EdWorkingPaper No. 22-650

Does High School Size Affect Rates of Risky Health Behaviors and Poor Mental Health Among Low-Income Teenagers? Evidence from New York City

Kai Hong
New York University

Syeda Sana Fatima
American Institutes for Research

Sherry Glied
New York University

Leanna Stiefel
New York University

Amy Ellen Schwartz
Syracuse University

There is increasing concern about risky behaviors and poor mental health among school-aged youth. A critical factor in youth well-being is school attendance. This study evaluates how school organization and structure affect health outcomes by examining the impacts of a popular urban high school reform -- "small schools" -- on youth risky behaviors and mental health, using data from New York City. To estimate a causal estimate of attending small versus large high schools, we use a two-sample-instrumental-variable approach with the distance between student residence and school as the instrument for school enrollment. We consider two types of small schools – "old small schools," which opened prior to a system-wide 2003 reform aimed at increasing educational achievement and "new small schools," which opened in the wake of that reform. We find that girls enrolled in older small schools are less likely to become pregnant, and boys are less likely to be diagnosed with mental health disorders than their counterparts in large schools. Both girls and boys enrolled in more recently opened small schools, however, are more likely to be diagnosed with violence-associated injuries and (for girls only) with mental health disorders. These disparate results suggest that improving a school's organization and inputs together is likely more effective in addressing youth risky behaviors than simply reducing school size.

VERSION: September 2022

Suggested citation: Hong, Kai, Syeda Sana Fatima, Sherry Glied, Leanna Stiefel, and Amy Ellen Schwartz. (2022). Does High School Size Affect Rates of Risky Health Behaviors and Poor Mental Health Among Low-Income Teenagers? Evidence from New York City. (EdWorkingPaper: 22-650). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/751q-er06

Does High School Size Affect Rates of Risky Health Behaviors and Poor Mental Health Among Low-Income Teenagers? Evidence from New York City[1]

Running Title: High School Size and Risky Behaviors

Kai Hong[a], Syeda Sana Fatima[b], Sherry Glied[c], Leanna Stiefel[d], Amy Ellen Schwartz[e]

[a]Wagner School of Public Service
295 Lafayette St., 2nd floor
New York, NY 10012
kai.hong1114@gmail.com

[b]American Institutes for Research
1400 Crystal Drive, 10th Floor
Arlington, VA 22202
sfatima@air.org

(Corresponding Author)
[c]Wagner School of Public Service
295 Lafayette St., 2nd floor
New York, NY 10012
Sherry.glied@nyu.edu

[d]Wagner School of Public Service
295 Lafayette St., 2nd floor
New York, NY 10012
leanna.stiefel@nyu.edu

[e]Maxwell School of Citizenship and Public Affairs
Syracuse University
Syracuse, NY 13244
amyschwartz@syr.edu

[1] This research was supported by the Robert Wood Johnson Foundation Policies for Action Program.

Abstract: There is increasing concern about risky behaviors and poor mental health among school-aged youth. A critical factor in youth well-being is school attendance. This study evaluates how school organization and structure affect health outcomes by examining the impacts of a popular urban high school reform -- "small schools" -- on youth risky behaviors and mental health, using data from New York City. To estimate a causal estimate of attending small versus large high schools, we use a two-sample-instrumental-variable approach with the distance between student residence and school as the instrument for school enrollment. We consider two types of small schools – "old small schools," which opened prior to a system-wide 2003 reform aimed at increasing educational achievement and "new small schools," which opened in the wake of that reform. We find that girls enrolled in older small schools are less likely to become pregnant, and boys are less likely to be diagnosed with mental health disorders than their counterparts in large schools. Both girls and boys enrolled in more recently opened small schools, however, are more likely to be diagnosed with violence-associated injuries and (for girls only) with mental health disorders. These disparate results suggest that improving a school's organization and inputs together is likely more effective in addressing youth risky behaviors than simply reducing school size.

Key words: small high school, instrumental variable, youth pregnancy, youth violence, mental health

JEL codes: H41, I12, I21, J13

1. Introduction

Risky health behaviors among teenagers, which encompass a wide range of unsafe activities such as teen pregnancy and school violence, and poor mental health among youth have been growing concerns in the United States for decades. Teen pregnancy and fertility, while decreasing, continue to impose substantial social costs (Lindberg, Santelli and Desai, 2016). For example, in 2019, women aged 15–19 had a birth rate of 16.7 per 1,000 persons (CDC, 2021). In 2022, the Supreme Court overturned Roe v. Wade (the constitutional right to an abortion), which may leave pregnant teens in some states facing even more complicated situations than previously. Nearly 20% of high school students reported being bullied at school in the last year, and youth homicides and assault-related injuries caused about $21 billion in medical and work loss (CDC, 2019). Youth risky behaviors affect both short-run safety and learning (Dee and Evans, 2001; Card and Lemieux, 2001) and long-run income and longevity (Farrell and Fuchs, 1982; Bhattacharya and Currie, 2001). There is also an increasing rate of poor mental health among youth (Panchal, Rudowitz and Cox, 2022). In 2019, about 15.7% of teenagers between 12 and 17 (or 3.8 million people) met the diagnostic criteria for major depression (Center for Behavioral Health Statistics and Quality, 2019).

Schools play important roles in regulating teenagers' behaviors. Some studies find attendance at school reduces teenagers' risky behaviors, such as youth crime and adolescent childbirth (Jacob and Lefgren, 2003; Black, Devereux and Salyanes, 2008; Berthelon and Kruger, 2011). The organization of schools, in particular their size and focus, may affect these outcomes. The existing empirical evidence, however, depicts a mixed picture of the effect of school size on risky behaviors and mental health outcomes. Studies that directly examine violence, crime and drug abuse are relatively consistent: large schools are associated with greater numbers of episodes of gang incidents, vandalism, drug usage and other discipline problems (Miller and Chandler, 2003; Leung and Ferris, 2008; Ferris and West, 2004; Cotton, 1996; Nathan and Thao, 2007; Darling-Hammond, Ancess and Ort, 2002). But studies also find that bullying (physical or cyber) is positively correlated with substance use, violent behavior, and unsafe sexual behavior (Litwiller and Brausch, 2013) and the evidence regarding the effect of school size on bullying is quite mixed. While some literature finds that students in small schools are more likely to be bullies (Ma, 2001) and be involved in conflicts (Brezina, Piquero and Mazerolle, 2001), other literature finds no significant or even a positive relationship between school size and bullying (Gottfredson et al., 2005; Wolke et al., 2001; Bowes et al., 2009; Schwartz, Stiefel and Wiswall, 2016).

School organization and inputs other than school size also may play important roles in engagement in risky behaviors and mental health outcomes. The literature finds, overall, that high-quality schools, e.g., charter schools in the Harlem Childrens Zone in New York City or lottery public high schools in Chicago, reduce youth risky behaviors, such as incarceration and teen pregnancy (Fryer and Katz, 2013; Dobbie and Fryer, 2015; Cullen, Jacob and Levitt, 2006). The effects are attributed, at least partially, to more educational investments, including more schooling hours (after-school tutoring and Saturday classes), high-quality teachers (measured by value-added in test scores and teacher turnover rates), data-driven monitoring systems for student progress and so on. A closer look at individual educational inputs indicates that high value-added teachers (Chetty, Friedman and Rockoff, 2014) and closer student-teacher relationships (Rudasill et al., 2010) are related to less risky behavior. Other schooling factors that affect risky behavior include peer interaction (Clark and Lohéac, 2007) and school climate (Astor et al., 2002).

School-based health services, which are typically provided via school-based health centers (SBHC), can be another important channel through which schooling reduces youth risky behaviors and improves mental health outcomes. The literature finds that SBHCs can increase students' health knowledge and access to health services (Kisker and Brown, 1996), and have positive effects on school attendance and educational attainment (Van Cura, 2010; Kerns et al., 2011). However, the impacts on risky behaviors are inconsistent. Some literature finds that the expansion of SBHCs in the United States since the early 1990s' has reduced adolescent fertility (Ricketts and Guernsey, 2006; Lovenheim, Reback and Wedenoja, 2016), while other research finds no or small and statistically insignificant effects of SBHCs on substance use, sexual activity, contraceptive use, pregnancies or births (Kirby, Waszak and Ziegler, 1991; Kisker and Brown, 1996).

An empirical challenge of studying how schooling is associated with youth risky behaviors and mental health outcomes is obtaining the effects of the mediator of most interest, e.g., school size, while disentangling other possible mediators of these outcomes. Small schools are often created with different motivations, designs, and resources than large schools. As a result, schools of different sizes differ not only in total enrollment but also in many other dimensions of school organization and inputs. Availability of services from SBHCs may also differ by school size, depending on, for example, space limitations on site and colocation with other schools.

The implementation of small high school policies in New York City (NYC), which introduced variations in size and organization across a large number of high schools, provides us an opportunity to estimate the effects of multiple mediators of risky behaviors and mental health outcomes separately, including small high school attendance, overall school quality proxied by multiple measures, and SBHC availability. Since about 1960, when Conant (1959) suggested that the size of high schools should be between 400 and 600 students and Barker and Gump (1964) argued that the supposed advantages of large schools are illusory, there has been a movement toward smaller schools. During the past several decades, NYC established numerous small high schools with enrollment less than or equal to 550; these now account for about 70% of public high schools. We distinguish old small schools from "new" small schools (established after 2002 as part of a series of reforms initiated by NYC Public Schools Chancellor Joel Klein), since the two types of small schools had different creation processes and regulatory environments. For example, post-2002, new small schools were established through a competitive process with some applications failed. Almost all the new small schools were supported by non-profit organizations. The new small schools also received some exemptions regarding serving special needs students and union rules for teacher hiring. Principals in new small schools were trained by an organization to embrace accountability and empowerment by schools.[2] Large schools, old, and new small schools also differ substantially in observed characteristics other than size and year of establishment, such as the probability of co-location and access to SBHCs (for details see descriptive statistics in the next section). The existing literature shows that old and new small schools differ in their school environments and have different impacts on student academic outcomes (Kahne, Sporte and Torre, 2006; Schwartz, Stiefel and Wiswall, 2013).[3]

In this paper, we use student-level data from the New York City Department of Education (NYCDOE) administrative and survey databases and Medicaid claim data to examine the effects of small high schools in NYC on teenage pregnancy, violent injury, and mental health diagnoses. We disentangle the effects of school size (total enrollment), educational set-up and the

[2] Details about the difference between the two types of schools can be found in Bloom, Thompson and Unterman (2010), Schwartz, Stiefel and Wiswall (2013, 2016), and Stiefel, Schwartz and Wiswall (2015).
[3] However, Stiefel, Schwartz, and Wiswall (2016) did not find consistent evidence that students attending small schools in NYC perceived better learning environment.

organization of schools (proxied by old versus new small schools), and a health organizational factor (availability of on-campus health services, proxied by the presence of an SBHC), by separately estimating the effects of large, old and new small schools, and the effects of SBHC utilization.

It is challenging to control for individual (self-selection) and neighborhood (geographical sorting) confounding effects when school enrollment is not randomly assigned. To address the possibility of endogenous enrollment in small schools, we use a distance instrumental variable strategy similar to that of Schwartz, Stiefel and Wiswall (2013, 2016). There is a small literature on education evaluations using distance as an instrumental variable, recognizing that distance between home residence and school address plays a significant role in school choice (Cullen, Jacob and Levitt, 2005; Currie and Morretti, 2003; Booker et al., 2011). This distance-based instrumental variable method, unlike that of Bloom, Thompson and Unterman (2010) who use a lottery design to estimate impacts for oversubscribed schools – that is, popular schools – allows us to include a wider range of schools – including "unpopular" schools.

Specifically, our paper starts with reduced-form analyses to investigate how the minimum Euclidean distances between a student's residential address and the nearest schools are associated with the examined outcomes. Then we use the instrumental variable approach to obtain insights into the effects of actual enrollment in schools of different sizes, using the Euclidean distances as instruments to account for endogenous selection into small schools. Because our measures of risky behaviors and mental health diagnoses and small school attendance are in two different datasets, which we cannot link due to data restrictions, we use a two-sample instrumental variable approach (Klevmarken, 1982; Angrist, 1990; Angrist and

Krueger, 1992, 1995; Inoue and Solon, 2010), by applying the distance IV method across the two datasets.

Our results from the reduced-form and IV analyses are consistent with each other. Overall, pooling old and new small schools together, we find few significant effects of attending small high schools on student risky behaviors and mental health, except for an increased likelihood of violent injury for girls attending smaller schools. Nevertheless, when we examine old and new small high schools separately, we find that the marginal girl attending an old small high school is less likely to become pregnant, while the marginal boy attending an old small school has a lower likelihood of mental health diagnoses. By contrast, the marginal student attending a new small high school has an increased likelihood of violent injury (for both boys and girls), and of a diagnosis of a mental health disorder (for girls). This heterogeneity in effects between new and old small schools implies that simply reducing school size will not completely address the problems of youth risky behaviors and poor mental health. Other characteristics of the school environment, such as teacher quality, student expenses, and teacher-student ratio, are likely more important. Finally, differential access to SBHCs does not fully explain our findings. Old small schools are more (less) likely to be served by SBHCs than large schools (new small schools), but their students have the lowest utilization of SBHCs among all three types of schools.

Our estimated impacts of small school enrollment on youth risky behaviors and mental health through multiple channels adds to the existing literature by shedding light on the potential mechanisms linking small schools and student academic advantages. The literature finds that small schools increase graduation rates (Schwartz, Stiefel and Wiswall, 2013, 2015; Bloom, Thompson and Unterman, 2010), student achievement (Howley, Strange and Bickel, 2000; McMillen, 2004), and long-term success in the educational system and the labor market

(Humlum and Smith, 2015). However, it is not clear what differences between small and large schools drive these observed impacts. By examining the differences in student non-academic outcomes between small and large schools, our study provides possible explanations for the evidence in the literature on small school enrollment and academic achievements and indicates that small schools may have long-term impacts if the findings on risky behaviors and youth mental health persist. Knowing the roles of key features of small schools may help policymakers better understand how to make small schools work more effectively in promoting student development.

