the onus of health-decision making on young children may only get them so far. The convenience of the SBHC is considered its biggest asset, but with parents no longer present at appointments, it may also be a reason why this negative effect persists.

A simpler mechanism to explain this, though, is that the SBHC actually just is not as good as wherever Billy was going before. Maybe the SBHC is overcrowded and less resourced and does not provide the same quality of care that he had previously been receiving. The results suggest that the quality is different enough that Billy is actively being made worse off with this transition to the SBHC. While the heterogeneous effects by existing community health services quality are suggestive, they also imply the need for future research in this field.

Beyond the empirical theories that I was able to test, there are a few other hypotheses for why I may see these consistently negative results on the test scores. Many SBHCs have a health education component, where educators from the health center come into the classroom to teach about things like nutrition and sex education. If these lessons are taking important instruction time away from things like math and reading, this could potentially decrease test scores in the long run. It is also possible that kids who had previously missed testing days due to absenteeism are now sitting for the standardized tests. Perhaps the test scores are lower because it is a different group of students who are showing up to school and taking the test on test day. Finally, it is possible that there are changes in funding when an SBHC opens which could take away resources from classroom teaching. This argument is a bit less convincing since funding of the SBHC comes primarily from the outside health partner but cannot be entirely ruled out.

VIe. Robustness Checks

I conduct two separate checks to show that my results are robust to different threats to identification. In my main sample, I include any school that has an SBHC in the treatment group. In my first check, I drop schools that had SBHCs that opened prior to the period for which I have data. The results are shown in Table 14.

Table 14: Robustness Check, Only SBHCs That Open 1997-2019

	Attendance (1)	Graduation (2)	Test Scores (3)
SBHC Opening	0.09	1.61**	-0.09**
	(0.23)	(0.77)	(0.04)
Observations	55.932	47,015	181,844
School Fixed Effects	Yes	Yes	Yes
State by Year Fixed Effects	Yes	Yes	Yes
Covariates	Yes	Yes	Yes

Notes: Each column presents results from a separate regression. The dependent variable is in the top row. Attendance rates are calculated for New York only from 2005-2019. Standard errors are clustered at the school level. * significant at 10%, ** significant at 5%, and *** significant at 1%.

The point estimates and statistical significance shown in Table 14 are very close to the main results shown in Table 3. For attendance rates and standardized test scores, the values between the two results are nearly identical. For graduation rates, this specification actually slightly increases the magnitude and significance of the positive effect of the SBHC.

My analysis assumes that schools who do not receive a school-based health center are an effective control group for the study. However, I have established schools that are receiving an SBHC look different than schools that are never treated. My initial sample includes all public schools in New York and California, including many suburban, wealthy, high-quality public schools which may not be an accurate benchmark with which to compare

SBHC schools. I use propensity score matching to create a control group that looks much more similar to the SBHC treated schools. These results are shown in Table 15.

Table 15: Robustness Check, Propensity Score Matching

	Attendance (1)	Graduation (2)	Test Scores (3)
SBHC Opening	-0.13	0.35	-0.09**
	(0.23)	(0.81)	(0.04)
Observations	7,778	7,975	9,420
School Fixed Effects	Yes	Yes	Yes
State by Year Fixed Effects	Yes	Yes	Yes
Covariates	Yes	Yes	Yes

Notes: Each column presents results from a separate regression. The dependent variable is in the top row. Attendance rates are calculated for New York only from 2005-2019. Standard errors are clustered at the school level. * significant at 10%, ** significant at 5**, and *** significant at 1%

Column 1 shows that the attendance rates remain not statistically significant in this check, and although the sign does flip, the standard errors are quite large and imprecise. Column 3 shows that using the propensity matching gives almost identical results for test scores compared to the full sample. Column 2 shows, however, that restricting the sample to just these schools greatly reduces the effect on graduation rates. The coefficient is about a quarter of the size and no longer statistically significant. This is relatively consistent with the attenuating effect on graduation rates I found across my different restrictions and heterogeneity tests.

