Title
Essays on the Distribution and Effectiveness of Educational Resources

Author
Lafortune, Julien

Publication Date
2018

Peer reviewed|Thesis/dissertation

Essays on the Distribution and Effectiveness of Educational Resources

by

Julien Lafortune

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jesse Rothstein, Chair
Professor David Card
Associate Professor Patrick Kline
Professor Hilary Hoynes

Spring 2018

Essays on the Distribution and Effectiveness of Educational Resources

Abstract

Essays on the Distribution and Effectiveness of Educational Resources

by

Julien Lafortune

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Jesse Rothstein, Chair

Public schools are the foundation of the American educational system, and the public K-12 education system is often idealized as being one of, if not *the*, "great equalizer" in American society. Despite this lofty ideal, educational resources are not equally distributed and there are tremendous discrepancies in student outcomes, both between and within schools. This dissertation examines the implications of such discrepancies in the provision of school resources, both financial and otherwise.

In Chapter 1, I link data on new facility openings to administrative student and real estate records in Los Angeles Unified School District (LAUSD) to provide new evidence on the importance of school capital expenditures for students and neighborhoods. Since 1997, LAUSD has built and renovated hundreds of schools as a part of the largest public school construction project in US history. Using an event-study design that exploits variation in the timing of new school openings, I find that spending 4 years in a new school increases test scores by 10% of a standard deviation in math, and 5% in English-language arts. This in part reflects non-cognitive improvements: Treated students attend four additional days per school year and teachers report greater effort. Effects do not appear to be driven by changes in class size, teacher composition, or peer composition, but reduced overcrowding plays a role. House prices increase by 6% in neighborhoods that receive new schools. Real estate capitalization is greater than program cost, implying a willingness-to-pay in the range of 1.2 to 1.6 per dollar spent.

In Chapter 2, I study the impact of post-1990 school finance reforms, during the so-called "adequacy" era, on absolute and relative spending and achievement in low-income school districts. Using an event study research design that exploits the apparent randomness of reform timing, I show that reforms lead to sharp, immediate, and sustained increases in spending in low-income school districts. Using representative samples from the National Assessment of Educational Progress, I find that reforms cause increases in the achievement of students in these districts, phasing in gradually over the years following the reform. The implied effect of school resources on educational achievement is large.

In Chapter 3, I consider educational discrepancies of a different sort: tracking and the segmentation of students into different curricular paths. In most U.S. schools, a significant track diversion occurs in 8th grade: higher- achieving students are tracked into Algebra, while lower-achieving students take Algebra in 9th grade or later. Using a fuzzy regression discontinuity design around prior-year test score proficiency thresholds, I examine the impact of tracking into Algebra in 8th grade rather than in high school. For students near the 80th percentile in the 7th grade state math distribution, advanced track enrollment leads to large increases in mathematics course-taking, AP course participation, and college entrance exam scores. However, for students near the 30th percentile in the 7th grade math distribution, advanced track enrollment is associated with large decreases in Algebra performance, with little indication of any longer-term gains. Results show that advanced math tracking in secondary schools has heterogeneous impacts on students based on prior math achievement. Expanding access to advanced math courses among high-achieving but not low-achieving students could yield large improvements in mathematics skills and college preparedness.

To my mother, Sharyl Politi, whose creativity, joy, and affection have shaped the person I am today; my father Stéphane Lafortune, whose intellectual curiosity and unconditional support provided the foundation for me to pursue my aspirations; and my wife and best friend, Emily Lafortune, who supported me every step of the way, and whose impact on my life goes beyond words.

Contents

List of Figures

List of Tables

Acknowledgments

A great number of people have helped my academic career in ways large and small, and I would not have made it through graduate school without these many individuals who lifted me along the way. In particular, I am forever indebted to my advisor, Jesse Rothstein, for his tireless support and encouragement, and for the countless opportunities that he presented to me. His continued support over these past six years has had a profound impact on my progression as a researcher. I wish to also thank my dissertation committee, David Card, Hilary Hoynes, and Pat Kline, for continued and invaluable guidance. I have also benefitted immensely from the support and advice of Chris Walters, Danny Yagan, Emmanuel Saez, Alex Gelber, Steven English, and Tom Rubin. I also wish to thank Bruce Fuller and Jeff Vincent in particular for many helpful conversations and for data assistance. I thank Jack Bragg, Divya Dhar, and Sophie Stimson for outstanding research assistance. I am grateful to the Los Angeles Unified School District, the National Center for Education Statistics, and the North Carolina Education Research Data Center for data access and support. This work was supported by a National Academy of Education / Spencer Foundation Dissertation Fellowship. Finally, I wish to thank my fellow classmates, from whom I have learned so much, in particular, Dmitri Koustas, Jen Kwok, Tomas Monnarez, Waldo Ojeda, Raffaele Saggio, Jon Schellenberg, David Schönholzer, Avner Shlain, and Ferenc Szucs.

Chapter 1

Do School Facilities Matter? Measuring the Effects of Capital Expenditures on Student and Neighborhood Outcomes

1.1 Introduction

There has been a longstanding debate among educational policymakers and researchers over the productivity of school spending, with little consensus as to whether and under what circumstances increased expenditures improve student outcomes (e.g. Hanushek (2006)). Much of the empirical literature has focused on instructional inputs, with considerably less attention paid to the role of capital expenditures. However, capital expenditures comprise an important component of US public school spending: in the 2013-2014 school year roughly 8% of total expenditures went towards direct capital outlays, and an additional 9% was spent on operation and maintenance of existing facilities and equipment (McFarland et al., 2017). Despite the magnitude of this spending, one-quarter of U.S. public schools are in fair or poor condition[1] (Alexander and Lewis, 2014), and estimates of the funding required to address substandard facilities conditions range in the hundreds of billions nationally (Crampton et al. (2001); Arsen and Davis (2006); M. Filardo (2016)). Substandard facilities are thought to be a particular problem in low-income districts, which have schools that are more likely to be in fair or poor condition and/or rely on temporary rather than permanent buildings (Alexander and Lewis, 2014), and on average spend 15% less on capital investments than do high-income districts.

In this paper, we address three fundamental unanswered questions. First, do increases

[1] "Fair" condition means that the facility meets minimum needs, but requires frequent maintenance and has other limitations. "Poor" means that the facility does not meet minimal requirements for normal school operation.

in school capital expenditures improve student outcomes? Second, how are these additional expenditures capitalized into local real estate prices? And third, are locally-financed school capital improvements welfare-improving, taking account of the taxes needed to pay for them? We investigate these questions in the context of the largest public school capital construction program in U.S. history. From 2002 to 2017, Los Angeles Unified School District (LAUSD) constructed over 150 new schools and renovated hundreds more. Using administrative student and property sale records, we provide precise and comprehensive estimates of the causal impact of school facility expenditures on student outcomes and neighborhood house prices. Finally, we use these estimates to evaluate the welfare consequences of the construction program for LAUSD residents.

The empirical literature on capital expenditures offers little guidance with regard to these questions. Several studies find no or imprecise effects of capital expenditures on student achievement (see Cellini et al. (2010), Bowers and Urick (2011), Goncalves (2015), Martorell et al. (2016)), while others find some evidence of positive impacts on student achievement, often only in reading and English-language arts (Welsh et al. (2012), Neilson and Zimmerman (2014), Hong and Zimmer (2016), Conlin and Thompson (2017)). Despite inconclusive evidence in the literature and general skepticism among economists, resource-based capital expenditure programs continue to be used by policymakers at the state and local level as tools to improve schools and reduce achievement gaps.

We find robust evidence that attending newly constructed schools in LAUSD leads to large, significant gains in cognitive and non-cognitive student outcomes. Relying on within-student variation in the timing of exposure to new facilities, we estimate that spending four years in a new school facility leads to a 0.1 standard deviation increase in standardized math scores and a 0.05 standard deviation increase in English-language arts (ELA) scores. In addition, students who attend newly constructed schools attend on average four additional days per academic year, and score 0.06 standard deviations higher on teacher-reported measures of student effort. We provide additional evidence of smaller indirect test score and attendance gains for students at existing facilities who experienced reductions in overcrowding induced by peer outflows to newly constructed schools. These indirect effects allow us to decompose the relative contribution of overcrowding reductions to the observed student gains at newly constructed schools. Examining the mechanisms through which these effects are mediated, we conclude that the majority of the effects were due to improved facility quality, while reduced overcrowding was also an important factor. We find no evidence that student sorting, changes in teacher quality, changes in peer quality, or changes in teacher-pupil ratios were contributing factors.

We find significant valuation of school quality improvements in the real estate market. Using administrative records on property sales, we find that house prices increase by 6% in neighborhoods that receive new school facilities. Effects accumulate in the first three years following construction, with little evidence of anticipatory house price increases. House prices in nearby neighborhoods are mostly unaffected, although we find some evidence of negative house price spillovers for properties very close to, but outside new school catchment areas. We use a simple model to assess the household valuation of a redistributive public education capital expenditure program. From this model, we derive an expression with a direct difference-in-differences analogue to assess the valuation of the spending program using relative price changes between neighborhoods. Reduced-form estimates of the change in relative house prices imply a household willingness to pay ratio in the range of 1.2 to 1.6 per dollar spent, providing evidence that the total real estate capitalization resulting from the program exceeds the total program cost, and that educational capital had been under-provided.[2]

Our study contributes to a few related literatures. First, we provide precise, large-sample estimates of student-level effects from facility improvements. We estimate direct effects on treated students, as well as indirect effects on students who are affected by cohort-peer outflows from existing to new school facilities. Most prior studies examine effects of capital expenditure programs on district-level average outcomes, often finding mixed and imprecise estimates of effects on student outcomes (Cellini et al. (2010), Martorell et al. (2016), Hong and Zimmer (2016), Conlin and Thompson (2017), Goncalves (2015)). These studies do not measure effects on directly treated students, and are generally underpowered to detect modest but meaningful effects. Most school districts consist of at least several school campuses, and thus programs to construct new schools or renovate existing ones often only affect a subset of students.[3] Other studies have looked at longer-run impacts of school construction programs that expand access to education (e.g. Duflo (2001), Duflo (2004)), measuring the effects of more generally increasing human capital accumulation.

The study most related to ours is Neilson and Zimmerman (2014). They examine a similar construction boom in New Haven, Connecticut, and using student-level administrative data find evidence of positive effects on reading but not math scores several years after school construction. Many fewer students and schools were impacted than in the LAUSD program. The scale of the LAUSD program allows us to carefully decompose effects and examine specific mechanisms, such as teacher quality, student/peer sorting, class sizes, and other school-level changes. In addition, we examine outcomes of students who experience cohort-level peer outflows induced by new school openings, providing new evidence of indirect effects of new facilities through reduced overcrowding at nearby existing school facilities.

Second, we contribute to the literature estimating the capitalization of school quality in the real estate market. We provide some of the first large-sample evidence of localized house price capitalization of dynamic changes in school quality. Much of the work in this literature has estimated the capitalization of *static differences* in school quality, and thus

[2]The efficient choice of local public expenditures is typically defined by the "Samuelson condition" (Samuelson, 1954): spending levels will equate the marginal rate of transformation of the public good and the sum of the marginal rates of substitution between numeraire consumption and the public good. Here, a WTP ratio greater (less) than one suggests under-provision (over-provision) of local educational capital expenditures. It is worth noting, however, that educational inputs are not pure public goods; schooling is both excludable and subject to congestion. In Sections 1.3 and 1.8 we provide a more detailed discussion of program efficiency and welfare implications.

[3]See Figure 1.11 for a comparison of the estimated test score effects by per-pupil spending change for prior studies of school capital expenditures.

does not provide direct estimates of how *changes* in school quality are valued in the real estate market. Several papers, most notably Black (1999) and Bayer et al. (2007) exploit boundary discontinuities within narrowly defined neighborhoods to estimate the market valuation of school quality. Other papers have used variation across district boundaries (e.g. Barrow (2002); Barrow and C. E. Rouse (2004)), within-district boundary changes (e.g. Ries and Somerville (2010); Collins and Kaplan (2017)), school "report-card" grades (Figlio and Lucas, 2004), and public reporting of teacher value-added scores (Imberman and Lovenheim, 2016). Static differences in house prices between school zones include parental preferences for school quality, peer quality, and racial composition; these estimates are less informative for understanding the dynamic effects from policy changes.

A handful of recent papers provide estimates of real estate capitalization of changes in school quality using variation induced by capital expenditure policies, generally finding positive effects after several years (see Cellini et al. (2010); Goncalves (2015); Conlin and Thompson (2017); Neilson and Zimmerman (2014)). We build upon these prior studies by more precisely examining the dynamics of these changes, over both time and space. Moreover, we study a (mostly) locally funded program that was inherently redistributive: local property taxes were raised districtwide to fund new schools in only one-third of neighborhoods. We directly relate our estimates of within-district relative price changes within a simple spatial equilibrium model, allowing us to directly assess the efficiency of program spending.

Finally, we contribute generally to the broad literature and debate over the efficacy and efficiency of resource-based education policies. Economists have long been skeptical of the productivity of such investments (e.g. Hanushek (1997)), although recent studies of state-level school finance reforms have provided evidence that broad based expenditure programs can improve educational outcomes (e.g. Jackson et al. (2016), Lafortune et al. (Forthcoming), Candelaria and Shores (2015), Hyman (Forthcoming)), labor market outcomes (Jackson et al., 2016), and intergenerational mobility (Biasi, 2017). Our study of the LAUSD school construction program provides additional evidence that: (1) school expenditures - even those dedicated to capital costs - can improve student cognitive and non-cognitive outcomes; (2) such programs can induce increases in aggregate real estate prices in excess of program cost.

There are two important caveats to these conclusions. First, as our study focuses on the outcomes of one large district, our results may not generalize to other districts or states. However, many large, urban districts as well as smaller districts serving disadvantaged students face consistently underfunded and worse quality facilities relative to more affluent districts (e.g. M. W. Filardo et al. (2006)). Our study is directly applicable to these contexts. Second, an important feature of the LAUSD program was the reduction of overcrowding and the expansion of available school facilities. We find some evidence of larger gains for students coming from previously overcrowded schools. Our estimates are therefore likely to represent upper bounds on possible effects in districts with stable or declining enrollment seeking to replace, rather than expand the school capital stock. Importantly, overcrowded school facilities are not unique to LAUSD; over 25% of California public schools are designated as overcrowded (Rogers et al., 2009), and thus our results are relevant to many school districts facing similar constraints.

The paper proceeds as follows. In Section 1.2 we detail the context for our study and discuss specific details of the LAUSD program. In Section 1.3 we outline a simple theoretical framework to motivate our analysis and interpretation of house price changes. In Section 1.4 we briefly describe each of the data sources we use. Section 1.5 outlines the empirical specifications and quasi-experimental setup we use to estimate program effects. In Section 1.6 we present the student-level results, and discuss mechanisms and indirect effects. In Section 1.7 we present house price results, and examine potential spatial spillover effects. Section 1.8 provides a discussion of results, and an assessment of the benefits, costs, and welfare implications of the program. Finally, in Section 1.9, we conclude with a brief summary of results, their implications, and their generalizability.

1.2 Context of Study

LAUSD is the second largest school district in the United States, serving 747,009 students at its peak in the 2003-2004 school year. It enrolls roughly 10% of all public K-12 students in California. Like nearly every large urban school district in the US, it is majority-minority, and serves students who are much more disadvantaged than the typical US public school student. The district itself encompasses 26 cities in total, including the City of Los Angeles, as well as other nearby "gateway" cities and some unincorporated areas within Los Angeles County. Some of the more affluent areas in LA County, including Beverly Hills and Santa Monica, operate separate school districts for their residents. Relative to the rest of California, students in LAUSD are underachieving: in 2002 the average student scored roughly 28% of a standard deviation below the state mean in English-Language Arts (ELA) and roughly 21% of a standard deviation below the state average in math.[4]

As of the early 2000s, LAUSDs capital stock had fallen well below current needs. As shown in Figure 1.1, no new schools were opened between 1975 and 1996, and the average student attended a school that was over 60 years old in 2000. Many were in extremely poor condition. In a 1999 review of the facilities practices of LAUSD and other California districts, the California "Little Hoover Commission", an independent oversight body, reprimanded the district for gross mismanagement and noted in particular that LAUSD school facilities were "overcrowded, uninspiring and unhealthy", and that "Researchers have attempted to gauge the link between the quality of school buildings and the quality of learning. In Los Angeles, however, this link is obvious. In some classrooms, there are twice as many children as there are desks." (Terzian, 1999) Classrooms were often non-functional, with broken and missing equipment, and school facilities sometimes lacked adequate restrooms.[5] Inadequate climate control was additionally a major source of distraction; classroom temperatures upwards of 90 degrees fahrenheit were not uncommon. One teacher noted that "... we had roaches, ants,

[4]Scores from the CST ELA exam in grades 2-11, and the CST math exam in grades 2-7.

[5]For example, one high school of nearly 2000 students had only one functioning bathroom.

an air conditioner that barely worked, no sink [...] and barely any storage for classroom materials." (Fuller et al., 2009)

The schools were also severely overcrowded, as the district's enrollment had increased roughly 10% since 1975 (Figure 1.1). Nearly 25,000 students were bused daily to faraway schools to relieve capacity constraints, and roughly half of students attended schools that operated on multi-track calendars that staggered the school year to use the facility year-round and thereby accommodate as many as 50% more students than could be served at any single time. Even with these measures, many schools relied on lower-quality portable classrooms, and even converted gymnasiums, libraries, and computer labs into classroom space. This also severely limited student access to extra-curricular opportunities. Rapid depreciation of facility condition due to continued overuse compounded these issues.

Between 1997 and 2007, voters in Los Angeles approved a series of bonds dedicating over $27 billion in local and state funding to the construction, expansion, and renovation of hundreds of schools. This was the largest public infrastructure program in the U.S. since the interstate highway system (Fuller et al., 2009). The first new school was completed in 2002, and over the next 15 years nearly 150 new school facilities were constructed in LAUSD, totaling over $10 billion in capital expenditures. Many more schools were renovated, modernized, or received additions that increased school capacity. By 2012, over 75,000 students attended a newly constructed school (see Figure 1.12), less than 1% of students remained on a multi-track calendar (see Figure 1.13), overcrowding had been effectively eliminated, and there was no longer widespread busing of students to distant schools.

After the first bond authorization in 1997, the district began by identifying overcrowded schools and attendance areas. Designated search areas were defined for each of these locations, and construction sites were selected from within these areas primarily based on site feasibility (e.g. size, location, accessibility), cost of acquiring land, environmental concerns, and local community engagement. By 2001, nearly all new school sites had been identified, although the process of acquiring land, securing adequate funding, negotiating with local stakeholders, meeting environmental regulations, and designing and constructing schools resulted in a staggered delivery of new facilities over the next decade. It is this plausibly random variation in the timing of openings, induced primarily through idiosyncrasies in the construction process, which we exploit to estimate the effect of new schools on student and neighborhood outcomes. We will provide a comprehensive discussion of our empirical approach and identification assumptions in Section 1.5.

In this paper we focus on new school facilities completed between 2002 and 2012, for which we have detailed project data matched to administrative student data. A database of capital projects in LAUSD, including measures of project cost, size, completion timeline, and location, was constructed from records listed publicly by the LAUSD Facilities Services Division (FSD). The data cover all major projects and new school constructions with a preferred site designated between 1997 and 2011,[6] and include over 500 capital projects

[6]Projects not yet constructed by the end of 2011, but that were already in the planning phase, are included.

totaling nearly \$17 billion in planned or realized spending. We restrict attention only to large new school construction projects, defined as those that created over 100 new seats and/or cost at least \$10 million.[7]

Summary statistics for the new school projects are presented in Table 1.1. In total, there were 143 new schools built as a part of 114 new school campuses. In some cases, a new school campus comprised several new schools, either because the site was combined to house both elementary and middle (or middle and high school students), or because magnet or alternative schools serving the same grade levels were housed on the same campus. The median project cost \$57 million and created about 800 new student seats, with several projects costing in the hundreds of millions of dollars.[8] Projects typically took two years to construct, and were complete roughly 5 years after the site had been designated by the district. In total, the projects we study in our data cost \$9.17 billion (roughly \$6,000 per household or \$15,000 per pupil), the majority being funded from the various local bonds that were passed in and after 1997.

Figure 1.2 shows the time series of educational spending in LAUSD relative to the other nearby districts in LA County. Panel A shows per-pupil capital expenditures, while panel B shows per-pupil instructional expenditures. Capital expenditures in LAUSD and in other LA County school districts increased similarly during the 1990s, and prior to the passage of the first school construction bond in 1997, capital expenditures were slightly lower in LAUSD (roughly \$500 per pupil) than in the rest of LA County (roughly \$750 per pupil). The magnitude of the program is clearly seen in panel A: expenditures rose rapidly in LAUSD during the construction boom, to a peak of nearly \$4000 per pupil in 2009. Capital expenditures increased much less dramatically in other LA County districts until 2005, before declining to roughly the same level in 2012 as in 1990. Conversely, instructional expenditures saw much smaller increases during the new construction boom from 2002-2012, and the relative difference between LAUSD and other LA County schools was essentially unchanged during this period. Overall, the sample period from 2002-2012 was marked by a large increase in capital expenditures, without a meaningful increase in instructional educational expenditures in absolute or relative terms.

Figure 1.3 shows the attendance zones for new and existing school facilities in 2012. As can be seen in the figure, new schools at all levels were concentrated in East Los Angeles, where students are predominantly Hispanic and schools were previously the most overcrowded and in need of repair. Schools in East LA serve students who are socioeconomically disadvantaged; for example, the median school in the areas most heavily affected by new school construction serves a student population where fewer than one-fifth of students have a parent with any level of postsecondary attainment.

[7]We do not examine effects for the small number of projects for school campuses that already existed in the first year of the student sample (e.g. major additions). In a few instances, students show up at a particular new school in either the year before or after the listed completion year; we adjust the completion year to correspond to the student administrative records in these cases.

[8]One controversial high school project, the Robert F. Kennedy Community Schools, cost nearly \$600 million to construct, becoming the most expensive public school ever built.

New schools were filled quickly, typically reaching close to steady state enrollment within 2 years after construction. Students from nearby schools were reassigned based on redrawn school assignment zones to the newly constructed schools. Switching students experienced drastic changes in facility quality: they switched from schools that were on average 70 years old and had substantial physical deficiencies.[9] These student outflows also generated substantial changes in school environments for those students who "stayed behind" at existing facilities. New school facilities enabled the district to reduce overcrowding and eliminate multi-track calendar schedules at both new and nearby existing schools. Our main analyses will focus on the students who switched to new facilities, as induced by the change in school assignment in the year of construction completion. Here, effects are estimated relative to a control group consisting of all other students in LAUSD, most of whom were unaffected by peer outflows to new facilities. Later, in Section 1.6, we will use an analogous identification strategy to examine changes in outcomes for the indirectly treated students who stayed behind at existing schools.

1.3 Theoretical Framework

The LAUSD school construction program induced dramatic changes in the physical and educational environment of district schools. By the end of 2012, roughly one-third of residential properties within district boundaries were assigned to a school zone for a newly constructed school facility at least one level (i.e. elementary, middle, or high school). Improved school facilities can affect students in a number of ways. Reduced overcrowding and improvements in the physical school environment can have direct effects on student learning through reduced distraction[10] and improved health. Better facilities may improve student motivation and effort, leading to indirect improvements in student learning. Improved facilities may also improve teacher motivation and health, as well as help to attract and retain higher quality teachers (e.g. Buckley et al. (2004); Uline and Tschannen-Moran (2008)). Beyond direct and indirect effects on student learning, educational capital investments affect dimensions of the school environment that improve the amenity value of a school to both students and parents. Insofar as parents value improvements in educational quality and the physical amenities of a school campus, local real estate prices will respond to these changes. To fully evaluate the many potential impacts induced by the construction program, we use local changes in equilibrium housing prices to identify revealed preferences for educational spending changes.

The school construction program was funded primarily through local bond initiatives, which increased property taxes throughout the district to fund new school constructions

[9] Among switching students, the average ratio of the cost of facility deficiencies to current replacement value of prior schools was 0.4.

[10] Student distraction due to externalities from disruptive peers has been proposed as a motivation for class size effects (Lazear, 2001). Poor facility condition could cause similar disruptions that impede classroom learning.

only in a subset of district neighborhoods. Typical analysis of the valuation of local public goods relates changes in real estate prices to changes in the provision and tax-price of public goods. When educational spending increases are valued more than associated changes in taxes, real estate prices will rise, and vice versa. Importantly, however, only a subset of households receive additional school spending (in the form of capital spending on new school construction), meaning that the implied valuation will necessarily be asymmetric: neighborhoods that pay increased taxes but do not receive new school facilities should see prices fall, whereas prices will rise in neighborhoods in areas that receive new facilities to the extent that the additional spending in valued greater than the loss in consumption induced by the tax increase. We formalize this notion by examining the comparative statics of a such a tax and expenditure change within a simple hedonic equilibrium model, borrowing heavily from the models presented in Brueckner (1979), Barrow and C. E. Rouse (2004) and Cellini et al. (2010).

We begin by assuming there are N households, who derive utility from school amenities A_j, and consumption c: $U_j(A_j, c)$. Households can live in one of two neighborhoods: $j \in \{0,1\}$. Households in neighborhood 1 receive new school facility spending, while households in neighborhood 0 do not. Denote the number of households in each neighborhood as N_j. New schools are funded by a tax τ on households, and the local government faces the budget constraint $R = \tau N$. The local government spends all of the revenues in neighborhood 1, and thus the per capita change in school funding (denoted R_j) is:

$$\begin{aligned} R_1(\tau) &= \frac{N}{N_1}\tau \\ R_0(\tau) &= 0 \end{aligned}$$

To understand how the spending policy affects the level of school amenities, it is helpful to write the school amenity value as a function of tax expenditures: $A_j = A(R_j(\tau))$. Households receive income y and face the budget constraint $c = y - \tau - p_j$ where p_j is the rental price of housing. We can therefore write the household's indirect utility function as $V(A_j(R_j(\tau)), y - \tau - p_j)$.