2. Institutional Background and Data Description

Small school reform dates to the late 1960s, beginning with targeted reform for students who were not successful in large comprehensive high schools. During the 1990s to 2000s, small school reform had become of more general interest. (See Iatarola and Stiefel, 2018 for a review of the history of small school reform.) In 2002, Michael Bloomberg, the newly elected mayor of NYC who was granted control over the schools by the New York State Legislature in 2002, hired Joel Klein as Chancellor, expecting him to begin reforms that would improve the performance of public school students. One of Klein's major reforms was to establish new small high schools to replace large dysfunctional ones. Bloom, Thompson and Unterman (2010) report that by 2010 more than 20 underperforming public high schools had been closed and more than 200 new secondary schools had been opened (starting from the 2003-2004 academic year). At the same time that small high schools were created, all NYC public school 8th graders began to apply to high schools through the High School Application Processing System (a new centralized single-offer system based on a deferred acceptance algorithm), which assigns students to high schools according to both students’ preferences and schools’ selection criteria (Abdulkadiroglu, Pathak and Roth, 2005; Stiefel, Schwartz, and Wiswall, 2015).[4]

We should note that there is no consensus on the definition of small schools. In this paper, we define small schools by the total enrollment in grades 9 to 12 when the student enters grade 9, and set the cutoff at 550, which is the standard adopted in previous studies of small high schools in NYC (Bloom, Thompson and Unterman, 2010; Schwartz, Stiefel and Wiswall, 2013, 2016). All in all, using our classification there were 107 large schools, 86 old small schools (established before 2003), and 187 new small schools in our data in 2009 and 2010.[5]

Despite having small enrollment sizes in common, old and new small schools have different emphases. Old small schools were established as alternatives for students who did not do well in traditional high schools (Stiefel et al, 2000; Quint et al., 2010). With limited resources, those old small schools needed to choose between offering a more nurturing schooling environment and academic performance. New small schools, as part of Bloomberg's high school reform and following accountability efforts with No Child Left Behind, focused on student achievement (Bloom et al., 2010). In terms of observed school characteristics, according to Schwartz, Stiefel and Wiswall (2016), who use the same samples of schools, both types of small schools receive more funding and have lower pupil-teacher ratios than large schools. But they also have lower percentages of teachers with master's degrees, higher percentages of inexperienced teachers and higher teacher turnover rates. Schwartz, Stiefel and Wiswall (2016) find no significant differences in the learning environment between large and small schools.

[4] The High School Application Processing System, which was introduced in NYC in the 2003-04 academic year, requires all 8th graders in public schools in NYC to submit a list of up to 12 high schools in order of preference. Then the NYCDOE will use a computerized process to assign students to their highest-ranked high schools with available space whose admission criteria have been met. See Schwartz, Stiefel and Wiswall (2013, 2016) and Stiefel, Schwartz and Wiswall (2015) for details about the high school application processing system in NYC.

[5] We exclude Staten Island, where there are three small high schools. Generally, students do not travel to or from Staten Island to attend a small school.

Overall, old small schools have advantages over new small schools in terms of resources and organizational factors: they have more funding, a higher percentage of teachers with master's degrees, a higher percentage of experienced teachers (though still lower than large schools), and a higher percentage of teachers with valid certifications (Schwartz, Stiefel and Wiswall, 2013). For instance, in 2009 and 2010, on average, per student total expenditures in old small schools were $1,396 higher than in large schools while expenditures in new small schools were only $511 higher. Moreover, old small schools are less likely to be co-located with large schools. More than two-thirds of new small schools are co-located with another school.[6] The differences in resources and organizational factors between old and new small schools are detailed in Table A1 in the appendix.

In NYC, SBHCs are like medical health centers on campus, providing primary care and preventive health services to students. In high schools, SBHCs also provide access to reproductive health services, including health education and counseling for the prevention of risk-taking behaviors such as drug abuse, pregnancy, and sexually transmitted infections; counseling and access to contraception and testing for pregnancy; and referral and follow-up for reproductive health conditions requiring further evaluation and treatment. As of October 2016, there were 147 SBHCs serving 347 schools in NYC. Small high schools, especially new ones, are more likely to have SBHC services, probably due to co-location with large schools. Given all the differences between old and new small schools, in our analyses, we will first evaluate all small schools as a whole, and then estimate the effects of old and new small schools separately, consistent with prior research.

[6] If the potential benefits of a small school are due to smaller school size that generates positive peer effects and desirable schooling environment both inside and outside classrooms, co-location with other schools, which increases actual school size, may offset those benefits, particularly the ones outside classrooms.

The background and enrollment information summarized in Table 1 come from NYCDOE administrative datasets (student-level) and public school-level data; the outcomes of risky behaviors and related diagnoses are from Medicaid claims data from the New York State Department of Health[7]. To make the two sources of data sets comparable, we restrict the sample in the first dataset (education dataset) to students who were certified eligible for reduced or free lunch.[8] In the education dataset, we include students who entered 9th grade for the first time in 2009 or 2010 (with the majority born in 1994 and 1995, respectively), and attended NYC public schools in 8th grade. In the second dataset (health dataset), we include students who were born in 1994 or 1995, lived in NYC in January 2009 or 2010, when they are supposed to be in grade 8, and were continuously enrolled in Medicaid in the four-year (46-month) period we examine. For girls who were born in 1994 (1995), we also exclude those who had pregnancy diagnoses between January and August 2009 (2010) (1,487 in total), to focus on new pregnancies after high school enrollment. Later, we will show that our results are not sensitive to our restrictions regarding continuous Medicaid enrollment and pregnancies in the first half of the year. We also collect information on census tracts from the ACS census data.

Descriptive statistics for students and census tracts, by gender and school size in the education dataset, can be found in Table 1. We first focus on two indicators of risky behaviors: pregnancy and violence-related injury, and an indicator of mental health diagnoses.[9] All main

[7] New York State Medicaid claims data include personal identifiers, address history, Medicaid enrollment information, and codes for the types of service received, diagnoses, procedures, and service dates. We used both fee-for-service and managed care encounter claims data, which were comparable in quality (source: Mathematica. Medicaid Managed Care and Integrated Delivery Systems: Technical Assistance to States and Strengthening Federal Oversight.; 2013.)

[8] From 2009 to 2014, the national income eligibility of guidelines for the reduced- and free-lunch program were 185% and 130% of the federal poverty line, respectively. Meanwhile, the NYC income eligibility of guidelines for Medicaid for children under 19 are 100-154% of the federal poverty line.

[9] We also examined substance use disorders. However, our results were implausibly large, which may have to do with discretion in identification of these disorders. Therefore, we excluded substance use disorders from the list of outcomes.

outcomes are binary variables, which equal one if the student was diagnosed with the corresponding outcome at least once during the four years when he/she was supposed to be in high school (September 1, 2009-June 30, 2013 for students born in 1994 and September 1, 2010-June 30, 2014 for the students born in 1995). We also explore several additional outcomes such as contraception and SBHC visits to interpret our findings.[10] Boys are more likely to have violence-related injuries. Girls are more likely to have health problems associated with being overweight and to visit SBHCs. There are apparent gender disparities in most examined mental health diagnoses, although the overall probabilities of being diagnosed with at least one of the examined mental health diagnoses are similar. The codes on claims we use to identify the outcomes are listed in Table A2 in the appendix.

Overall, boys and girls have similar student characteristics. About 70% of high schools are small (see Table A3 in the appendix for school characteristics), with more than one-third of the total 53,976 students enrolled in these schools. Students of both genders in small schools are disproportionately black and Hispanic, and live outside Queens, in a census tract with smaller than average percentages of Asian and white residents, residents with college degrees or higher, or who were foreign-born. There are also substantial differences in the characteristics of students attending new and old small schools. Asian students and students living in Manhattan or a census tract with a larger percentage of Asian and white residents were more likely to attend old small schools, while black students and students living in Brooklyn or a census tract with smaller percentages of Asian and white residents were more likely to attend new small schools.

Table 1: Descriptive Statistics of Students by Gender and School Size

	Boys			Girls		
	Large	**Small**	**Small**	**Large**	**Small**	**Small**

[10] We also use overweight/obesity as outcomes for a validity check. However, as we will note later, we have inconsistent definitions of overweight/obesity across datasets. In the education data we define overweight/obesity by body mass index, while in the health data what we have are overweight diagnoses.

			New	Old			New	Old
	Main Outcomes							
Pregnancy	-				0.378			
Violent injury	0.037				0.022			
Mental health diagnosis	0.301				0.306			
	Secondary Outcomes							
Contraception	-				0.301			
Most effective (sterilization or long-acting reversible)	-				0.035			
Moderately effective (injectable, oral pill, patch, ring, or diaphragm)	-				0.286			
Overweight	0.164				0.206			
SBHC visit	0.087				0.122			
Number of SBHC claims	0.394 (2.369)				1.074 (4.867)			
	Student characteristics							
Distance to large high school	1.227 (0.982)	1.263 (1.066)	1.302 (1.120)	1.177 (0.932)	1.220 (0.970)	1.243 (1.044)	1.298 (1.116)	1.137 (0.881)
Distance to new small high school	1.119 (0.830)	0.725 (0.591)	0.711 (0.573)	0.757 (0.628)	1.088 (0.817)	0.732 (0.592)	0.734 (0.596)	0.729 (0.585)
Distance to old small high school	1.845 (1.221)	1.395 (1.255)	1.511 (1.320)	1.139 (1.055)	1.828 (1.218)	1.386 (1.231)	1.521 (1.316)	1.128 (1.002)
Asian	0.266	0.088	0.071	0.126	0.241	0.083	0.065	0.116
Black	0.261	0.386	0.407	0.339	0.286	0.417	0.428	0.395
Hispanic	0.372	0.480	0.484	0.471	0.383	0.459	0.474	0.431
Age	15.39 (0.552)	15.47 (0.630)	15.47 (0.633)	15.46 (0.623)	15.33 (0.500)	15.40 (0.582)	15.42 (0.594)	15.37 (0.557)
Manhattan	0.081	0.154	0.115	0.240	0.078	0.159	0.115	0.243
Brooklyn	0.379	0.336	0.395	0.207	0.389	0.345	0.394	0.251
Queens	0.404	0.168	0.158	0.190	0.382	0.178	0.178	0.179
Cohort 2009	0.463	0.382	0.372	0.403	0.458	0.383	0.364	0.419
	Census Tract Characteristics							
Population (ln)	8.332 (0.511)	8.424 (0.510)	8.406 (0.500)	8.464 (0.527)	8.340 (0.503)	8.424 (0.512)	8.407 (0.510)	8.457 (0.515)
Percent Asian	0.162 (0.186)	0.074 (0.134)	0.065 (0.119)	0.096 (0.162)	0.154 (0.185)	0.074 (0.133)	0.066 (0.120)	0.088 (0.153)
Percent White	0.226 (0.251)	0.121 (0.189)	0.113 (0.179)	0.141 (0.208)	0.212 (0.245)	0.123 (0.192)	0.115 (0.181)	0.138 (0.209)
Percent Hispanic	0.321 (0.245)	0.408 (0.262)	0.405 (0.263)	0.415 (0.260)	0.327 (0.249)	0.401 (0.260)	0.405 (0.262)	0.394 (0.257)
Percent age 65 and up	0.115 (0.053)	0.102 (0.047)	0.100 (0.047)	0.105 (0.049)	0.114 (0.053)	0.102 (0.047)	0.101 (0.047)	0.104 (0.047)
Percent with high school diploma	0.364 (0.111)	0.349 (0.117)	0.357 (0.117)	0.331 (0.116)	0.365 (0.112)	0.350 (0.117)	0.356 (0.116)	0.339 (0.120)
Percent with college degree or higher	0.229 (0.129)	0.194 (0.136)	0.185 (0.120)	0.214 (0.163)	0.224 (0.129)	0.194 (0.137)	0.184 (0.121)	0.214 (0.161)
Percent foreign born	0.438 (0.158)	0.367 (0.144)	0.368 (0.141)	0.365 (0.150)	0.432 (0.159)	0.364 (0.146)	0.366 (0.145)	0.361 (0.150)
Employment rate	0.635 (0.087)	0.601 (0.098)	0.598 (0.095)	0.606 (0.104)	0.631 (0.088)	0.599 (0.101)	0.597 (0.098)	0.603 (0.104)
Average household income (ln)	11.00 (0.366)	10.84 (0.391)	10.83 (0.369)	10.87 (0.433)	10.98 (0.351)	10.84 (0.420)	10.83 (0.415)	10.87 (0.427)
Sample size	17,102	8,470	5,818	2,652	18,245	10,159	6,662	3,497

Note: Standard deviations for the continuous variables are shown in parentheses. The omitted baseline category of race is white, and the omitted baseline borough is the Bronx. Age is the age on September 1st, 2009 (2010) for the

2009 (2010) cohort. We are not able to summarize outcome variables by school size because we do not have school enrollment information in the health dataset.

3. Methodology

We start with a reduced-form analysis as the following:

$$y_i = \tau_0 + \tau_1 DS_i + \tau_2 DL_i + X_i^s \tau_3 + X_i^b \tau_4 + X_i^c \tau_5 + \epsilon_i, \quad (1)$$

where y_i is the risky behavior outcome of student i, measured during the study period. DS_i and DL_i are the Euclidean distance from the nearest small school and large school to student home address, respectively. X_i^s is a vector of student characteristics, including the cohort effect. X_i^b is a vector of borough (NYC county) fixed effects. X_i^c is a vector of census tract characteristics. A full list of X_i can be found in Table 1 in the appendix.[11] ϵ_i is the error term.

The estimates of τ_1 and τ_2 shed light on the relationship between risky behaviors and distances to the nearest schools. We further adopt an instrumental variable strategy to obtain insights into the effects of actual small school enrollment, using the distances to the nearest schools as the instruments to solve the potential school selection issue, i.e., students self-select into small schools based on unobserved characteristics. Specifically, we estimate the following model by two-stage-least-square estimation (2SLS):

$$\begin{cases} Small_i = \beta_0 + \beta_1 DS_i + \beta_2 DL_i + X_i^s \beta_3 + X_i^b \beta_4 + X_i^c \beta_5 + e_i \\ y_{it} = \alpha_0 + \alpha_1 Small_i + X_i^s \alpha_2 + X_i^b \alpha_3 + X_i^c \alpha_4 + \varepsilon_{it} \end{cases}, \quad (2)$$

where $Small_i$ is the indicator of small school enrollment.

Existing educational literature has shown that distance is an important predictor of school choice (Schneider and Buckley, 2002; Burgess and Briggs, 2010; Hastings, Kane and Staiger, 2006) and it has been used as an instrument to address selection bias in multiple settings (Cullen,

[11] To save notations, we use lowercase rather than uppercase letter to denote the coefficient vector of X_i. This rule applies throughout the entire paper.

Jacob and Levitt, 2005; Barrow, Schanzenbach and Claessens, 2015; Booker et al., 2011). In the current NYC setting, after controlling for a set of background variables, especially those at the census tract level that proxy geographic sorting of students, the distances from the nearest schools are likely to be valid instruments for small school enrollment. There are several reasons that school location is likely to be exogenous across neighborhoods for risky behavior (and health) determinants. First, more than 80 percent of the small schools were co-located with large schools in buildings that had existed for a long period. Small schools were placed in these buildings mainly because of the historical circumstances of those co-located large schools, which are likely exogenous for risky behaviors of current students. Second, small schools were also located because of space considerations, as well as other administrative concerns. Moreover, the high school application-assignment system in NYC, which is based on student preferences for schools citywide, makes it unlikely that residents base residential location primarily on school options. Students can move, but mobility is not high, with fewer than 10% of students changing their location-based zip codes between 7th and 8th grade. Finally, our health data include only low-income students who were on Medicaid. This group is relatively homogeneous in terms of socio-economic status. As a result, after controlling for a rich set of characteristics at both individual and census tract levels, as well as cohort effects and borough effects, the possibility of student systematic sorting around a certain type of school is unlikely.[12] We describe the results of sensitivity tests on the validity of our assumptions later in the paper.[13]

[12] Ideally, we could include census tract fixed effects. However, we would not have enough variation in school enrollment within census tract to estimate the first stage model. Our reduced-form results with census tract fixed effects (available upon request), especially the correlation between distance to old small schools and teen pregnancy, are similar with the results with census tract characteristics.