VII. Conclusion

The existing literature on child health and education shows there is a well-established link between the two. Healthier children have higher educational attainment. Furthermore, socioeconomic status is also closely related to both of these outcomes, where poor children suffer from worse health, and by extension, have lower academic achievement. School-based health centers offer a potential intervention to break this cycle of poverty by offering care to the most vulnerable student populations.

The literature on school-based health centers is small but growing. The research comes mainly from the public health and medical fields and shows consistently positive effects on health outcomes through use of SBHCs, albeit using non-causal, case study analyses. However, the question of whether SBHCs have a causal impact on younger children's education remained unanswered. My research contributes to the growing literature on SBHCs by focusing on an expanded number of education outcomes, including test scores for elementary and middle school students, attendance rates, and graduation rates.

I find that while opening a school-based health center seems to have no statistically or economically significant impact on attendance rates, there is a 1.48 marginally significant percentage point increase in graduation rates. However, this increase in graduation rates is not robust and attenuates across a number of specifications and robustness checks. Additionally, there is an unexpected decrease of 0.10 standard deviations in test scores for elementary and middle school students. In exploring several different aspects of potential heterogeneity, I find that the largest decreases are coming from more advantaged groups. "Not Poor" schools see significantly larger decreases in test scores than "Poor" schools, and White and Asian students see much greater decreases than their Black and Hispanic peers. Of course, decreased test scores for students are never the goal, but it does suggest that maybe Black and Hispanic students, who have been historically excluded from the health care system, are benefitting more from the SBHCs than their peers.

Additionally, I find that SBHCs opening in areas with high levels of existing clinical services seem to also be driving negative trends in standardized test scores. One possible explanation is that SBHCs are of lower quality than the existing health services. The results suggest that SBHCs may be an improvement over no care at all, but may not be as effective as well-established, high-quality care. If students are utilizing the SBHC instead of their existing primary care services, I find they may be worse-off after this substitution.

While much of this paper has been spent exploring possible reasons for the negative effect on standardized test scores, the small increase in graduation rates should not go unnoticed. Certainly, graduating from high school determines much more in life than how you scored on a standardized math test in the fourth grade. While the negative effects on test scores are disappointing, they do not seem to be having long-run detrimental impacts on the most important education outcomes. Even the propensity score matching robustness check, my most restrictive specification, finds a small and positive, though not statistically significant, effect on graduation. I can rule out with confidence an economically significant negative impact on graduation rates. Given the mixed, somewhat contradictory results on graduation rates and test scores, future research is needed to explore the role of SBHCs in the school environment, particularly with regards to quality.

One of the major drawbacks of my analysis is that it does not have any data on the actual health impacts of the SBHC, which is of course the intended target of the intervention. Future research should be directed at causally examining the health benefits of SBHCs for younger children. In particular, dedicating more attention to the impacts of SBHCs on mental health is critical, given improvements here may have the largest spillovers into education Additionally, some sort of cost-benefit analysis of SBHCs as a mechanism of health care delivery compared to FQHCs and other community centered interventions would be useful in determining their effectiveness. School-based health centers are still a relatively unknown and uncommon intervention; just 2.5% of public schools in the US have access to one. However, given my somewhat perverse test-score results and the potential that SBHCs are actually making some students worse off, this future research is important before expanding the number of SBHCs across the country.

In the meantime, however, perhaps schools could limit which students are allowed to use the SBHC. If only students who qualify for free and reduced lunch are allowed to use the elementary school centers, the intervention becomes more specifically targeted at students who need it the most, without making the wealthier students worse off. Furthermore, with more time and resources being spent on the most underserved students, this may lead to actual improvements in their education outcomes.

While my results show somewhat mixed effects on education outcomes after the opening of a school-based health center, my findings should not be used to diminish their overall importance. In fact, based on the existing research, both from the medical literature and Lovenheim et al (2016), SBHCs appear to be effective in delivering health care to the most underserved patient populations. My results simply suggest that the effects do not have positive spillovers that are consistently large enough to be seen in education outcomes.