When neighborhood j provides higher utility than alternatives, willingness to pay for housing there will be higher, prices will therefore be bid up. With homogeneous households, the equilibrium market price of housing will equalize utility in all neighborhoods.[11] A household's willingness to pay, or "bid", for a given neighborhood is therefore implicitly defined by function $P_j = P(\tau)$. Using the implicit function theorem, we can derive the change in neighborhood house prices, for a marginal increase in τ:

[11] This is true in equilibrium because if a household would achieve higher utility elsewhere, it would move. More generally, if we were to allow heterogeneity in preferences and/or income, the market price of a neighborhood would be equal to the bid of the marginal consumer, and marginal households with the same preferences and income would achieve identical utility.

$$\frac{\partial P_j}{\partial \tau} = \frac{\frac{\partial V_j}{\partial A_j}}{\frac{\partial V_j}{\partial c}} \frac{\partial A_j}{\partial R_j} \frac{\partial R_j}{\partial \tau} - 1 \tag{1.1}$$

Equation (1.1) shows that the change in prices for a tax increase is a function of the marginal rate of substitution between the educational amenity and consumption, the marginal product of educational amenities with respect to educational expenditures, and the concentration of total tax receipts spent in a given neighborhood. In neighborhood 0, where $R_0 = 0$, $\frac{\partial R_0}{\partial \tau} = 0$. In neighborhood 1, $R_1 = \frac{N}{N_1}\tau$, so $\frac{\partial R_1}{\partial \tau} = \frac{N}{N_1}$. Therefore we have:

$$\begin{aligned}
\frac{\partial P_0}{\partial \tau} &= -1 \\
\frac{\partial P_1}{\partial \tau} &= \frac{N}{N_1}\left[\frac{\frac{\partial V_j}{\partial A_j}}{\frac{\partial V_j}{\partial c}} \frac{\partial A_j}{\partial R_j}\right] - 1
\end{aligned}$$

Intuitively, as households in neighborhood $j = 0$ receive no additional educational expenditures, their marginal willingness to pay is exactly equal to the negative of the tax increase. For neighborhoods that receive the additional spending, their willingness to pay is equal to the product of the MRS and the marginal product of educational amenities with respect to expenditures, multiplied by the per-capita increase in expenditures, minus 1. Taking the difference in the two price changes yields:

$$\left(\frac{\partial P_1}{\partial \tau} - \frac{\partial P_0}{\partial \tau}\right) = \frac{N}{N_1}\left[\frac{\frac{\partial V_j}{\partial A_j}}{\frac{\partial V_j}{\partial c}} \frac{\partial A_j}{\partial R_j}\right] \tag{1.2}$$

Equation (1.2) shows that for a one unit increase in τ, relative prices will rise by the concentration of spending multiplied by the marginal valuation of the additional educational expenditures. For example, if households are evenly split between the two neighborhood types, relative prices will rise by two times the marginal valuation of the additional expenditures for households. When spending is at the efficient level, i.e. when the "Samuelson condition" holds, the aggregate marginal rates of substitution over all households will equal the marginal rate of transformation, and relative prices in neighborhoods that get public investments will rise by the concentration of spending per tax dollar: $\left(\frac{\partial P_1}{\partial \tau} - \frac{\partial P_0}{\partial \tau}\right) = \frac{N}{N_1}$. If prior spending levels were inefficiently low (i.e., if the marginal rate of transformation of funding into amenity value was higher than the marginal rate of substitution between amenities and consumption) and educational facilities had been under-provided, $\left(\frac{\partial V_j}{\partial A_j} / \frac{\partial V_j}{\partial c}\right)\frac{\partial A_j}{\partial R_j}$ will be greater than one and prices will rise by greater than $\frac{N}{N_1}$, as marginal households value the increase in expenditures more than the forgone consumption. Alternatively, a relative price change of less than $\frac{N}{N_1}$ implies that the additional spending is inefficiently high, and that there had been over-provision of educational facilities.[12] Equation (1.2) therefore motivates an evaluation of the efficiency of the construction program using *relative price changes* between neighborhoods that received new schools and those that did not. Difference-in-differences estimates of price changes in response to school constructions will approximate (2) and provide a useful benchmark for evaluation of the program, which we will return to in Section 1.8.

Changes in real estate prices are informative about the product of the MRS between educational amenities and consumption, and the marginal product of additional capital expenditures. Examining the direct impacts of capital expenditures on student outcomes allows us to further understand $\partial A_j/\partial R_j$, the productivity of additional school resources. Assuming all of the amenity value of new schools comes through test score improvements, estimates of the treatment effect of attending new schools on test scores can be directly interpreted as a (non-marginal) approximation of this marginal product. Under this assumption, a direct comparison of difference-in-differences estimates of (2) and estimates of $\partial A_j/\partial R_j$ using student data allow us to recover plausible estimates the MRS for improvements in school quality for marginal parents. However, as discussed earlier, test scores likely only capture a portion of the amenity value associated with new school facilities; any such estimates will therefore represent upper bounds on the parental valuation of test score improvements.[13]

1.4 Data

Student data

To study the effects of increased capital expenditures on student outcomes, we use administrative records from LAUSD from the 2002-2003 school year to the 2012-2013 school year. Every student who attended LAUSD during this time period is included, and the data allow for longitudinal links across years for students who remain in the district. These data provide one record per student-year with information on student grades, test scores, demographics, attendance records, school assignment, and teacher assignments.[14] Demographics include gender, race, language spoken at home, parental education, and eligibility for free or reduced price lunch. Students in grades 2-11 are administered the California Standards

[12] Over- or under-provision of educational facilities may result from allocative inefficiencies, where the district provides an inefficient level of facilities, or from productive inefficiencies, where the district does not minimize costs. In this paper we will abstract from this distinction when evaluating the overall efficiency of the expenditure program.

[13] Recent work using revealed preferences from school choice applications suggests no relationship between parental school preferences and school productivity (measured by school test score value-added), once peer quality is taken into account (Abdulkadiroğlu et al., 2017). This need not be inconsistent with our findings: to the extent that parents value non-test score school improvements and/or new schools are a more salient signal of school treatment effects, increased local education expenditures would generate positive relative price changes in equation (1.2).

[14] For some years and grade levels, data are included from both the fall and spring semesters; we collapse these data to the annual level for comparability.

Test (CST) annually in math and English-language arts (ELA). In each of grades 2-7, students take the same grade-level math exam; however, beginning in grade 8 the particular test depends on the student's particular math course enrollment. For the CST ELA exam, exams do not depend on a student's enrollment.[15] To ensure comparability of scores across students, we focus only on CST math scores for grades 2-7 and CST ELA scores for grades 2-11. Test scores are normalized relative to the California-wide mean and standard deviation reported in the California Standardized Testing and Reporting (STAR) documentation provided by the California Department of Education.

Total annual attendance, measured in days, is recorded for each student.[16] For elementary school students, report card data contain teacher-reported measures of both achievement and effort in different classroom subjects. These are reported on an ordinal scale from 1 to 4 for over one dozen subjects. Scores pertaining to student effort are averaged within each student-year record to construct a "effort" index. Scores pertaining to student achievement or proficiency are averaged within each student-year record to construct a teacher-reported "marks" index. These indices are then normalized to have mean zero and a standard deviation of one within each grade-year cell.

Data on teacher education, experience, age, and gender are available in all years, except 2009 and 2011. Teacher identifiers are also available for all years in the student data, and teachers can be linked longitudinally using unique teacher IDs. However, teacher IDs are scrambled between the secondary student and teacher demographic datasets, meaning that secondary school students are less reliably linked to teacher demographic variables. In the student data, each elementary record contains a single teacher identifier. Teacher-student links for secondary school are constructed using student-level course data. Class size is constructed for elementary school students by measuring the total number of students associated with a particular teacher ID in a given year. For students in secondary school we do not compute class size as direct classroom identifiers are unavailable.

Summary statistics for students are presented in Table 1.2. Column 1 shows the average demographic characteristics for all student-year observations in the sample. Column 2 reports means for students who never attend a newly constructed school during the sample period (i.e. "never treated"). Column 3 reports means for "always treated" students, that is, those whose first year in the data sample is at a newly constructed school. In practice, these are almost always kindergarten students, although this also includes students who show up in LAUSD for the first time in other grades. Columns 4 and 5 show means for switchers and "stayers", respectively. The former are students who switch to a newly constructed school at some point during the sample period, while the latter are defined as students at schools where more than 10% of grade-year cohort switches to a newly constructed school in the following year.

[15]Some students with limited English proficiency and/or individual education programs take alternative exams. These students are excluded from all test score analyses.

[16]Data on absences are more frequently missing and/or inconsistent in some years, thus we limit attention to only to total days attended when examining effects on student attendance

Over 85% of students in LAUSD are black or hispanic, and most students speak a language other than English at home with their parents. Students in LAUSD are also much more socioeconomically disadvantaged than the typical California school district: over three-quarters of students are eligible for free and reduced price lunch and do not have a parent who attended any level of postsecondary education. Importantly, treated students who attend newly constructed schools are even more likely to be black or hispanic, low-income, and speak a language other than English at home. Comparing students who switch to new schools against their peers who stay behind at old schools, the same pattern of selection emerges: student switchers are slightly more likely to be low income and score more than 10% of a standard deviation lower in both math and ELA than those students who stay behind at old schools. This selection pattern was a deliberate feature of the construction program: new school facilities were targeted toward neighborhoods with the most overcrowded and depreciated schools, and these school zones were overwhelming located in the most underprivileged areas of the district. Comparing the stayers and switchers shows that even within disadvantaged neighborhoods, new schools were located in slightly worse areas and their catchment areas encompassed slightly lower performing and slightly more disadvantaged students.

Real estate data

To analyze the effects of increased capital expenditures on the real estate market in Los Angeles, we use administrative records from the Los Angeles County Assessor's Office. Records contain information for each property in Los Angeles county, and includes data on the three most recent sales,[17] as well information on property characteristics from the most recent assessment. Properties are matched to the assigned school district, school attendance assignment (for elementary, middle, and high school) in each year, city, and tax rate area (TRA). The TRA is defined as the specific geographical area within a county wherein each parcel is subject to the same combination of taxing entities; the tax rate is therefore uniform for all properties in a given TRA. We limit attention only to the years 1995 to 2012. Our database of LAUSD school assignment zones is only comprehensive up to 2012; moreover, our project database of post-2012 school constructions is also incomplete. For this reason, we exclude the years 2013-2016 from the baseline real estate analyses, although results are robust to including these later years.[18]

We focus only on sales of residential properties with non-missing sales prices. We limit attention to single-family residences. We exclude large parcels with greater than 1 acre of usable area. We then drop the less than 1% of properties with missing information on property characteristics. Data on property characteristics is available only for the most recent assessment; we therefore drop to-be rebuilt properties (i.e. those sales with a negative building age) to avoid biases arising from incorrect valuation of property characteristics. This final restriction is non-trivial; roughly 2.8% of sales are excluded. Finally, we exclude

[17]As of April 2017, when we retrieved the data.

[18]See Appendix Table 1.16, where we compare results using all years to pre-2013 years.

the top 1% and bottom 1% of property sales in each year to avoid results being affected by outliers or non-market-rate transactions.[19]

Table 1.3 summarizes these data. Column 1 reports means for all property sales in the sample within LAUSD district boundaries. Column 2 restrict to only those properties that ever reside in a new school attendance zone, while column 3 reports means for those properties that never receive a newly constructed school facility during the sample period. The average single-family residence in the district was $565,801 (in 2015$) during the sample period. Comparisons of columns 2 and 3 show that new school neighborhoods are generally negatively selected in terms of house prices: houses in new school zones sold for over $200,000 less than those in areas that did not receive new schools. Overall, after sample restrictions, the assessor dataset covers 505,835 property sales for 350,299 unique properties, roughly one-third of which are located in neighborhoods that received new schools during the construction program.

1.5 Empirical Strategy

Student Outcomes

To estimate the effect of attending a newly constructed school on student outcomes we use a generalized difference-in-differences strategy that relies on variation in the year a student begins at a new facility. Importantly, we only observe the school a student attends and not her actual neighborhood school assignment. Moreover, families may systematically sort between neighborhoods based on differences in preferences for educational quality and/or school amenities. If residential sorting or school assignment non-compliance are correlated with underlying student-level characteristics, estimates of the effect of attending a newly constructed school facility may suffer from selection bias. To address this, we rely only on *within-student* changes in outcomes over time, controlling for a student fixed-effect to eliminate any biases due to time-invariant differences between students who matriculate at different schools. The key identification assumption is that the timing of student switching to newly constructed school facilities is as good as random, after accounting for fixed differences between students, grades, and years. This leads to a flexible event-study specification that allows for differential effects of attending a new school for each year a student outcome is observed:

$$y_{it} = \alpha_i + \gamma_{g(i,t)} + \delta_t + \sum_{k=\underline{K}}^{\overline{K}} \beta_k \mathbb{1}(t = t_i^* + k) + \epsilon_{it} \tag{1.3}$$

for an outcome y_{it}, for student i in year t and grade $g(i,t)$. We include fixed effects for student (α_i), grade ($\gamma_{g(i,t)}$), and year (δ_t). Here, the coefficient β_k captures the effect

[19] See Appendix Table 1.17 for a comparison of estimates with relaxed sample restrictions.

of attending a newly constructed facility k years after the first year a student attends, t_i^*. $k = 0$ in a student's first year attending a school, and thus β_k estimates the effect of $k+1$ years of exposure to a new facility. Effects are measured relative to year $k = -1$, which is excluded in estimation. Endpoints are binned at $\underline{K} = -3$ and $\overline{K} = 3$,[20] which represent the average of student outcome y_{it} three or more years prior to attending a new school, or three or more years after first attending (i.e. after four or more years of exposure to a new facility), respectively. Standard errors are two-way clustered by both school and student, to account for any serial correlation within school and/or within student outcomes over time. This design builds in placebo tests that identify violations of the identification assumption that the timing of student switching is as good a random: for $k < 0$, nonzero coefficients would be an indication of non-randomness in the timing of student switching.

Equation (1.3) estimates the effects of attending a new school separately by year. We can approximate the dynamics of these effects by estimating a more parametric version of (1.3) where we allow for a new school to have an immediate effect, and for effects to phase in gradually over time. Imposing linearity in the growth rate of student outcomes and defining $\tilde{t}_i \equiv t - t_i^*$, we can estimate the following generalized difference-in-differences specification:

$$y_{it} = \alpha_i + \gamma_{g(i,t)} + \delta_t + \beta_1 \mathbb{1}(\tilde{t}_i \geq 0) + \beta_2 \mathbb{1}(\tilde{t}_i \geq 0) * \tilde{t}_i + \beta_3 \tilde{t}_i + \epsilon_{it} \tag{1.4}$$

Here β_1 captures the immediate effect of a new school facility in the first year a student attends, t_i^*. We include a linear trend in "event time", $\tilde{t}_i$, to control for any selection on trends into schools opening in a particular year. β_3 captures this selection, while β_2 reflects effects of the new school that accrue gradually over the time a student is exposed to a new school.[21] As a student is repeatedly exposed to improved facilities in each year she attends a new school, we would expect effects to cumulate and increase over time with continued exposure: $\beta_2 > 0$.

Estimates from equations (1.3) and (1.4) are presented in Section 1.6. Event-study estimates from equation (1.3) indicate that the parametric specification in (1.4) does a good job of capturing the dynamics of the effects on various student outcomes. Later, to more parsimoniously examine heterogeneity, mechanisms, and robustness, we will focus on the estimates from equation (1.4), and on even simpler versions that constrain $\beta_1 = \beta_3 = 0$. In our baseline estimation we use all student-year observations in the relevant grades for a given outcome, with the sole exception of those students who attend multiple new facilities, who are excluded to avoid any confounds in the dynamics of estimated treatment effects. Students who never attend new school facilities are included in the regressions as controls,[22] as are students who we observe at newly constructed schools in their first year in the data

[20] We choose $\overline{K} = 3$ as few students attend a new school facility for more than 4 years in the data.

[21] We can directly interpret β_2 as an impact on the gain score, often an outcome of interest in many studies of educational interventions.

[22] Event-time indicators are set to zero for these students, who contribute only to the estimation of the year (δ_t) and grade ($\gamma_{g(i,t)}$) effects in the regressions.

(e.g. students who begin elementary at a newly constructed school, or transfer from another school district). Inclusion of the latter group of students may induce bias if students on different trajectories in outcome y_{it} sort into LAUSD to attend school at a newly constructed facility. Furthermore, students who "stay behind" at existing school facilities and see significant changes in their school and peer environments are also included as controls. In Section 1.6 we compare estimates where "stayers", never treated, and always treated students are excluded; reassuringly, results are very robust to the inclusion or exclusion of these students.

Real Estate Capitalization

As expected due to the design of the construction program detailed in Section 1.2, the location of the new schools is negatively selected: areas that received new schools had lower house prices, lower average incomes and educational attainment, and lower student test scores. However, conditional on a neighborhood receiving a new school, the timing of new school constructions is plausibly exogenous relative to any underlying neighborhood characteristics or trends. Thus, parallel to our estimation of student effects, we estimate house price effects of the program in a dynamic setting by examining changes in school quality induced by new constructions, relying on variation in the exact timing of completion.

Specifically, we compare changes in house prices over time in neighborhoods that received new schools, relying on variation in the exact year of school construction between these neighborhoods, and controlling for neighborhood effects to account for any time invariant neighborhood characteristics. Changes in prices reflect the present discounted value of current and future benefits of new schools to households. Thus, we estimate the mean difference in house prices before and after construction with following difference-in-differences specification:

$$ln(P_{it}) = \alpha_{j(i)} + \delta_t + \beta N_{j(i),t} + X'_{it}\Gamma + \epsilon_{it} \tag{1.5}$$

where $N_{j(i),t} = \mathbb{1}[NewSchoolZone_{j(i),t} = 1]$ is an indicator for a property sale occurring in a new school attendance zone, after the date of the new school opening, for a given property i in neighborhood $j(i)$ that is sold at time t. X'_{it} is a vector of property characteristics that includes the number of bathrooms, the number of bedrooms, building square footage, square footage squared, building age, age squared, effective age, effective age squared, usable lot area, usable lot area squared, an indicator for the specific tax rate area, and an indicator for number of sales observed in the data for specific parcel. $\alpha_{j(i)}$ and δ_t are fixed effects for neighborhood and year, respectively.

We define neighborhoods as the elementary-middle-high school assignment triplet in the 2000-2001 academic year, prior to the construction of any new facilities.[23] In all house price specifications, standard errors are clustered by neighborhood. Baseline specifications include only those parcels that are ever assigned to the attendance zone of a newly constructed school.

[23]See Figure 1.18 for a map of these neighborhoods.

As long as the exact timing of school construction within the set of receiving neighborhoods is uncorrelated with time-varying neighborhood trends, estimation of equation (1.5) will yield an unbiased estimation of β. In addition, we estimate specifications that also include "never-treated" properties as controls, and specifications that control for year-by-high school zone fixed effects,[24] to flexibly account for differential trends in house prices between local areas.

If capitalization occurs prior to construction due to anticipatory effects,[25] neighborhood house prices may diverge prior to construction between those soon to receive new schools and those receiving new schools in later years. Conversely, initial uncertainty by parents as to the quality of a new school could lead to house price effects that gradually cumulate post-completion. Thus, we also estimate more flexible event-study models, akin to equation (1.3), that estimate the difference in house prices relative to the year prior to building occupancy:

$$ln(P_{it}) = \alpha_{j(i)} + \delta_t + \sum_{k=\underline{K}}^{\overline{K}} \beta_k \mathbb{1}(t = t_i^* + k) + X_{it}'\Gamma + \epsilon_{it} \tag{1.6}$$

In these non-parametric event study models, β_k measures the effect of receiving a new school in year t_i^* k years after construction (or prior, where $k < 0$). Effects are measured relative to year $k = -1$, which is excluded in estimation. We focus on a ten-year window, binning endpoints at $\underline{K} = -6$ and $\overline{K} = 3$, which represent average house prices six or more years prior to construction or three or more years after the year of construction, respectively.

In equations (1.5) and (1.6), identification of β assumes that trends in house prices are uncorrelated with the exact timing of school construction, conditional on property-specific controls and controls for time-invariant differences between neighborhoods. This assumption could be potentially violated if unobserved differences in the characteristics of those properties sold in a given year are correlated with the timing of switching; for example, if houses with positive unobserved characteristics are more likely to be sold within a given neighborhood post-construction than pre-construction. To account for this potential source of bias, we can estimate equation (1.5) with property fixed effects, controlling for time-invariant unobserved differences between individual properties:

$$ln(P_{it}) = \alpha_t + \alpha_i + \beta N_{j(i),t} + \epsilon_{it} \tag{1.7}$$

In equation (1.7), estimation of β relies only on properties with repeat sales in the sample window. Repeat sales indices are commonly used in papers estimating dynamic capitalization in real estate prices (e.g. Figlio and Lucas (2004)) to account for unobserved differences in property and neighborhood characteristics. In practice, estimates of β are very similar in

[24]Here we use the high school zones from the 2004 school year, the year before the first new high school construction.

[25]Recall: new school locations were announced on average 5 years prior to school completion.

both equations (1.5) and (1.7), implying that differences unobserved property characteristics are uncorrelated with timing of construction and do not drive the estimated results. We find little evidence of differential house price trends in the years prior to school construction. Moreover, effects accrue quickly, typically within 2 or 3 years following construction. Therefore, we emphasize the simple linear differences-in-differences estimate of β from equation (1.5).

1.6 Student Results

Student achievement

Table 1.4 reports estimates of equation (1.4) for math (columns 1-3) and ELA (columns 4-6) standardized test scores. Cumulative four-year test score effects estimates are reported in row 4. In columns 1 and 4, a simple one-parameter specification is reported where only the change in the slope of student growth is included ($\mathbb{1}(\tilde{t_i} \geq 0) * \tilde{t_i}$). Here, the estimate on β_2 is 0.029 (SE 0.007) for math and 0.019 (SE 0.004) for ELA, implying that for each additional year a student attends a newly constructed school facility her test score increases by 3% and 2% of a standard deviation in math and ELA, respectively. The implied test score effect for a student who attends a new school for four years is 0.086 (SE 0.021) for math and 0.058 (SE 0.011) for ELA. Columns 2 and 4 add in indicators for attending a newly constructed school ($\mathbb{1}(\tilde{t_i} \geq 0)$). Student achievement declines in the first year of attending a new school, although these coefficients are small and insignificant for both math and ELA test scores. Notably, the coefficient on the slope of student growth (β_2) and the implied 4-year test score effect are essentially unchanged. Columns 3 and 6 add in a linear trend in student event time. The coefficient on the linear trend is marginally significant for math, and statistically significant for ELA. However, these coefficients are both minuscule: less than one-half of one percent of a standard deviation per year in both math and ELA. More importantly, the inclusion of the linear trend in the specification does little to affect the magnitude or statistical significance of the coefficient on the change in trend, while the total implied 4-year effect declines somewhat due to initial effects (β_1) that are slightly more negative.

Figure 1.4 reports estimates of the event study coefficients, β_k, from equation (1.3) for both math and ELA test scores. Standard errors are two-way clustered by both school and student. Time $k = -1$ is excluded; all effects are relative to the year before a student begins attending a new school facility. Panel A reports estimated coefficients on standardized math scores. There is little indication that students who switch to new schools have rising (or falling) scores relative to other students prior to the switch. Then, in the first year at a new school, there is a small but significant decrease of 3.8% of a standard deviation. This decline is short-lived, however: scores increase nearly linearly with each successive year a student attends school in a newly constructed facility, relative to other students who did not switch to a new school. After four or more years of attending a newly constructed school, students score 10% (SE 2.6%) of a standard deviation higher. Estimates for standardized ELA tests, reported in panel B, are quite similar. Students who attend a new facility for 4 or more years score 5.2% (SE 1.4%) of a standard deviation higher in ELA. For both math and ELA scores, the event-study figures indicate that the parametric specification in equation (1.4) fits the data quite well: after an initial decline in the year a student transitions to a new facility, test scores gradually increase, roughly linearly in years of exposure.

Both event study and linear difference-in-differences specifications show that student test score gains accumulate gradually, after a slight decline in student performance in the year of the switch. This pattern of gradual improvement is different from many other educational interventions considered in the literature, where effects tend to fade out over time. Improvements in school facility quality are not a one-time intervention, however: students are continuously exposed to improved facility conditions for every year in which they attend a given school. We would therefore expect that achievement gains accumulate over time with additional years of exposure,[26] even in the absence of initial disruption effects due to student-level switching costs[27] or school-level inefficiencies in the first few years post-construction.

Student non-cognitive effects

Table 1.5 reports analogous estimates for attendance (columns 1-3) and effort (columns 4-6). Unlike test score outcomes, which measure a stock of accumulated knowledge, student effort is a flow, and thus we would expect effects to occur immediately rather than accrue over time with continued exposure. For this reason, in columns 1 and 4 we begin with one-parameter specifications where only the coefficient for mean difference in the outcome post matriculation at a new facility ($\mathbb{1}(\tilde{t_i} \geq 0)$) included. Columns 2 and 5 add a phase-in coefficient ($\mathbb{1}(\tilde{t_i} \geq 0) * \tilde{t_i}$), and columns 3 and 6 include a linear trend in student event-time. For student attendance, estimates in columns 1-3 imply that most of the effect occurs immediately upon switching to a new school. In column 1, the estimate of β_1 is 3.97 (S.E. 0.55), meaning that student attendance increases by 4 days per year at newly constructed schools. Adding the phase-in coefficient in column 2 picks up some of this effect, reducing the coefficient on β_1 slightly. Column 3 adds in a linear trend in event-time, which does little to affect the estimates of β_1 and β_2. Estimates in columns 4-6 show a similar pattern for teacher-reported student effort, which increases immediately upon a student's switch to a new school. In column 1, the point estimate is 0.061 (SE 0.017), implying a 6% of a standard deviation increase in student effort at new schools. In columns 5 and 6 estimates

[26] The closest analogue is perhaps the STAR class size experiment, in which treated students were assigned to small classes for up to four consecutive years. In STAR, the treatment effects grew after the first year, like here, but at a slower rate. We see no sign here that the treatment effect is concentrated in the first year.

[27] Event study estimates for non-facility related student switches are reported in Figure 1.16. Estimates suggest that "normal" switches are associated with disruption effects of similar magnitudes, which fade out over time. Importantly, these switches are not associated with any short or long run student test score improvements. These findings are consistent with results in Hanushek, Kain, et al. (2004), who find evidence of short-run disruption effects with no-long run gains for students who switch schools within-district.

of the pre- and post-trends (β_2 and β_3) are both small and insignificant, and the inclusion of these coefficients has little impact on estimates of β_1.

Figure 1.5 reports event study estimates for student attendance and teacher-reported effort. Panel A shows the change in annual days attended for students who switch to new schools. Upon switching to a new school, students attend an additional three days per year. In the second year a student attends a new school facility, this jumps to seven days. The effect tapers off somewhat in subsequent years, although after four or more years of attending a new school facility, students attend on average more than four additional days per academic year. Again, as with the student cognitive test score effects, there is no indication of a prior trend in student attendance in the years prior to switching to a new school facility; if anything, attendance appears to be declining slightly, although this trend is minuscule and insignificant - a decline in annual attendance of less than one-third of a day per year in the three years prior to switching.

Panel B shows the effect of switching to a new school facility on teacher-reported student effort for elementary students. Upon matriculation into a new school facility, student effort increases by greater than 6% of standard deviation.[28] As with attendance effects, the estimated increase in effort occurs immediately upon switch with no indication of an increasing trend in effort in the years prior to switching. This effect remains roughly constant with additional years of exposure, and is statistically significant for the first three years a student attends a newly constructed facility. After 4 or more years of exposure to new elementary school facilities, the estimated effect on effort is slightly smaller, around 5% of a standard deviation, and no longer statistically significant. Notably, two years before attending a new facility, effort marks are roughly 3% of a standard deviation higher than in the year prior to attending a new facility, which is significant at the 10% level. Of all baseline event-study estimates, this is the only estimated pre-effect that is marginally significant, providing additional justification for the identification assumption that the timing of student switching is as good as random.

Robustness to sample treatment

Baseline estimates from one-parameter models for cognitive and non-cognitive outcomes in Tables 1.4 and 1.5 (columns 1 and 4) are reported in Table 1.6 for different sample definitions, varying the set of students used as the control group for students switching to new schools. As test score effects reflect the cumulative impact of multiple years of exposure to new schools, we compare one-parameter estimates of the phase-in coefficients (β_2) from models where we constrain $\beta_1 = \beta_3 = 0$. Reassuringly, implied cumulative 4-year effects from parametric estimates in columns 1 and 4 of Table 1.4 are indeed very similar to point estimates reported in Figure 1.4 for students who attended new schools for four or

[28] Figure 1.15 also reports similar event-study estimates for teacher-reported student grades (in elementary school). Effects are noisier and insignificant, but suggest improvements of similar magnitude in report card grades only after the first year at a new school, which is qualitatively consistent with the observed patterns for test score effects.

more years. On the other hand, as we expect the flow of student effort and attendance to increase immediately upon matriculation to a new school, we report one-parameter estimates of the mean difference post-new school matriculation (β_1) from models where we constrain $\beta_2 = \beta_3 = 0$.

Column 1 repeats baseline estimates reported in Tables 1.4 and 1.5. Column 2 excludes students who stay behind at existing schools when 10% or more of their cohort switches to a new school. Estimated coefficients for ELA and days attended are only slightly larger, while estimates for math and effort standardized scores are essentially identical.[29] In column 3, we drop all students who never attend new schools, using only "ever-treated" students. If students who switch to new schools are systematically different from those who do not, inclusion of never-treated students as controls may induce bias (though our inclusion of student fixed effects would absorb differences in outcome levels). However, this does not appear to be the case, as estimates are nearly identical for all outcomes. Column 4 further excludes students who appear in the data sample in their first year at a new school. Inclusion of these "always treated" students could be problematic if new school constructions systematically induce students of different ability to enter LAUSD, perhaps from private schools or from outside the district. As shown in column 4, estimated treatment effects are, if anything, slightly larger when only switching students are included in the estimation sample, implying effects are not generated by a resorting of students entering LAUSD to attend newly constructed school facilities.