[13] More details on instrument validity can be found in Schwartz, Stiefel and Wiswall (2013, 2016) that estimate the effects of small schools on educational outcomes using similar (one-sample) instrumental variable approaches.

Since the information on small high school enrollment ($Small_i$) and risky behavior outcomes (y_{it}) are from two data sets (except for overweight/obesity), we use a two-sample, two-stage-least-square (TS2SLS) estimation (Inoue and Solon, 2010; Pacini, Windmeijer, 2016), which is related to two-sample IV estimation (Angrist and Krueger, 1992) but is more efficient (Inoue and Solon, 2010). Specifically, we first run the same first-stage estimation as equation (2) (the first equation) and obtain the predicted values $\widehat{Small}_i$, as well as the estimated asymptotic variances of all coefficients $\hat{V}_1$. Then we estimate

$$y_{it} = \theta_0 + \theta_1 \widehat{Small}_i + X_i^s \theta_2 + X_i^b \theta_3 + X_i^c \theta_4 + \varepsilon_{it} \qquad (3)$$

and obtain the TS2SLS estimator of the effects of small school enrollments as $\hat{\theta}_1$. Finally, we calculate the (robust) standard errors, taking into account information from the two samples, following Pacini and Windmeijer (2016). Specifically, we also estimate the regression of distance instrumental variables (and exogenous covariates) on the predicted school enrollment (and background controls):

$$\begin{cases} DS_i = \rho_0 + \rho_1 \widehat{Small}_i + X_i^s \rho_2 + X_i^b \rho_3 + X_i^c \rho_4 + e_i \\ DL_i = \pi_0 + \pi_1 \widehat{Small}_i + X_i^s \pi_2 + X_i^b \pi_3 + X_i^c \pi_4 + u_i \end{cases}. \qquad (4)$$

We then obtain the estimated asymptotic variances of all coefficients from equation (1), the first stage of (2), and equation (4) as $\hat{V}_1$, $\hat{V}_2$ and $\hat{V}_c$, respectively. Then the estimated asymptotic variances of all TS2SLS estimators ($\hat{\theta} = (\hat{\theta}_0, \hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4)$) are given by

$$\hat{V}(\hat{\theta}) = \hat{V}_c \hat{V}_1 \hat{V}_c{}' + \hat{\theta}' \otimes \hat{V}_c \hat{V}_2 (\hat{\theta} \otimes \hat{V}_c{}'),$$

where $\otimes$ denotes Kronecker product. $\hat{V}(\hat{\theta})$ is robust to heteroskedasticity when $\hat{V}_1$ and $\hat{V}_2$ are both robust. The standard error of the coefficient of the j-th regressor is the square root of the corresponding j-th elements on the diagonal of $\hat{V}(\hat{\theta})$. We estimate all regressions using linear probability models.

The model above does not distinguish old and new small schools in the baseline model, which provides the pooled effects of small school enrollment. Later we will separate old and small schools using a modified model that is similar to the baseline model but has two indicators for small school enrollment (old and new) and three instrumental variables (distances from the nearest old small school, new small school, and large school).

Like all instrumental variables analysis results, our TS2SLS findings are local average treatment effects, identified from students whose decision to attend a school of a particular type was driven by geographic proximity. From a policy perspective, our results can provide evidence of the effect of opening new schools of a particular type on students who live close to such new schools.

4. Empirical Results

4.1 Does School Size Matter?

We first estimate the associations between distances to the nearest small and large schools from students' 8th-grade residence addresses and their risky behaviors and mental health via the reduced-form model, for boys and girls separately. Table 2, panel (A) presents the results from the baseline pooled models. Overall, we find that distance rarely matters to risky behavior or mental health, with a few weak exceptions: Minimum distance to larger schools is negatively associated with the likelihood of mental health diagnosis for boys, and the minimum distance to small schools is negatively associated with violence-related injury for girls. Table 2, panel (B) presents the results from the separate models. The effects of the minimum Euclidean distance from the nearest high schools on the likelihood of risky behaviors and mental health diagnoses differ substantially between new and old schools. While for new schools the distance is only negatively correlated with the likelihood of violent injury for girls (-0.4 percentage points per mile), for old schools the distance is positively correlated with almost all outcomes except for violent injury for girls, ranging from 0.3 percentage points per mile for violent injury for boys to 1.4 percentage points per mile for pregnancy. Full details about the reduced-form estimation results are presented in Table A4 and A5 in the appendix, for the pooled and separate models respectively.

Table 2: Reduced-Form Estimation, Baseline (Pooled) and Separate Models: Likelihood of Risky Behavior Engagement and Mental Health Diagnosis

	Boys		**Girls**		
	Violent injury	**Mental Health Diagnosis**	**Pregnancy**	**Violent injury**	**Mental Health Diagnosis**
A: Baseline (Pooled) Model					
Distance to large high school	-0.001 (0.003)	-0.013* (0.007)	-0.002 (0.008)	-0.001 (0.003)	-0.006 (0.007)
Distance to small high school	-0.001 (0.002)	0.008 (0.005)	0.004 (0.006)	-0.003* (0.002)	-0.005 (0.006)
***p*-value for F test (both distances = 0)**	0.957	0.033	0.593	0.473	0.896
B: Separate Model					
Distance to large high school	-0.002 (0.003)	-0.017** (0.007)	-0.010 (0.008)	-0.001 (0.002)	-0.011 (0.008)
Distance to new small high school	-0.003 (0.002)	0.005 (0.005)	-0.003 (0.006)	-0.004*** (0.002)	-0.009 (0.006)
Distance to old small high school	0.003** (0.001)	0.007** (0.003)	0.014*** (0.003)	0.001 (0.001)	0.009*** (0.003)
***p*-value for F test (all distances = 0)**	0.136	0.015	0.000	0.023	0.024
Sample size	19,148		16,585		

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively.

We then estimate the effects of actual school enrollment on risky behaviors and mental health using the TS2SLS model. The main first-stage results are presented in Table 3 (Column (A) for the pooled model and Column (B) for the separate model), showing the effects of minimum Euclidean distance between the nearest school address and students' residence address on the likelihood of small school enrollment. As shown, the results are largely consistent with previous literature studying small schools: a student is less likely to attend a small high school if she/he lives farther from a small school. For example, the probability of attending a small school of any type as distance increases by one mile decreases by 5.2 percentage points for both boys and girls. The effects of distance on enrollment differ by the type of small school. Living one more mile away decreases the probability of enrollment by 4.3–4.8 percentage points for new small schools but only by 1.6–1.9 percentage points for old small schools. The multivariate F-statistic test and the Sanderson-Windmeijer multivariate F test show that the instruments are jointly strong for the first stage as a whole and each endogenous variable separately.[14] Full details of the first-stage estimation results are presented in Table A6 in the appendix.

Table 3: First-Stage Estimation, Baseline (Pooled) and Separate Models: Likelihood of Small High School Enrollment

	Baseline (Pooled) Model (A)		**Separate Model (B)**			
	Boys	**Girls**	**Boys**		**Girls**	
High School Enrollment	**Small**	**Small**	**New Small**	**Old Small**	**New Small**	**Old Small**
Distance						
Distance to large high school	0.010*** (0.003)	0.009*** (0.003)	0.009*** (0.003)	0.001 (0.002)	0.014*** (0.003)	-0.001 (0.002)
Distance to small high school	-0.052*** (0.004)	-0.052*** (0.004)	-	-	-	-
Distance to new small high school	-	-	-0.048*** (0.004)	-0.005** (0.003)	-0.043*** (0.004)	-0.008*** (0.003)
Distance to old small high school	-	-	0.018*** (0.003)	-0.016*** (0.002)	0.011*** (0.003)	-0.019*** (0.002)
Multivariate F of excluded instruments [*p*-value]	77.04 [0.000]	77.05 [0.000]	63.74 [0.000]	28.74 [0.000]	47.80 [0.000]	50.93 [0.000]
Sanderson-Windmeijer multivariate F test of excluded instruments [*p*-value]	77.04 [0.000]	77.05 [0.000]	99.50 [0.000]	41.37 [0.000]	75.50 [0.000]	76.84 [0.000]
Sample size	25,572	28,404	25,572		28,404	

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively.

Table 4 shows TS2SLS estimates of the effects of attending small schools on the likelihood of engaging in risky behaviors and mental health diagnoses, for the pooled models in panel (A)

[14] In our first stage, the F-statistics are above the recommended rule-of-thumb threshold of 10 (Staiger and Stock, 1997; Stock, Wright and Yogo, 2002), indicating that overall the instruments are strong. It is suggested that the exclusive first-stage F may not be appropriate when there are multiple endogenous variables (Angrist and Pischke, 2008): The F statistic can be large when only a part of instrumental variables is strong, in which case the model may still be weakly identified. To account for that possibility in the separate models that distinguish old and small schools, we also conduct the modified first stage F test - the Sanderson-Windmeijer multivariate F test (Angrist and Pischke, 2008; Sanderson ad Windmeijer, 2016) and get similar conclusions. We should note that for the baseline model the multivariate F-statistic test or the Sanderson-Windmeijer multivariate F test are equivalent because there is only one endogenous variable (small school enrollment).

and the separate models in panel (B). Overall, the TS2SLS results are consistent with the reduced-form results and provide more insights into the effects of actual school enrollment on the examined risky behaviors and mental health. For example, from the pooled model the only significant effect is that small school enrollment increases the likelihood of violence-related injury for girls (5.8 percentage points). When estimating the effects of attending a new and old small school separately for boys, we find that new small school enrollment weakly increases the likelihood of violence-related injuries (7.6 percentage points), but old small school enrollment significantly decreases the likelihood of a mental health diagnosis (41.5 percentage points, respectively). For girls, old small school enrollment decreases the likelihood of being pregnant at least once during the 4 years of high school (55.3 percentage points), whereas attending new small schools increases the probability of violence-related injuries (9.4 percentage points) and mental health diagnoses (22.0 percentage points). We note that the standard errors on these estimates are quite large and encompass a range of outcomes – importantly, however, for most outcomes, the effects of attending old and new small schools are substantially different (all p-values < 0.1 at least except for violent injury for girls). Full details about the TS2SLS estimation results for the pooled and separate models are presented in Table A7 and A8 in the appendix, respectively.

Table 4: TS2SLS Estimations, Baseline (Pooled) and Separate Models: Likelihood of Risky Behavior Engagement and Mental Health Diagnosis

	Boys		Girls		
	Violent injury	**Mental Health Diagnosis**	**Pregnancy**	**Violent injury**	**Mental Health Diagnosis**
A: Baseline (Pooled) Model					
Small school	0.015 (0.040)	-0.136 (0.105)	-0.071 (0.117)	0.058** (0.029)	0.106 (0.111)
B: Separate Model					
New small school	0.076* (0.042)	-0.063 (0.109)	0.150 (0.142)	0.094** (0.038)	0.220* (0.134)
Old small school	-0.065 (0.074)	-0.415*** (0.184)	-0.553*** (0.159)	0.003 (0.050)	-0.200 (0.148)
***p*-value for F test**	0.084	0.065	0.001	0.185	0.029

(new small=old small)					
Sample size	19,148		16,585		

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively.

The null effects on most outcomes in the pooled model indicate that school size alone does not matter much for teenage risky behaviors or mental health. Once we distinguish between new and old small schools, which have different set-ups and organizations, we do find that students in old small schools are less likely to report risky behaviors and mental health problems. The different organizational histories of these two types of schools may explain the disparities in findings. Old small schools, which were created as alternatives to traditional high schools and aim to promote a nurturing environment, may be good for marginal students at risk. New small schools, which focus on student achievement, may not be that beneficial for students at the margin of engaging in risky behaviors, who are likely to be at the lower end of the distribution of academic performance.[15]

4.2 Conception Use and School-Based Health Center as Explanations of Risky Behaviors

It is possible that the pregnancy results we observe occur because students in old small schools receive more contraceptive treatments (possibly from SBHCs) but do not have less sexual intercourse. Table 5 reports the reduced-form and TS2SLS effects of attending new and old small schools on contraception (as well as SBHC utilization) for girls. The results confirm that the effect mainly comes from new small schools: while students in new small schools are more likely to use contraceptive treatments, which are moderately effective methods such as oral pills, the effect on contraception utilization for students in old small schools is negative and

[15] New small schools are also much more likely to be co-located with other schools (Schwartz, Stiefel and Wiswall, 2013, 2016). More than two-thirds of new small schools co-locate with another school. One may argue that co-location with other schools, which may serve to increase size of the school community, may offset the potential benefits of a small school that are due to smaller school size that generates positive peer effects and desirable schooling environment both inside and outside classrooms. However, the disparity in the estimated effects between new and old small schools is unlikely to be driven by such a co-location effect because with distances to the nearest schools as instruments our estimated effects are mostly identified by schools that do not co-locate with others.

insignificant. There is no significant difference in the utilization of the most effective methods, e.g., long-acting reversible contraceptives (LARC), between students in large and small schools. The results confirm that more female contraception is not the reason that we observe a lower likelihood of teen pregnancy in old small schools.[16]

Table 5: Reduced-Form and TS2SLS Results of the Effects of Attending Small Schools on Contraception and SBHC Utilization

	Contraception			SBHC	
	Any	**Most effective[1]**	**Moderately effective[2]**	**Use at least once**	**Total claims**
A: Reduced form					
Distance to large high school	0.000 (0.007)	-0.001 (0.003)	-0.001 (0.007)	0.010** (0.005)	0.180** (0.083)
Distance to new small high school	-0.012** (0.005)	-0.001 (0.002)	-0.012** (0.005)	-0.004 (0.003)	-0.006 (0.040)
Distance to old small high school	0.008*** (0.003)	0.001 (0.001)	0.009*** (0.003)	0.002 (0.002)	0.049** (0.023)
***p*-value for F test (all distances = 0)**	0.020	0.740	0.017	0.085	0.006
B: TS2SLS					
New Small	0.316** (0.128)	0.027 (0.042)	0.310** (0.127)	0.136* (0.072)	1.039 (0.965)
Old Small	-0.206 (0.145)	-0.023 (0.049)	-0.219 (0.144)	-0.107 (0.085)	-3.266*** (1.118)
***p*-value for F test (new=old)**	0.006	0.428	0.005	0.029	0.003

Note: The sample size is 16,585. Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively. [1] Sterilization and long-acting reversible methods. [2] Injectables, oral pills, patch, ring, and diaphragm.