References

Aizer, Anna, Laura Stroud, and Stephen Buka. 2016. "Maternal Stress and Child Outcomes: Evidence from Siblings." *Journal of Human Resources* 51 (3): 523–55.

Allison, Mandy A., Lori A. Crane, Brenda L. Beaty, Arthur J. Davidson, Paul Melinkovich, and Allison Kempe. 2007. "School-Based Health Centers: Improving Access and Quality of Care for Low-Income Adolescents." *Pediatrics* 120 (4): E887-E894.

Almond, Douglas. 2006. "Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population." *Journal of Political Economy* 114 (4): 672–712.

Almond, Douglas, and Bhashkar Mazumder. 2011. "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance during Pregnancy." *American Economic Journal: Applied Economics* 3 (4): 56–85.

Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* 25 (3): 153–72.

Bharadwaj, Prashant, Mallesh M. Pai, and Agne Suziedelyte. 2015. "Mental Health Stigma." w21240. National Bureau of Economic Research.

Bailey, Martha J., and Andrew Goodman-Bacon. 2015. "The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans." *American Economic Review* 105 (3): 1067–1104.

Bailey, Martha J., Brenden D. Timpe, and Shuqiao Sun. 2020. "Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency." w28268. National Bureau of Economic Research.

Barr, Andrew, and Chloe R Gibbs. 2018. "Breaking the Cycle? Intergenerational Effects of an Anti-Poverty Program in Early Childhood." Chloe Gibb's Personal Website.

Brown, David W, Amanda E Kowalski, and Ithai Z Lurie. 2020. "Long-Term Impacts of Childhood Medicaid Expansions on Outcomes in Adulthood." *The Review of Economic Studies* 87 (2): 792–821.

Bureau of Primary Health Care. 2018. "Health Center Compliance Manual- Chapter 20: Board Composition."

Case, Anne, Darren Lubotsky, and Christina Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review* 92 (5): 1308–34.

Case, Anne, Angela Fertig, and Christina Paxson. 2005. "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics* 24 (2): 365–89.

Cohodes, Sarah R., Daniel S. Grossman, Samuel A. Kleiner, and Michael F. Lovenheim. 2016. "The Effect of Child Health Insurance Access on Schooling: Evidence from Public Insurance Expansions." *Journal of Human Resources* 51 (3): 727–59.

Cronin, Christopher J., Matthew P. Forsstrom, and Nicholas W. Papageorge. 2020. "What Good Are Treatment Effects without Treatment? Mental Health and the Reluctance to Use Talk Therapy." w27711. National Bureau of Economic Research.

Currie, Janet. 2009. "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development." *Journal of Economic Literature* 47 (1): 87–122.

Currie, Janet, and Jonathan Gruber. 1996. "Health Insurance Eligibility, Utilization of Medical Care, and Child Health*." *The Quarterly Journal of Economics* 111 (2): 431–66.

Daley, Matthew F., C. Robinette Curtis, Jennifer Pyrzanowski, Jennifer Barrow, Kathryn Benton, Lisa Abrams, Steven Federico, et al. 2009. "Adolescent Immunization Delivery in School-Based Health Centers: A National Survey." *Journal of Adolescent Health* 45 (5): 445–52.

Darling-Hammond, Linda. 1998. "Unequal Opportunity: Race and Education." *Brookings Institute*.

Dereuddre, Rozemarijn, Bart Van de Putte, and Piet Bracke. 2016. "Ready, Willing, and Able: Contraceptive Use Patterns across Europe." *European Journal of Population* 32 (4): 543–73.

Geierstanger, Sara Peterson, Gorette Amaral, Mona Mansour, and Susan Russell Walters. 2004. "School-Based Health Centers and Academic Performance: Research, Challenges, and Recommendations." *Journal of School Health* 74 (9): 347–52.

Guo, J. J., R. Jang, K. N. Keller, A. L. McCracken, W. Pan, and R. J. Cluxton. 2005. "Impact of School-Based Health Centers on Children with Asthma." *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine* 37 (4): 266–74.