In column 5, we restrict the sample to include a balanced panel of students in event time. As discussed in Section 1.4, ELA test scores are recorded for students in grades 2-11, and attendance is measured for all grades. Math test scores are only included for grades 2-7, and effort marks are only measured in elementary school (grades KG-5). Thus, for math and effort we include students who have outcome data both one year before and one year after switching to a new school facility. For ELA and attendance we need not be as restrictive, and use a balanced panel of students with non-missing outcomes both 2 years before and after switching to a new facility. Estimated treatment effects in column 5 are less precise, as expected given the reduction in sample size, but point estimates are if anything slightly larger than those in columns 1-4. Results are robust to these sample permutations, and we therefore conclude that baseline estimates including all students are not biased by differential sample selection in event time.

Mechanisms

The pattern of estimated student effects provides consistent evidence that cognitive and non-cognitive student outcomes improved at new school facilities. Are these improvements due to the increased facility quality itself, or due to other changes in the school environment

[29] In Section 1.6 we specifically examine indirect effects on these students, finding evidence of small positive effects on ELA scores and attendance. Since these students make up only a small fraction of the overall "never treated" group in baseline regressions, we would therefore expect the magnitude of differences between columns 1 and 2 to be very small in the presence of small indirect effects.

associated with new school constructions? A thorough understanding of the mechanisms underlying student gains is important if the LAUSD construction program is to inform school capital expenditure decisions in other districts and institutional contexts. In this section we detail several facility and non-facility related changes associated with new school facilities (Figure 1.6 and Table 1.12). We examine heterogeneity in the results by prior school conditions to test whether these changes are systematically related to the observed student gains (Table 1.7).[30]

We find little evidence that changes in class size, peer quality, or teacher quality at newly constructed schools can explain student improvements. In fact, we find that moving to a new school is associated with slightly larger class sizes and slightly lower teacher and peer quality. We do find that switching to a new school is associated with large reductions in overcrowding and increases in total instructional days. We find suggestive evidence that student gains are larger for students who switched from schools that were on multi-track calendars, for students who switched from more overcrowded schools, and for students who came from schools with a high share of portable classroom buildings. We find inconclusive evidence that students gains are larger for those coming from older or more deteriorated facilities, although the magnitude of these prior differences is small relative to the total change in facility quality for switching students.

Peer Composition

If students who switch to newly constructed school facilities are exposed to higher quality peers, changes in peer quality could explain some of the observed effects. As discussed earlier and shown in Table 1.2, students who attend newly constructed schools are more disadvantaged relative to students in the rest of LAUSD. However, new schools could offer better peer groups than do other schools in nearby neighborhoods. This could occur if new school boundaries were drawn within receiving neighborhoods in a such a way as to increase the concentration of more advantaged students at new school facilities. In addition, insofar as parents have some discretion to override school residential assignment, one might expect that higher-SES parents from outside redrawn boundaries would be more likely to petition to enroll their children in at schools with new and improved facilities. However, empirically, this does not appear to be the case. Panel A of Figure 1.6 shows event study estimates of peer quality, measured as the school (leave-out) mean predicted test score.[31] Average peer predicted scores fall significantly upon switching to a new school, and after 4 years at a new school average peer predicted scores are (insignificantly) below their level prior to switching. Columns 5 and 6 of Table 1.12 report estimates of the average peer differences associated with switching to a new school. Column 5 shows the change in school proportion black and

[30]Note: results in Table 1.7 are estimated using only switching students; baseline estimates correspond to those presented in column 3 of Table 1.6

[31]Predicted scores are generated at the student-year level from a regression of contemporaneous ELA test scores on a vector of demographic characteristics. Leave-out mean school-year predicted scores are then computed for each student-year observation.

hispanic, while column 6 reports the mean difference in peer predicted scores. Estimates show that students who switch to new facilities attend more segregated schools, with a 2.7 percentage point higher share of black and/or hispanic students. Consistent with Figure 1.6 panel A, peer predicted scores are on average 2.4% of a standard deviation lower.

Class size

Panel B of Figure 1.6 reports event study estimates for elementary school students where class size is the dependent variable. At new schools, class sizes were actually somewhat larger: class sizes increased by less than one-half of a student per teacher.[32] Column 4 of Table 1.12 reports analogous difference-in-differences estimates of the change in class size at new schools. On average, teachers at new schools taught classes with 0.31 more students per teacher. The magnitude of this difference, however, is quite small; roughly speaking, the district was approximately able to maintain similar pupil-teacher ratios at new school facilities by transferring teachers to new facilities in roughly equal proportion to students.

Multi-track calendar

One of the stated goals of the LAUSD school construction program was to eliminate the use of multi-track academic calendars that required schools to continuously operate year-round. Schools on multi-track calendars operate year-round and divide the students and staff into separate tracks, which are staggered throughout the school year in an effort to increase overall facility capacity. Moreover, in LAUSD, students at multi-track schools often had fewer instructional days per academic year.[33]

Before the construction program, half of LAUSD students attended multi-track schools. By reducing overcrowding in neighborhood schools, district officials were able to begin new schools on traditional two-semester calendars, as well as convert existing schools from multi-track back to traditional calendars. Column 1, panel A of Table 1.12 and panel C of Figure 1.6 report difference-in-differences and event study estimates of the likelihood of being exposed to a multi-track calendar. Switching to a new school was accompanied by a 27 percentage point reduction in the likelihood that a student was exposed to a multi-track calendar. This conversion also meant that many students in new schools experienced additional instructional days: as reported in panel D of Figure 1.6 (and column 2 of Table 1.12) students switching to a new school had on average nearly 2 additional instructional days per year, relative to the prior year at an existing school. Taking the baseline estimate of 4 additional days attended per year from Table 1.5, this implies that almost half of the observed attendance effect is mechanically due to a change in school calendar.

[32] In fact, due to budget cuts in California during the Great Recession, LAUSD laid off roughly 25% of teachers between the 2008 and 2010 school years, increasing class sizes across the district, particularly in grades K-3.

[33] Many of the year-round district schools operated on a multi-track calendar known as "Concept 6", which increased school capacity by up to 50% but at the cost of 17 instructional days (out of 180). The loss in instructional days was made up by increased instructional time per day.

Student gains at new schools may be driven by increased instructional days and the conversion back to traditional two-semester calendars. To examine this, we estimate separate treatment effects by prior school calendar schedule (multi- or single-track) for each main outcome variable. Rows 2 and 3 (panel 2) of Table 1.7 report estimated effects for students who switched from a school on a multi-track calendar or a single-track calendar, respectively. Results show that student test scores and effort effects are larger for students who switched from multi-tracked schools, yet none of these differences are statistically significant. For attendance, gains are larger for those students who came from multi-track schools, and the difference in highly significant ($p < 0.01$). As previously mentioned, this is driven largely by mechanical changes in the total number of instructional days.[34] However, the overall attendance effect is not entirely due to the calendar change: even those students who switched from single-track schools and saw no increase in total instructional days attended 2.7 additional days of school per year after switching to a newly constructed school. It is worth noting that students switching from schools on multi-track calendars also faced greater overcrowding and worse facility conditions on average than those coming from single track schools; these differences are therefore likely to represent an upper bound on the effect of converting from a year-round multi-track calendar school back to a traditional two-semester calendar.

Overcrowding

Students who switched to new schools also experienced reduced overcrowding, which was another primary motivation of the district construction program. Panels 3 and 4 of Table 1.7 present heterogeneity in estimated effects by two measures of overcrowding, where treatment effects are split by whether a student is above or below the median among treated students on either measure in the year prior to switching to a new school facility. The first measure, the number of students per square foot of classroom space, gives a direct measure of the physical classroom capacity of a school. The second, the share of permanent classrooms, measures the extent to which portable classrooms are used to accommodate a school's student population. Portable facilities are also often of much worse quality, and have less functionality than traditional classroom space. The share of classrooms in permanent vs portable structures therefore relates both to the level of overcrowding of a school, and the underlying facility quality.

Results split by prior SQFT per pupil indicate mixed and generally insignificant differences: gains are larger for ELA and attendance when coming from more crowded prior schools (i.e. a low SQFT per pupil), but smaller for math and effort scores. Only the difference in attendance is statistically significant; gains are large and significant for students coming from either above or below median overcrowded schools for all but ELA scores. Estimates split by the share of permanent classrooms show a more consistent pattern: effects for students coming from schools with a low share of permanent classroom structures are

[34]For students switching from multi- to single-track calendar schools, the average gain in total attended days was approximately 4 days per year.

larger for all outcome measures, and the differences in ELA and effort scores are statistically significant. In fact, estimates for math, ELA, and effort scores are statistically insignificant for students who came from schools with relatively more permanent classroom space. This pattern of results suggests that reductions in overcrowding are important but alone do not fully explain observed test score gains: there are few systematic differences by prior SQFT per pupil, while differences by the prior share of permanent classrooms additionally reflect fundamental improvements in facility quality beyond overcrowding.

Facility Condition

New school constructions induced drastic changes in the facility quality for students who switched. Students who switched to new schools came from a school that was, on average, 70 years old (panel A, Table 1.12) and had substantial deficiencies. Results split by the share of permanent classrooms provided indirect evidence that observed student gains were larger for students switching from schools that were of poor quality. In panels 5 and 6 of Table 1.7 we examine heterogeneity in treatment effects by the age and physical condition (measured by FCI[35]) of a student's prior school. Results indicate that all student effects are larger for students switching from older schools, although these differences are insignificant. Student test score and effort effects are very similar between students switching from schools in above or below median condition, and the differences are insignificant. Only for student attendance is the difference significant; students switching from schools in relatively better condition (low FCI) actually saw larger attendance gains.

With the exception of estimates by the share of permanent classrooms, results presented in Table 1.7 provide inconclusive evidence of heterogeneity in student effects by prior facility quality. This does not necessarily imply that facility quality improvements themselves were not important: the variation in facility improvements within treated students is small relative to the change experienced for any student switching to a new school. Moreover, these variables are imperfect proxies for "true" facility quality, which we cannot directly quantify.

Overall, estimates in Table 1.7 suggest reductions in overcrowding and multi-track calendars may explain up to half of observed student effects. Later, in Section 1.6 we will examine students who stayed behind at existing school facilities and experienced significant peer outflows. These students experience very similar reductions in overcrowding and multi-track calendars, yet for these students we find much smaller effects, and only for ELA and attendance. Thus, taken together with results presented here, we argue that at least half of the observed test score effects for switching students are therefore attributable to the direct improvement in the physical school environment.

[35]Recall: the FCI is the ratio of deficiencies to current replacement value. An FCI close to zero indicates a facility is in good physical condition, whereas an FCI of greater than one indicates that a facility has deteriorated to the point where the total cost of deficiencies is greater than the total replacement cost of the facility.

Teacher quality

Student gains at new schools could in part be due to systematic differences in teacher quality. New school facilities provide improved working environments for teachers, and these amenities could attract better quality teachers to these schools from either within or outside the district.[36] Due to budget cuts following the Great Recession, the district effectively stopped hiring new teachers: prior to 2009, roughly 10% of the teachers in LAUSD in any given year were new entrants, while afterwards this decreased to 4% or less. Even for those new facilities that opened before this reduction in teacher hiring, the teaching staff was composed of greater than 80% existing teachers who switched from elsewhere in the district. Following the reduction in teacher hiring, schools opening in 2009 or later this proportion increased to over 90%. Thus, any differences in the quality of new teachers is unlikely to explain a large share of the observed effects. However, new facilities may have attracted relatively better teachers from within the district. Improved non-wage amenities at new school facilities could have led to sorting of higher quality teachers into new schools. On the other hand, priority for intra-district teacher transfers within LAUSD was allocated using a tenure-based point system, which may not be systematically correlated with underlying teacher quality (broadly defined).

Systematic teacher resorting would imply that student gains at new schools came at the expense of students at existing schools; any within-district resorting of existing teachers would be zero-sum in aggregate. To empirically assess whether differential sorting of higher quality teachers into new school facilities explains any of the observed student gains, we compare differences in teacher observables and test score value-added in Table 1.8. Panel A reports differences in teacher observables at new schools. Students who switch to new school facilities have teachers who are, on average, less experienced, younger, and slightly more likely to have a masters degree. Students at new school facilities are also 5.4 percentage points more likely to have a new teacher in either math or ELA. Observable teacher characteristics, however, are generally not highly correlated with test-score based measures of quality. Thus, in panel B, we examine differences in test score value-added for teachers at new schools.[37]

Standard value-added models can confound school and teacher effects. For example, new school facilities could generate improvements in student attentiveness and/or teacher productivity, both of which would result in gains in estimated teacher valued-added. However, student gains resulting from school improvements would reflect improvements resulting from the new facility itself, and not from variation in underlying (prior) teacher quality. Thus, to directly assess whether teacher resorting explains any of the student gains, we focus specifically on switching teachers, for whom we have an estimate of value-added based on student

[36]Complementarities between facility quality and teacher effort and/or performance could also result in improved teacher productivity at new schools. Unfortunately, we cannot directly assess this using our data, as any such improvements could not be separately distinguished from general school- or student-level improvements.

[37]See Appendix B for an explanation of how teacher value-added scores are calculated at the teacher-year level.

test score observations from their prior, existing school facilities.

For these switching teachers, we compute the student-weighted average of prior value-added scores, using only data from years a teacher taught at an existing school facility. Specifically, we define $\text{VA}_j^{prior} \equiv \sum \frac{n_{jt}}{n_j}\text{VA}_{jt}$, where VA_{jt} is the estimated value-added for teacher j in year t, n_{jt} is the number of student observations for contributing to teacher$j$'s value-added score in year t, and n_j is the total number of students taught by teacher j (prior to switching to a newly constructed facility). For each student-year observation, we assign the mean prior value-added score, averaged over all teachers in a given school-year.[38] Columns 1 and 2 of panel B report difference-in-differences estimates of the change in mean prior value-added for students attending newly constructed school facilities. Results indicate that students who switched to new schools experienced teachers with *lower* test-score value-added scores than prior to switching. The point estimates are for both math and ELA are small, although the estimate is more negative and statistically significant for ELA.

We find little evidence of positive restoring of existing teachers into new schools, but it could still be the case that the new teachers hired into new schools were of differential quality. We cannot directly compare contemporaneous value-added scores of new teachers at new and existing schools, as this would confound student gains due to school-level facility improvements with improvements in new teacher quality. However, under the assumption that new facilities affect novice and experienced teachers identically, we can assess the quality of new teachers by testing whether the school-level gap in value-added scores between new and existing teachers is larger or smaller at new facilities. We can decompose the estimated teacher effect to include the true teacher effect, a new-school specific shock, and an unobserved error term:[39]

$$\text{VA}_{jst} = \mu_{jt} + \theta_{st} + \eta_{jst}$$

Insofar as the effect of a new school in a given year, θ_{st}, is constant for all teachers, we can use the gap between experienced and novice teachers at new schools to difference out the any differential new school effects at the school by year level:

$$\begin{aligned}\overline{\text{VA}}_{st}^{GAP} &\equiv \overline{\text{VA}}_{st}^{New} - \overline{\text{VA}}_{st}^{Old} \\ &= \overline{\mu}_{jt}^{New} - \overline{\mu}_{jt}^{Old} + \tilde{\eta}_{st}\end{aligned}$$

We therefore assign each student the difference between the school-year average value-added of new teachers and existing teachers. A positive school-level gap between new and existing teachers would indicate that the new teachers at a school have higher value-added than the

[38]Results are nearly identical if we instead assign a student the prior value-added score of her specific teacher in a given year.

[39]In Appendix B teacher-year value-added is defined as the average residual from a regression of student test scores on polynomials in lagged test scores, demographic variables, and school variables: $\text{VA}_{jt} \equiv \overline{\nu}_{j(i,t)t}$ where $\nu_{it} = y_{i,t} - \alpha_{t,g(i,t)} - X_{it}'\beta$.

existing teachers, and vice-versa. Thus, holding existing teacher quality constant, if new teachers hired into new facilities are of higher quality, we would expect a positive coefficient on the gap.

Columns 3 and 4 report these estimates, where the dependent variable is the school-year mean gap in value-added between novice and experienced teachers, $\overline{\text{VA}}_{st}^{GAP}$. The point estimate for math is small, negative, and insignificant. For ELA, the point estimate is positive and of larger magnitude, but insignificant. Given that we find evidence of negative sorting of existing teachers on value-added ($\overline{\mu}_{jt}^{Old} \leq 0$ in columns 1-2), the difference in point estimates between columns 1 and 3 and columns 2 and 4 would need to be positive to support an interpretation that newly hired teachers were of higher quality at new facilities. As the estimated coefficients in columns 3 and 4 are small and noisily estimated we do not report a formal test of these differences. Results from Table 1.8 panel B therefore provide little evidence that newly hired teachers were of higher quality at new schools.

Overall, the evidence presented in Table 1.8 reveals that systematic differences in teacher quality cannot account for observed student test score gains. As the overwhelming majority of students at new schools were taught by existing teachers, point estimates from columns 1 and 2 of panel B imply that student test score gains at new schools would have been roughly 15% larger in math and 50% larger in ELA had teacher quality remained constant. The upper bound of the 95% confidence interval for the math effect can rule out positive teacher sorting explaining more than 30% of the total effect from column 1 of Table 1.4. In the longer-run, it is still possible that higher-quality facilities could attract and retain better teachers, although further research is necessary to determine if this channel to improve teacher quality is empirically relevant.

Effects on staying students

Students who switched to new school facilities were not the only students to experience significant school-level changes: student switches to new facilities induced cohort-level outflows from existing facilities. Those students who stayed behind experienced reductions in overcrowding, conversion from year-round multi-track calendars back to traditional two-semester calendars, and changes in peer composition, but not improvements in facility quality. Thus, examining the effects of new facility openings on the outcomes of students who stayed behind at existing facilities can shed light on the relative importance of crowding vs direct facility quality effects in producing the aforementioned estimated impacts on students at new schools.

We define "stayers" to be students for whom 10% or more of their school-grade cohort switched to a newly constructed school facility.[40] We then define event-time analogously for these students: year "0" is the year in which a school cohort experienced a large outflow induced by a nearby new school construction. We estimate effects for these students using

[40]Appendix Figure 1.17 reports analogous event study estimates using a 20% threshold. This reduces the sample considerably, but results are robust to alternative thresholds.

the same event study methodology for the main student effects presented in equations (1.3) and (1.4); because these cohort outflows were induced by new facilities, estimates rely on the same variation in the timing of construction between different students.

Panel B of Table 1.12 presents estimates of the changes staying students experienced after they experienced a cohort outflow, analogous to estimates for switching students presented in panel A. Students who switched to new schools are excluded from estimation; estimates are relative to a control group of students in the same grade and year who have yet to experience a cohort outflow shock, and never-treated students who experienced no significant peer outflow. Results indicate that stayers experienced a significant decline in multi-track calendar usage and a significant increase in the total number of instructional days per year. Both staying and switching students experienced a roughly equivalent decline in multi-track calendars, while staying students actually experienced a slightly larger increase in the total number of instructional days than switching students (2.3 days vs 1.7 days). Class sizes decreased slightly for students who stay behind, by about one-third of a student per teacher. Though significant, the magnitude of the effect is negligible.[41] Columns 5 and 6 report changes in the average peer group. Consistent with the fact that switching students were slightly more disadvantaged and lower-scoring than staying students, stayers see reductions in peer minority shares and increases in predicted scores of peers due to cohort outflows to new facilities. Taken together, these results suggest that small indirect effects would be likely, even in the absence of facility improvements, due to the reductions in overcrowding, increase in instructional days, improved peer quality, and slightly decreased class size.

Figure 1.7 shows event-study estimates of cognitive and non-cognitive outcomes for stayers. Stayers see small increases in math (panel A) and ELA (panel B) test scores, although the math effects show some indication of a pre-trend prior to the year of the cohort outflow to a new facility. The increase in days attended (panel C) is immediate and significant - students attend roughly 4 more days relative to the year prior to the cohort outflow. As was the case for switching students, much of this increase derives from the reduced use of multi-track schedules in stayers' schools. Panel D shows estimates for standardized effort scores, for which the point estimates are all very close to zero and insignificant.

Parametric versions of the estimates corresponding to equation (1.4) are reported in Table 1.10. For each outcome, both one- and three-parameter estimates are shown. Columns 1 and 2 report estimates for math test scores. Estimates in column 1 show no change in test score growth in the years following the cohort outflow, while estimates in column 2 show that once pre-existing trends are included, there is a small effect immediate effect that fades out within the following year. For ELA (columns 3 and 4), the pattern is different, and the parametric estimates more closely align with the event study estimates. Column 3 shows an 0.01 standard deviation increase in ELA test score growth in the years following the cohort outflow. However, once the post indicator and trend variable are included in column 4, all of the effect loads onto the post coefficient, with no ensuing growth or fade-out of effects. This

[41]Note that the effect is similar in (absolute) magnitude to the increased class size documented for switching students in panel A.

pattern of cognitive effects differs from that of students attending new schools: effects accrue immediately, and either fade out (math), or remain constant (ELA). Columns 5 and 6 report estimates for days attended. Stayers see a roughly 3.5 day increase in days attended, which is robust to the inclusion of trend variables. Comparing these estimates to the estimated 2.3 day increase in total instructional days from column 2 of Table 1.12 Panel B implies that roughly two-thirds of the attendance effect is mechanically driven by increased number of days. Columns 7 and 8 show no effects on teacher-reported effort.

Taken together, these results are suggestive of positive indirect effects induced by peer outflows to new school facilities, but only for ELA test scores[42] and total days attended.[43] Attendance effects are mostly driven by an increase in the total number of instructional days, and the residual non-mechanical effect is roughly half the size as for switching students (1.3 vs 2.3 additional days). These indirect effects are likely driven by reductions in overcrowding, improved peer quality, and the switch from multi-track calendars to traditional schedules.

The small magnitude of effects relative to baseline effects on switching students implies that reductions in multi-track calendars and overcrowding alone cannot explain the bulk of baseline effects, as these changes were similar for students who stayed behind at existing schools. Moreover, other notable changes in the school environment (peers, class sizes, and teachers) all went *against* finding positive test score effects. This supports the conclusion that direct facility quality effects – e.g. Increased concentration due to reduced distractions from inadequate heating, cooling, or other aspects of the physical environment – account for a substantial portion of the new school effect seen earlier.

1.7 Real estate capitalization

Next, we turn to the analysis of the impact of new school openings on local housing prices. In Table 1.10 we present difference-in-differences estimates corresponding to equations (1.5) and (1.7), while in Figure 1.8 we report event study estimates corresponding to equation (1.6). Panel A of Table 1.10 reports estimates of the effect of new school constructions on house prices. Columns 1-5 report estimates using fixed effects for school zone and property-specific control variables, which corresponds to the specification in equation (1.5). Columns 6 and 7 report estimates of equation (1.7) using property fixed effects.

Column 1 reports estimates using all properties in LAUSD and basic year and neighborhood effects. The point estimate is negative and insignificant, which indicates that neighborhoods in areas that did not receive new schools saw, if anything, larger increases in house prices during the sample period. However, using uniform year effects for all of LAUSD may

[42]McMullen and K. E. Rouse (2012) also find that reading, but not math test scores are adversely affected by school facility overcrowding and congestion.

[43]Note that this is consistent with the evidence reported in Table 1.6, which showed only small increases in the ELA and attendance estimates when the stayers were excluded from the control group (column (2)). Given that stayers represent a small fraction of the control group, we would expect that the inclusion of stayers in baseline regressions only produces a small downward bias in the presence of small positive indirect effects.

confound the effects of differential price trends and shocks in different areas of the city and surrounding areas. Recall that the new schools are concentrated in East LA, where baseline prices were low and poverty rates high relative to the rest of the district. For example, if house prices in more affluent areas were already growing at a higher rate than those in less affluent areas (where the new schools were mainly built), difference-in-differences estimates of the effects of new schools could be biased downwards. Rather than impose parametric trends for each neighborhood, in column 2 we substitute year effects for year-by-high school zone[44] effects to more flexibly account for any differential local house price trends or changes. The point estimate flips sign and is statistically significant, implying that house prices rise 6.0% (SE 1.8%) post construction in neighborhoods that receive new schools, relative to nearby property sales in the same year within the same initial high school attendance area.

Properties that are very far from new school zones are included as "never-treated" controls in columns 1 and 2, and even with the inclusion of year-by high school zone effects we may still be worried about bias from the inclusion of these properties. To account for this, in column 3 we drop "never-treated" properties further than one kilometer from a new school zone, and in column 4 we further restrict the sample to only those properties that ever receive a new school.[45] Results in columns 2 and 3 are nearly identical, and the estimated coefficient drops slightly to 4.4% (SE 1.1%) in column 4. Column 5 substitutes year effects for the year-by-high school zone effects introduced in column 2 – now unnecessary as we have limited the control group to properties near the new schools – and the point estimate increases slightly to 5.5% (SE 1.5%).

To address additional concerns that within-neighborhood difference-in-differences results may be biased by fixed unobserved property-level differences, we rely on repeat sales and estimate effects *within-property*, using property fixed effects to account for any differences. Columns 6 and 7 report estimates analogous to columns 4 and 5 using property fixed effects; property controls and neighborhood fixed effects are excluded. Here, variation comes only from properties sold multiple times during the sample window, resulting in a sample size reduction of nearly half. In column 6, estimation includes year-by-high school zone effects, while column 7 shows estimates where only year-specific effects are included. Estimated effects are very similar to analogous neighborhood fixed effects estimates in columns 4 and 5. Overall, estimates imply that house prices increase by roughly 4-6% post-construction in new school attendance areas.

Difference-in-differences coefficients correspond only to the mean difference in house prices pre vs post construction. Pre-existing differential trends between neighborhoods in the same initial high school zone could still induce bias, even with the inclusion of flexible year-by-high school zone effects. More importantly, difference-in-differences estimates ob-

[44]Recall, here we define school zones using pre construction boundaries from 2000, to eliminate concerns over endogenous new school attendance boundary formation. Reassuringly, this distinction makes no quantitative difference, as results are nearly identical when post construction boundaries are used instead (Table 1.16 panel B).

[45]By restricting to these properties, identification is coming solely off variation in the timing of when a specific neighborhood receives a new school facility.

scure the dynamics of effects, which could result in downward bias if capitalization occurs gradually, and/or in anticipation of construction. New school locations were announced on average 5 years before completion: real estate capitalization may occur in advance of school completion insofar as parents and other homebuyers are forward-looking and are able to anticipate whether a given property falls within the school assignment zone for the new school. On the other hand, initial uncertainty by parents in the actual improvements generated by a new school may lead to more gradual capitalization post-construction, as the quality of the new school is revealed.

To account for flexibly for any dynamics in the timing of capitalization effects, in Figure 1.8 we report event study estimates of the effects of new school constructions, corresponding to the specification in equation (1.6). In panel A, estimation includes only those properties ever within any new school zone and year-by-high school fixed effects, corresponding to the specification in column 4 of panel A of Table 1.10. In panel B, we include all never-treated properties in LAUSD as controls, corresponding to column 2. Effects are estimated relative to the year before school occupancy, which is omitted from the regression. Results in both panels of Figure 1.8 show little sign of pre-existing trends or dynamic anticipatory effects pre-construction: all estimated pre-construction effects are practically zero. Capitalization occurs somewhat gradually upon completion, with nearly all of the effect coming in the first two years after school completion, before stabilizing after three or more years. Three or more years after the new school construction, house prices in the new school attendance areas were 7% higher, slightly larger than the point estimates presented in Table 1.10.