Table 1 shows that students in small schools, particularly those in new small schools, are more likely to be served by school-based health centers. As a result, girls in small schools may have convenient access to health care education and services. We use two measures as proxies for access to SBHCs: whether a girl visits an SBHC at least once during the four years in high school and her total number of claims from SBHCs during the same period. Table 5 (the last two columns) shows the effects of attending small schools on those two SBHC utilization outcomes.

[16] However, we should note that more male contraception may still contribute to the observed lower likelihood of teen pregnancy.

Girls in new small schools use more services via SBHCs. Compared with girls in large schools, students in new small schools are more likely to visit SBHCs, but girls in old small schools have fewer SBHC visits. The disparity of teen pregnancy in small schools cannot be explained by variation in the receipt of reproductive health services via SBHCs.[17] Full details about the reduced-form and TS2SLS estimation results on contraception and SBHC utilization for the separate models are presented in Table A9 and A10 in the appendix, respectively.

5. Validity of Distances to Schools as Instrumental Variables

The Euclidean distances that we use as instruments are based on a broad literature on school choice that adopts similar instruments. One threat to the validity of the instruments is that the location of small schools is correlated with the underlying risk of local teenagers becoming pregnant or being involved in violence. For example, small schools might locate or open in neighborhoods with dysfunctional schools, poor student academic performance, and/or a high prevalence of teenage risky behaviors. In an earlier paper that also examines small schools in NYC, Schwartz, Stiefel and Wiswall (2013) argue that school location might not be an issue because many small schools (about 60%), especially new ones, share buildings with other large schools that had existed for a long time. As a result, the location of small schools is likely to be correlated with long-term historical neighborhood characteristics that are shared by large schools, rather than current unobserved student characteristics. They perform a falsification test and find that the distance variables are not correlated with the performance of earlier cohorts that attended school before the new small high schools were opened, suggesting that the new schools are not placed strategically.

[17] Boys are more likely to visit SBHC once and report more claims in new but not old small schools, compared with those in large schools.

In this paper we do not have risky behavior data for the period before 2003 when the new small schools opened, thus we are not able to perform that same falsification test. However, since academic performance is the main goal of most school systems; if the new schools are not placed strategically for academic performance, it is not likely that they are so placed for risky behaviors. Moreover, we have controlled for observed neighborhood socioeconomic characteristics at the borough level and at the census tract level, including demographic components (ethnicity and age), local economy (average household income and employment rate), education (percentage of high school and college degrees), and so on. Since in NYC there are more than 2,000 census tracts, but only 380 high schools, controlling for census tract characteristics should help us eliminate potential biases due to school locations.

In addition, we check whether our results are robust to omitted variable bias that is associated with school locations, following the idea that the results are unlikely to be driven by omitted variable bias if the estimated coefficients are stable after the inclusion of observed controls, especially those controls that are supposed to be important in explaining outcomes (Chiappori, Oreffice and Quintana-Domeque, 2012; Oster, 2017). For each gender, we estimate four reduced-form separate models with different subsets of controls (see Table A11 in the appendix for all results). The comparison between these results indicates that controlling for borough fixed effects and census tract characteristics significantly reduces potential omitted variable bias due to student sorting around certain types of schools. With the current set of controls, the estimated results are not sensitive to the inclusion of census tract average household income, which is an important indicator of the overall socioeconomic status of the neighborhood. The stable estimates after the inclusion of census tract average household income suggest that the main model accounts for most of the selection bias associated with school locations.

Student mobility is another related potential threat to the validity of the distance instruments. If a savvy student moved closer to a small (or large) school strategically, in hopes of gaining from a particular type of school, there might be correlations between distance to schools and unobserved student characteristics. Again, Schwartz, Stiefel and Wiswall (2013) argue that this situation is not likely, as the high school admission system in NYC does not create a strong incentive for such strategic sorting. They find that only a small percent of students changed zip codes between 7th and 8th grade, and these moves were uncorrelated to school locations. In our main analysis, we use students' home addresses in January when they are 8th graders. We check whether student mobility matters in this study by checking if our results are robust to student residencies at alternative times: students who were born in 1994 (1995) and lived in NYC in September 2008 (2009), the beginning of grade 8, or who lived in NYC in September 2010 (2011), the beginning of grade 9. The reduced-form results from separate models, which are similar to our main results, are presented in Appendix Table A12.[18]

6. Validity of Two-Sample IV Approach

The TS2SLS requires that the samples used in the first and second stages represent the same population. In this paper, the students in the two samples are from approximately the same cohorts in NYC. We also restrict the first sample to students who were eligible for the free and reduced lunch program to match the second sample on Medicaid (during the study period the income eligibility was 130-185% and 100-154% of the federal poverty line, respectively).

---

[18] Some families might make residential location decisions many years prior to student entry into grade 8, e.g., for school preferences for older siblings. However, that seems not matter in our case of cohort 2009 and 2010 in NYC, according to the two falsification tests conducted in Schwartz, Stiefel and Wiswall (2013) that examine cohort 2001 and 2002, and cohort 2007 and 2008, respectively. Moreover, given the cost of moving, including breaking family and social ties, such strategic moving in the past seems unlikely for low-income families studied here.

Hence, the two samples used in this study should be quite similar with substantial overlap.[19] For most of the observed characteristics of students and census tracts, the two samples are descriptively similar (Table A13 in the appendix), indicating that the common population assumption is likely to hold.

Nevertheless, our samples for the two stages are still not perfectly consistent with each other. Since there is no school enrollment information in the Medicaid data, for the second stage we choose the teenagers born in 1994 or 1995 who were the majority of 8th graders in the 2009 or 2010 academic year. To check whether our results are sensitive to such inconsistency, we impose the same birth year restriction in the first stage to restrict the education sample to those who made normal progress. The results using this restricted sample to estimate the first stage are largely in line with our main results (Appendix Table A14).

Another potential threat to the common population assumption is dropouts. Our estimated effects may incorporate the effects of attending small schools on dropouts. It has been shown that for students who are eligible for free lunch, enrollment in new small schools increases the probability of high school graduation in four years, but enrollment in old small schools decreases this probability (Schwartz, Stiefel and Wiswall, 2013). Students in NYC can drop out when they turn 17. To exclude the potential effects on dropouts for a clear interpretation of the effects of attending small schools, we impose an additional timing restriction on the outcomes: we define outcomes as whether a student has a claim related to risky behaviors or mental health by the student's 17-year birthday.[20] The results are summarized in Table A13 in the appendix. While

[19] We should also note that starting from the 2012-2013 academic year, students in NYC whose families qualified for Medicaid received free school meals automatically, per the Medicaid direct certification program (Maurice, 2018).
[20] Note that since we control for age in all regressions, we still compare outcomes with the same length of measurement periods.

we lose the effect on violence-related injury for boys, other results are roughly consistent with the main ones.

A direct test for the assumption would be available if the same outcomes were available in both data sets: we could estimate the effects of small schools on the outcome using standard 2SLS with one data set and then compare the results from 2SLS and TS2SLS. The common population assumption would more plausibly hold if the results were similar. We are not able to conduct the test for the main outcomes because they are available in the Medicaid data set only. But we do have one common outcome, which is overweight/obesity, in the two data sets, although the definitions differ somewhat.[21] Thus, we estimate the effects of attending small schools on overweight/obesity indicators by a standard 2SLS and compare them with our main effects on overweight diagnoses. Table A15 in the appendix compares the effects on overweight and obesity estimated by a standard 2SLS with our main effects on the overweight diagnosis. Overall, the effects show some similarities.[22] However, we caution that due to the difference in outcome definitions, the results of the comparison are not sufficient to fully confirm our assumption.

The actual concern about the common population assumption is whether our first-stage results, i.e., the relationship between distance and school enrollment, significantly differ from the

[21] In the education data set there is information on BMI (Body Mass Index), from which we can generate overweight and obesity indicators. We use the cutoffs from the National Center for Chronic Disease Prevention and Health Promotion to define overweight and obesity. A child/teen is described as overweight if BMI is at or above the 85th percentile and below the 95th percentile for children and teens of the same age and sex, and as obesity if BMI is at or above the 95th percentile for children and teens of the same age and sex. Source: https://www.cdc.gov/obesity/childhood/defining.html (for definitions of overweight and obesity) and https://www.cdc.gov/healthyweight/assessing/bmi/childrens_bmi/about_childrens_bmi.html (for US children BMI percentile). In the Medicaid data we have diagnoses related to overweight/obesity.

[22] We should note that the overweight variable in this study is likely to be closer to obesity (rather than overweight) by BMI, because our outcome is defined as whether the student had been diagnosed with problems related to overweight/obesity. Generally, a slight overweight is unlikely to cause health problems for teens or to be noticed by physician.

true results for students on Medicaid. Schwartz, Stiefel and Wiswall (2013, 2016) estimate the first stage using the full sample rather than the sample eligible for free lunch only. Their results are very similar to those in this paper, which indicates that the first-stage results estimated in this paper and other previous papers may hold for a broad range of subpopulations. Although the two data sets in this study do not reflect precisely the same population, the first-stage estimations from the education data may still be a good proxy of the true first-stage estimations for the Medicaid population. Nevertheless, we are not able to directly confirm the validity of the assumption, leaving room for potential bias in our TS2SLS results. Since our reduced-form and TS2SLS results are consistent, it is unlikely that our TS2SLS results are false. Combining the reduced-form and TS2SLS results, our main conclusions regarding the effects of small schools on youth risky behaviors and mental health are likely to hold, although the TS2SLS results may be less robust than the reduced-form results, with the magnitudes deviating somewhat from the true values.

7. Conclusions

Optimal school size, in terms of effects on educational outcomes and overall well-being, has been debated for decades. One strand of literature argues that larger-sized schools benefit from economies of scale (Kenny, 1982; Hanushek, 1986; Card and Krueger, 1996). Other studies argue that larger school size makes students more likely to be isolated and frustrated (Ferris and West, 2002), which may offset benefits from economies of scale. Given empirical evidence that smaller schools increase graduation rates (Schwartz, Stiefel and Wiswall, 2013, 2015; Bloom, Thompson and Unterman, 2010), student achievement (Howley, Strange and Bickel, 2000; McMillen, 2004) and long-term success in the educational system and the labor market (Humlum and Smith, 2015), many major urban school districts have implemented small high school reforms in the past two decades. Despite the movement toward smaller schools, we do not know much about how smaller-sized schools affect youth risky behaviors or mental health, on which the small body of existing literature lacks consensus.

One possible explanation for the inconsistency between findings in the literature on the relationship between school size and youth risky behaviors and mental health is that there is heterogeneity in the mission and features of small schools. Small schools built with different purposes can significantly differ from each other and large schools in many aspects other than size. Evaluating a pooled group of small schools of different types may lead to inconsistent conclusions, depending on specific policies in practice. Moreover, ignoring other relevant confounding policies, such as SBHCs in this case, may lead to misinterpretation of the estimated results. We disentangle school size from other schooling organizations and inputs by distinguishing between old and new small schools and estimating their effects separately.

Another possible explanation for the inconsistency between findings in the literature is that self-selection of school size is not always addressed. To estimate the causal effects of attending small schools on youth risky behaviors and mental health rather than correlations between small school attendance and risky behaviors or mental health that are linked by other unobserved pre-existing characteristics, we use a two-sample-instrumental-variable estimator, which exploits exogenous distance from student residence to schools together with characteristics of census tracts as a correction for the self-selection issue and the fact that the information of high school enrollment and youth risky behaviors are from two separate data sets.

Using student-level data from NYCDOE administrative and survey databases and Medicaid claims data, we examine the effect of attending small high schools in NYC on teenage pregnancy, violent injury, and mental health diagnoses. Overall, we find few effects of small schools when we pool old and new small schools. When we examine old and new small schools separately, we find negative effects of attending old small high schools on teen pregnancy and mental health diagnoses for boys. We also find positive effects of attending new small high schools on the injury from violence, for both boys and girls, and on mental health diagnoses for girls. Those effects cannot be fully explained by the utilization of contraception treatment or access to on-campus health facilities such as school-based health centers.

Our results indicate that reducing school size alone is not likely to be an effective solution for youth risky behaviors such as teen pregnancy, for low-income students in an urban environment. Combined with other investments in educational inputs, such as teacher quality and student expenditures, small school policies may be a promising way to address youth risky behaviors and mental health. While research shows that some small schools (new) have positive effects on educational attainment, we argue that future studies should explore impacts outside of education. For instance, our use of health data in this study provided an opportunity to understand how students enrolled in small schools fare differently from their counterparts in large schools. In turn, these findings may contribute to the current understanding of the role of school size in shaping student outcomes and, ultimately, optimal school size.

This study has some limitations, particularly in our ability to fully understand the channels, through which old small schools work. We have explored several potential channels such as the availability of school-based health centers and the utilization of contraception, but they fail to explain our findings. Lower dropout rates in new small schools may help explain the results partially, but we still find positive effects of new small schools on pregnancy before age 17, when students in NYC can leave school legally (without parental consent). Small schools have lower enrollment, lower student-teacher ratios, and higher per-student expenditures. Since those three advantages are correlated in the small schools, it is not clear whether the effects come from more desirable peer interactions (peer effects) due to smaller school sizes or stronger behavioral interventions due to lower student-teacher ratios, among others. Moreover, all students in this analysis came from low-income families that were eligible for the free and reduced lunch (first stage) or the Medicaid program (second stage), and small schools might lead to different outcomes for higher-income students. Future studies could continue to investigate the effects of small schools on broader student populations and their channels, such as social networks, and exploit more measures of risky behavior, such as gunshots and bullying, with richer data sets. Finally, due to data limitations, we are not able to directly test the assumptions for the TS2SLS approach. Despite being in line with each other, our TS2SLS results might be less robust than the reduced-form results.

Reference

Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth. "The New York City high school match." *American Economic Review* 95, no. 2 (2005): 364-367.

Angrist, Joshua D. "Lifetime earnings and the Vietnam era draft lottery: evidence from social security administrative records." *The American Economic Review* (1990): 313-336.

Angrist, Joshua D., and Alan B. Krueger. "The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples." *Journal of the American statistical Association*87, no. 418 (1992): 328-336.

Angrist, Joshua D., and Alan B. Krueger. "Split-sample instrumental variables estimates of the return to schooling." *Journal of Business & Economic Statistics* 13, no. 2 (1995): 225-235.

Angrist, Joshua D., and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2008.

Astor, Ron Avi, Rami Benbenishty, Anat Zeira, and Amiram Vinokur. "School climate, observed risky behaviors, and victimization as predictors of high school students' fear and judgments of school violence as a problem." *Health Education & Behavior* 29, no. 6 (2002): 716-736.

Barker, Roger G., and Paul V. Gump. *Big school, small school: High school size and student behavior*. Stanford university press, 1964.