Gustafson, Elaine M. 2005. "History and Overview of School-Based Health Centers in the US." *Nursing Clinics of North America* 40 (4): 595–606.

Kerns, Suzanne E. U. 2011. "Adolescent Use of School-Based Health Centers and High School Dropout." *Archives of Pediatrics & Adolescent Medicine* 165 (7): 617.

Love, Hayley, Samira Soleimanpour, John Schlitt, Nirmita Panchal, and Caroline Behr, "2016-17 National School-Based Health Care Census Report." National School-Based Health Alliance. Accessed November 11, 2020

Love, Hayley E., John Schlitt, Samira Soleimanpour, Nirmita Panchal, and Caroline Behr. 2019. "Twenty Years of School-Based Health Care Growth And Expansion." *Health Affairs* 38 (5): 755–64.

Lovenheim, Michael, Randall Reback, and Leigh Wedenoja. 2016. "How Does Access to Health Care Affect Teen Fertility and High School Dropout Rates? Evidence from School-Based Health Centers." w22030. Cambridge, MA: National Bureau of Economic Research.

Maruthappu, Mahiben, Rele Ologunde, and Ayinkeran Gunarajasingam. 2012. "Is Health Care a Right? Health Reforms in the USA and Their Impact Upon the Concept of Care." *Annals of Medicine and Surgery* 2 (1): 15–17.

New York Department of Health. 2017. "Principles and Guidelines for School-Based Health Centers in New York State,"

Oliver, Kristin, Colleen McCorkell, Ilana Pister, Noora Majid, Denise H. Benkel, and Jane R. Zucker. 2019. "Improving HPV Vaccine Delivery at School-Based Health Centers." *Human Vaccines & Immunotherapeutics* 15 (7–8): 1870–77.

Porter, Philip J. 1981. "Realistic Outcomes of School Health Service Programs." *Health Education Quarterly* 8 (1): 81–87.

Rand, Cynthia M., and Nicolas P.N. Goldstein. 2018. "Patterns of Primary Care Physician Visits for US Adolescents in 2014: Implications for Vaccination." *Academic Pediatrics* 18 (2): S72–78.

Wade, Terrance J., Mona E. Mansour, Kristin Line, Tracy Huentelman, and Kathryn N. Keller. 2008. "Improvements in Health-Related Quality of Life among School-Based Health Center Users in Elementary and Middle School." *Ambulatory Pediatrics: The Official Journal of the Ambulatory Pediatric Association* 8 (4): 241–49.

Walker, Sarah Cusworth, Suzanne E. U. Kerns, Aaron R. Lyon, Eric J. Bruns, and T. J. Cosgrove. 2010. "Impact of School-Based Health Center Use on Academic Outcomes." *Journal of Adolescent Health* 46 (3): 251–57.

Webber, Mayris P., Kelly E. Carpiniello, Tosan Oruwariye, Yungtai Lo, William B. Burton, and David K. Appel. 2003. "Burden of Asthma in Inner-City Elementary Schoolchildren: Do School-Based Health Centers Make a Difference?" *Archives of Pediatrics & Adolescent Medicine* 157 (2): 125–29.

Data References

California Department of Education. 1997-2010. "Student & School Data Files (Downloadable)" Accessed October 2020. https://www.cde.ca.gov/ds/sd/sd/

California School-Based Health Alliance. “Active SBHC Record Details” Accessed September 2020. Personal communication with Tracy Macdonald Mendez.

“County Health Rankings & Roadmaps.” 2021. University of Wisconsin Population Health Institute. Accessed April 2021 https://www.countyhealthrankings.org/

New York Department of Education. 1999-2019. “Report Card Database.” Accessed December 2020. https://data.nysed.gov/downloads.php

New York Department of Health. “SBHC Opening Dates 9.30.2020” Accessed October 2020. Personal communication with Denise Hernas.

U.S. Department of Education. “Common Core of Data”. National Center for Education Statistics. Accessed January 2021. https://nces.ed.gov/ccd/