As discussed in Section 1.6, schools that experienced large student outflows to new schools saw significant reductions in overcrowding and multi-track calendar utilization, and small but significant increases in the share of more advantaged students. Students at these schools also experienced gains in ELA scores and attendance. To what extent were these gains at existing "sending" schools capitalized into local house prices? In panel B of Table 1.10 we report difference-in-differences estimates where treatment is similarly defined for existing "sending" schools that experienced student outflows to newly constructed facilities.[46] Specifications in columns 1 and 2 correspond to those in columns 1 and 2 of panel A; specifications in columns 3-6 correspond to those in columns 4-7 in panel A. Overall, results provide little indication that house prices increased in the sending school neighborhoods. In column 1 the coefficient is positive and significant, but this result is not robust to the inclusion of year-by-high school zone effects in column 2, nor the exclusion of never-treated properties in columns 3-6. These results suggest that (a) parental valuation of new schools is driven by non-test score/amenity improvements at new schools, independent of the school calendar or level of overcrowding, and/or (b) improvements in school quality due to reductions in overcrowding and multi-track calendar utilization are less salient to prospective homebuyers, who may instead rely on school facility condition as a signal for underlying school quality.

[46] "Sending" schools are defined as schools that have a non-trivial share (greater than 10%) of student enrollment that experienced a substantial cohort outflow to a newly constructed school. The treatment year for sending schools is analogously defined as the treatment year for stayers; i.e. the year in which the peer outflow occurred.

By neighborhood price

While new school quality was similar across treated neighborhoods,[47] the tax price of the new facilities faced by district residents was greater in areas with higher property values.[48] Later, in Section 1.8, we use the estimated house price effect for a welfare calculation, applying the coefficient to the mean home value in LAUSD. But insofar as home prices capitalize local investment, one might expect larger percentage effects on prices in low-price neighborhoods than in high-price neighborhoods. If so, applying the average percentage treatment effect to the average house price could overstate the aggregate impact. Empirically this does not appear to be the case. In Figure 1.9 we report heterogeneity in estimated treatment effects by neighborhood prior mean house prices. We define neighborhood prior mean house prices as the average house price in a neighborhood over all pre-treatment years in the sample, 1995-2001. Estimates of β from equation (1.5) are shown interacted with \$100,000 bins of neighborhood prior mean house prices.[49] With the exception of the \$500,000 -\$600,000 bin, all effects are similar and statistically significant, providing little evidence of smaller estimated treatment effects in areas with higher property values.

Local boundary and spillover effects

Increased demand for neighborhoods receiving new schools could have differential effects on house prices near the boundaries of new school neighborhoods. School assignment boundaries do not stay constant in perpetuity, and due to uncertainty over future boundary locations, capitalization effects may be smaller near the boundaries within new school zones. In addition, if home buyers substitute housing in existing school zones for housing purchases in new school zones, prices could decline in other LAUSD neighborhoods. On the other hand, new school constructions and changing neighborhood composition could lead to spillovers that increase house prices both within and near new school zones. Prices in nearby neighborhoods that did not receive new schools could increase due to positive externalities from neighborhood upgrading (e.g. Hornbeck and Keniston (2017)). Moreover, new schools could act as a direct amenity that generates positive benefits (e.g. increased park/playground space) both within and outside the actual attendance areas. Estimates in Figure 1.10 and Table 1.11 assess the extent to which the effect of new school constructions varies by distance to the attendance boundary, and whether new schools generate spillover effects beyond the attendance zone.

Table 1.11 reports estimates of treatment effects by distance to the school attendance boundary. Column 1 repeats baseline estimates from column 2 of Table 1.10 panel A. In

[47] In conversations with district officials, it was stated that much of the variation in project cost was due to site-specific acquisition expenses, and not systematic differences in new facility quality.

[48] Unlike in the model presented in Section 1.3, which assumed a constant lump sum tax for all households, property owners in higher-priced areas contributed a greater dollar amount towards district bond revenues.

[49] Note: the \$100K bin includes a small number of properties in neighborhoods with mean house prices below \$100K; the \$600K bin includes properties in all neighborhoods with mean house prices greater than \$600K in 1995-2001.

column 2, we add a coefficient for distance from the boundary (in kilometers) and distance to the boundary interacted the treatment dummy. Both added coefficients are small and insignificant, and do little to affect the point estimate on the treatment indicator in the first row. Column 3 adds an indicator for being within 2 kilometers of a new school zone, after completion. For these properties, we assign the treatment date of the nearest new school construction. Effects are estimated relative to properties greater than 2 kilometers from any new school boundary. The estimated effect on properties just outside the new school's attendance zone is -1.3% (SE 1.0%) and insignificant. These estimates provide little evidence of substitution patterns that indicate decreased demand for housing in existing school attendance zones within 2 kilometers of a new school zone, nor that new school zones generate positive spillovers in nearby neighborhoods, as would be expected if new schools induced general neighborhood amenity upgrading. Column 4 adds in controls for distance and the interactions with the treatment indicators. Here the treatment dummy for being outside the zone is highly negative and significant, while the interaction with distance is positive and significant, implying a large negative effect on house prices immediately after crossing a new school attendance boundary that fades out within 1 kilometer outside the boundary.

Figure 1.10 provides a non-linear visualization of the pattern reported in column 4. Each point represents a difference-in-differences treatment effect estimate interacted with distance to the new school attendance zone boundary, in 400 meter bins. Properties with positive distance are located within new school boundaries, while those with negative distance are in school zones where the residential assignment is to an existing school. Results indicate that within the new school zones, capitalization is roughly constant at approximately 5% for all distance bins. We find no evidence of smaller effects closer to the boundary. Properties within 400 meters but outside of the boundary actually see statistically significant declines in house prices of 4.9% (SE 1.7%) post-construction, providing suggestive evidence of negative spillovers for properties that are "unlucky" enough to fall just outside the new school zone. These negative spillover effects quickly diminish however; point estimates for distances greater than 1.2 km are positive, though insignificant, consistent with the findings from Table 1.11. This pattern is consistent with cross-neighborhood substitution within very narrowly defined markets, wherein demand for properties located marginally outside the new school zones decreases for prospective homebuyers searching within the vicinity a new school.

1.8 Welfare Analysis

Thus far we have shown that new school constructions in LAUSD generated large student cognitive and non-cognitive gains. These improvements in school quality - physical and educational - were capitalized into the real estate market, as properties in new school attendance areas saw large and significant increases in prices post construction. In this section we use our estimated price effects to compute the implied willingness-to-pay for residents who received new schools. As outlined in Section 1.3, the magnitude of difference-in-differences estimates of the relative price change induced by new school constructions provides a benchmark to assess the economic efficiency of the spending program. One-third of households in LAUSD reside in a new school attendance zone. Thus, if the estimated relative price change is less than the per household cost of the program, multiplied by three, then we can infer that homebuyers value the new schools less than the cost of building them, and therefore that using taxpayer money to build new schools reduced welfare. Conversely, if the estimated price change exceeds this we can infer that the additional expenditures were valued in excess of the total program cost by homebuyers.

This computation relies on strong assumptions. Most notably, we assume that the observed price change affects all household units in LAUSD, although we only estimate on the subsample of single-unit properties that sold during the sample window. According to the 2005-2009 American Community Survey (ACS), there are 1.52 million non-vacant housing units in LAUSD. The total cost of the program was \$9.17 billion, meaning that the average cost to a housing unit of the program is approximately \$6,045 in present value. During the treatment period from 2002-2012, the average sale price (within-sample) of properties in zones that received new schools was \$494,650. Using the estimates in Table 1.10 panel A, this implies a price change in the range of \$21,765 to \$29,679, where the preferred estimates from column 2 using all properties in LAUSD are the upper bound of that range. Comparing this to the program cost per housing unit in a new school attendance zone, $3 * \$6,045 = \$18,135$, implies a willingness-to-pay ratio in the range of 1.2 to 1.64.[50] Put differently, each additional dollar of capital expenditures by the district generated 1.2 to 1.64 additional dollars in the real estate market. These results suggest that the value to families of the school capital expenditure program was greater than the program cost, implying the program raised welfare.

The real estate valuation of the program incorporates the market valuation of all potential benefits generated by the new school program, beyond simply the effects related to increased academic performance of students. However, many studies of educational interventions rely on extrapolations of test score effects to assess a program's efficiency. Using the estimates presented in Chetty, John N Friedman, et al. (2011), we can project forward the gain in future earnings from the observed test score gains. Chetty et al. use experimental variation in classroom quality to estimate that a 0.1 standard deviation increase in test scores[51] leads to a 1.3% increase in earnings at age 27. To extrapolate our estimates forward, we first compute the present discounted value of future earnings for future cohorts:

$$PDV_{cohort} = N_c \sum_{t=16}^{56} \frac{E_t}{(1+\delta)^t}$$

[50] Using instead the average price over all properties in the treatment period, as suggested by the model in Section 1.3, would increase this WTP ratio by roughly 33%, to a range of 1.62 to 2.21.

[51] Notably, this is for kindergarten scores. However, non-experimental estimates in the same paper show that the correlation between test scores and earnings grows with age, suggesting that these effects may underestimate the effects of improvements in later grades.

where E_t = earnings gain at each age, which we compute under the assumption of a constant percentage gain of 1.3% per 0.1 SD increase in test scores, using age-earnings profiles from the March CPS.[52] The average elementary school student is 11 years old, therefore we discount forward 16 years to age 27, and count benefits until retirement at age 67. From our data, roughly 16% of students entering elementary school, 13% of students entering middle school, and 25% of students entering high school in LAUSD were in a newly constructed school facility. Plugging this in and using the estimated effects on math test scores, assuming a 3% discount rate, yields a present discounted value of future earnings per cohort of $150 million. From our facilities data, we estimate that a brand new facility would take roughly 35 years to depreciate to the mean condition of existing facilities in LAUSD. Assuming the effects are constant for this 35 year horizon and discounting the earnings of future cohorts implies a gain in future earnings of $3.8 billion in present discounted value. The total program cost was $9.17 billion, implying that the gain in future earnings from test score improvements covers roughly 40% of the total program cost.[53]

Real estate capitalization greatly exceeds the estimated increases in future earnings from test score improvements, providing strong evidence that parental valuation of educational expenditures exceeds benefits captured by test scores alone.[54] New schools also generated improvements in student non-cognitive outcomes, improvements in school safety and health, and allowed for increased access to extra-curricular opportunities, among many other benefits. While test score improvements provide a useful benchmark for interpreting the efficiency of educational interventions, they are likely to severely understate the true benefits of capital infrastructure investments.

1.9 Conclusion

In this paper we provide robust and comprehensive estimates of the effects of educational capital investments on student outcomes and neighborhood house prices. To date, the literature on the effects of school capital investments has been mixed and inconclusive; many prior studies are underpowered to detect modest effects, often relying on district-level average outcomes to study the impacts of capital expenditure programs that impact only a subset of students (Figure 1.11). Studying the largest school construction program in US history, we provide robust new evidence that school facility investments lead to modest, gradual improvements in student test scores, large immediate improvements in student attendance, and significant improvements in student effort. New facilities also generated

[52] We compute the age-earnings profiles using data from 2012-2016, and use the average earnings, including those with zero earnings. This follows the procedure in Chetty, John N Friedman, et al. (2011), but may overstate impacts if earnings of LAUSD students are below average over the life cycle.

[53] Here we are not counting any indirect improvements for students who stayed behind at existing schools. Including these would slightly increase aggregate future earnings gains, but would not change the qualitative conclusion that future earnings gains from test score improvements do not cover total program costs.

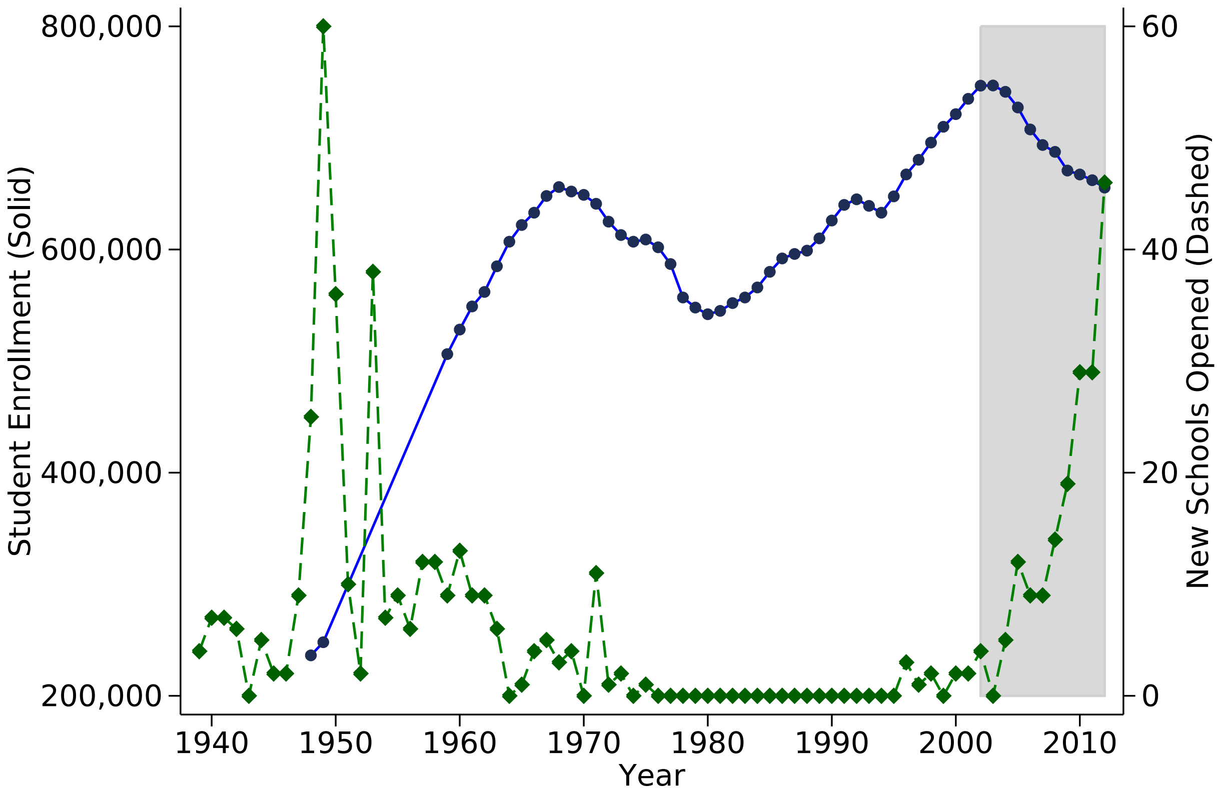
[54] Appendix Figure 1.20 plots school-level test score treatment effects against school-level house price effects, showing little systematic relationship between the two in both math and ELA.

indirect improvements for students elsewhere in the district who did not attend new facilities, but nonetheless saw improvements in their school environments due to peer outflows to new facilities. Reductions in overcrowding and the elimination of "multi-track" academic calendars only account for some of the observed gains, implying that capital improvements themselves were responsible for student gains beyond reductions in congestion.

New school constructions induced large increases in neighborhood house prices upon completion, implying significant parental valuation of improvements in school quality, generally defined. House prices increased substantially in areas that received new schools, although we find no evidence of similar price increases in existing school zones that sent students to new schools and experienced corresponding reductions in overcrowding. Overall, house price estimates imply that the total real estate capitalization exceeded program cost, and suggest an implied willingness-to-pay on behalf of district residents of 1.2 to 1.6 for one dollar of per-household school capital investment. Willingness-to-pay estimates provide evidence that prior capital spending had been inefficiently low in the distinct, and that the targeted program to improve facilities for the most disadvantaged students in the district generated aggregate welfare increases in the district. These findings are especially relevant for large, urban districts and other districts serving low-income students with a history of facilities underinvestment, and imply that policies to improve school capital can be productive and efficient uses of public funding.

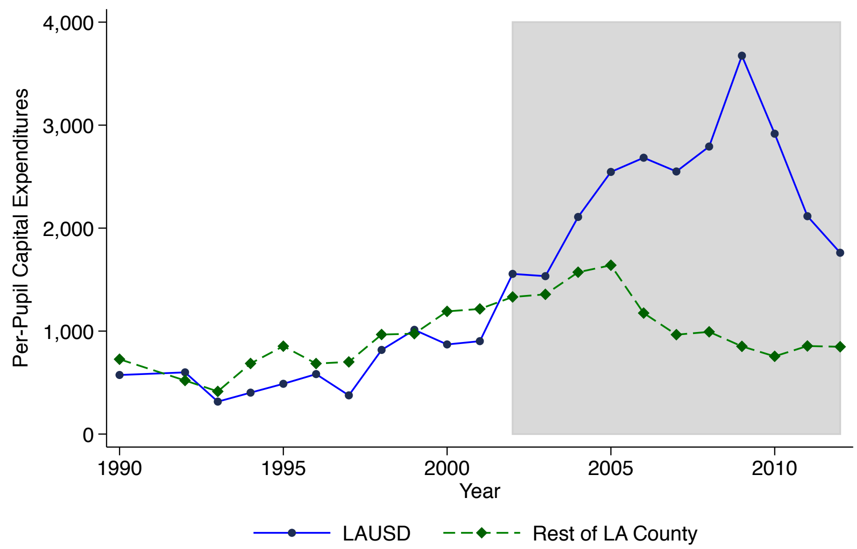
1.10 Figures

Figure 1.1: School construction and enrollment, LAUSD 1940-2012

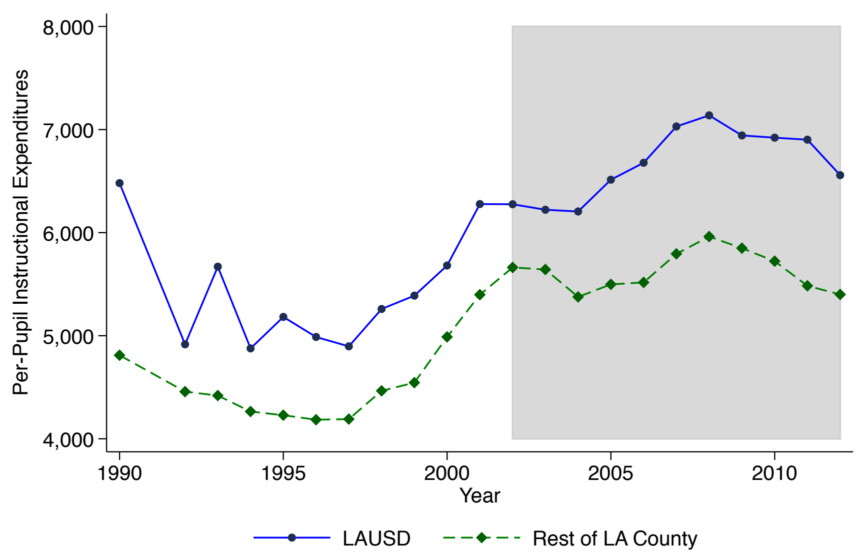


Notes: Solid blue line depicts student enrollment by year (left axis) and dashed green line depicts the number of new school facilities opened in a given year (right axis). Shaded area from 2002-2012 shows the treatment period covered in the main analysis. The number of new school openings only includes facilities still open in 2008, and is computed as the minimum age over all buildings that comprise a given school. Historical student enrollment data were obtained from the California Department of Education.

Figure 1.2: Spending per pupil, LAUSD vs LA County



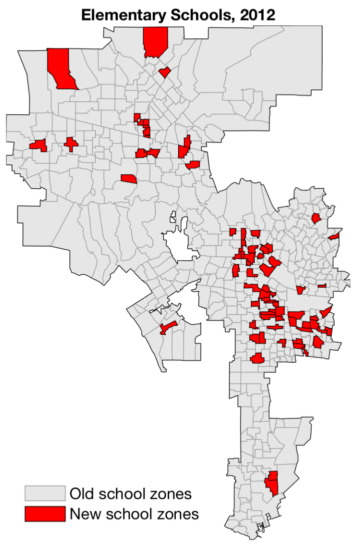
(a) Capital



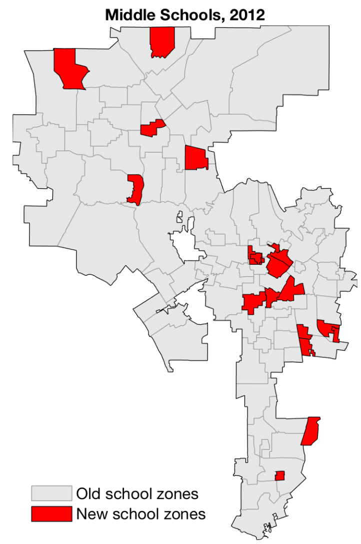
(b) Instructional

Notes: Panel A shows per-pupil capital expenditures and panel B shows per-pupil instructional expenditures. Expenditures are expressed in real 2013 dollars. In both panels, the expenditures for LAUSD (solid blue line) and the student-weighted average of all other LA County public school districts (dashed green line) are shown. The shaded area from 2002-2012 shows the treatment period covered in the main analysis. Expenditure data were from the National Center for Education Statistics (NCES) annual census of school districts and from the Census of Governments.

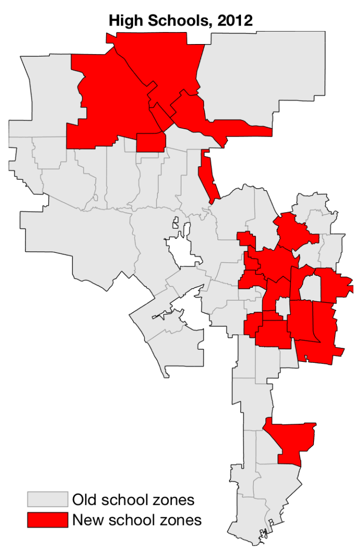
Figure 1.3: LAUSD school attendance zones, 2012



(a) Elementary



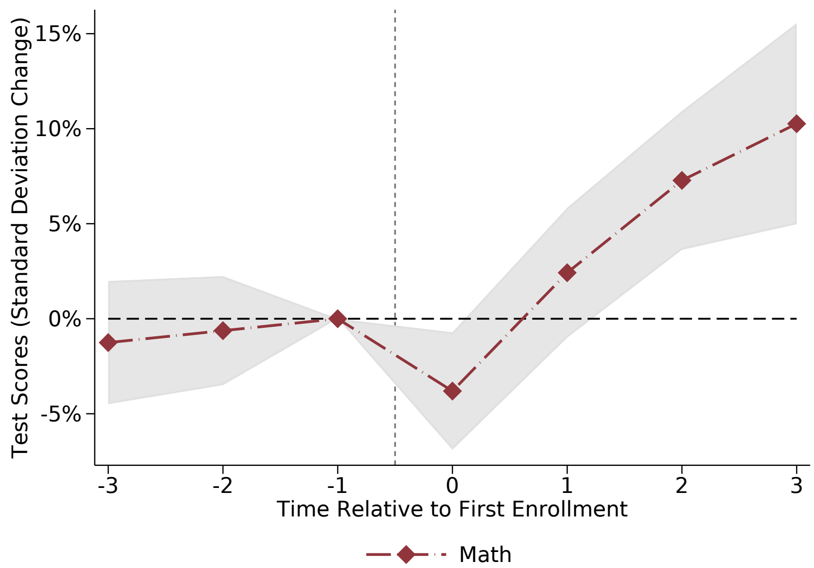
(b) Middle



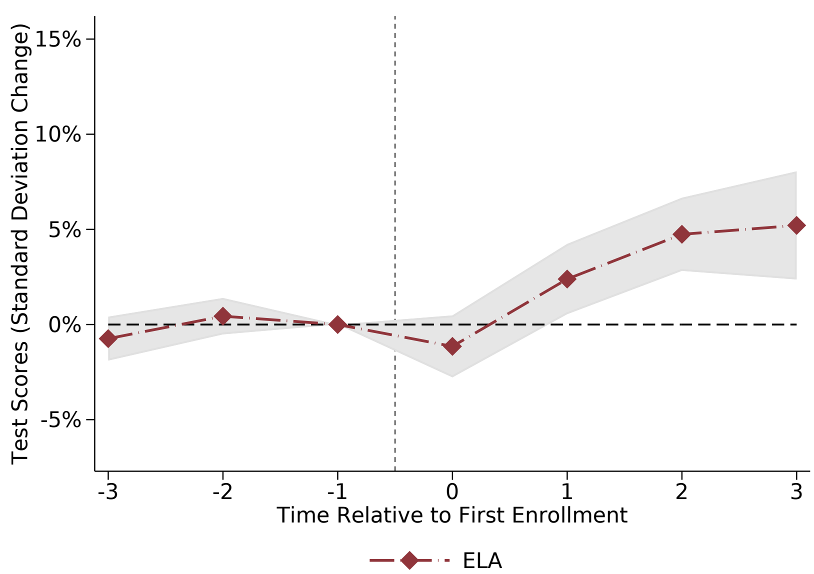
(c) High

Notes: Figure displays school attendance boundaries for elementary schools (panel A), middle schools (panel B), and high schools (panel C) in LAUSD in 2012. Shaded areas in red denote attendance zones that correspond to schools newly constructed during the sample period from 2002-2012.