Barrow, Lisa, Diane Whitmore Schanzenbach, and Amy Claessens. "The impact of Chicago's small high school initiative." *Journal of Urban Economics*87 (2015): 100-113.

Berthelon, Matias E., and Diana I. Kruger. "Risky behavior among youth: Incapacitation effects of school on adolescent motherhood and crime in Chile." *Journal of public economics* 95, no. 1 (2011): 41-53.

Bhattacharya, Jay, and Janet Currie. "Youths at Nutrition Risk: Malnourished or Misnourished?." In *Risky behavior among youths: An economic analysis*, pp. 483-522. University of Chicago Press, 2001.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. "Staying in the classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births." *The Economic Journal* 118, no. 530 (2008): 1025-1054.

Bloom, Howard S., Saskia Levy Thompson, and Rebecca Unterman. "Transforming the high school experience." *How New York City's New Small Schools Are Boosting Student Achievement and Graduation Rates. New York: MDRC. Recuperado de http://www. mdrc. org/sites/default/files/full_589. pdf, consultado el* 20 (2010).

Bowes, Lucy, Louise Arseneault, Barbara Maughan, Alan Taylor, Avshalom Caspi, and Terrie E. Moffitt. "School, neighborhood, and family factors are associated with children's bullying involvement: A nationally representative longitudinal study." *Journal of the American Academy of Child & Adolescent Psychiatry* 48, no. 5 (2009): 545-553.

Booker, Kevin, Tim R. Sass, Brian Gill, and Ron Zimmer. "The effects of charter high schools on educational attainment." *Journal of Labor Economics*29, no. 2 (2011): 377-415.

Brezina, Timothy, Alex R. Piquero, and Paul Mazerolle. "Student anger and aggressive behavior in school: An initial test of Agnew's macro-level strain theory." *Journal of research in crime and delinquency* 38, no. 4 (2001): 362-386.

Brooks, Traci L., Sion Kim Harris, Jeannie S. Thrall, and Elizabeth R. Woods. "Association of adolescent risk behaviors with mental health symptoms in high school students." *Journal of Adolescent Health* 31, no. 3 (2002): 240-246.

Burgess, Simon, and Adam Briggs. "School assignment, school choice and social mobility." *Economics of Education Review* 29, no. 4 (2010): 639-649.

Card, David, and Alan B. Krueger. "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina." *The Journal of Economic Perspectives* 10, no. 4 (1996): 31-50.

Card, David, and Thomas Lemieux. "Dropout and enrollment trends in the postwar period: What went wrong in the 1970s?." In *Risky behavior among youths: An economic analysis*, pp. 439-482. University of Chicago Press, 2001.

Center for Behavioral Health Statistics and Quality. "2019 National Survey on Drug Use and Health: Mental Health Detailed Tables. " Rockville, MD: Substance Abuse and Mental Health Services Administration; 2020.

Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. "Preventing Youth Violence." (2019). https://www.cdc.gov/violenceprevention/pdf/yv-factsheet508.pdf

Centers for Disease Control and Prevention. "About Teen Pregnancy." (2021). https://www.cdc.gov/teenpregnancy/about/index.htm

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review* 104, no. 9 (2014): 2633-79.

Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque. "Fatter attraction: anthropometric and socioeconomic matching on the marriage market." *Journal of Political Economy* 120, no. 4 (2012): 659-695.

Clark, Andrew E., and Youenn Lohéac. "“It wasn’t me, it was them!” Social influence in risky behavior by adolescents." *Journal of health economics* 26, no. 4 (2007): 763-784.

Conant, James Bryant, and John W. Gardner. *The American high school today: A first report to interested citizens*. New York: McGraw-Hill, 1959.

Cotton, Kathleen. "Affective and Social Benefits of Small-Scale Schooling. ERIC Digest." (1996). https://www.ericdigests.org/1997-2/small.htm

Cullen, Julie Berry, Brian A. Jacob, and Steven D. Levitt. "The impact of school choice on student outcomes: an analysis of the Chicago Public Schools." *Journal of Public Economics* 89, no. 5 (2005): 729-760.

Cullen, Julie Berry, Brian A. Jacob, and Steven Levitt. "The effect of school choice on participants: Evidence from randomized lotteries." *Econometrica* 74, no. 5 (2006): 1191-1230.

Currie, Janet, and Enrico Moretti. "Mother's education and the intergenerational transmission of human capital: Evidence from college openings." *The Quarterly Journal of Economics* 118, no. 4 (2003): 1495-1532.

Darling-Hammond, Linda, Jacqueline Ancess, and Susanna Wichterle Ort. "Reinventing high school: Outcomes of the coalition campus schools project." *American educational research journal* 39, no. 3 (2002): 639-673.

Dee, Thomas S., and William N. Evans. "Teens and traffic safety." In *Risky behavior among youths: An economic analysis*, pp. 121-166. University of Chicago Press, 2001.

Dobbie, Will, and Roland G. Fryer Jr. "The medium-term impacts of high-achieving charter schools." *Journal of Political Economy* 123, no. 5 (2015): 985-1037.

Farrell, Phillip, and Victor R. Fuchs. "Schooling and health: the cigarette connection." *Journal of health economics* 1, no. 3 (1982): 217-230.

Ferris, J. Stephen, and Edwin G. West. "Education vouchers, the peer group problem, and the question of dropouts." *Southern Economic Journal* (2002): 774-793.

Ferris, J. Stephen, and Edwin G. West. "Economies of scale, school violence and the optimal size of schools." *Applied Economics* 36, no. 15 (2004): 1677-1684.

Fryer Jr, Roland G., and Lawrence F. Katz. "Achieving escape velocity: neighborhood and school interventions to reduce persistent inequality." *American Economic Review* 103, no. 3 (2013): 232-37.

Gottfredson, Gary D., Denise C. Gottfredson, Allison Ann Payne, and Nisha C. Gottfredson. "School climate predictors of school disorder: Results from a national study of delinquency prevention in schools." *Journal of research in crime and delinquency* 42, no. 4 (2005): 412-444.

Gruber, Jonathan. "Introduction to" Risky Behavior among Youths: An Economic Analysis"." In *Risky Behavior Among Youths: An Economic Analysis*, pp. 1-28. University of Chicago Press, 2001.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. *Preferences and heterogeneous treatment effects in a public school choice lottery*. No. w12145. National Bureau of Economic Research, 2006.

Hanushek, Eric A. "The economics of schooling: Production and efficiency in public schools." *Journal of economic literature* 24, no. 3 (1986): 1141-1177.

Howley, Craig, Marty Strange, and Robert Bickel. "Research about School Size and School Performance in Impoverished Communities. ERIC Digest." (2000).

Humlum, Maria Knoth, and Nina Smith. "Long-term effects of school size on students' outcomes." Economics of Education Review 45 (2015): 28-43.

Iatarola, Patrice and Leanna Stiefel, "Student Achievement and Costs of Small Schools," *Handbook on the Economics of Education,* Geraint Johnes et. al, editors, Edward Elgar. 2018

Inoue, Atsushi, and Gary Solon. "Two-sample instrumental variables estimators." *The Review of Economics and Statistics* 92, no. 3 (2010): 557-561.

Jacob, Brian A., and Lars Lefgren. "Are idle hands the devil's workshop? Incapacitation, concentration, and juvenile crime." *The American Economic Review* 93, no. 5 (2003): 1560-1577.

Kahne, Joseph E., Susan E. Sporte, and Marisa De La Torre. "Small High Schools on a Larger Scale: The First Three Years of the Chicago High School Redesign Initiative. Research Report." *Consortium on Chicago School Research* (2006).

Kenny, Lawrence W. "Economies of scale in schooling." *Economics of Education Review* 2, no. 1 (1982): 1-24.

Kerns, Suzanne EU, Michael D. Pullmann, Sarah Cusworth Walker, Aaron R. Lyon, T. J. Cosgrove, and Eric J. Bruns. "Adolescent use of school-based health centers and high school dropout." *Archives of pediatrics & adolescent medicine*165, no. 7 (2011): 617-623.

Kirby, Douglas, Cynthia Waszak, and Julie Ziegler. "Six school-based clinics: their reproductive health services and impact on sexual behavior." *Family Planning Perspectives*(1991): 6-16.

Kisker, Ellen E., and Randall S. Brown. "Do school-based health centers improve adolescents' access to health care, health status, and risk-taking behavior?." *Journal of Adolescent Health* 18, no. 5 (1996): 335-343.

Klevmarken, Anders. *Missing variables and two-stage least-squares estimation from more than one data set*. No. 62. IUI Working Paper, 1982.

Leung, Ambrose, and J. Stephen Ferris. "School size and youth violence." *Journal of Economic Behavior & Organization* 65, no. 2 (2008): 318-333.

Lindberg, Laura, John Santelli, and Sheila Desai. "Understanding the decline in adolescent fertility in the United States, 2007–2012." *Journal of Adolescent Health* 59, no. 5 (2016): 577-583.

Litwiller, Brett J., and Amy M. Brausch. "Cyber bullying and physical bullying in adolescent suicide: the role of violent behavior and substance use." *Journal of youth and adolescence* 42, no. 5 (2013): 675-684.

Lovenheim, Michael F., Randall Reback, and Leigh Wedenoja. *How does access to health care affect teen fertility and high school dropout rates? Evidence from school-based health centers*. No. w22030. National Bureau of Economic Research, 2016.

McMillen, Bradley J. "School size, achievement, and achievement gaps." *education policy analysis archives* 12 (2004): 58.

Miller, Amanda K., and Kathryn A. Chandler. "Violence in US public schools: 2000 school survey on crime and safety." In *Violence in US public schools: 2000 school survey on crime and safety*. Government Printing Office, 2003.

Oster, Emily. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics* (2017): 1-18.

Pacini, David, and Frank Windmeijer. "Robust inference for the Two-Sample 2SLS estimator." *Economics Letters* 146 (2016): 50-54.

Panchal, Nirmita, Robin Rudowitz, and Cynthia Cox. "Recent Trends in Mental Health and Substance Use Concerns Among Adolescents" (2022). KFF published June 28, 2022. https://www.kff.org/coronavirus-covid-19/issue-brief/recent-trends-in-mental-health-and-substance-use-concerns-among-adolescents/ accessed September 15, 2022.

Quint, Janet C., Janell K. Smith, Rebecca Unterman, and Alma Moedano. "New York City's Changing High School Landscape High Schools and Their Characteristics, 2002-2008." *Manpower Demonstration Research Corporation* (2010).

Ricketts, Sue A., and Bruce P. Guernsey. "School-based health centers and the decline in black teen fertility during the 1990s in Denver, Colorado." *American journal of public health*96, no. 9 (2006): 1588-1592.

Rudasill, Kathleen Moritz, Thomas G. Reio, Natalie Stipanovic, and Jennifer E. Taylor. "A longitudinal study of student–teacher relationship quality, difficult temperament, and risky behavior from childhood to early adolescence." *Journal of school psychology* 48, no. 5 (2010): 389-412.

Sanderson, Eleanor, and Frank Windmeijer. "A weak instrument F-test in linear IV models with multiple endogenous variables." *Journal of Econometrics* 190, no. 2 (2016): 212-221.

Santelli, John S., Laura Duberstein Lindberg, Lawrence B. Finer, and Susheela Singh. "Explaining recent declines in adolescent pregnancy in the United States: the contribution of abstinence and improved contraceptive use." *American journal of public health* 97, no. 1 (2007): 150-156.

Schwartz, Amy Ellen, Leanna Stiefel, and Matthew Wiswall. "Do small schools improve performance in large, urban districts? Causal evidence from New York City." *Journal of Urban Economics* 77 (2013): 27-40.

Schwartz, Amy Ellen, Leanna Stiefel, and Matthew Wiswall. "Are all schools created equal? Learning environments in small and large public high schools in New York City." *Economics of Education Review* 52 (2016): 272-290.

Schneider, Mark, and Jack Buckley. "What do parents want from schools? Evidence from the Internet." *Educational evaluation and policy analysis* 24, no. 2 (2002): 133-144.

Staiger, Douglas, and James H. Stock. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65, no. 3 (1997): 557-586.

Stiefel, Leanna, Amy Ellen Schwartz, and Matthew Wiswall. "Does small high school reform lift urban districts? Evidence from New York City." *Educational Researcher* 44, no. 3 (2015): 161-172.

Stiefel, Leanna, Robert Berne, Patrice Iatarola, and Norm Fruchter. "High school size: Effects on budgets and performance in New York City." *Educational Evaluation and Policy Analysis* 22, no. 1 (2000): 27-39.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo. "A survey of weak instruments and weak identification in generalized method of moments." *Journal of Business & Economic Statistics* 20, no. 4 (2002): 518-529.

Van Cura, Maureen. "The relationship between School-Based health centers, rates of early dismissal from school, and loss of seat time." *Journal of School Health* 80, no. 8 (2010): 371-377.

Wolke, Dieter, Sarah Woods, Katherine Stanford, and Henrike Schulz. "Bullying and victimization of primary school children in England and Germany: Prevalence and school factors." *British journal of psychology* 92, no. 4 (2001): 673-696.

Appendix

Table A1: Differences in Resources and Organizational Factors by School Size

	Compared with large schools					
	Total expenses per student ($)	Expenses on classroom instruction per student ($)	% master or more	% less than three-year experience	% no valid certification	% colocation
New small schools	+511	nsd.	-22	+18	+3	+57
Old small schools	+1,396	+559	-11	+5	nsd.	+24

Source: selected measures of resources and organizational factors in Schwartz, Stiefel and Wiswall (2013). + (-) indicates significantly higher (lower) than large schools. nsd. stands for no significant difference.