Figure 1.4: Test score effects



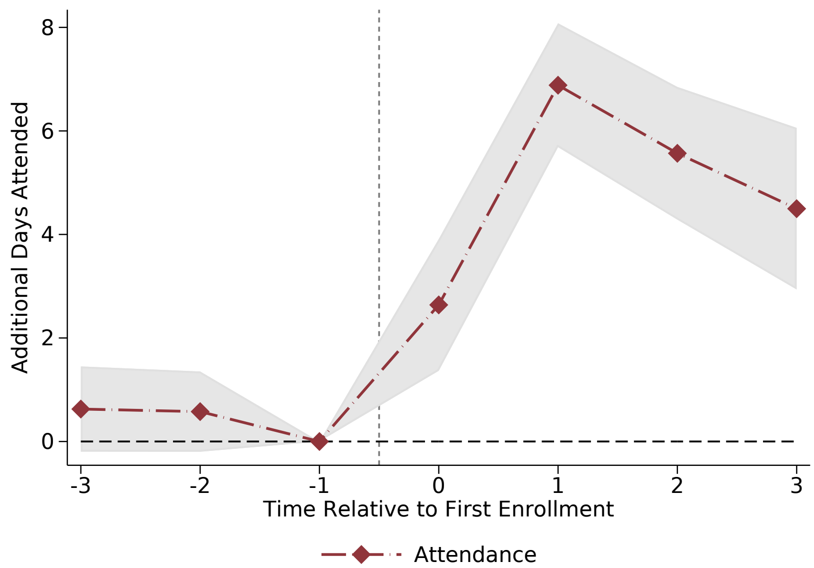
(a) Math



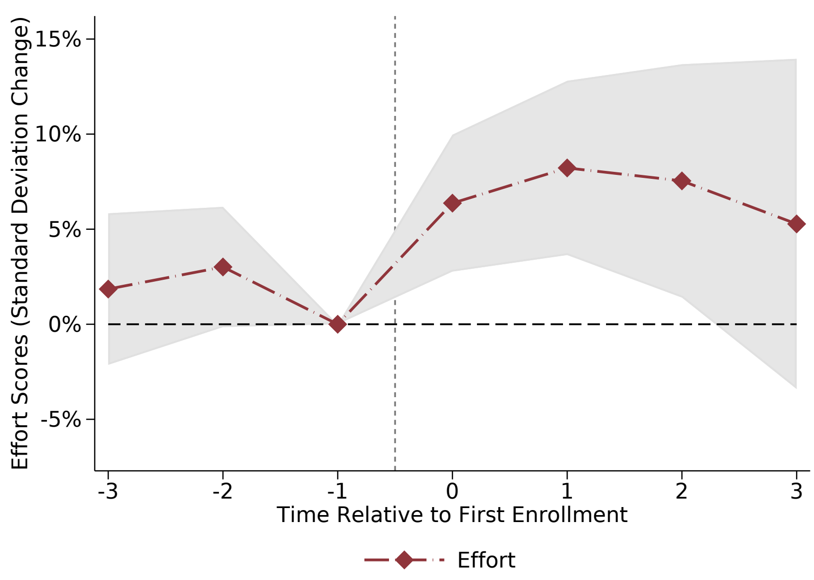
(b) ELA

Notes: Figures shows estimated coefficients from event study regressions following equation (1.3). Dependent variables are standardized math test scores for students in grades 2-7 (panel A) and standardized english-language arts test scores for students in grades 2-11 (panel B). Test scores are standardized relative to the statewide mean and standard deviation for each year-grade-subject exam. The shaded areas denote 95% confidence intervals for the estimated coefficients. Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Figure 1.5: Non-cognitive effects



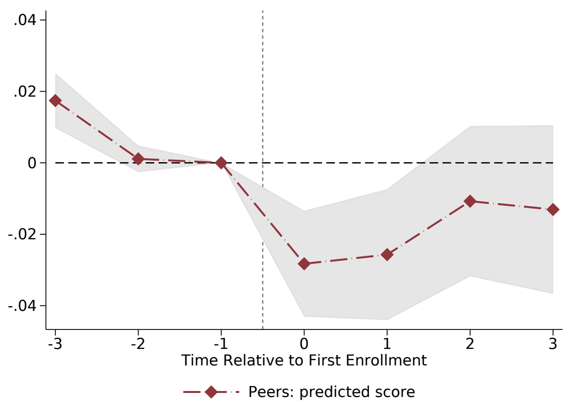
(a) Attended days



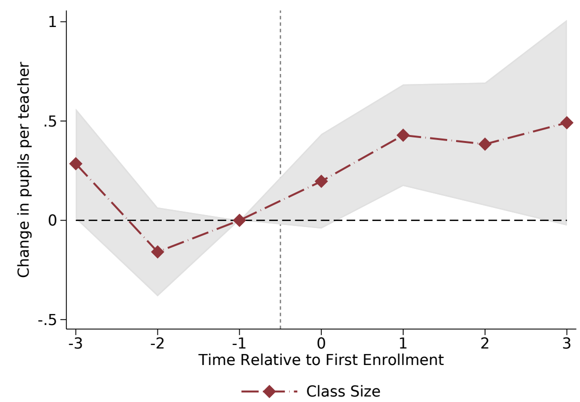
(b) Teacher-reported student effort

Notes: Figures show estimated coefficients from event study regressions following equation (1.3). Dependent variables are annual days attended (panel A) and standardized teacher-reported effort scores for students in grades K-5 (panel B). The shaded areas denote 95% confidence intervals for the estimated coefficients. Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

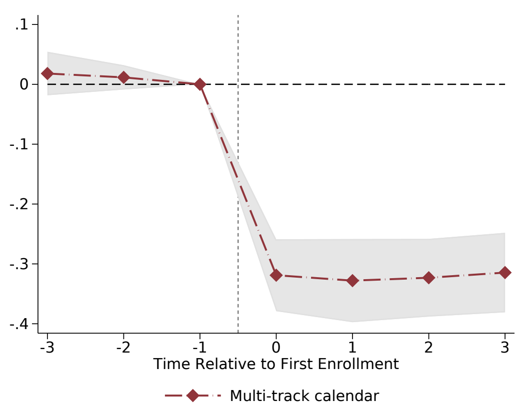
Figure 1.6: School effects



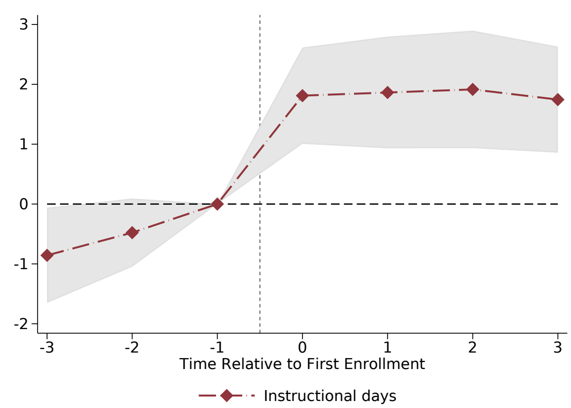
(a) Peers: predicted score



(b) Class size



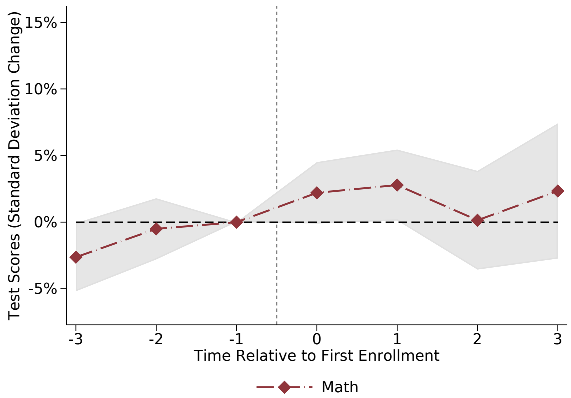
(c) Multi-track calendar



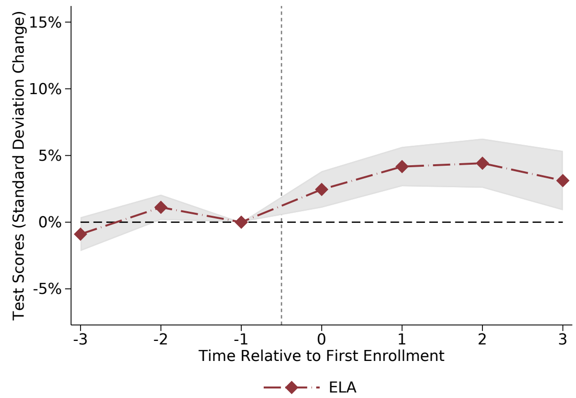
(d) Total instructional days

Notes: Figures show estimated coefficients from event study regressions following equation (1.3). Dependent variables are leave-out school mean predicted test scores (panel A), class size for students in grades K-5 (panel B), multi-track calendar status (panel C), and total instructional days for a given school-year (panel D). The shaded areas denote 95% confidence intervals for the estimated coefficients. Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

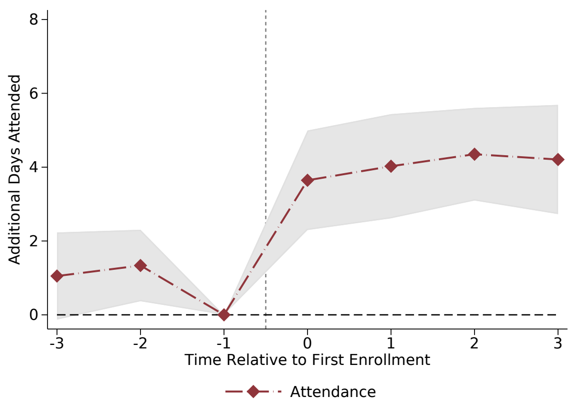
Figure 1.7: Student effects: Stayers



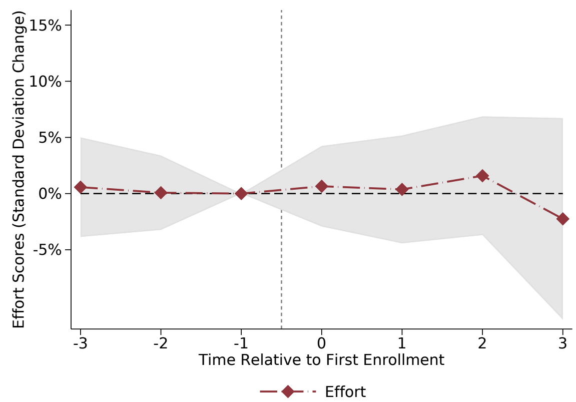
(a) Test scores: Math



(b) Test scores: ELA



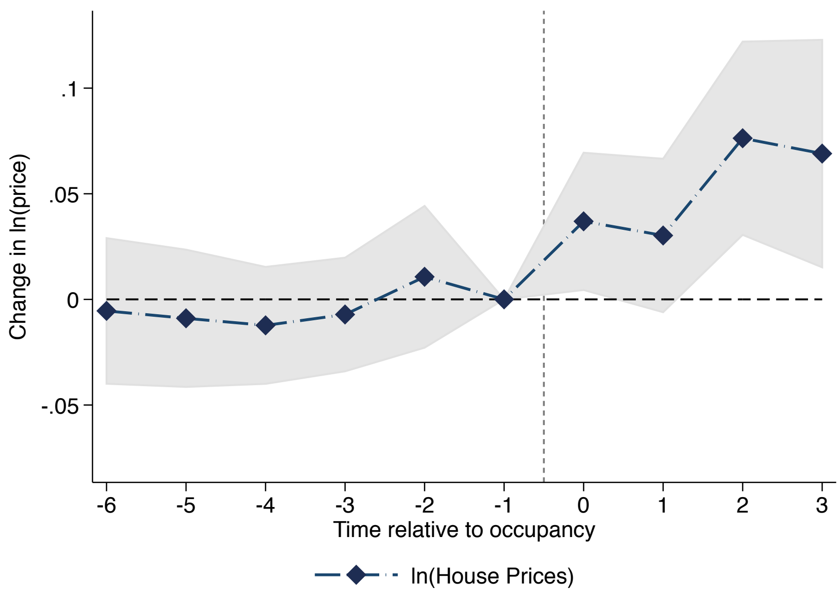
(c) Attended days



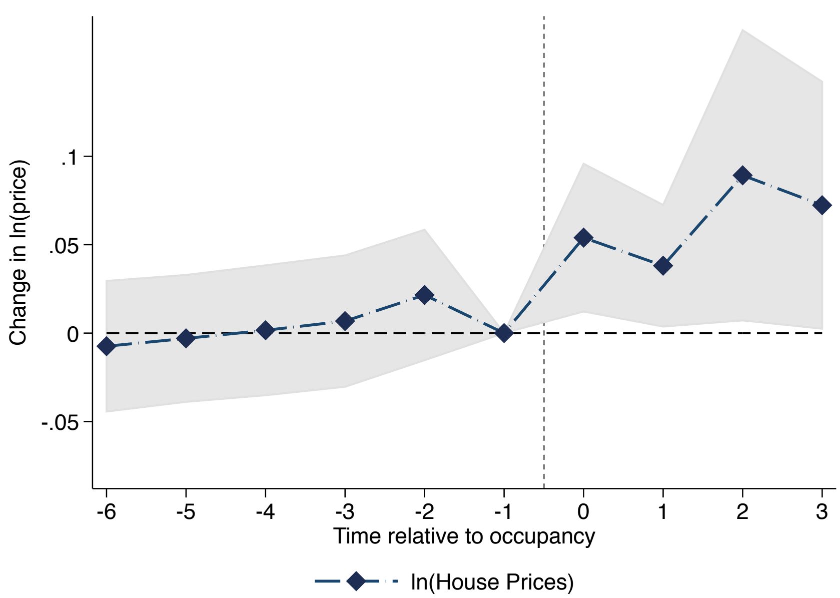
(d) Teacher-reported student effort

Notes: Figures show estimated coefficients from event study regressions following equation (1.3) for students that had 10% or more of their school-grade cohort exit to a newly constructed school. Event time is centered relative to the year of the peer outflow. Dependent variables are standardized math test scores for students in grades 2-7 (panel A), standardized english-language arts test scores for students in grades 2-11 (panel B), annual days attended (panel C), and standardized teacher-reported effort scores for students in grades K-5 (panel D). The shaded areas denote 95% confidence intervals for the estimated coefficients. Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Figure 1.8: House price effects



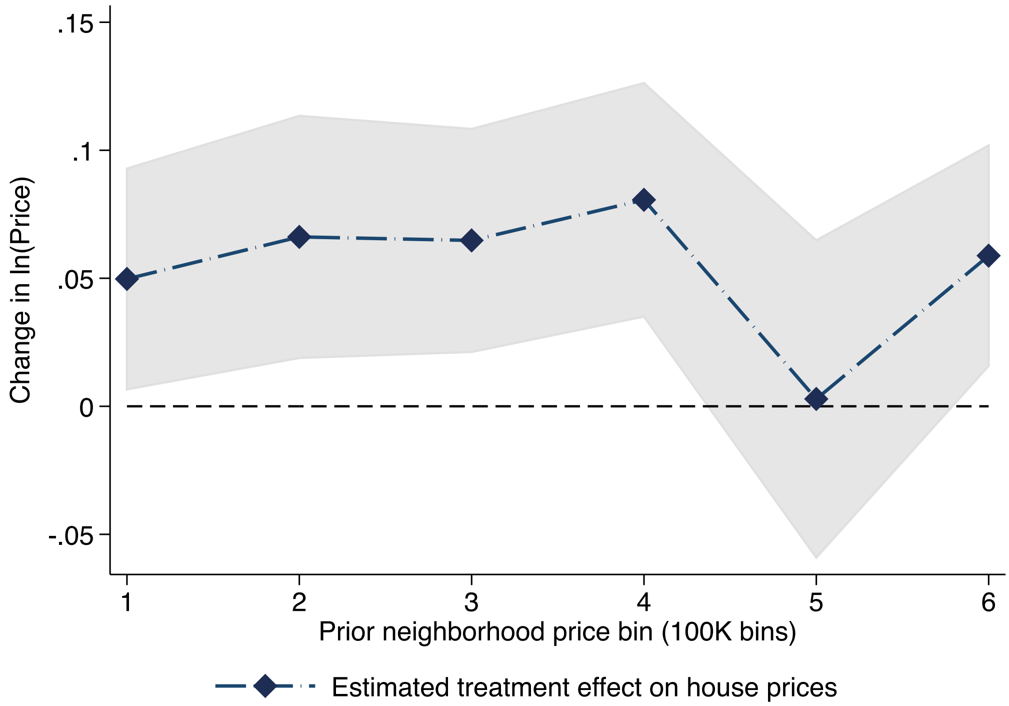
(a) House prices: Only treated



(b) House prices: All LAUSD

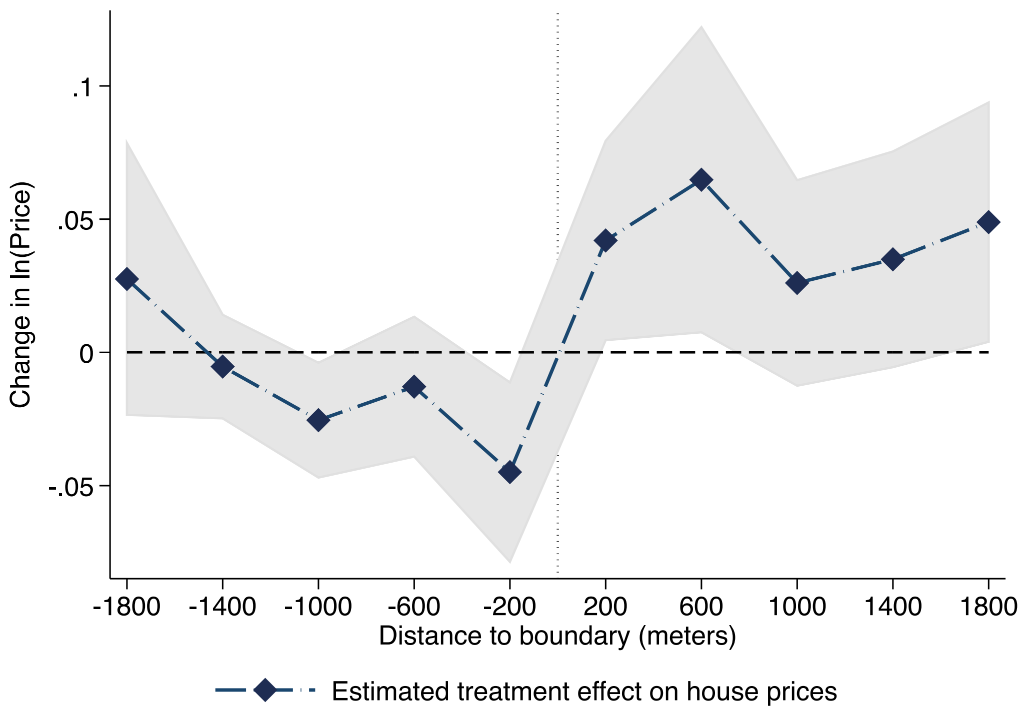
Notes: Figures show estimated coefficients from event-study regressions following equation (1.6). Dependent variable in both panels is the ln(sale price). In panel A, only properties that are ever in a new school attendance zone are included in the estimation, corresponding to baseline estimates presented in column 4 of Table 1.10. In panel B, all properties in LAUSD in the data sample are included in estimation, corresponding to baseline estimates presented in column 2 of Table 1.10. Specifications include property-specific controls, year-by-high school zone fixed effects, neighborhood fixed effects, and month fixed effects. Standard errors are clustered by neighborhood.

Figure 1.9: Heterogeneity: By neighborhood mean prior house prices



Notes: Figure shows estimated coefficients from a difference-in-differences regression based on equation (1.5), where the treatment indicator is interacted with indicators for $100,000 bins of prior neighborhood average prices. Bin 1 also includes average neighborhood house prices less than $100K, while bin 6 includes all neighborhoods with average house prices above $600K; all other bins only include a $100K range. Prior neighborhood average house prices are calculated using data from pre-construction property sales from 1995-2001. All properties in LAUSD in the data sample are included in estimation, corresponding to baseline estimates presented in column 2 of Table 1.10. All specifications include property-specific controls, year-by-high school zone fixed effects, neighborhood fixed effects, and month fixed effects. Standard errors are clustered by neighborhood.

Figure 1.10: Spillovers: Effects by distance to school attendance boundary



Notes: Figure shows estimated coefficients from a difference-in-differences regression based on equation (1.5), where the treatment indicator is interacted with indicators for 400 meter bins of distance to the new school attendance zone in 2012. Properties with positive (negative) distance are inside (outside) the new school attendance zones. Properties outside the attendance zone and within 2 km of a new school attendance zone are assigned the construction date corresponding to the nearest new school attendance zone boundary. Each point reports the estimated coefficient for the treatment indicator interacted with the corresponding distance bin. Points are located at the midpoint of each distance bin (i.e. the estimate at 200m corresponds to the 0-400m distance bin). All properties in LAUSD in the data sample are included in estimation, corresponding to baseline estimates presented in column 2 of Table 1.10. Specifications include property-specific controls, year-by-high school zone fixed effects, neighborhood fixed effects, and month fixed effects. Standard errors are clustered by neighborhood.

1.11 Tables

Table 1.1: Summary statistics, new school projects

	Mean	Median	Min	Max
Total cost (million USD)	81.9	56.5	11.1	578.7
New student seats	1,050	800	162	3,440
New classrooms	40.3	32	6	130
Building SQFT	100,585	70,115	12,507	391,840
Completion year	2008	2008	2002	2012
Site designation to completion (yrs)	5.18	5	2	9
Construction to completion (yrs)	2.11	2	1	5
New School Codes	1.25	1	1	5
Total New School Campuses	114			
Total New School Codes	143			

Notes: Table reports summary statistics for new school project data, at the project level.

Table 1.2: Summary statistics, LAUSD student data

	All LAUSD	Never Treated	Always Treated	Switchers	Stayers
Free/reduced-price lunch	0.82	0.80	0.93	0.95	0.91
Hispanic	0.73	0.71	0.85	0.89	0.83
Black	0.11	0.11	0.05	0.05	0.07
White	0.09	0.10	0.03	0.02	0.05
Asian	0.04	0.04	0.04	0.02	0.03
Parent: any college	0.26	0.27	0.23	0.16	0.21
English spoken at home	0.32	0.34	0.27	0.17	0.21
Predicted test score	-0.26	-0.24	-0.29	-0.40	-0.34
Math score ($t = -1$)				-0.35	-0.19
ELA score ($t = -1$)				-0.52	-0.37
Days attended ($t = -1$)				156.73	155.37
N student-years	7,284,175	6,471,912	108,611	703,652	1,307,071

Notes: Table reports summary statistics for LAUSD student data, at the student-year level.

Table 1.3: Summary statistics, LA County assessor data

	All LAUSD	New School Zones	Existing School Zones
Sale price (2015$)	565,801	416,509	636,010
Building SQFT	1,664	1,539	1,722
Number of bedrooms	2.9	2.9	2.8
Number of bathrooms	2.2	2.1	2.3
Building age	44	45	44
Effective age	39	40	39
Useable lot SQFT	5,238	5,704	5,018
N property sales	505,835	161,795	344,040
N properties	350,299	115,247	235,052

Notes: Table reports summary statistics for LA County Assessor data, at the property sale level.

Table 1.4: Student effects, cognitive

	Math Score			ELA Score		
	(1)	(2)	(3)	(4)	(5)	(6)
New School * Trend	0.029***	0.034***	0.031***	0.019***	0.020***	0.017***
	(0.007)	(0.008)	(0.008)	(0.004)	(0.004)	(0.004)
New School		-0.021	-0.028		-0.003	-0.014
		(0.017)	(0.017)		(0.008)	(0.009)
Trend			0.004*			0.004***
			(0.002)			(0.002)
Cumul. 4yr Effect	0.086***	0.080***	0.064***	0.058***	0.058***	0.037***
	(0.021)	(0.022)	(0.024)	(0.011)	(0.011)	(0.012)
Grade FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Stu FEs	X	X	X	X	X	X
N student-years	2,935,156	2,935,156	2,935,156	4,716,377	4,716,377	4,716,377
N students	735,811	735,811	735,811	971,568	971,568	971,568
N treated students	87,132	87,132	87,132	99,685	99,685	99,685
N treated schools	78	78	78	126	126	126
R2	0.82	0.82	0.82	0.84	0.84	0.84

Notes: Table reports estimates of parametric event study models corresponding to equation (1.4). Columns 1 and 4 include only the coefficient for the change in growth (β_2); β_1 and β_3 are constrained to be zero. Columns 2 and 5 include coefficients for both the immediate effect (β_1) and the change in growth (β_2); β_3 is constrained to be zero. Columns 3 and 6 include all coefficients, corresponding exactly to the specification in equation (1.4). Row 4 reports the implied cumulative test score effect after four years, equal to $3\beta_2$ in columns 1 and 4, and $\beta_1 + 3\beta_2$ in columns 2-3 and 5-6. Dependent variable is the standardized math test score (grades 2-7) in columns 1-3. In columns 4-6 the dependent variable is the standardized ELA test score (grades 2-11). Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Table 1.5: Student effects, non-cognitive

	Days Attended			Effort Score		
	(1)	(2)	(3)	(4)	(5)	(6)
New School	3.973***	3.398***	3.314***	0.061***	0.063***	0.056***
	(0.551)	(0.599)	(0.633)	(0.017)	(0.017)	(0.018)
New School * Trend		0.777***	0.745***		-0.002	-0.008
		(0.196)	(0.204)		(0.012)	(0.013)
Trend			0.036			0.006
			(0.085)			(0.004)
Grade FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Stu FEs	X	X	X	X	X	X
N student-years	5,350,867	5,350,867	5,350,867	1,924,572	1,924,572	1,924,572
N students	1,121,933	1,121,933	1,121,933	552,855	552,855	552,855
N treated students	116,947	116,947	116,947	71,636	71,636	71,636
N treated schools	143	143	143	75	75	75
R2	0.51	0.51	0.51	0.63	0.63	0.63

Notes: Table reports estimates of parametric event study models corresponding to equation (1.4). Columns 1 and 4 include only the coefficient for the immediate new school effect (β_1); β_2 and β_3 are constrained to be zero. Columns 2 and 5 include coefficients for both the immediate effect (β_1) and the change in growth (β_2); β_3 is constrained to be zero. Columns 3 and 6 include all coefficients, corresponding exactly to the specification in equation (1.4). Dependent variable is the annual days attended in columns 1-3. In columns 4-6 the dependent variable is the standardized average teacher-reported effort score (grades K-5). Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Table 1.6: Student effects, robustness

	Baseline	No Stayers	Only Treated	Only Switchers	Balanced
ELA Score					
New School * Trend	0.019***	0.022***	0.018***	0.016***	0.027*
	(0.004)	(0.004)	(0.005)	(0.005)	(0.014)
Math Score					
New School * Trend	0.029***	0.029***	0.034***	0.035***	0.059*
	(0.007)	(0.007)	(0.011)	(0.012)	(0.033)
Days Attended					
New School	3.97***	4.33***	4.02***	4.43***	8.54***
	(0.55)	(0.57)	(0.78)	(0.79)	(1.65)
Effort Score					
New School	0.061***	0.061***	0.077***	0.089***	0.045
	(0.017)	(0.017)	(0.024)	(0.027)	(0.060)

Notes: Table reports estimates of parametric event study models corresponding one-parameter versions of equation (1.4). Panels A and B include only the coefficient for the change in growth (β_2); β_1 and β_3 are constrained to be zero. Panels C and D include only the coefficient for the immediate new school effect (β_1); β_2 and β_3 are constrained to be zero. Dependent variables are standardized english-language arts test scores (panel A), standardized math test scores (panel B), annual days attended (panel C), and standardized average teacher-reported effort scores (panel D). Estimates in column 1 repeat baseline one-parameter estimates from columns 1 and 4 of Tables 1.4 and 1.5. Column 2 excludes "staying" students that had 10% or more of their school-grade cohort exit to a newly constructed school. Column 3 excludes never-treated students. Column 4 restricts estimation only to those students observed at an existing school prior to attending a school at a new facility. Column 5 restricts to a balanced sample with 5 years of data in panels A and C, or 3 years of data in panels B and D. Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Table 1.7: Student effects heterogeneity, mechanisms

	Math	ELA	Attendance	Effort
Pooled (switchers only)	0.035***	0.016***	4.432***	0.089***
	(0.012)	(0.005)	(0.789)	(0.027)
By multi-track:				
Multi track	0.039***	0.017***	8.214***	0.097***
	(0.014)	(0.006)	(0.788)	(0.029)
Single track	0.019	0.011*	2.741***	0.042
	(0.013)	(0.006)	(0.728)	(0.031)
p-value	0.15	0.37	0.00	0.11
By prior SQFT pp:				
Low prior SQFT pp	0.031**	0.017***	6.254***	0.071***
	(0.013)	(0.006)	(0.744)	(0.027)
High prior SQFT pp	0.037**	0.007	4.668***	0.098***
	(0.017)	(0.006)	(0.797)	(0.033)
p-value	0.66	0.15	0.03	0.43
By share permanent classrooms:				
Low share permanent	0.038***	0.019***	6.133***	0.092***
	(0.013)	(0.005)	(0.666)	(0.024)
High share permanent	0.020	0.003	4.955***	-0.013
	(0.018)	(0.007)	(0.903)	(0.044)
p-value	0.29	0.02	0.14	0.01
By prior building age:				
Below median age	0.026**	0.013**	5.625***	0.075**
	(0.013)	(0.005)	(0.647)	(0.032)
Above median age	0.045***	0.017**	6.095***	0.092***
	(0.017)	(0.007)	(0.863)	(0.035)
p-value	0.23	0.53	0.52	0.68
By prior building FCI:				
Low FCI	0.030*	0.016***	6.806***	0.125**
	(0.018)	(0.006)	(0.922)	(0.052)
High FCI	0.035**	0.013**	4.910***	0.068**
	(0.014)	(0.006)	(0.632)	(0.027)
p-value	0.78	0.79	0.02	0.29

Notes: Table reports estimates of one parameter event study models. Dependent variables are ELA scores (column 1), math scores (column 2), annual days attended (column 3), and standardized teacher-reported effort scores (column 4). Panel A repeats one-parameter estimates from column 4 of Table 1.6. Panel B reports estimates of coefficients interacted with prior school multi-track status. Remaining panels show coefficients on the interactions for being below or above the median in terms of prior school SQFT per pupil (panel C), prior school share permanent classrooms (panel D), prior school age (panel E), and prior school FCI (panel F). Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Table 1.8: Teacher changes at new schools

(a) Demographics

	(1) Age	(2) Experience	(3) MA+	(4) Pr(New)
New School	-3.289***	-2.659***	0.042***	0.054***
	(0.336)	(0.227)	(0.012)	(0.005)
Grade FEs	X	X	X	X
Year FEs	X	X	X	X
Stu FEs	X	X	X	X
N student-years	3,935,106	3,927,063	3,931,757	5,902,165
N students	926,501	925,300	926,203	1,140,815
N treated students	108,323	108,124	108,252	121,887
N treated schools	137	137	137	143
R2	0.32	0.36	0.28	0.29