Table A2: Claim Codes for the Outcomes

Pregnancy	
ICD-9-CM	63*, 64*, 65*, 66*, 67*, V22*, V23*, V24*, V27*, V28*, V91*, V616*, V617*, V7242*
CPT	72*, 73*, 740, 741, 742, 744, 7440, 7499, 59812, 59820, 59821, 59830, 59120, 59121, 59130, 59135, 59136, 59140, 59840, 59841, 59850, 59851, 59852, 59855, 59856, 59857, 59400, 59409, 59410, 59412, 59425, 59426, 59510, 59514, 59515, 59610, 59612, 59614, 59618, 59620, 59622
Violent injury	
ICD-9-CM	E96*
Mental health diagnosis	
ICD-9-CM	29*, 30*, 31*
Contraception, most effective (sterilization or long-acting reversible)	
ICD-9-CM	V25.11, V25.13, V25.42, V45.51, 996.32, V25.5, V25.43, V45.52, V25.2, V26.51
CPT	697, 0UH97HZ, 0UH98HZ, 0UHC7HZ, 0UHC8HZ, 58300, J7300, J7301, J7302, J7306, J7307, S4989, Q0090, S4981, 11981, 11983, 662, 0U574ZZ, 0U578ZZ, 0UL74CZ, 0UL74DZ, 0UL74ZZ, 0UL78DZ, 0UL78ZZ, 58600, 58605, 58615, 58611, 58670, 58671, 58565, A4264
NDC	50419042101, 50419042201, 50419042301, 51285020401, 00052027201, 00052027401, 00052433001
Contraception, moderately effective (injectable, oral pill, patch, ring, or diaphragm)	
ICD-9-CM	V25.01, V25.41
CPT	J1050, S4993, J7304, J7303, 57170, A4266
NDC	00009074630, 00009074635, 00009470901, 00009470913, 00009737604, 00009737607, 00009737611, 00247210801, 00703680101, 00703680104, 00703681121, 23490585401, 54569370100, 54569490400, 54569552700, 54569561600, 54569621900, 54868361300, 54868410000, 54868410001, 54868525700, 55045350501, 59762453701, 59762453702, 59762453801, 59762453802, 59762453809, 00008111720, 00008111730, 00008251402, 00008253505, 00008253601, 00008253605, 00052026106, 00052028306, 00052028308, 00062125100, 00062125115, 00062125120, 00062133220, 00062141116, 00062141123, 00062171400, 00062171415, 00062176100, 00062176115, 00062178100, 00062178115, 00062179600, 00062179615, 00062190120, 00062190320, 00062190700, 00062190715, 00062191000, 00062191015, 00093214062, 00093209028, 00093209058, 00093313482, 00093532862, 00093542328, 00093542358, 00093566128, 00093566158, 00093614882, 00247052028, 00247069028, 00247069128, 00247069228, 00247139828, 00247151328, 00247151628, 00247151728, 00247176404, 00247176421, 00247176521, 00247198621, 00247198628, 00247200828, 00247201004, 00247201008, 00247201028, 00247201228, 00247201328, 00247214728, 00247216928, 00247217028, 00247223028, 00247223528, 00247226028, 00247226828, 00378728153, 00378655053, 00378727253, 00378729253, 00378730153, 00378730853, 00430000531, 00430001005, 00430042014, 00430048214, 00430053014, 00430053550, 00430054050, 00430057014, 00430057045, 00430058014, 00430058045, 00430058114, 00430058514, 00430058545, 00555034458, 00555071558, 00555900867, 00555900942, 00555901058, 00555901258, 00555901467, 00555901658, 00555901858, 00555902058, 00555902542, 00555902557, 00555902658, 00555902742, 00555902757, 00555902858, 00555903270, 00555903458, 00555904358, 00555904558, 00555904758, 00555904958, 00555905058, 00555905158, 00555905167, 00555906458, 00555906467, 00555906558, 00555906658, 00555906667, 00555912366, 00555913167, 00555913179, 00603359017, 00603359049, 00603751217, 00603751249, 00603752117, 00603752149, 00603752517, 00603752549, 00603754017, 00603754049, 00603760615, 00603760648, 00603760715, 00603760748, 00603760817, 00603760917, 00603761017, 00603761049, 00603762517, 00603762549, 00603763417, 00603763449, 00603764017, 00603764217, 00603766317, 00603766517, 00781405815, 00781406015, 00781406215, 00781558307, 00781558315, 00781558336, 00781558436, 00781558491, 00781565615, 00781565815, 16714033003, 16714034004, 16714034604, 16714034704, 16714034804, 16714035904, 16714036004, 16714036304, 16714036504, 16714037003, 16714040703, 16714044004, 16714044104, 21695076928, 21695077028, 23490765301, 23490767001, 23490769901, 24090080184, 24090096184, 35356001468, 35356001568, 35356002168, 35356025528, 35356037028, 50102010048, 50102012048, 50102012803, 50102013048, 50102015403, 50419040201, 50419040203, 50419040303, 50419040503, 50419040701, 50419040703, 50419041112, 50419041128, 50419043306, 50419043312, 50419048203, 50419048303, 50452025115, 50458017115, 50458017615, 50458017815, 50458019115, 50458019411, 50458019416, 50458019615, 50458019715, 50458025115, 51285005866, 51285007997, 51285008070, 51285008198, 51285008297, 51285008370, 51285008498, 51285008787, 51285009158, 51285009287, 51285011458, 51285012058, 51285012570, 51285012698, 51285012797, 51285012998, 51285013197, 51285043165, 51285054628, 51660012786, 51660057286, 52544006431, 52544014331, 52544017572, 52544020431, 52544021028, 52544021928, 52544022829, 52544023328, 52544023528, 52544023531, 52544024531, 52544024728, 52544024828, 52544024928, 52544025428, 52544025928, 52544025988, 52544026528, 52544026531,

	52544026829, 52544026884, 52544027428, 52544027431, 52544027621, 52544027928, 52544029021, 52544029128, 52544029231, 52544029241, 52544029528, 52544038328, 52544038428, 52544055028, 52544055228, 52544055428, 52544062928, 52544063028, 52544063128, 52544084728, 52544084828, 52544089228, 52544093628, 52544094028, 52544094928, 52544095021, 52544095121, 52544095328, 52544095428, 52544095931, 52544096691, 52544096728, 52544098131, 52544098231, 54569067900, 54569068500, 54569068501, 54569068900, 54569068901, 54569143900, 54569384400, 54569422200, 54569422201, 54569426900, 54569427301, 54569481700, 54569487800, 54569487801, 54569489000, 54569498400, 54569499700, 54569499800, 54569516100, 54569534900, 54569549300, 54569549302, 54569579600, 54569579700, 54569579800, 54569581600, 54569582600, 54569603200, 54569612800, 54569614400, 54569627200, 54569628000, 54569628100, 54868042800, 54868044300, 54868050200, 54868050700, 54868050801, 54868050901, 54868051600, 54868151200, 54868156400, 54868231600, 54868260600, 54868270100, 54868377200, 54868386300, 54868394800, 54868409300, 54868423900, 54868436900, 54868453800, 54868459000, 54868460700, 54868473000, 54868473100, 54868474200, 54868474500, 54868475400, 54868477600, 54868481400, 54868482800, 54868485100, 54868486000, 54868491100, 54868502800, 54868528600, 54868532600, 54868535600, 54868582600, 54868582800, 54868594200, 55045348506, 55045349701, 55045349801, 55045378106, 55045378206, 55289024708, 55289088704, 55887005228, 55887028628, 58016474701, 58016482701, 66993061128, 66993061528, 68180084313, 68180084413, 68180084613, 68180084813, 68180085413, 68180087611, 68180087613, 68180089213, 68180089713, 68180089813, 68180089913, 68180090213, 68462030329, 68462030529, 68462030929, 68462031629, 68462031829, 68462038829, 68462039429, 68462055629, 68462056529, 68462063729, 68462064693, 00378728053, 00378728353, 00378728753, 00378729653, 00430053750, 16714007304, 16714035903, 16714036704, 16714040402, 16714040404, 16714040501, 16714040504, 16714040601, 16714040604, 16714040803, 16714041304, 50419040903, 65162031684, 65162034784, 68180087513, 68180087711, 68180087713, 68180088213, 68180088613, 68180089211, 68180089313, 75854060101, 00062192001, 00062192015, 00062192024, 50458019201, 50458019215, 54569541300, 54868467000, 00378334053, 00052027301, 00052027303, 54569586500, 54868483201, 55887075401, 00027013160, 00027013180, 00062330100, 00062330200, 00062330300, 00062330400, 00062330500, 00062330600, 00062330700, 00062330800, 00062330900, 00062331000, 00062331100, 00062331200, 00062331300, 00062334100, 00062334200, 00062334300, 00062334400, 00062334500, 00062334600, 00062334700, 00062334800, 00062334900, 00062335000, 00062335100, 00062335200, 00062338100, 00062338200, 00062338300, 00062338400, 00062338500, 00062338600, 00062338700, 00062338800, 00062338900, 00062364103, 00062364300, 00234005100, 00234013100, 00234013150, 00234013155, 00234013160, 00234013165, 00234013170, 00234013175, 00234013180, 00234013185, 00234013190, 00234013195, 00234013600, 00234013660, 00234013665, 00234013670, 00234013675, 00234013680, 00234013685, 00234013690, 00234013695, 00396401065, 00396401070, 00396401075, 00396401080
Overweight	
ICD-9-CM	278*, 649.1*

ICD-9-CM: The International Classification of Diseases, Nineth Revision, Clinical Modification. CPT: Current Procedural Terminology. NDC: National Drug Code. * indicates including all subcodes.

Table A3: Descriptive Statistics of School Characteristics

	All	Large	Small	Small	
				New	Old
Number of schools	380	107	273	187	86
Enrollment	760	1812	348	323	402
Age	17	44	7	4	13
SBHC (%)	26.3	17.8	29.7	31.6	25.6
Colocation (%)	71.1	37.4	84.2	95.2	60.5

Note: three schools in Staten Island are excluded. The enrollment cutoff between large and small schools is 550.

Table A4: Reduced-Form Estimation, Pooled Model, Risky Behaviors and Mental Health

	Boys		Girls		
	Violent injury	**Mental health diagnosis**	**Pregnancy**	**Violent injury**	**Mental health diagnosis**
Distance					
Distance to large high school	-0.001 (0.003)	-0.013* (0.007)	-0.002 (0.008)	-0.001 (0.003)	-0.006 (0.007)
Distance to small high school	-0.001 (0.002)	0.008 (0.005)	0.004 (0.006)	-0.003* (0.002)	-0.005 (0.006)
Student characteristics					
Asian	-0.013*** (0.004)	-0.084*** (0.011)	-0.132*** (0.012)	-0.003 (0.002)	-0.087*** (0.011)
Black	0.015*** (0.004)	0.048*** (0.009)	0.034*** (0.011)	0.019*** (0.004)	0.035*** (0.010)
Hispanic	-0.001 (0.004)	0.042*** (0.010)	0.060*** (0.011)	0.002 (0.003)	0.054*** (0.011)
Age	-0.003 (0.005)	-0.015 (0.011)	0.036*** (0.013)	0.004 (0.004)	-0.025** (0.012)
Manhattan	0.006 (0.006)	0.037*** (0.013)	0.014 (0.014)	-0.006 (0.005)	0.007 (0.014)
Brooklyn	-0.010** (0.005)	-0.009 (0.011)	0.010 (0.012)	0.002 (0.004)	-0.021* (0.012)
Queens	0.003 (0.005)	0.049*** (0.012)	0.020 (0.014)	0.004 (0.005)	0.024* (0.013)
Cohort 2009	0.006 (0.004)	0.026*** (0.010)	0.005 (0.011)	0.006* (0.004)	0.017 (0.011)
Census tract characteristics					
Population (ln)	-0.013*** (0.003)	-0.020*** (0.007)	-0.012 (0.008)	0.002 (0.003)	-0.001 (0.008)
Percent Asian	0.006 (0.012)	0.055* (0.030)	-0.047 (0.034)	-0.006 (0.009)	0.011 (0.033)
Percent White	-0.031*** (0.007)	-0.151*** (0.018)	-0.141*** (0.021)	-0.020*** (0.006)	-0.130*** (0.019)
Percent Hispanic	0.004 (0.010)	0.047** (0.022)	0.091*** (0.025)	0.010 (0.008)	0.095*** (0.024)
Percent age 65 and up	0.017 (0.027)	0.357*** (0.071)	0.134* (0.080)	0.022 (0.024)	0.365*** (0.076)
Percent with high school diploma	0.032 (0.024)	-0.101* (0.053)	0.075 (0.062)	-0.002 (0.020)	-0.092 (0.060)
Percent with college degree or higher	0.031 (0.020)	0.114** (0.052)	0.209*** (0.059)	0.027* (0.016)	0.048 (0.057)
Percent foreign born	-0.026** (0.013)	-0.222*** (0.033)	-0.113*** (0.037)	-0.021* (0.011)	-0.106*** (0.036)
Employment rate	-0.033 (0.025)	0.037 (0.058)	-0.089 (0.066)	-0.032 (0.020)	-0.087 (0.064)
Average household income (ln)	-0.011 (0.007)	-0.006 (0.018)	-0.018 (0.021)	-0.001 (0.006)	0.035* (0.020)
Constant	0.314*** (0.109)	0.773*** (0.263)	0.119 (0.295)	-0.034 (0.094)	0.373 (0.284)
Multivariate F of excluded instruments [p-value]	0.957	0.033	0.593	0.473	0.896

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively. The sample size is 19,148 for boys and 16,585 for girls.

Table A5: Reduced-Form Estimation, Separate Model, Risky Behavior and Mental Health

	Boys		Girls		
	Violent injury	Mental health diagnoses	Pregnancy	Violent injury	Mental health diagnoses
Distance					
Distance to large high school	-0.002 (0.003)	-0.017** (0.007)	-0.010 (0.008)	-0.001 (0.002)	-0.011 (0.008)
Distance to new small high school	-0.003 (0.002)	0.005 (0.005)	-0.003 (0.006)	-0.004*** (0.002)	-0.009 (0.006)
Distance to old small high school	0.003** (0.001)	0.007** (0.003)	0.014*** (0.003)	0.001 (0.001)	0.009*** (0.003)
Student characteristics					
Asian	-0.013*** (0.004)	-0.086*** (0.011)	-0.135*** (0.012)	-0.004 (0.002)	-0.089*** (0.011)
Black	0.015*** (0.004)	0.047*** (0.009)	0.033*** (0.011)	0.018*** (0.004)	0.034*** (0.010)
Hispanic	-0.002 (0.004)	0.042*** (0.010)	0.060*** (0.011)	0.002 (0.003)	0.053*** (0.011)
Age	-0.003 (0.005)	-0.014 (0.011)	0.037*** (0.013)	0.004 (0.004)	-0.024** (0.012)
Manhattan	0.007 (0.006)	0.038*** (0.013)	0.019 (0.014)	-0.004 (0.005)	0.012 (0.014)
Brooklyn	-0.013** (0.005)	-0.017 (0.011)	-0.006 (0.013)	0.001 (0.004)	-0.031** (0.012)
Queens	0.002 (0.005)	0.043*** (0.012)	0.010 (0.014)	0.003 (0.005)	0.018 (0.013)
Cohort 2009	0.007 (0.004)	0.028*** (0.010)	0.010 (0.011)	0.007* (0.004)	0.020* (0.011)
Census tract characteristics					
Population (ln)	-0.012*** (0.003)	-0.020*** (0.007)	-0.012 (0.008)	0.002 (0.003)	-0.001 (0.008)
Percent Asian	0.007 (0.012)	0.053* (0.030)	-0.046 (0.034)	-0.006 (0.009)	0.012 (0.033)
Percent White	-0.032*** (0.007)	-0.156*** (0.018)	-0.153*** (0.021)	-0.021*** (0.006)	-0.137*** (0.019)
Percent Hispanic	0.005 (0.010)	0.046** (0.022)	0.092*** (0.025)	0.010 (0.008)	0.096*** (0.024)
Percent age 65 and up	0.001 (0.029)	0.317*** (0.074)	0.044 (0.083)	0.012 (0.025)	0.309*** (0.080)
Percent with high school diploma	0.029 (0.024)	-0.111** (0.054)	0.057 (0.062)	-0.004 (0.020)	-0.103* (0.060)
Percent with college degree or higher	0.030 (0.020)	0.112** (0.052)	0.207*** (0.059)	0.026 (0.016)	0.047 (0.057)
Percent foreign born	-0.030** (0.013)	-0.233*** (0.033)	-0.139*** (0.038)	-0.024** (0.011)	-0.123*** (0.036)
Employment rate	-0.031 (0.025)	0.045 (0.059)	-0.073 (0.067)	-0.031 (0.020)	-0.078 (0.064)
Average household income (ln)	-0.010 (0.007)	-0.007 (0.018)	-0.018 (0.021)	-0.001 (0.007)	0.034* (0.020)
Constant	0.312*** (0.109)	0.782*** (0.263)	0.123 (0.295)	-0.035 (0.095)	0.376 (0.284)
Multivariate F of excluded instruments [p-value]	0.136	0.015	0.000	0.023	0.024

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively. The sample size is 19,148 for boys and 16,585 for girls.