(b) Value-added

	VA: Average (pre-switch)		VA: Novice/Experienced gap	
	(1) Math	(2) ELA	(3) Math	(4) ELA
New School	-0.005	-0.010**	-0.003	0.015
	(0.007)	(0.004)	(0.012)	(0.012)
Grade FEs	X	X	X	X
Year FEs	X	X	X	X
Stu FEs	X	X	X	X
N student-years	2,443,716	4,265,444	1,267,199	2,347,897
N students	689,206	955,346	432,813	672,731
N treated students	82,315	94,956	60,155	75,073
N treated schools	69	119	54	83
R2	0.61	0.56	0.38	0.33

Notes: Table reports estimates of difference-in-differences models corresponding one-parameter versions of equation (1.4), where only the coefficient for the immediate new school effect (β_1) is included; β_2 and β_3 are constrained to be zero. In panel A, dependent variables are teacher age (column 1), teacher years experience (column 2), an indicator for having a masters degree or higher (column 3), and an indicator for having a new teacher in either math or ELA (column 4). Panel B reports estimates where dependent variables are school-year averages of teacher value-added: in columns 1 and 2 dependent variables are average value-added scores based on prior-year observations at existing school facilities in math and ELA, respectively. In columns 3 and 4 dependent variables are the school year gap in mean value-added between novice and experienced teachers in math and ELA, respectively. See Appendix B for further detail on computation of value-added measures. All specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Table 1.9: Student effects, "staying" students

	Math		ELA		Days Attended		Effort	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post*Trend:	-0.001	-0.013**	0.009***	-0.001		0.212		-0.013
	(0.006)	(0.006)	(0.003)	(0.003)		(0.183)		(0.012)
Post		0.014		0.014**	3.653***	3.049***	0.007	-0.018
		(0.012)		(0.007)	(0.494)	(0.652)	(0.021)	(0.021)
Trend		0.010***		0.006***		0.125		0.015***
		(0.003)		(0.002)		(0.107)		(0.004)
Gr FEs	X	X	X	X	X	X	X	X
Yr FEs	X	X	X	X	X	X	X	X
Stu FEs	X	X	X	X	X	X	X	X
N stu-yrs	2,562,332	2,562,332	4,161,767	4,161,767	4,729,758	4,729,758	1,650,087	1,650,087
N stu	654,687	654,687	883,676	883,676	1,019,337	1,019,337	480,544	480,544
N trt stu	144,220	144,220	164,644	164,644	171,870	171,870	109,717	109,717
N trt coh	22,753	22,753	28,795	28,795	34,530	34,530	19,221	19,221
R2	0.82	0.82	0.84	0.84	0.52	0.52	0.63	0.63

Notes: Table reports estimates of parametric event study models corresponding to equation (1.4), for students that had 10% or more of their school-grade cohort exit to a newly constructed school. Event time is centered relative to the year of the peer outflow. Columns 1 and 2 include only the coefficient for the change in growth (β_2); β_1 and β_3 are constrained to be zero. Columns 5 and 7 include coefficients only the coefficient for the immediate effect (β_1); (β_2) and β_3 are constrained to be zero. Columns 2, 4, 6, and 8 include all coefficients, corresponding exactly to the specification in equation (1.4). Dependent variable is the standardized math test score (grades 2-7) in columns 1-2, the standardized ELA test score (grades 2-11) in columns 3-4, annual days attended in columns 5-6, and the standardized average teacher-reported effort score in columns 7-8. All specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Table 1.10: House price effects

(a) New school zones

	Neighborhood Fixed Effects					Repeat Sales	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
New School	-0.011	0.060***	0.059***	0.044***	0.055***	0.045***	0.059***
	(0.014)	(0.018)	(0.016)	(0.011)	(0.015)	(0.013)	(0.016)
Yr FEs	X				X		X
Yr-HSZ FEs		X	X	X		X	
Month FEs	X	X	X	X	X	X	X
Sch Zone FEs	X	X	X	X	X		
Prop Controls	X	X	X	X	X		
Prop FEs						X	X
New Sch Zones	X	X	X	X	X	X	X
w/in 1km			X				
All LAUSD	X	X					
Number of sales	505,781	505,781	255,481	161,775	161,782	87,523	87,551
R2	.81	.82	.79	.78	.75	.91	.9

(b) "Stayers" school zones

	Neighborhood Fixed Effects				Repeat Sales	
	(1)	(2)	(3)	(4)	(5)	(6)
Post: School	0.042***	-0.008	0.023	-0.009	-0.010	-0.014
	(0.012)	(0.017)	(0.018)	(0.019)	(0.031)	(0.025)
Yr FEs	X			X		X
Yr-HSZ FEs		X	X		X	
Month FEs	X	X	X	X	X	X
Sch Zone FEs	X	X	X	X		
Prop Controls	X	X	X	X		
Prop FEs					X	X
All LAUSD	X	X				
Number of sales	343,997	343,997	180,504	180,504	107,458	107,458
R2	.82	.83	.82	.81	.93	.93

Notes: Tables report estimates from difference-in-differences regressions following equations (1.5) and (1.7). Dependent variable is the ln(sale price). In panel A the coefficient of interest is an indicator for being in a new school zone, whereas in panel B it is an indicator for being in an existing school zone that was affected by student outflows to a new school. In panel A, columns 1-5 include neighborhood fixed effects and property specific controls; columns 6-7 include property fixed effects. Columns 1, 5, and 7 report estimates using year effects; remaining columns include year-by-high school zone fixed effects. Columns 1-2 include all properties in LAUSD. Column 3 restricts the sample to only properties within a new school zone or within a 1km of a new school zone. Columns 4-7 include only properties within a new school zone by 2012: "never-treated" properties are excluded. In panel B, properties in new school zones are excluded from estimation; columns 1-4 report estimates corresponding to equation (1.5), with neighborhood fixed effects and property specific controls. Columns 5-6 show estimates with property fixed effects. Columns 1, 4, and 6 of panel B include year effects, while remaining columns include year-by-high school zone effects. Columns 1-2 include all properties in LAUSD, while columns 3-6 restrict the sample to only those properties in school zones affected by student outflows. All specifications include month effects. Standard errors are clustered by neighborhood.

Table 1.11: House price effects, by distance to school assignment zone boundary

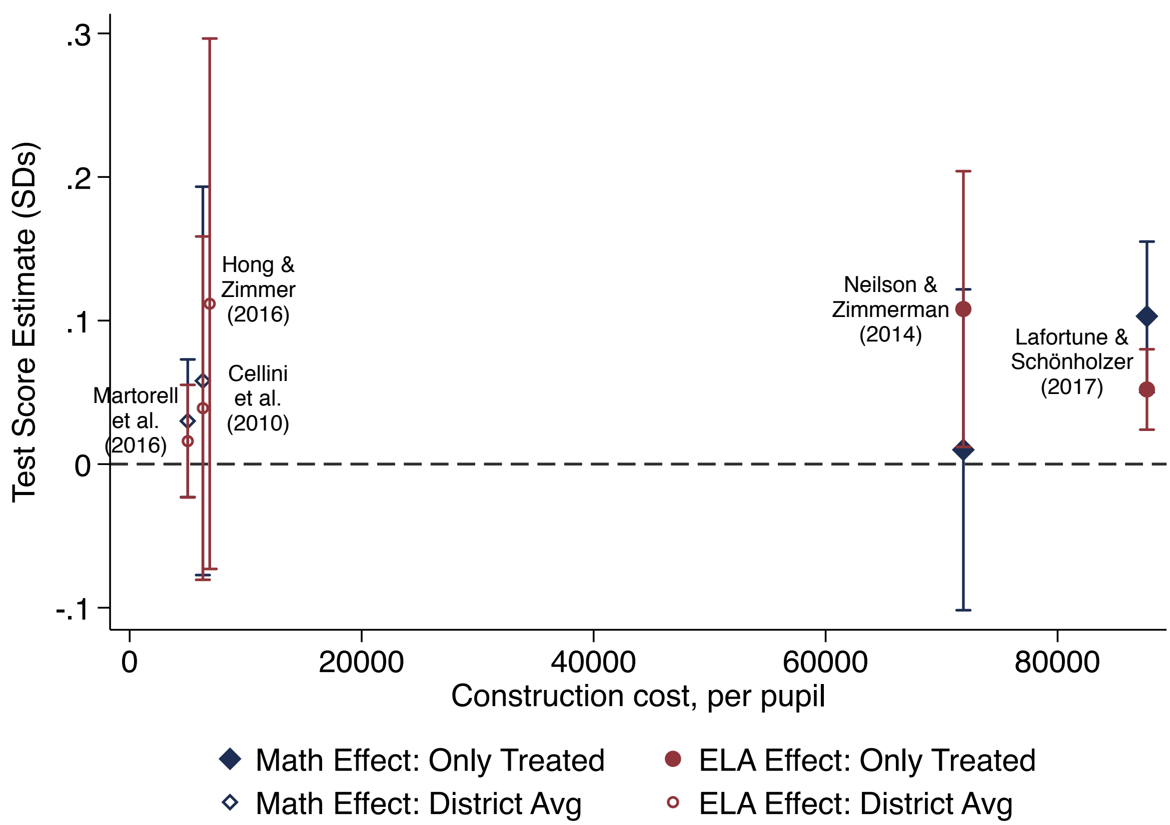
	(1)	(2)	(3)	(4)
New School: inside zone	0.060***	0.061***	0.054***	0.051***
	(0.018)	(0.018)	(0.017)	(0.018)
Distance to boundary		0.019		0.017
		(0.014)		(0.013)
Inside zone * dist to boundary		-0.005		-0.004
		(0.010)		(0.010)
New School: outside w/in 2km			-0.013	-0.046***
			(0.010)	(0.017)
Outside w/in 2km * dist to boundary				0.035**
				(0.014)
Yr-HSZ FEs	X	X	X	X
Month FEs	X	X	X	X
Sch Zone FEs	X	X	X	X
Prop Controls	X	X	X	X
New Sch Zones	X	X	X	X
All LAUSD	X	X	X	X
Number of sales	505,781	505,781	505,781	505,781
R2	.82	.82	.82	.82

Notes: Table reports estimated coefficients from difference-in-differences regressions based off of equation (1.5). Dependent variable is the ln(sale price). Column 1 repeats baseline estimates reported in column 2 of Table 1.10. Column 2 adds coefficients for property-level distance to the school assignment boundary and the interaction between distance to the boundary and the new school zone treatment variable. Column 3 includes an additional treatment variable for properties outside but within 2km of the new school attendance zone, where the completion date assigned to these properties corresponds to that of the nearest new school attendance zone. Column 4 combines columns 2 and 3, and adds an interaction with distance to the boundary for properties outside but within 2km of the new school zone. All properties in LAUSD in the data sample are included in estimation. All specifications include property-specific controls, year-by-high school zone fixed effects, neighborhood fixed effects, and month fixed effects. Standard errors are clustered by neighborhood.

1.12 Appendix Tables and Figures

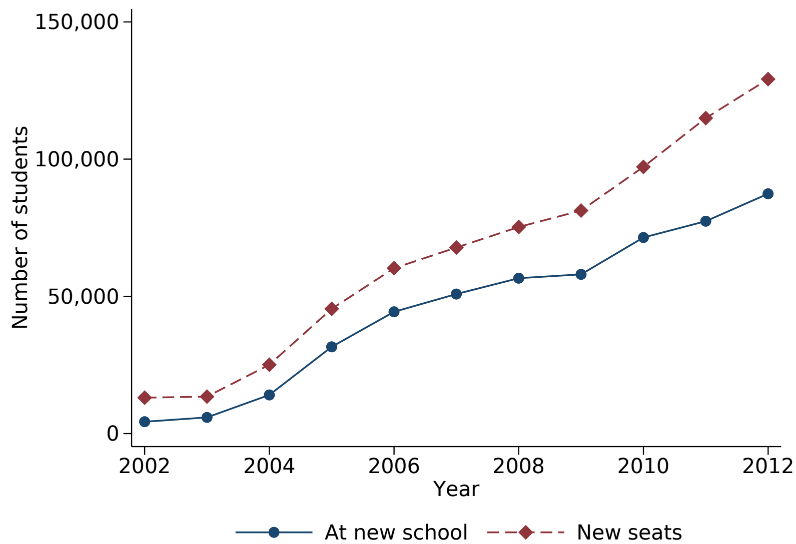
Appendix Figures:

Figure 1.11: Student effects comparison from capital expenditure literature



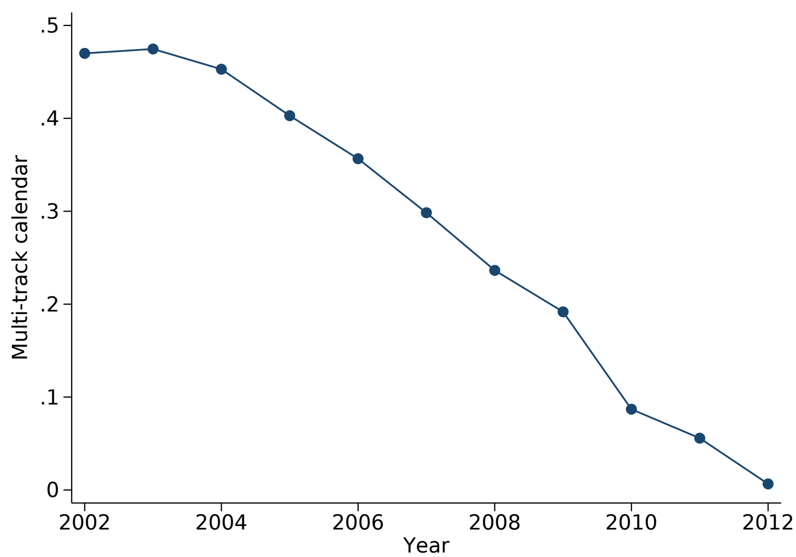
Notes: Figure plots estimated coefficients from related papers in economics evaluating the effects of school capital expenditures (y axis) against per-pupil expenditures in each study (x axis). Blue diamond shaped markers denote math test score estimates whereas red circular markers denote English / Language Arts test score estimates (both in standard deviation units). Solid markers denote estimates on directly treated students from Neilson and Zimmerman (2015) and Lafortune and Schönholzer (2017), 4 years after school construction or student occupancy, respectively. For these studies, construction cost is calculated per treated pupil. Hollow markers denote estimates from studies examining district average test scores after passage of a capital construction bond, where construction cost per pupil is the average over all students in the district. For these studies, estimates 6 years after bond passage are reported.

Figure 1.12: Students at newly constructed schools



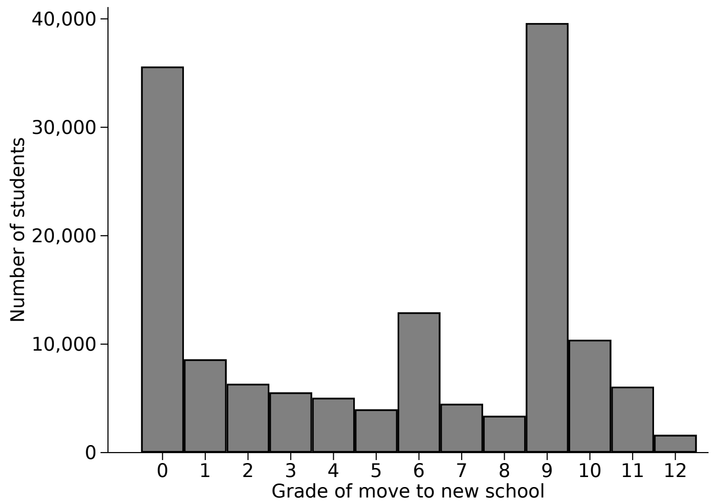
Notes: Figure shows time series of total new seats (from new construction project database) and the number of students attending newly constructed school facilities (from the student microdata).

Figure 1.13: School age and multi-track calendars in LAUSD



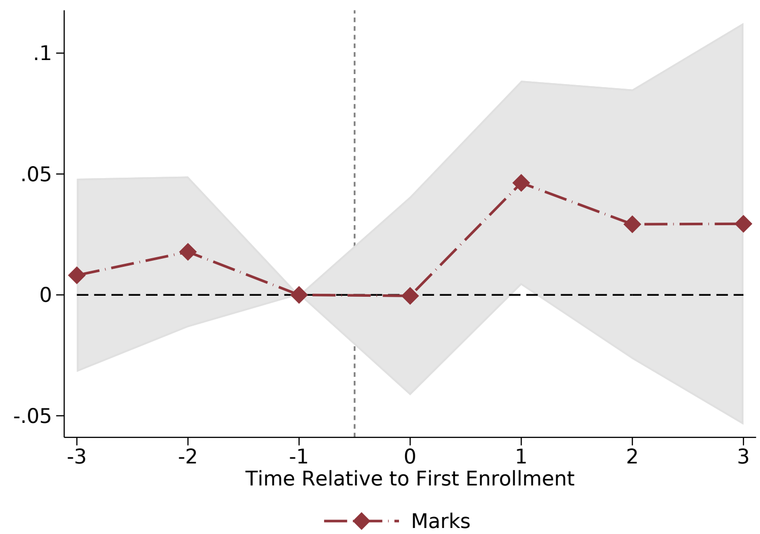
Notes: Figure reports proportion of students attending a school on a multi-track calendar, by year.

Figure 1.14: Grade of switch to new school facilities



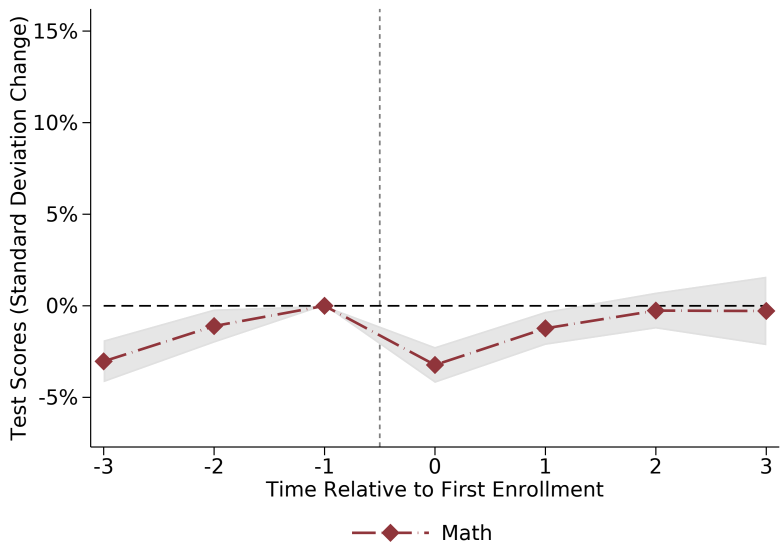
Notes: Figure shows grade of switch for students switching to new facilities. Y-axis reports number of students.

Figure 1.15: Event study estimates, teacher-reported marks

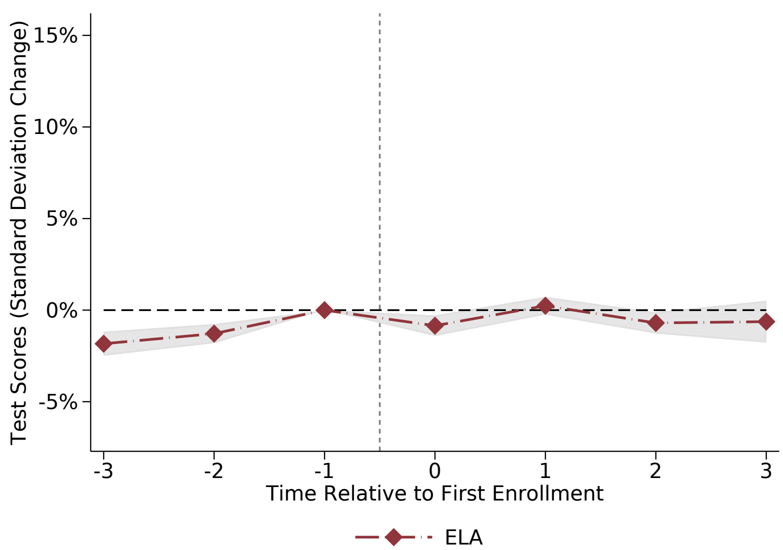


Notes: Figure shows estimated coefficients from event study regressions following equation (1.3). Dependent variable is the standardized teacher-reported marks (averaged over all subjects) for students in grades K-5. The shaded areas denote 95% confidence intervals for the estimated coefficients. Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Figure 1.16: Student switching, non-new facility related



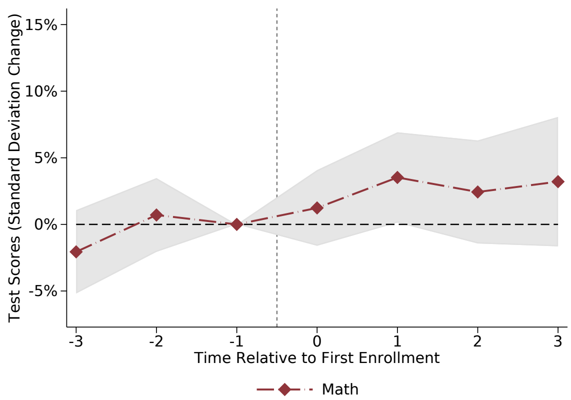
(a) Math



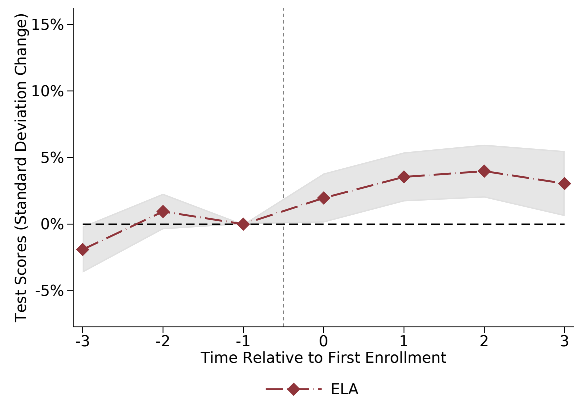
(b) ELA

Notes: Figures show estimated coefficients from event study regressions following equation (1.3), for students who switch schools for reasons unrelated to new school facilities. Dependent variables are standardized math test scores for students in grades 2-7 (panel A) and standardized english-language arts test scores for students in grades 2-11 (panel B). Test scores are standardized relative to the statewide mean and standard deviation for each year-grade-subject exam. The shaded areas denote 95% confidence intervals for the estimated coefficients. Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

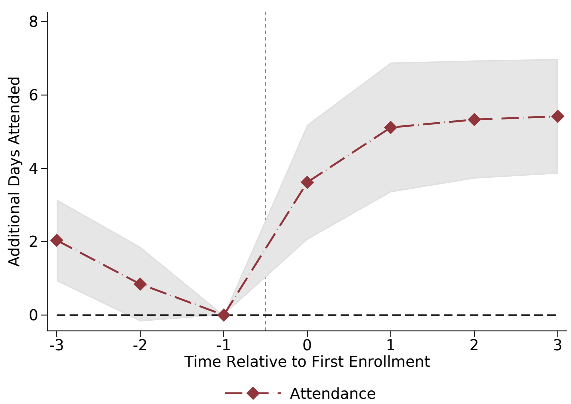
Figure 1.17: Student effects: Stayers, 20% cohort exit threshold



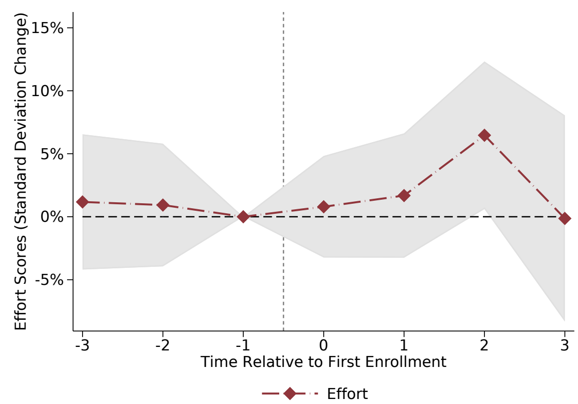
(a) Test scores: Math



(b) Test scores: ELA



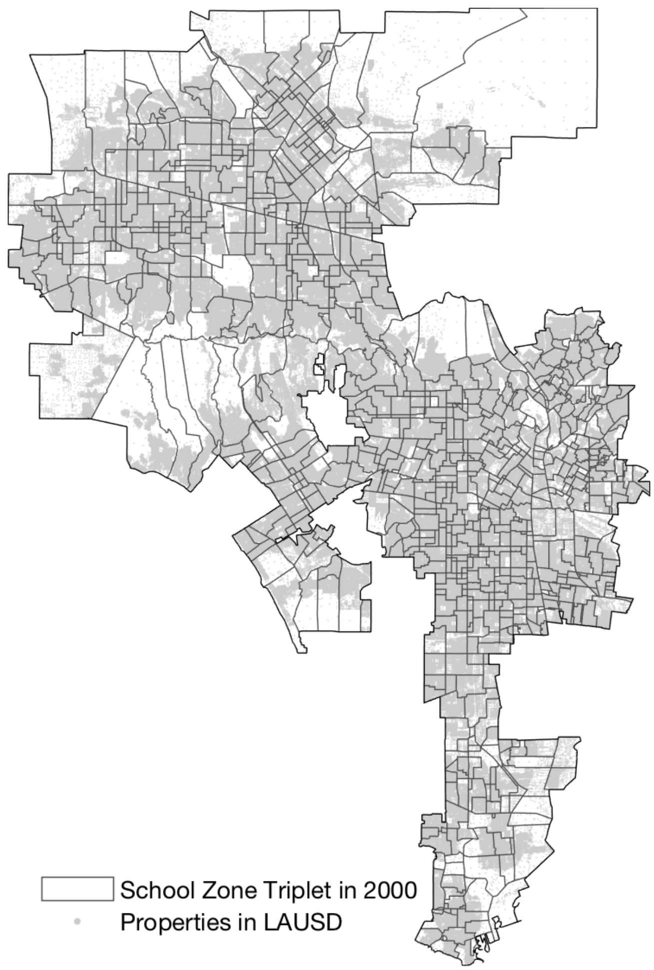
(c) Attended days



(d) Teacher-reported student effort

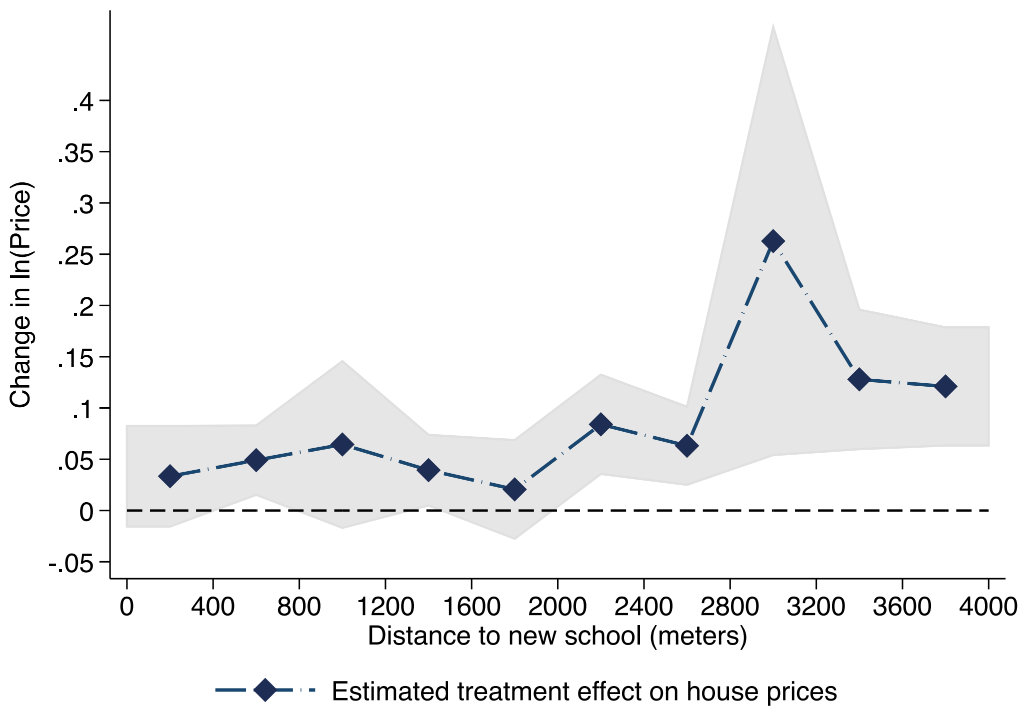
Notes: Figures show estimated coefficients from event study regressions following equation (1.3), for students that had 20% or more of their school-grade cohort exit to a newly constructed school. Figures are analogous to Figure 1.7, with the threshold for "stayers" raised from 10% to 20% of a student's cohort. Event time is centered relative to the year of the peer outflow. Dependent variables are standardized math test scores for students in grades 2-7 (panel A), standardized english-language arts test scores for students in grades 2-11 (panel B), annual days attended (panel C), and standardized teacher-reported effort scores for students in grades K-5 (panel D). The shaded areas denote 95% confidence intervals for the estimated coefficients. Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Figure 1.18: Neighborhood boundaries in LAUSD, based on 2000 school zones



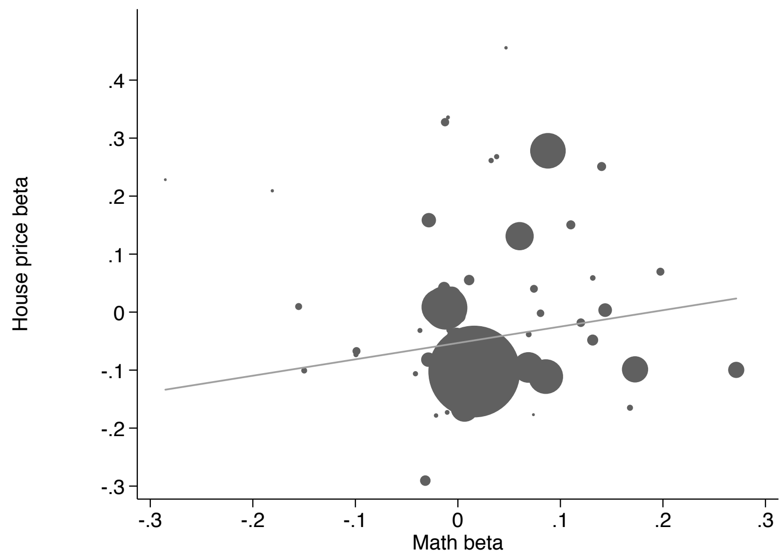
Notes: Figure shows school assignment zone triplets in LAUSD using 2000 assignment boundaries, which are used to define neighborhoods in the estimation of real estate effects. Solid lines denote neighborhood boundaries. Each gray dot represents one property from the LA County Assessor data.