Table A6: First-Stage Estimation, Baseline (Pooled) and Separate Models

	Baseline (Pooled) Model (A)		Separate Model (B)			
	Boys	**Girls**	**Boys**		**Girls**	
High School Enrollment	**Small**	**Small**	**New Small**	**Old Small**	**New Small**	**Old Small**
Distance						
Distance to large high school	0.010*** (0.003)	0.009*** (0.003)	0.009*** (0.003)	0.001 (0.002)	0.014*** (0.003)	-0.001 (0.002)
Distance to small high school	-0.052*** (0.004)	-0.052*** (0.004)	-	-	-	-
Distance to new small high school	-	-	-0.048*** (0.004)	-0.005** (0.003)	-0.043*** (0.004)	-0.008*** (0.003)
Distance to old small high school	-	-	0.018*** (0.003)	-0.016*** (0.002)	0.011*** (0.003)	-0.019*** (0.002)
Student characteristics						
Asian	-0.056*** (0.010)	-0.058*** (0.011)	-0.021*** (0.008)	-0.034*** (0.007)	-0.024*** (0.009)	-0.035*** (0.008)
Black	0.044*** (0.013)	0.066*** (0.013)	0.054*** (0.011)	-0.011 (0.009)	0.072*** (0.011)	-0.007 (0.009)
Hispanic	0.032*** (0.011)	0.034*** (0.011)	0.057*** (0.009)	-0.026*** (0.007)	0.064*** (0.009)	-0.032*** (0.008)
Age	0.004 (0.005)	0.017*** (0.005)	0.005 (0.005)	-0.001 (0.003)	0.023*** (0.005)	-0.007* (0.004)
Manhattan	-0.036*** (0.013)	0.040*** (0.012)	-0.081*** (0.012)	0.049*** (0.010)	-0.042*** (0.011)	0.087*** (0.010)
Brooklyn	-0.140*** (0.010)	-0.114*** (0.010)	-0.043*** (0.010)	-0.099*** (0.007)	-0.014 (0.009)	-0.091*** (0.007)
Queens	-0.219*** (0.011)	-0.163*** (0.010)	-0.143*** (0.010)	-0.076*** (0.008)	-0.083*** (0.010)	-0.070*** (0.008)
Cohort 2009	-0.043*** (0.009)	-0.049*** (0.008)	-0.047*** (0.008)	0.008 (0.006)	-0.055*** (0.008)	0.003 (0.006)
Census tract characteristics						
Population (ln)	0.020*** (0.006)	0.018*** (0.006)	0.017*** (0.005)	0.002 (0.004)	0.013** (0.005)	0.005 (0.004)
Percent Asian	-0.051** (0.024)	-0.042* (0.023)	-0.109*** (0.021)	0.053*** (0.016)	-0.064*** (0.021)	0.026 (0.016)
Percent White	-0.249*** (0.019)	-0.182*** (0.019)	-0.194*** (0.017)	-0.054*** (0.012)	-0.128*** (0.017)	-0.050*** (0.013)
Percent Hispanic	0.025 (0.020)	0.008 (0.020)	0.041** (0.019)	-0.023* (0.014)	0.071*** (0.018)	-0.062*** (0.014)
Percent age 65 and up	-0.134** (0.063)	-0.239*** (0.061)	-0.081 (0.057)	-0.066 (0.043)	-0.034 (0.056)	-0.192*** (0.043)
Percent with high school diploma	-0.047 (0.047)	-0.051 (0.045)	0.087** (0.044)	-0.151*** (0.031)	0.084** (0.042)	-0.123*** (0.031)
Percent with college degree or higher	0.155*** (0.044)	0.087** (0.042)	0.175*** (0.039)	-0.024 (0.030)	0.131*** (0.038)	-0.039 (0.030)
Percent foreign born	-0.253*** (0.026)	-0.284*** (0.025)	-0.119*** (0.024)	-0.134*** (0.017)	-0.154*** (0.023)	-0.131*** (0.017)
Employment rate	0.047 (0.051)	0.116** (0.046)	-0.052 (0.044)	0.099*** (0.034)	0.044 (0.043)	0.063* (0.034)
Average household income (ln)	-0.013 (0.014)	-0.040** (0.012)	0.000 (0.011)	-0.011 (0.009)	-0.036*** (0.012)	-0.000 (0.009)
Constant	0.518*** (0.179)	0.617*** (0.154)	0.122 (0.150)	0.406*** (0.116)	0.171 (0.153)	0.414*** (0.117)
Multivariate F of excluded instruments [*p*-value]	77.04 [0.000]	77.05 [0.000]	63.74 [0.000]	28.74 [0.000]	47.80 [0.000]	50.93 [0.000]

Sanderson-Windmeijer multivariate F test of excluded instruments [*p*-value]	77.04 [0.000]	77.05 [0.000]	99.50 [0.000]	41.37 [0.000]	75.50 [0.000]	76.84 [0.000]
Sample size	25,572	28,404	25,572		28,404	

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively.

Table A7: TS2SLS Estimation, Pooled Model, Risky Behavior and Mental Health

	Boys		Girls		
	Violent injury	**Mental health diagnoses**	**Pregnancy**	**Violent injury**	**Mental health diagnoses**
Enrollment					
Small School	0.015 (0.040)	-0.136 (0.105)	-0.071 (0.117)	0.058** (0.029)	0.106 (0.111)
Student characteristics					
Asian	-0.012*** (0.004)	-0.091*** (0.012)	-0.136*** (0.014)	0.000 (0.003)	-0.081*** (0.013)
Black	0.014*** (0.005)	0.053*** (0.011)	0.039*** (0.013)	0.015*** (0.004)	0.028** (0.013)
Hispanic	-0.002 (0.004)	0.046*** (0.011)	0.062*** (0.012)	-0.000 (0.004)	0.050*** (0.012)
Age	-0.003 (0.005)	-0.014 (0.011)	0.037*** (0.013)	0.003 (0.004)	-0.026** (0.012)
Manhattan	0.006 (0.006)	0.035*** (0.013)	0.017 (0.015)	-0.008* (0.005)	0.005 (0.015)
Brooklyn	-0.007 (0.008)	-0.027 (0.020)	0.002 (0.020)	0.009 (0.006)	-0.008 (0.019)
Queens	0.006 (0.011)	0.016 (0.028)	0.008 (0.025)	0.013* (0.007)	0.039 (0.024)
Cohort 2009	0.006 (0.004)	0.020* (0.011)	0.001 (0.012)	0.009** (0.004)	0.022* (0.012)
Census tract characteristics					
Population (ln)	-0.013*** (0.004)	-0.018** (0.007)	-0.011 (0.008)	0.001 (0.003)	-0.004 (0.009)
Percent Asian	0.007 (0.013)	0.052 (0.032)	-0.049 (0.035)	-0.004 (0.009)	0.019 (0.034)
Percent White	-0.027** (0.013)	-0.183*** (0.035)	-0.153*** (0.033)	-0.009 (0.009)	-0.109*** (0.032)
Percent Hispanic	0.004 (0.009)	0.052** (0.022)	0.092*** (0.025)	0.009 (0.008)	0.095*** (0.024)
Percent age 65 and up	0.018 (0.028)	0.332*** (0.073)	0.117 (0.085)	0.035 (0.025)	0.387*** (0.082)
Percent with high school diploma	0.033 (0.024)	-0.108** (0.054)	0.071 (0.062)	0.001 (0.021)	-0.085 (0.060)
Percent with college degree or higher	0.030 (0.021)	0.144** (0.056)	0.216*** (0.061)	0.022 (0.017)	0.045 (0.059)
Percent foreign born	-0.022 (0.018)	-0.257*** (0.044)	-0.133** (0.052)	-0.005 (0.014)	-0.075 (0.050)
Employment rate	-0.034 (0.025)	0.046 (0.059)	-0.081 (0.068)	-0.039* (0.020)	-0.099 (0.065)
Average household income (ln)	-0.011 (0.008)	-0.013 (0.019)	-0.021 (0.022)	0.001 (0.007)	0.036* (0.021)
Constant	0.311*** (0.114)	0.894*** (0.273)	0.167 (0.309)	-0.066 (0.099)	0.334 (0.298)

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively. The first-stage sample size is 25,572 for boys and 28,404 for girls. The second-stage sample size is 19,148 for boys and 16,585 for girls.

Table A8: TS2SLS Estimation, Separate Model, Risky Behavior and Mental Health

	Boys		Girls		
	Violent injury	Mental Health Diagnoses	Pregnancy	Violent injury	Mental Health Diagnoses
Enrollment					
New Small School	0.076* (0.042)	-0.063 (0.109)	0.150 (0.142)	0.094** (0.037)	0.220* (0.134)
Old Small School	-0.065 (0.074)	-0.415** (0.184)	-0.553*** (0.159)	0.003 (0.050)	-0.200 (0.148)
Student characteristics					
Asian	-0.013*** (0.005)	-0.100*** (0.013)	-0.150*** (0.014)	-0.001 (0.003)	-0.090*** (0.013)
Black	0.010* (0.005)	0.046*** (0.012)	0.018 (0.016)	0.012** (0.005)	0.016 (0.015)
Hispanic	-0.008 (0.005)	0.034*** (0.013)	0.032** (0.016)	-0.004 (0.005)	0.033** (0.015)
Age	-0.003 (0.005)	-0.015 (0.011)	0.030** (0.013)	0.002 (0.004)	-0.031** (0.013)
Manhattan	0.017** (0.008)	0.056*** (0.018)	0.075*** (0.022)	-0.000 (0.007)	0.040* (0.021)
Brooklyn	-0.015 (0.011)	-0.058** (0.026)	-0.053** (0.023)	0.003 (0.007)	-0.043** (0.022)
Queens	0.007 (0.011)	0.000 (0.029)	-0.019 (0.025)	0.011 (0.007)	0.020 (0.023)
Cohort 2009	0.011** (0.005)	0.028** (0.012)	0.018 (0.014)	0.011*** (0.004)	0.032** (0.013)
Census tract characteristics					
Population (ln)	-0.014*** (0.004)	-0.019** (0.008)	-0.012 (0.009)	0.001 (0.003)	-0.004 (0.008)
Percent Asian	0.020 (0.014)	0.074** (0.035)	-0.019 (0.038)	0.001 (0.010)	0.035 (0.036)
Percent White	-0.020 (0.013)	-0.187*** (0.034)	-0.159*** (0.033)	-0.008 (0.008)	-0.116*** (0.031)
Percent Hispanic	0.001 (0.010)	0.041* (0.023)	0.048* (0.029)	0.004 (0.009)	0.069** (0.027)
Percent age 65 and up	0.003 (0.003)	0.286*** (0.079)	-0.051 (0.098)	0.017 (0.029)	0.286*** (0.092)
Percent with high school diploma	0.013 (0.028)	-0.166*** (0.064)	-0.020 (0.070)	-0.011 (0.023)	-0.141** (0.066)
Percent with college degree or higher	0.017 (0.022)	0.124** (0.058)	0.175*** (0.065)	0.016 (0.018)	0.022 (0.061)
Percent foreign born	-0.030 (0.019)	-0.295*** (0.048)	-0.185*** (0.052)	-0.009 (0.015)	-0.111** (0.049)
Employment rate	-0.020 (0.027)	0.084 (0.064)	-0.046 (0.072)	-0.036* (0.021)	-0.077 (0.067)
Average household income (ln)	-0.012 (0.008)	-0.017 (0.019)	-0.018 (0.022)	0.002 (0.007)	0.037* (0.021)
Constant	0.338*** (0.119)	1.019*** (0.285)	0.370 (0.318)	-0.044 (0.101)	0.471 (0.300)

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively. The first-stage sample size is 25,572 for boys and 28,404 for girls. The second-stage sample size is 19,148 for boys and 16,585 for girls.

Table A9: Reduced-Form Estimation, Separate Model, Contraception and SBHC Utilization

	Contraception			SBHC	
	Any	Most effective[1]	Moderately effective[2]	Use at least once	Total claims
Distance					
Distance to large high school	0.000 (0.007)	-0.001 (0.003)	-0.001 (0.007)	0.010** (0.005)	0.180** (0.083)
Distance to new small high school	-0.012** (0.005)	-0.001 (0.002)	-0.012** (0.005)	-0.004 (0.003)	-0.006 (0.040)
Distance to old small high school	0.008*** (0.003)	0.001 (0.001)	0.009*** (0.003)	0.002 (0.002)	0.049** (0.023)
Student characteristics					
Asian	-0.111*** (0.010)	-0.014*** (0.003)	-0.107*** (0.010)	-0.027*** (0.006)	-0.287*** (0.072)
Black	0.021** (0.010)	-0.002 (0.004)	0.021** (0.010)	0.012 (0.008)	0.035 (0.117)
Hispanic	0.027** (0.011)	0.009* (0.005)	0.024** (0.011)	-0.007 (0.008)	-0.081 (0.128)
Age	0.065*** (0.012)	0.020*** (0.005)	0.057*** (0.012)	-0.013 (0.008)	-0.181 (0.129)
Manhattan	0.033** (0.014)	-0.006 (0.006)	0.040*** (0.014)	-0.004 (0.012)	-0.073 (0.190)
Brooklyn	-0.056*** (0.012)	-0.023*** (0.005)	-0.046*** (0.012)	-0.136*** (0.009)	-1.501*** (0.141)
Queens	-0.064*** (0.013)	-0.029*** (0.005)	-0.049*** (0.013)	-0.165*** (0.009)	-1.739*** (0.136)
Cohort 2009	-0.001 (0.010)	0.003 (0.004)	-0.001 (0.010)	-0.015** (0.007)	-0.049 (0.103)
Census tract characteristics					
Population (ln)	-0.000 (0.008)	-0.003 (0.003)	0.002 (0.008)	-0.006 (0.006)	-0.105 (0.085)
Percent Asian	-0.064** (0.031)	-0.031*** (0.011)	-0.049 (0.030)	-0.126*** (0.018)	-1.011*** (0.266)
Percent White	-0.142*** (0.019)	-0.030*** (0.007)	-0.131*** (0.019)	-0.100*** (0.011)	-0.855*** (0.154)
Percent Hispanic	0.122*** (0.023)	0.019** (0.009)	0.114*** (0.023)	0.046*** (0.017)	0.379 (0.237)
Percent age 65 and up	0.110 (0.077)	-0.018 (0.029)	0.137* (0.076)	-0.111** (0.049)	-1.303* (0.682)
Percent with high school diploma	0.047 (0.057)	-0.023 (0.023)	0.060 (0.057)	0.014 (0.039)	-0.247 (0.573)
Percent with college degree or higher	0.181*** (0.055)	0.041** (0.020)	0.162*** (0.054)	0.100*** (0.035)	0.843 (0.547)
Percent foreign born	-0.128*** (0.035)	0.005 (0.014)	-0.141*** (0.034)	0.038 (0.024)	0.097 (0.373)
Employment rate	0.029 (0.062)	-0.054** (0.027)	0.057 (0.061)	-0.030 (0.043)	-0.098 (0.666)
Average household income (ln)	-0.046 (0.019)	0.003 (0.008)	-0.045*** (0.019)	-0.022* (0.013)	-0.288 (0.201)
Constant	-0.194 (0.278)	-0.230** (0.114)	-0.135 (0.275)	0.712*** (0.191)	8.860*** (3.111)
Multivariate F of excluded instruments [p-value]	0.020	0.740	0.017	0.085	0.006

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively. The sample size is 16,585. [1] Sterilization and long-acting reversible methods. [2] Injectables, oral pills, patch, ring and diaphragm.