Figure 1.19: Spillovers: Effects by distance to new school

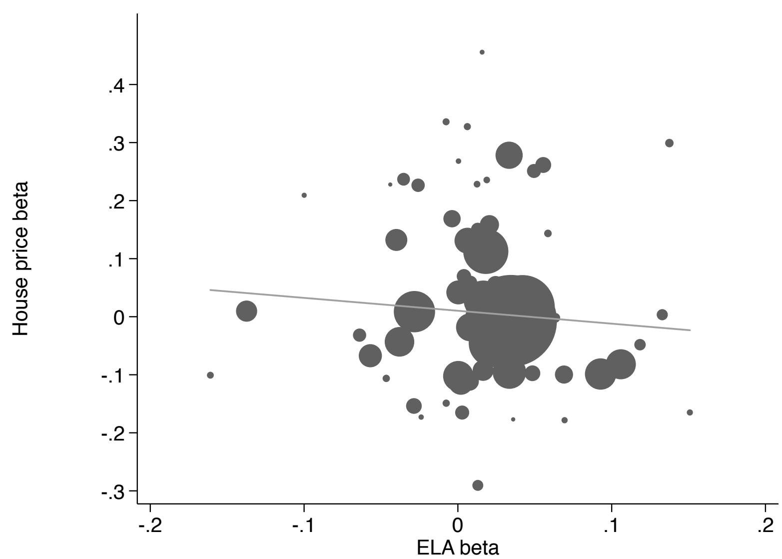


Notes: Figure shows estimated coefficients from a difference-in-differences regression based on equation (1.5), where the treatment indicator is interacted with indicators for 400 meter bins of distance to the new school in 2012. Each point reports the estimated coefficient for the treatment indicator interacted with the corresponding distance bin. Points are located at the midpoint of each distance bin (i.e. the estimate at 200m corresponds to the 0-400m distance bin). All properties in LAUSD in the data sample are included in estimation, corresponding to baseline estimates presented in column 2 of Table 1.10. Specifications include property-specific controls, year-by-high school zone fixed effects, neighborhood fixed effects, and month fixed effects. Standard errors are clustered by neighborhood.

Figure 1.20: Correlation between house price and test score effects



(a) Math



(b) ELA

Notes: Figures show scatterplots of estimated school-level test score gains in math (panel A) and ELA (panel B) against estimated house price effects in the corresponding school attendance zone. The solid line in each figure displays the bivariate regression line. Points and regression lines are weighted by the product of the inverse sampling variances of the estimated test score gain and the estimated house price change for a given school. The size of each point is proportional to the weight. In panel A the point estimate on the regression line is 0.28 (SE 0.29) and in panel B the point estimate is -0.22 (SE 0.30).

Appendix Tables

Table 1.12: School-level changes

(a) Switching students

	Calendar		School		Peers	
	(1) Multi-track	(2) Max days	(3) Age	(4) Stu/tch	(5) Peers: Bl/Hisp	(6) Peers: pred
New School	-0.267***	1.762***	-71.086***	0.312***	0.027***	-0.024***
	(0.029)	(0.262)	(1.132)	(0.084)	(0.004)	(0.006)
Grade FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Stu FEs	X	X	X	X	X	X
N stu-years	6,601,535	5,898,902	6,519,509	3,155,009	6,594,053	4,601,340
N stu	1,224,566	1,186,057	1,217,043	779,669	1,222,196	939,620
N treated stu	122,172	120,164	122,112	96,526	122,043	97,856
N treated sch	143	140	143	79	143	126
R2	0.68	0.51	0.79	0.75	0.88	0.85

(b) Staying students

	Calendar		School		Peers	
	(1) Multiple	(2) Max days	(3) Age	(4) Stu/tch	(5) Peers: Bl/Hisp	(6) Peers: pred
Post: Stayers	-0.249***	2.305***	1.572	-0.278**	-0.016***	0.022***
	(0.027)	(0.304)	(1.074)	(0.123)	(0.003)	(0.004)
Grade FEs	X	X	X	X	X	X
Year FEs	X	X	X	X	X	X
Stu FEs	X	X	X	X	X	X
N stu-years	5,837,507	5,214,065	5,759,302	2,737,469	5,830,319	4,053,776
N stu	1,119,399	1,081,590	1,111,323	686,610	1,117,171	853,479
N treated stu	178,022	176,086	177,607	132,170	177,847	161,177
N treated sch	801	791	752	500	802	787
R2	0.70	0.53	0.68	0.75	0.88	0.85

Notes: Table reports estimates of difference-in-differences models corresponding one-parameter versions of equation (1.4), where only the coefficient for the immediate new school effect (β_1) is included; β_2 and β_3 are constrained to be zero. Dependent variables are multi-track status in a given school-year (column 1), total instructional days in a given school-year (column 2), school age (column 3), class size (i.e. pupils per teacher) for students in grades K-5 (column 4), school leave-out mean proportion black and/or hispanic (column 5), and school leave-out mean predicted test scores (column 6). Panel A reports estimates for students attending new school facilities. Panel B reports analogous estimates for staying students: here β_1 is an indicator for having experienced a 10% or greater school-grade cohort exit to a newly constructed school. Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Table 1.13: Teacher changes at existing schools

(a) Demographics

	(1) Age	(2) Experience	(3) MA+	(4) Pr(New)
Post: Stayers	1.117***	0.768***	0.007	-0.000
	(0.248)	(0.182)	(0.010)	(0.004)
Grade FEs	X	X	X	X
Year FEs	X	X	X	X
Stu FEs	X	X	X	X
N student-years	3,935,106	3,927,063	3,931,757	5,902,165
N students	926,501	925,300	926,203	1,140,815
N treated students	156,306	156,183	156,282	176,213
N treated schools	797	797	797	802
R2	0.32	0.35	0.28	0.29

(b) Value-added

	VAM: Average (pre-switch)		VAM: Experienced/Novice gap	
	(1) Math	(2) ELA	(3) Math	(4) ELA
Post: Stayers	-0.009	0.002	-0.007	0.016
	(0.007)	(0.003)	(0.022)	(0.014)
Grade FEs	X	X	X	X
Year FEs	X	X	X	X
Stu FEs	X	X	X	X
N student-years	2,175,273	3,722,137	1,267,199	2,347,897
N students	640,572	887,238	432,813	672,731
N treated students	132,846	156,769	104,810	134,051
N treated schools	609	785	585	731
R2	0.33	0.35	0.38	0.33

Notes: Table reports estimates of difference-in-differences models corresponding one-parameter versions of equation (1.4), for students that had 10% or more of their school-grade cohort exit to a newly constructed school. Only the coefficient for having experienced a 10% or greater school-grade cohort exit is included (β_1); β_2 and β_3 are constrained to be zero. In panel A, dependent variables are teacher age (column 1), teacher years experience (column 2), an indicator for having a masters degree or higher (column 3), and an indicator for having a new teacher in either math or ELA (column 4). Panel B reports estimates where dependent variables are school-year averages of teacher value-added: in columns 1 and 2 dependent variables are average value-added scores based on prior-year observations at existing school facilities in math and ELA, respectively. In columns 3 and 4 dependent variables are the school year gap in mean value-added between novice and experienced teachers in math and ELA, respectively. See Appendix B for further detail on computation of teacher value-added. All specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Table 1.14: Student effects, heterogeneity

	Math	ELA	Attendance	Effort
Pooled	0.029***	0.019***	3.973***	0.061***
	(0.007)	(0.004)	(0.551)	(0.017)
By Sex:				
Female	0.036***	0.025***	3.823***	0.073***
	(0.007)	(0.004)	(0.547)	(0.018)
Male	0.022***	0.014***	4.131***	0.053***
	(0.007)	(0.004)	(0.567)	(0.019)
p-value	0.00	0.00	0.04	0.19
By parental education:				
No college	0.029***	0.021***	4.284***	0.050***
	(0.007)	(0.004)	(0.598)	(0.017)
Any college	0.026***	0.014***	3.278***	0.107***
	(0.010)	(0.004)	(0.494)	(0.023)
p-value	0.69	0.03	0.00	0.00
By school level:				
Elementary	0.028***	0.017***	1.608***	0.061***
	(0.007)	(0.004)	(0.333)	(0.017)
Middle	0.038	-0.002	3.368***	
	(0.026)	(0.007)	(0.526)	
High		0.030***	5.464***	
		(0.007)	(1.055)	
p-value	0.72	0.00	0.00	
By grade of switch:				
Reg (KG,G6,G9)	0.019**	0.018***	5.305***	0.040
	(0.009)	(0.004)	(0.643)	(0.025)
Irregular	0.038***	0.022***	1.976***	0.071***
	(0.009)	(0.005)	(0.532)	(0.020)
p-value	0.08	0.49	0.00	0.31

Notes: Table reports estimates of parametric event study models corresponding one-parameter versions of equation (1.4). Columns 1 and 2 include only the coefficient for the change in growth (β_2); β_1 and β_3 are constrained to be zero. Columns 3 and 4 include only the coefficient for the immediate new school effect (β_1); β_2 and β_3 are constrained to be zero. Dependent variables are standardized english-language arts test scores (column 1), standardized math test scores (column 2), annual days attended (column 3), and standardized average teacher-reported effort scores (column 4). Panel A repeats baseline one-parameter estimates from columns 1 and 4 of Tables 1.4 and

1.5. The remaining panels report estimates of coefficients interacted with student gender (panel A), parental education (panel B), school level (panel C), and whether a student switched in a typical (KG, G6, G9) or atypical grade (panel D). P-values for the test of equality of the coefficient(s) are reported in the third row of each panel. Specifications include fixed effects for student, year, and grade. Standard errors are two-way clustered by school and student.

Table 1.15: House price effects, by school level

	(1)	(2)	(3)	(4)	(5)
New Elementary	0.051***			0.026*	
	(0.015)			(0.014)	
New Middle		0.031		0.003	
		(0.023)		(0.016)	
New High			0.071**	0.065**	
			(0.030)	(0.029)	
Only New Elementary					0.065***
					(0.021)
Only New Middle					0.008
					(0.018)
Only New High					0.072**
					(0.034)
p, Elem effects =0	.00063			.064	.0027
p, Mid effects =0		.18		.87	.66
p, HS effects =0			.019	.027	.034
p, All effects =0				.04	.0036
p, All effects equal				.17	.024
Yr-HSZ FEs	X	X	X	X	X
Month FEs	X	X	X	X	X
Sch Zone FEs	X	X	X	X	X
Prop Controls	X	X	X	X	X
All LAUSD	X	X	X	X	X
Number of sales	381,407	374,915	480,967	505,781	471,528
R2	.83	.83	.82	.82	.83

Notes: Table reports estimated coefficients from difference-in-differences regressions by school level, based off of equation (1.5). Columns 1, 2, and 3 report estimates of the effects of new elementary, new middle, and new high schools, respectively. Properties in new school zones for schools at the other two levels are excluded from the control group in estimation in columns 1-3 (i.e. column 1 excludes properties that received new middle and/or new high school zones but not elementary schools from the control group). Column 4 includes coefficients for all three school levels. Column 5 restricts estimation to include only those properties in the attendance area of a single new school level. P-values for the tests that the effect at each level equals zero are included, as are p-values for the omnibus hypothesis tests that effects for all levels are equal to zero and that effects for all levels are equal. All specifications include property-specific controls, year-by-high school zone fixed effects, neighborhood fixed effects, and month fixed effects. Standard errors are clustered by neighborhood.

Table 1.16: House price effects, using post-2012 data or post-construction neighborhood definitions

(a) Including post-2012 data

	Neighborhood Fixed Effects					Repeat Sales	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
New School	-0.021*	0.053***	0.049***	0.034***	0.049***	0.043***	0.046***
	(0.013)	(0.015)	(0.013)	(0.011)	(0.013)	(0.012)	(0.014)
Yr FEs	X				X		X
Yr-HSZ FEs		X	X	X		X	
Month FEs	X	X	X	X	X	X	X
Sch Zone FEs	X	X	X	X	X		
Prop Controls	X	X	X	X	X		
Prop FEs						X	X
New Sch Zones	X	X	X	X	X	X	X
w/in 1km			X				
All LAUSD	X	X					
Number of sales	593,414	593,414	298,507	188,222	188,229	114,519	114,542
R2	.81	.82	.79	.77	.74	.91	.9

(b) Neighborhoods based on 2012 boundaries

	Neighborhood Fixed Effects					Repeat Sales	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
New School	-0.010	0.068***	0.067***	0.046***	0.055***	0.045***	0.059***
	(0.014)	(0.019)	(0.017)	(0.013)	(0.016)	(0.014)	(0.017)
Yr FEs	X				X		X
Yr-HSZ FEs		X	X	X		X	
Month FEs	X	X	X	X	X	X	X
Sch Zone FEs	X	X	X	X	X		
Prop Controls	X	X	X	X	X		
Prop FEs						X	X
New Sch Zones	X	X	X	X	X	X	X
w/in 1km			X				
All LAUSD	X	X					
Number of sales	505,795	505,779	255,481	161,773	161,779	87,523	87,551
R2	.81	.82	.8	.78	.75	.91	.9

Notes: Table reports estimates from difference-in-differences regressions following equations (1.5) and (1.7). Panel A includes additional data from 2013-2015, while panel B uses neighborhood effects based on 2012 school assignment zones in lieu of 2000 school zones. Dependent variable is the ln(sale price). Columns 1-5 report estimates from equation (1.5), including neighborhood effects and property specific controls. Columns 6 and 7 report estimates from equation (1.7), including property fixed effects. Columns 1, 5, and 7 report estimates using year effects; the remaining columns include year-by-high school zone effects. In columns 1 and 2, all properties in LAUSD in the data sample are included. Column 3 restricts the sample to include only properties within a new school zone or within a 1km of a new school zone (by 2012). Columns 4-7 include only properties within a new school zone by 2012; "never-treated" properties are excluded. All specifications include month effects. Standard errors are clustered by neighborhood.

Table 1.17: House price effects, robustness to sample restrictions

		Relaxing sample restrictions for:			
	(1) Baseline	(2) Price outliers	(3) Renovated/torn-down	(4) Large/multi-unit	(5) Non-residential
New School	0.060***	0.088***	0.056***	0.058***	0.048***
	(0.018)	(0.028)	(0.020)	(0.018)	(0.014)
Yr FEs					
Yr-HSZ FEs	X	X	X	X	X
Month FEs	X	X	X	X	X
Sch Zone FEs	X	X	X	X	X
Prop Controls	X	X	X	X	X
Baseline sample	X	X	X	X	X
Price outliers		X			
Renovated			X		
Large/multi-unit				X	X
Non-residential					X
Number of sales	505,780	512,577	525,469	513,039	625,632
R2	.82	.75	.75	.8	.72

Notes: Table reports estimated coefficients from difference-in-differences regressions corresponding to estimates of equation (1.5). Dependent variable is the ln(sale price). Column 1 repeats baseline estimates presented in Table 1.10 panel A column 2. Column 2 makes no restriction on sale price, including the top and bottom 1% of sales based on price. Column 3 relaxes the restriction on renovated and/or torn-down properties, including these properties with an additional indicator variable for having been renovated and/or torn-down in the controls. Column 4 includes large properties, with greater than one acre of space. Column 5 includes non-residential properties. All specifications include neighborhood fixed effects, property specific controls, and month fixed effects. Standard errors are clustered by neighborhood.

1.13 Appendix: Computing Teacher Value-Added

To estimate teacher value-added scores, we use a subsample of students for which the following criteria are met: (1) Student-year observations have non-missing test scores and are currently in grades 3-7 in math, and 3-11 in ELA; value-added scores are not computed for grade 2 teachers so as to have at least one prior score for a student; (2) Student-year observations have non-missing teacher assignment;[55] (3) Student-year observations are in classrooms with at least 7 students. Consider the following data-generating process for test scores, closely following Kane and Staiger (2008) and Chetty, John N. Friedman, et al. (2014):

$$y_{i,t} = \alpha_{t,g(i,t)} + X'_{it}\beta + \nu_{it} \tag{1.8}$$

$$\nu_{it} = \mu_{j(i,t),t} + \epsilon_{it} \tag{1.9}$$

where $y_{i,t}$ is student i's test score in a given subject in year t, $g(i,t)$ denotes a student's grade in a given year, $j(i,t)$ denotes a student's teacher in a given year, and X'_{it} is a vector of controls. Here, $\mu_{j(i,t),t}$ is a teacher's effect on student test scores in year t and $\epsilon_{j(i,t),t}$ captures unobserved error in test scores unrelated to teacher quality.

To compute value-added for a given teacher-year, we estimate equation (1.8), and then compute the average residual within each teacher-year cell: $\text{VA}_{jt} \equiv \overline{\nu}_{jt}$. Unlike many prior studies, we do not use an Empirical Bayes or similar procedure to "shrink" these noisy estimates of value-added, as we will only use these measures as dependent variables and are therefore less concerned about measurement error (and potentially more concerned about biased estimates).[56]

In estimation, X'_{it} includes third-degree polynomials in lagged student test scores (for both subjects), demographics (race, gender, parental education, free/reduced-price lunch status, limited English status), class size (only available for elementary students), and school-level variables (school leave-out means of the share black/hispanic, share with any parental postsecondary education, share who speak English at home, and the share eligible for free or reduced-price lunch). We do not include school fixed effects in estimation, meaning estimated teacher effects are relative to all other teachers within LAUSD.

1.14 Appendix: Treatment Effect Heterogeneity

Student Effects: Heterogeneity

Heterogeneity in estimated student effects is presented in Table 1.14. Row 1 reports pooled estimates using the entire sample, which correspond to baseline estimates presented in

[55] Nearly every student in K-5 has a non-missing assignment; teacher IDs in later grades were assigned to a student-subject pair based on the teacher associated with a student's math and/or ELA class

[56] See Jacob and Rothstein (2016) for a more detailed discussion of potential problems using estimated posterior means of student test scores as dependent variables in regression models.

column 1 of Table 1.4. In the remaining rows, the one-parameter treatment effect coefficients are interacted with student demographic and other characteristics.[57] Estimated cognitive effects are nearly twice as large for girls than boys, and the differences are statistically significant ($p < 0.01$) for both math and ELA. Effects on student effort are also larger for girls, although the magnitude of the difference is smaller and not significant. The pattern is the opposite for attendance, as effects on the number of days attended are larger for boys than girls, although the magnitude of the difference is small. These differences suggest that substandard classroom facilities may inhibit girls' learning more than boys, although the mechanisms underlying this difference are unclear.

When results are split by level of parental education, a mixed picture emerges. Estimated effects on math scores, ELA scores, and attendance are larger for students with parents who did not attend any level of postsecondary education, although the difference in math scores is small and insignificant. For student effort, estimated effects are over twice as large for students with parents who have any level of postsecondary education than for those whose parents have a high school education or less. Overall, the results provide little evidence that improvements in school facilities systematically benefit students from lower socio-economic backgrounds.[58] Recall however, as shown in Table 1.2, that there is little variation in socio-economic status in LAUSD: nearly 90% of treated students are eligible for free or reduced-price lunch and less than one-fifth have parents with any level of post-secondary education.

Table 1.14 also shows estimates split by school level. Cognitive effects are insignificant for students who attend new middle schools, although for math, the difference between elementary and middle school effects is insignificant. For ELA, effects are large and significant in both elementary and high school, and are essentially zero for students who switch to a new middle school. For attendance, a clear pattern emerges: effects increase monotonically with school level, and are the largest for students in new high schools. Insofar as student motivation is impacted by new facilities and drives changes in student attendance, we would expect effects to grow with grade level as older students have greater autonomy over attendance decisions than younger students, whose daily attendance is more directly dictated by parental influence.

Finally, estimated effects are also split by whether a student switches schools during a "regular" grade transition (KG, G6, G9) or switches to a new school in another grade. "Irregular" grade transitions in off-grades occurred immediately following school construction, when students were transferred between schools to fill enrollment at the new school. Overall, effects are similar for both types of switching students, with only a large and significant difference in estimated attendance gains. Estimated effects on cognitive outcomes and student effort are somewhat larger for initial switchers who switch during an irregular grade transition, although only the difference in math scores is statistically significant at the 10% level. For student attendance, effects are significantly larger (5.2 days vs 1.9 days) for regular grade

[57]Note that this constrains grade and year effects to be equivalent for each group, as opposed to running separate regressions or also interacting fixed effects with demographic indicators.

[58]Analogous breakdowns by race and free lunch status (not reported) show only small and insignificant differences.

switchers. Students switching at a typical grade transition are mostly switching in grades 6 and 9, which explains most of the difference in days attended, as attendance gains are larger for middle and high school students than elementary school switchers.[59]

Real Estate Effects: By School Level

Estimates reported in Table 1.10 and Figure 1.8 include properties that received multiple new schools. The average treated property in the sample was in the school attendance area of 1.1 new school constructions, implying the the effect of receiving a single school (elementary, middle or high) would be 9% lower than the baseline estimates, roughly a 5.5% increase in house prices per new school construction using baseline estimates from column 2. In Table 1.15 we report house price effects separately by school level. Results indicate that effects are largest for new elementary and high schools, although we cannot statistically reject differences in estimated coefficients in all specifications. Qualitatively, results are consistent with student effect heterogeneity reported in Table 1.14, which provided evidence that test score effects were larger and more significant for newly constructed elementary and high school than for new middle schools. As middle schools represent the shortest duration of student attendance (3 years, vs 4 for high school and up to 6 for elementary), it is unsurprising that the effects may be smaller.

[59]See Figure 1.14 for the distribution of student switching grades to new schools.

Chapter 2

School Finance Reform and the Distribution of Student Achievement

2.1 Introduction

Economists have long been skeptical of resource-based education policies, based in part on observational studies showing small or zero effects of additional funding (see, e.g., Coleman et al. (1966), Hanushek (1986), Hanushek (2006)).[1] Hanushek, for example, writes: "Simply providing more funding or a different distribution of funding is unlikely to improve student achievement (even though it may affect the tax burdens of school financing across the citizens of a state)" (Hanushek (1997), p. 153). Accordingly, recent policy discussions have focused on ways to improve the productivity of existing inputs rather than on changes in school resource levels.

Nevertheless, states have continued to implement aggressive resource-based policies, aimed in part at reducing achievement gaps. Figure 2.1 shows the evolution of average revenues per pupil, in 2013 dollars, in the lowest- and highest-income school districts in each state (defined as the bottom and top fifths of the states district-level mean household income distribution).[2] Between 1990 and 2012, real per-pupil revenues rose by roughly 30 percent in the highest-income districts, and by over 50 percent in the lowest-income districts. Thus, while low-income districts collected about 20 percent less than high-income districts in 1990, they have been in rough parity since around 2001.

Much of this change came via reforms to state education funding formulas, many implemented in response to court orders. Figure 2.2 shows revenues of low-income districts relative to high-income districts, each defined as in Figure 2.1, separately for the 26 states

[1] There are also observational (Card and Krueger, 1992a) and experimental (Krueger (1999); Dynarski et al. (2013)) studies pointing to positive school resource effects. There is no consensus about how to reconcile these (see, e.g., Burtless (1996); Hanushek (2003); Krueger (2003)).

[2] Hawaii and the District of Columbia are excluded. Districts are weighted by log enrollment in computing state quintile means, which are then averaged without weights in Figure 2.1. We discuss data sources and definitions in Section 2.4.

that have implementedor at least been ordered to implement by courtsschool finance reforms since 1990 and for 23 states that have not. Growth in low-income districts relative revenues has been more than twice as rapid in the former states than in the latter.

There are two primary types of school finance reforms (SFRs). In the 1970s and 1980s, SFRs were primarily "equity" reforms, aimed at reducing resource disparities across districts. Since 1990, the pace of reforms has quickened, and most have been adequacy reforms, aimed at achieving sufficient funding in low income districts regardless of implications for equity.[3]

SFRs are arguably the most substantial national policy effort aimed at promoting equality of educational opportunity since the turn away from school desegregation in the 1980s. But there is little evidence about their effects on student achievement. What evidence there is derives from non-representative data on students who took the SAT college entrance exam (Card and Payne, 2002); from long-run outcomes measured in the relatively small Panel Study of Income Dynamics sample (Jackson et al., 2016); or from case studies of individual reforms (Guryan (2001); Clark (2003); Hyman (Forthcoming)).[4] These studies primarily examine pre-1990, equity-based SFRs, and generally find positive effects on student outcomes. But funding levels were much higher by 1990 than earlier, and the most severe inequities in school resources had been addressed. Thus, there may have been less scope for more recent, adequacy-based SFRs to benefit students.

The impacts of SFRs on student achievement are closely related to the impact of additional resources. The literature regarding whether "money matters" in education (Hanushek (1986), Hanushek (2003); Hanushek (2006); Card and Krueger (1992a); Burtless (1996)) is contentious and does not offer clear guidance. State funding formulas are the main policy tool available to address inequities in academic outcomes, so funding shifts deriving from changes in these formulas are the most policy-relevant variation in school resources.