Table A10: TS2SLS Estimation, Separate Model, Contraception and SBHC Utilization

	Contraception			SBHC	
	Any	Most effective[1]	Moderately effective[2]	Use at least once	Total claims
Enrollment					
New Small School	0.316** (0.128)	0.027 (0.042)	0.310** (0.127)	0.136** (0.071)	1.039 (0.965)
Old Small School	-0.206 (0.145)	-0.023 (0.049)	-0.219 (0.144)	-0.107 (0.085)	-3.266*** (1.118)
Student characteristics					
Asian	-0.110*** (0.013)	-0.014*** (0.004)	-0.107*** (0.013)	-0.028*** (0.007)	-0.387*** (0.090)
Black	-0.003 (0.015)	-0.004 (0.006)	-0.003 (0.015)	0.002 (0.010)	-0.062 (0.145)
Hispanic	-0.000 (0.015)	0.006 (0.006)	-0.003 (0.015)	-0.019* (0.010)	-0.252* (0.149)
Age	0.056*** (0.013)	0.019*** (0.006)	0.048*** (0.012)	-0.017** (0.009)	-0.228* (0.130)
Manhattan	0.065*** (0.021)	-0.002 (0.008)	0.072*** (0.021)	0.010 (0.015)	0.235 (0.234)
Brooklyn	-0.070*** (0.021)	-0.024*** (0.008)	-0.061*** (0.021)	-0.145*** (0.014)	-1.818*** (0.201)
Queens	-0.053** (0.022)	-0.029*** (0.008)	-0.040* (0.022)	-0.160*** (0.013)	-1.856*** (0.193)
Cohort 2009	0.016 (0.013)	0.005 (0.005)	0.016 (0.013)	-0.006 (0.008)	0.032 (0.119)
Census tract characteristics					
Population (ln)	-0.004 (0.008)	-0.004 (0.003)	-0.010 (0.008)	-0.007 (0.006)	-0.090 (0.086)
Percent Asian	-0.038 (0.034)	-0.028** (0.011)	-0.022 (0.033)	-0.117*** (0.019)	-0.905*** (0.276)
Percent White	-0.111*** (0.030)	-0.027*** (0.010)	-0.101*** (0.030)	-0.089*** (0.017)	-0.919*** (0.235)
Percent Hispanic	0.087*** (0.027)	0.016 (0.010)	0.079*** (0.027)	0.029 (0.019)	0.089 (0.270)
Percent age 65 and up	0.084 (0.090)	-0.021 (0.033)	0.108 (0.089)	-0.131** (0.057)	-1.977*** (0.788)
Percent with high school diploma	-0.003 (0.064)	-0.027 (0.025)	0.008 (0.064)	-0.013 (0.043)	-0.782 (0.626)
Percent with college degree or higher	0.135** (0.060)	0.038* (0.021)	0.117** (0.059)	0.072* (0.037)	0.448 (0.578)
Percent foreign born	-0.105** (0.047)	0.007 (0.018)	-0.120** (0.047)	0.043 (0.031)	-0.221 (0.471)
Employment rate	0.028 (0.066)	-0.054** (0.027)	0.057 (0.065)	-0.029 (0.045)	0.080 (0.697)
Average household income (ln)	-0.036* (0.021)	0.004 (0.008)	-0.037* (0.020)	-0.015 (0.013)	-0.192 (0.208)
Constant	-0.148 (0.297)	-0.220* (0.117)	-0.079 (0.294)	0.705*** (0.200)	9.464*** (3.262)

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively. The first-stage sample size is 28,404. The second-stage sample size is 16,585. [1] Sterilization and long-acting reversible methods. [2] Injectables, oral pills, patch, ring and diaphragm.

Table A11: Reduced-Form Estimation, Separate Model, Risky Behavior and Mental Health, Omitted Bias

	Boy					Girl				
	Explanatory Variables Controlled					Explanatory Variables Controlled				
	X^s	X^s and X^b	X^s, X^b and census income	X^s, X^b and X^c (w/o income)	Main (X^s, X^b and X^c)	X^s	X^s and X^b	X^s, X^b and census income	X^s, X^b and X^c (w/o income)	Main (X^s, X^b and X^c)
Pregnancy										
Distance to large high school	-	-	-	-	-	0.002 (0.007)	-0.003 (0.008)	-0.002 (0.008)	-0.010 (0.007)	-0.010 (0.008)
Distance to new small high school	-	-	-	-	-	-0.016*** (0.006)	-0.015** (0.006)	-0.012** (0.006)	-0.004 (0.006)	-0.003 (0.006)
Distance to old small high school	-	-	-	-	-	-0.001 (0.003)	0.006* (0.003)	0.007** (0.003)	0.014*** (0.003)	0.014*** (0.003)
Violent injury										
Distance to large high school	0.001 (0.003)	-0.002 (0.003)	-0.001 (0.003)	-0.002 (0.003)	-0.002 (0.003)	0.001 (0.002)	-0.000 (0.002)	0.000 (0.002)	-0.001 (0.002)	-0.001 (0.002)
Distance to new small high school	-0.006*** (0.002)	-0.006*** (0.002)	-0.005*** (0.002)	-0.004* (0.002)	-0.003 (0.002)	-0.006*** (0.001)	-0.006*** (0.001)	-0.005*** (0.001)	-0.004*** (0.002)	-0.004*** (0.002)
Distance to old small high school	-0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.003** (0.001)	0.003** (0.001)	-0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.001 (0.001)	0.001 (0.001)
Mental health diagnoess										
Distance to large high school	-0.005 (0.007)	-0.015** (0.007)	-0.014** (0.007)	-0.017** (0.007)	-0.017** (0.007)	0.006 (0.007)	-0.006 (0.007)	-0.005 (0.007)	-0.010 (0.008)	-0.011 (0.008)
Distance to new small high school	0.001 (0.005)	-0.003 (0.005)	-0.003 (0.005)	0.005 (0.005)	0.005 (0.005)	-0.014** (0.005)	-0.015*** (0.006)	-0.014** (0.006)	-0.008 (0.006)	-0.009 (0.006)
Distance to old small high school	-0.007*** (0.002)	0.002 (0.003)	0.003 (0.003)	0.007** (0.003)	0.007** (0.003)	-0.007*** (0.003)	0.004 (0.003)	0.004 (0.003)	0.009*** (0.003)	0.009*** (0.003)

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively. The sample size is 19,148 for boys and 16,585 for girls. X^s: student characteristics. X^b: borough fixed effects. X^c: census tract characteristics. See the methodology section for more detail on the notations.

Table A12: Reduced-Form Estimation, Separate Model, Risky Behavior and Mental Health, Robustness Checks

	Boy			Girl		
	Residency in September of Grade 8	**Residency in September of Grade 9**	**Main, Residency in January of Grade 8**	**Residency in September of Grade 8**	**Residency in September of Grade 9**	**Main, Residency in January of Grade 8**
Pregnancy						
Distance to large high school	-	-	-	-0.012 (0.008)	-0.014* (0.008)	-0.010 (0.008)
Distance to new small high school	-	-	-	-0.002 (0.006)	-0.006 (0.006)	-0.003 (0.006)
Distance to old small high school	-	-	-	0.014*** (0.003)	0.014*** (0.003)	0.014*** (0.003)
Violent injury						
Distance to large high school	-0.002 (0.003)	-0.001 (0.003)	-0.002 (0.003)	-0.001 (0.003)	-0.001 (0.002)	-0.001 (0.002)
Distance to new small high school	-0.004* (0.002)	-0.004** (0.002)	-0.003 (0.002)	-0.005*** (0.002)	-0.004** (0.002)	-0.004*** (0.002)
Distance to old small high school	0.003** (0.001)	0.002 (0.001)	0.003** (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Mental health diagnosis						
Distance to large high school	-0.016** (0.007)	-0.015** (0.007)	-0.017** (0.007)	-0.013 (0.008)	-0.021*** (0.007)	-0.011 (0.008)
Distance to new small high school	0.004 (0.006)	0.003 (0.005)	0.005 (0.005)	-0.010 (0.006)	-0.007 (0.006)	-0.009 (0.006)
Distance to old small high school	0.007** (0.003)	0.006** (0.003)	0.007** (0.003)	0.010*** (0.003)	0.009*** (0.003)	0.009*** (0.003)
Sample size	18544	19395	19148	16021	16942	16585

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively.

Table A13: Descriptive Statistics of Characteristics of Students and Census Tracts

	Education Data (a)		Health Data (b)	
	Boys	**Girls**	**Boys**	**Girls**
Student characteristics				
Distance to large high school	1.239 (1.011)	1.228 (0.997)	0.744 (0.523)	0.750 (0.537)
Distance to new small high school	0.988 (0.781)	0.961 (0.763)	0.913 (0.753)	0.911 (0.752)
Distance to old small high school	1.696 (1.250)	1.670 (1.241)	1.797 (1.582)	1.830 (1.621)
Asian	0.207	0.184	0.142	0.145
Black	0.302	0.333	0.298	0.300
Hispanic	0.408	0.410	0.186	0.186
Age	15.42 (0.580)	15.35 (0.532)	15.15 (0.288)	15.15 (0.289)
Manhattan	0.105	0.107	0.134	0.131
Brooklyn	0.365	0.373	0.395	0.392
Queens	0.326	0.309	0.222	0.222
Cohort 2009	0.436	0.431	0.484	0.477
Census Tract Characteristics				
Population (ln)	8.363 (0.513)	8.370 (0.508)	8.387 (0.515)	8.399 (0.499)
Percent Asian	0.133 (0.175)	0.125 (0.172)	0.112 (0.170)	0.115 (0.174)
Percent White	0.191 (0.237)	0.180 (0.231)	0.208 (0.273)	0.199 (0.267)
Percent Hispanic*	0.350 (0.255)	0.354 (0.255)	0.396 (0.263)	0.399 (0.264)
Percent age 65 and up	0.111 (0.052)	0.109 (0.051)	0.102 (0.051)	0.102 (0.050)
Percent with high school diploma	0.359 (0.114)	0.360 (0.114)	0.359 (0.111)	0.357 (0.111)
Percent with college degree or higher	0.218 (0.132)	0.214 (0.132)	0.196 (0.134)	0.195 (0.131)
Percent foreign born	0.414 (0.157)	0.408 (0.158)	0.378 (0.163)	0.382 (0.161)
Employment rate	0.624 (0.092)	0.620 (0.094)	0.595 (0.100)	0.596 (0.098)
Average household income (ln)	10.95 (0.381)	10.93 (0.383)	10.83 (0.369)	10.83 (0.366)
Sample size	25,572	28,404	19,148	16,585

Note: Standard deviations for the continuous variables are shown in parentheses. The omitted baseline category of race is white and the omitted baseline borough is the Bronx. Age is the age on September 1st, 2009 (2010) for 2009 (2010) cohort. *Note that the Medicaid system and the DOE use different approaches to race/ethnicity classification.

Table A14: TS2SLS Estimation, Separate Model, Risky Behavior and Mental Health, Robustness Checks

	Boy			Girl		
	Birth Cohort in the 1st Stage	**Outcome by Age 17 in the 2nd Stage**	**Main**	**Birth Cohort in the 1st Stage**	**Outcome by Age 17 in the 2nd Stage**	**Main**
Pregnancy						
New small	-	-	-	0.136 (0.139)	0.027 (0.118)	0.150 (0.142)
Old small	-	-	-	-0.523*** (0.153)	-0.478*** (0.135)	-0.553*** (0.159)
Violent injury						
New small	0.074* (0.044)	0.009 (0.031)	0.076* (0.042)	0.090** (0.037)	0.045* (0.027)	0.094** (0.037)
Old small	-0.058 (0.071)	-0.060 (0.056)	-0.065 (0.074)	0.006 (0.048)	-0.017 (0.034)	0.003 (0.050)
Mental health diagnoses						
New small	-0.089 (0.115)	-0.074 (0.098)	-0.063 (0.109)	0.206 (0.131)	0.210* (0.119)	0.220* (0.134)
Old small	-0.395** (0.176)	-0.300* (0.165)	-0.415** (0.184)	-0.183 (0.142)	-0.215 (0.133)	-0.200 (0.148)

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively. The sample size is 19,148 for boys and 16,585 for girls.

Table A15: IV Results of the Effect of Attending Small Schools on Overweight/Obesity

	2SLS		TS2SLS
	Overweight	**Obesity**	**Overweight problem**
	Boys		
Small	0.040 (0.039)	0.014 (0.036)	0.053 (0.085)
New Small	0.018 (0.036)	-0.018 (0.033)	0.002 (0.087)
Old Small	0.273*** (0.076)	0.227*** (0.072)	0.110 (0.140)
Sample size	25,572		19,148
	Girls		
Small	0.033 (0.035)	-0.019 (0.032)	0.054 (0.092)
New Small	-0.029 (0.037)	-0.057* (0.033)	-0.018 (0.111)
Old Small	0.215*** (0.053)	0.121** (0.048)	0.070 (0.122)
Sample size	28,404		16,585

Note: Robust standard errors are reported in parentheses. *, **, and *** indicates statistical significance at 10%, 5% and 1% confidence level, respectively.