We provide the first evidence from nationally representative data regarding the impact of SFRs on student achievement. We exploit little-used data from the National Assessment of Educational Progress (NAEP), also known as "the Nations Report Card." State-representative samples of 100,000-200,000 students in the fourth and eighth grades have taken math and reading tests every two to four years since 1990. Importantly, the tests have been uniform across states and over time, facilitating comparisons.

We use the NAEP data to construct a state-by-year panel of relative achievement in low-income school districts, covering 1990 to 2011. Conveniently, the beginning of our NAEP panel coincides with the onset of the adequacy era of school finance, which dates to the 1990 Kentucky Education Reform Act (KERA).

To distinguish the causal impacts of SFRs from other potential determinants of spending and test score trends, we use an event study framework, taking advantage of plausibly random

[3]Studies of the implications of SFRs for school finance, mostly examining equity reforms, include S. E. Murray et al. (1998); Card and Payne (2002); Hanushek and Lindseth (2009); Berry and Wysong (2012); Ladd and Fiske (2015)); Sands (2015).

[4]Cascio, Gordon, et al. (2013) and Cascio and Reber (2013) examine the introduction of federal Title I funding to low-income schools via the 1965 Elementary and Secondary Education Act.

variation in the location and timing of post-1990 SFRs.[5] We find no sign of systematic changes in either funding or test scores in the period leading up to a reform, supporting our assumption that reform timing is exogenous. Following reforms, we document sharp increases in state revenues, with larger increases in low-income districts and smaller but still positive increases in high-income districts.[6] These changes occur quickly after reform events, persist for many years, and are not offset by reductions in local revenues. Absolute and relative funding in low-income districts rises by approximately $1,200 and $700 per pupil per year, respectively. We find that, on average, schools use the additional funds on instructional spending, to reduce class size, and for capital outlays.

We also find clear changes in achievement trends following events. These cumulate over subsequent years: Ten years after a reform, relative achievement of students in low-income districts has risen by roughly 0.1 standard deviation, approximately one-fifth of the baseline gap between high- and low-income districts. The implied impact is between 0.12 and 0.24 standard deviations per $1,000 per pupil in annual spending. This is at least twice the impact per dollar that is implied by the Tennessee Project STAR class size experiment.[7] Given existing estimates of the relationship between test scores and students subsequent earnings, our results imply that a $1 increase in funding to low-income school districts will raise students eventual earnings by more than $1 in present value.

Nevertheless, we find no discernible effect of reforms on statewide achievement gaps between high- and low-income students or between minority and white students. This is not inconsistent with our results on the impacts on scores in low-income districts, nor does it indicate that only the high-income students in those districts benefit. Rather, we show that low-income and minority students are not very highly concentrated in school districts with low mean incomes. As a result, SFRs lead to only small increases in the funding to which the average low-income or minority student is exposed. Thus, while our analysis suggests that finance reforms can be quite effective at reducing between-district inequities, other policy tools aimed at closing *within*-district achievement gaps will be needed to address overall equity concerns.

[5] A simple long-difference analysis of test score gaps between low-income and high-income districts, similar to the analysis of finance in Figure 2.2, shows that gaps have shrunken in states that implemented reforms relative to states that have not. See Figure 2.10 in the Appendix.

[6] Anecdotally, legislators facing court orders to increase funding to low-income districts often respond by increasing overall funding, as a way of disguising the resulting redistribution. Reforms are associated with sharp increases in total state education expenditures and tax collections.

[7] STAR raised costs by about 30% in K-3, and raised early grade test scores by 0.17 SDs (Krueger (1999); Krueger (2003); Krueger and Whitmore (2001)). Current spending per pupil in Tennessee is around $9,000, so comparable proportional class size reductions would cost around $2,700 per pupil per year. The implied effect is thus around 0.06 SDs per $1,000 per (early elementary) pupil per year.

2.2 School Finance Reforms[8]

Historically, American public schools were locally managed and financed primarily via local property taxes. As school districts vary widely in both their tax bases and their voters willingness to tax themselves to fund schools, this meant that school spending and quality varied substantially across districts.

In the 1960s, a group of legal scholars argued that local school finance violates federal and state constitutional provisions that guarantee equal access to public services (see, e.g., Krueger and Whitmore (1967); Horowitz (1965); Kirp (1968); and Coons et al. (1970)). Advocates brought and won suits in many states demanding more equitable school finance systems; in other states, legislatures acted without court decisions, often to stave off potential rulings.[9] The resulting finance regimes often involved substantial increases in state transfers to districts with low property tax bases. An extensive "fiscal federalism" literature examines the effects of these reforms on the distribution of school funding (see, e.g., S. E. Murray et al. (1998); Card and Payne (2002); Hanushek and Lindseth (2009); Corcoran and Evans (2015)).

We focus on a second wave of finance reforms, which began with a 1989 Kentucky Supreme Court ruling that the state constitution, which as in many other states dictates an efficient system of public schools, requires that "[e]ach child, *every child*, must be provided with an equal opportunity to have an adequate education" (*Rose v. Council for Better Education*[10]; emphasis in original). The Court emphasized that equal funding was not sufficient, and articulated a standard closer to equality of outcomes for students in low-income districts ("sufficient levels of academic or vocational skills to enable public school students to compete favorably with their counterparts in surrounding states, in academics or in the job market"). The Kentucky legislature responded with the Kentucky Education Reform Act of 1990 (KERA), which revamped the states educational finance, governance, and curriculum. Clark (2003) and Flanagan and S. Murray (2004) find KERA substantially increased spending in low-income districts.

Since 1990, courts in many other states have found adequacy requirements in their own constitutions. In many cases reforms have aimed at higher spending in low-income than in high-income districts, to compensate for the out-of-school disadvantages that low-income students face.[11]

We have attempted to identify all major SFRs between 1990 and 2011. We began with lists of court-ordered reforms compiled by Corcoran and Evans (2015) and Jackson et al.

[8] Our discussion here draws heavily on Koski and Hahnel (2015).

[9] The U.S. Supreme Court held in 1973 that education is not a fundamental right under the U.S. Constitution (*San Antonio Independent School District v. Rodriguez*, 411 US 1, 1973). Subsequent suits focused on state constitutions, which often mandate adequate and/or equitable systems of public education.

[10] 790 SW 2d 186. *Rose* was not the first adequacy ruling, but earlier rulings attracted less attention.

[11] A small industry has developed to calculate the spending level needed to satisfy an adequacy standard. See, e.g., Downes and Stiefel (2015) and Duncombe and Yinger (2015). Sims (2011a) and Corcoran and Evans (2015) contrast fiscal effects of adequacy and equity reforms. Each relies on a sample ending in 2002, early in the adequacy era.

(2016). We supplemented these with our own research into case histories, and updated them through 2011. We also tabulated major legislative SFRs. In some important cases (e.g., Colorado, California), legislatures reformed finance systems without prior court decisions, often to forestall adverse judgments in threatened or ongoing lawsuits. Our primary analyses include these, though we also present results that focus exclusively on court orders. Some of the reforms were accompanied by governance, curriculum, or accountability changes, though our assessment is that these additional changes were typically not very important or impactful.

Appendix Table 2.9 presents a complete list of our events and compares it to those used in other studies. We identify a total of 64 school finance reform events in 26 states between 1990 and 2011.[12] 39 (61 percent) involve court orders; the remainder are legislative actions without a major court order in the same year. States with events are quite geographically diverse, though reforms are rare in the Deep South and upper Midwest.

18 states had multiple events in our period. These were generally closely spaced: 60 percent were three or fewer years apart. In these cases, we suspect that only one generated a major change in the states finance rules and that others were procedural steps (e.g., court orders that were disregarded or legislation changes that were later found inadequate). Our analytical strategy is built with this idea in mind, though our results are robust to alternative models of the impact of multiple reform events in the same state.

2.3 Analytic Approach

To identify the causal effect of school finance reforms, we leverage variation in the timing of reform events in an event-study framework. Our strategy is based on the idea that states without events in a particular year form a useful counterfactual for states that do have events in that year, after accounting for fixed differences between the states and for common time effects. The key assumption is that the exact timing of events is as good as random. We think this is plausible, given the idiosyncrasies of judicial processes. An attractive feature of our approach is that it builds in placebo tests that should identify likely violations of this assumption.

Our simplest event study specification models events as permanent, immediate shifts in outcomes relative to other states:

$$\theta_{st} = \delta_s + \kappa_t + \mathbb{1}[t > t_s^*]\beta^{jump} + \epsilon_{st} \tag{2.1}$$

Here, θ_{st} represents some summary of the distribution of funding or achievement in state s in year t. We discuss our particular measures below. δ_s and κ_t represent state and year effects, respectively. t_s^* is the date on which state ss event occurred. (For now, we assume that each state has just one event; this term is set to zero for states without events.) The coefficient estimate β^{jump} represents the change in the outcome following the event. In all of

[12] Our panel excludes the 1989 *Rose* decision but includes KERA, the legislatures response in 1990.

our analyses, we use standard errors that are clustered at the state level to allow for arbitrary dependence of ϵ_{st} across t within s.

SFRs may not affect θ_{st} immediately, but may develop more gradually. This is particularly true for student achievement outcomes, as the achievement of a student in year t likely depends in part on the quality of the schooling she received in prior years. In addition, if event timing is non-random, states with events may diverge from states without events even before the date of the event. To accommodate these ideas, we add two trend terms to (2.1):

$$\theta_{st} = \delta_s + \kappa_t + \mathbb{1}[t > t_s^*]\beta^{jump} + \mathbb{1}[t > t_s^*](t - t_s^*)\beta^{phasein} + (t - t_s^*)\beta^{trend} + \epsilon_{st} \tag{2.2}$$

$\beta^{phasein}$ captures delayed event effects and represents the annual change in outcomes in state s after t_s^*, relative to the same state prior to the event. β^{trend}, which is identified from changes in s relative to other states in years prior to t_s^*, represents a falsification test: $\beta^{trend} \neq 0$ would indicate that event timing is meaningfully non-random.

We also estimate non-parametric models that do not constrain the phase-in and prior trend effects to be linear:

$$\theta_{st} = \delta_s + \kappa_t + \sum_{r=k_{min}}^{k_{max}} \mathbb{1}[t = t_s^* + r]\beta_\tau + \epsilon_{st} \tag{2.3}$$

Here, β_r represents the effect of an event in year t_s^* on outcomes r years later (or previously, for $r < 0$). These effects are measured relative to year $r = 0$, which is excluded. We censor r at $k_{min} = -5$, so β_{-5} represents average outcomes five or more years prior to an event, relative to those in the event year.

Comparisons of the parametric and non-parametric estimates indicate that the simple specification (2.2) does a good job of capturing dynamics in finances and student achievement surrounding events, though the post-event "jump" is sometimes spread out over a few years following the event. In only one of the specifications that we estimate do we reject the null hypothesis that the pre-event coefficients (β^{trend} in (2.2) and $\{\beta_{-k}, ..., \beta_{-1}\}$ in (2.3)) are all zero, and in this case it appears to be an idiosyncratic blip in a single β_{-r} coefficient (see Figure 2.7, below). This supports our identifying assumption.

When we examine finance outcomes, all of the post-event effect appears to be nearly immediate, so we focus on the simpler specification (2.1). By contrast, in our student achievement analysis, the "jump" is never distinguishable from zero, and all of the effect that we estimate operates through the $\beta^{phasein}$ coefficient. We thus emphasize specifications that allow for a phase-in effect but no post-event jump. In each case, these simple specifications fit the non-parametric results quite well.

Our event study methodology is a form of difference-in-differences (DD). The identifying assumption is that without finance reforms, outcomes would have moved in parallel in treated and untreated states. While we view this as plausible, it may not be correct (Hanushek et al. (1996a), Hanushek et al. (1996b)). We can weaken the assumption by shifting our focus from the absolute level of test scores to the relative scores of different students in the same states. Given the emphasis in adequacy rulings on districts serving disadvantaged students, a natural contrast is between students in high- and low-income districts. When we use as a dependent variable the gap in test scores between low-income and high-income districts in a state, the event study strategy is robust to arbitrary state-by-year shocks to achievement, so long as they have similar effects on districts at different income levels. The identifying assumption is that the *relative* outcomes of low-income districts would have followed parallel trends across states in the absence of SFRs.

We consider two measures of relative outcomes in low-income districts. First, we use the gap between districts in the top and bottom quintiles of the state income distribution. These quintile gaps can be noisy, in part because they discard information on the middle 60 percent of districts. We thus emphasize a second measure, the slope of district-level outcomes with respect to log average income across all districts in the state.[13] A more negative slope corresponds to higher relative outcomes in low-income districts. For both finance and achievement outcomes, the slope and quintile gaps are highly (negatively) correlated, and all of our results are robust to the choice of relative outcome measure.

Event Studies with Multiple Events

Many states had multiple events (court orders or legislation) over our period. Unfortunately, there is no accepted strategy for conducting event studies with multiple events per unit. Our primary estimates are based on a single event in each state. The intuition here is that when states have multiple events, they often represent jockeying between the legislature and the courts with only minor changes in school finance until the legislature finally enacts a major reform, and then continued jockeying afterward as advocates continue to push for additional changes. To identify the most consequential reform, we use data on state aid to districts to identify a regime change in the progressivity of a states finance system, relying on methods for the identification of change points in time series data (e.g., Bai (1997); see also Card, Mas, et al. (2008)). We then use that as the date of the event for our analyses of student achievement.

Specifically, let θ_{st} be our slope measure of the progressivity of state aid. For each state and each potential event date t_s^* – that is, each year that we observe a major court order or legislative change we estimate a time series regression using as the only explanatory variable an indicator for observations after that date:

$$\theta_{st} = \alpha + \mathbb{1}[t > t_s^*]\kappa + \epsilon_{st} \tag{2.4}$$

We select the event date that yields the largest t statistic for κ or, equivalently, the smallest mean squared error for this time series regression.[14] We treat the selected date as

[13]Specifically, we regress district-level spending per pupil or mean achievement on log mean income, controlling for log enrollment. The regression is estimated separately for each state and year, and in achievement models for each subject and grade. The district log income coefficients are used as θ_{st} for subsequent analyses at the state-year-(subject-grade) level. See the Appendix for further detail.

[14]We restrict attention to $t*$ for which the estimated κ has the expected sign.

the single event in state s.

Bai (1997) shows that if there really is a structural break in the time series (with a non-zero true κ) this method is super-consistent for the location of the break, permitting inference regarding κ to treat its location as known. However, in the event that there is no structural break (i.e., that each court order and legislative change in the state was ineffective, with $\kappa = 0$), our method will nevertheless pick one of the potential events. This could lead us to overstate the effect of a true reform on the progressivity of state aid. Our main outcome, however, is student achievement, and we do not use achievement data in selecting events. Thus, the potential inclusion of some non-reforms in our event study analysis might lead us to understate the effect of a true SFR on student achievement, since our estimates would combine the effects of true reforms with those of spurious non-events.

We also present estimates from two additional approaches to multiple events. One includes all events, without judgment about their relative importance. To implement this approach, we create a separate copy of the time series for the state for each apparent event, using a different value of t^*_s for each copy. We then stack the copies, replacing the state effects in equations (2.1)-(2.3) with state-by-event effects.[15] In Monte Carlo simulations (see Appendix), this method works well to identify the average effect of events both when each event has the same effect and when only one event in a state has a non-zero effect. Our final approach follows the prior literaturewhich generally emphasizes simple specifications analogous to (2.1)by focusing on the initial court order in each state, even if this was not implemented for many years. Here, we treat states without court orders as untreated, though in some cases they saw legislative reforms. Results are extremely similar across all three methods. Accordingly, we do not view multiple events as a major issue in practice.

2.4 Data

Our analysis draws on data from several sources. We begin with our database of state SFR events, discussed above. We merge this to district-level finance data, from the National Center for Education Statistics (NCES) annual census of school districts and the Census of Governments; mean household income by district from the 1990 Census; and the NAEP achievement measures, aggregated to the district-year level.

The district finance data report enrollment, revenues and expenditures annually for each local education agency.[16] We convert all dollar figures to 2013 dollars per pupil, and exclude very small districts and those with highly volatile enrollment or implausible per-pupil funding. Details are in the appendix.

We construct student achievement measures from the restricted-use “State NAEP” microdata. The state NAEP began in 1990, with 42 states participating. It has been admin-

[15]Results are unchanged when data are reweighted to offset the overrepresentation of states with multiple events.

[16]Census data are available in 1989-90 and 1991-92, and annually since 1994-95. We use samples from the Census Bureaus Annual Survey of Government Finances for 1992-93 and 1993-94.

istered roughly every two years since. Since 2003, all states have participated in 4th and 8th grade assessments in math and reading in every odd-numbered year.[17] Table 2.1 shows the schedule. Tests are administered to around 100,000 students (more in later years) in each subject-grade-year. These consist of representative samples of about 3,500 students per state, spread across about 140 schools in 80 districts.

The NAEP uses a consistent scoring scale across years for each subject and grade in order to permit time-series comparisons. We standardize scores to have mean zero and standard deviation one in the first year that the test was given for the grade and subject, but allow both the mean and variance to evolve afterward. We then aggregate to the district-year-grade-subject level and merge to the district finance and demographics data.[18]

Table 2.2 presents district-level summary statistics, pooling data from 1990-2011. The rightmost columns show means for districts in the top (Q5) and bottom (Q1) quintiles by average family income in each state.

2.5 Finance Reforms and School Finance

We begin our empirical analysis by documenting the implications of SFR events for school finance. We use the approach discussed in Section 2.3 to select a single SFR event that best explains the time series of the state aid log district income slope in each state.

Figure 2.3 graphs event study results for state transfers per pupil in the lowest-income (Q1) quintile of districts. We present several plots of this basic form. The solid line represents estimates from the non-parametric event study specification (2.3), while dotted lines show pointwise 95 percent confidence intervals. The dashed line shows the parametric specification (2.2). There is a small upward trend in state revenues prior to the finance reform events, but this is not statistically significant in either the parametric or the nonparametric specification. Following reforms, state revenues increase substantially, by roughly $1,300 in the 4th post-event year. Though out-year estimates are noisy, impacts appear to persist through the end of our sample. Figure 2.4 repeats the same analyses for the highest income (Q5) districts. Estimated changes in funding following reforms are much smaller here; while the nonparametric post-event effects are jointly significant, the parametric estimates are not and in any event the magnitudes are quite small.

We report coefficients from our parametric specifications for state revenues in the lowest and highest income districts in columns 1 and 2 of Table 2.3; column 3 shows estimates for average revenues across all districts for comparison. In Panel A, we report the simple specification (2.1), while Panel B adds the pre-event and post-event trends from specification (2.2). (It is these that are shown in Figures 2.3 and 2.4.) The former indicates that average state funding rises by $1,225 following events in first quintile districts and by $527 (not

[17]The NAEP also tests 12th graders, but samples are smaller, and other subjects.

[18]The pre-2000 NAEP data do not use the same district codes as the CCD. We are grateful to Bruce Kaplan, Kate Pashley, and Fatih Unlu for their assistance in locating the crosswalk from the older NAEP data to schools and districts.

significant) in fifth quintile districts. The upward trends preceding events seen in Figures 2.3 and 2.4 are reflected in the point estimates in Panel B, but are small and not distinguishable from zero. Similarly, point estimates indicate that the post-event jumps fade slightly over subsequent years, but these trends are again small and insignificant.

Panels C and D of Table 2.3 repeat the specifications from Panels A and B, this time taking total district revenues, inclusive of state aid and other revenues, as the dependent variable. These are quite similar to those for state revenues in both low- and high-income districts. There is no indication that declines in local revenues offset increases in state funding in low income districts, nor in (panel D) of pre-trends or erosion of initial impacts. The more flexible nonparametric specifications (Appendix Figure 2.11) are also similar.

In additional analyses of state budgets (Appendix Table 2.10), we have found no indication that growth in educational spending following events crowds out state spending on other programs; rather, SFRs are associated with increases in state tax collections large enough to fully fund the increase in state transfers to districts.

As noted above, our analysis of student achievement impacts of SFRs focuses on contrasts between low- and high-income districts, to abstract from unrelated shocks to overall average achievement that might be correlated with the timing of these reforms. Columns 4 and 5 of Table 2.3 show estimates for these contrasts, first using the difference in funding between bottom- and top-quintile districts (column 4) and then the slope of funding with respect to log district income (column 5; this is shown graphically in Figure 2.5). Using each measure, we see sharp increases in relative state funding for low-income districts following events that show no sign of eroding thereafter. In no case is there any sign of a pre-event trend that would suggest a violation of our quasi-random timing assumption, nor is there any sign that increased progressivity of state aid is offset by local revenues.[19]

Table 2.3 makes clear that SFRs are associated with large increases in funding in low-income school districts. A natural question is how the additional funds are spent. Table 2.4 presents event-study coefficients from our simple model (2.1) for per-pupil revenues and spending in various categories. There is no apparent impact of SFRs on local or federal revenues. We see substantial impacts of SFRs on average instructional spending, both overall and in Q1 districts (columns 2 and 3). We also see effects on teachers per pupil and total teacher salaries but not on average teacher pay, suggesting that districts use additional funds to reduce class size.[20] Finally, we see large effects on non-instructional expenditures, particularly capital outlays.

Columns 4 and 5 show results for *relative* spending in low-income districts. Little of the increase in relative funding goes to instructional expenditures, while roughly half goes to capital spending. The capital spending effect is not surprising; many lawsuits specifically

[19]When we estimate specifications similar to the closely related analysis of earlier SFRs in Card and Payne (2002) (Appendix Table 2.11), estimated SFR effects are slightly larger but imprecise, and well within the earlier confidence intervals. Where Card and Payne find that total revenues rise by about \$0.50 per extra \$1 in state aid, our estimates indicate much more stickiness for the recent reforms.

[20]Using a different research design, Sims (2011b) finds effects of SFRs on teacher pay.

concern dreadful conditions in low-income schools, and SFR remedies often created funds to support renovation of schools in poor shape.[21]

2.6 Finance Reforms and District-level Student Achievement

The above results establish that reform events are associated with sharp, immediate improvements in the progressivity of school finance, with absolute and relative revenue increases in low-income school districts. We now turn to our main analysis, examining the effect of SFRs on student achievement.

Where the θ_{st} school finance measures formed a state-by-year panel, for test scores we have two additional dimensions: Grade and subject. We replace the year fixed effects (κ_t) in (2.1)-(2.3) with subject-grade-year effects. These capture any differences in tests between administrations, as well as changes in student performance by grade and/or subject that are common across states. To avoid confounding from state-level shocks, we focus on triple-difference specifications that use the achievement gap between low- and high-income districts as the dependent variable.

Sharp, permanent changes in funding, if used productively, should increase the flow of educational services. Achievement is cumulative, so these services are unlikely to have immediate impacts on test scores, but should raise scores gradually as students are exposed for longer. Effects should grow at least until students have been exposed to the new funding levels for their entire careers. They may even continue to grow beyond this point. For example, consider a state that responds to a court order by creating a new permanent facility to fund several school renovation and construction projects each year. Initially, only a few students benefit, but over time growing shares of students are exposed to funded projects. Insofar as better facilities promote student learning, achievement effects would continue to grow until several years after the last project is complete, potentially decades after the initial policy change. We thus emphasize the phase-in coefficient from equation (2.2) as the primary measure of SFR effects on test scores.

Figure 2.6 presents our event-study analysis of the slope of achievement with respect to district income. Recall that improvements in the relative achievement of students in low-income districts reduce this slope. As before, we present non-parametric results (equation (2.3)) as a solid line and estimates of our three-parameter model (equation (2.2)) as a dashed line. As before, there is no indication of a differential trend in reform states prior to events. Following events, the non-parametric series does not react immediately, but begins trending noticeably downward starting in about the fifth post-event year (though the immediate trend

[21] Neilson and Zimmerman (2014) find that school reconstruction causes increases in student achievement. Cellini et al. (2010) and Martorell et al. (2016) fail to find significant effects, but each study is under-powered to detect effects of plausible magnitude.

break encoded in (2.2) fits the data nearly as well). The downward trend continues through the end of our sample.[22]

Table 2.5 presents the parametric estimates. We begin in Column 1 with our three-parameter model, as shown in Figure 2.6. The estimated pre-event trend is essentially zero and the post-event jump is also small, but the post-event change in trend is large and statistically significant. Column 2 presents a specification that discards the other two coefficients. Results are quite similar. The estimated change in the slope is -0.010 per year. This implies that each year after an event, a district with log mean income one unit (about two-thirds) below the state average sees its scores rise relative to the state average by 0.010 standard deviations, accumulating to 0.10 SDs over ten years. This is quantitatively meaningful on average in our sample the slope of test scores with respect to log income is 0.96 so SFRs reduce this gradient by approximately one-tenth within ten years.

As discussed above, the pattern of gradually growing effects in Figure 2.6 is consistent with a view of achievement as a stock reflecting accumulated past input flows. The pattern deviates from expectations in one respect, however: There is no indication that the phase-in of the effect slows five or nine years after the event, when the 4th and 8th graders, respectively, will have attended school solely in the post-event period.[23] This may reflect the use of some additional funds for durable investments, as discussed above. We do not have enough precision, however, to rule out a flattening of the effect at the expected time.

Figures 2.7 and 2.8 present estimated test score impacts for the lowest- and highest-income districts, respectively. The effects on the income gradient are driven by dramatic increases in test scores in the lowest-income districts.[24] In higher-income districts, there is little sign of a systematic post-event change. Parametric estimates are shown in Columns 3 and 4 of Table 2.5; Column 5 shows that the impact of events on the test score gap between bottom- and top-quintile districts is 0.008 SDs per year, or 0.013 SDs in the more flexible model (column 6). The gap in mean log incomes between the top and bottom quintiles averages 0.65, so the quintile point estimate is a bit larger than what we obtain for our income slope measure in columns 1-2. Our earlier finance analyses also indicated larger effects for quintile gaps than for slopes.

Table 2.6 presents estimates separately by subject and grade. We cannot reject the null hypothesis of equal effects across each dimension. Appendix Figure 2.12 presents estimates of the phase-in coefficient for all five quintiles. Only the first quintile effect is large or distinguishable from zero. The ratio of test score effects to spending effects is larger at the bottom of the income distribution, consistent with the idea that funding is more productive in low-income districts, but equal ratios cannot be ruled out.

[22]The sawtooth pattern at the end of the sample likely reflects the biannual NAEP testing schedule.

[23]We have estimated separate non-parametric models for 4th and 8th grade scores. Both sets of effects grow roughly linearly through the end of our panels. See Lafortune et al. (2016), Appendix Figure 4.

[24]For the lowest-income districts (Figure 2.7), we can reject the null hypothesis of zero pre-event effects. This is driven by a temporary drop two years prior to events. A similar, though statistically insignificant, blip is apparent for high-income districts in Figure 2.8. There is no sign of systematic pre-event trends.

Robustness

Table 2.7 presents estimates of our key specifications from our two alternative approaches to event multiplicity. Column 1 repeats the estimates from our preferred approach from Tables 2.3 and 2.5. In Column 2, we include all identified events, creating separate panels for each; in Column 3, we focus only on the first court order in each state. Results are similar to those from our main specifications, though the initial court order approach yields less precise, insignificant estimates of finance effects in panel B.

One potential explanation for the achievement impacts that we identify is that they reflect changes in population stratification rather than changes in educational production. SFRs that flatten the gradient of school funding with respect to district income and that reduce the local share of school finance reduce the value of living in a high-income district, and may lead some high-income families to relocate to previously low-income districts. This could lead to rising achievement in these districts with no change in school effectiveness.

We assess this possibility in three ways. First, we have tested whether between-district income gaps narrow in the years following SFRs. We have found no evidence for thisdistrict log incomes in 2011 are highly correlated with those in 1990, and there is no sign that gaps narrow in states that had reforms relative to those that didnt. Second, we have conducted event study analyses, parallel to those for test scores, for district income or the district non-white or free- or reduced-price lunch eligible share (Appendix Table 2.12). In only one specificationfor the between-quintile gap in the free lunch sharedo we find evidence that the demographic composition of (initially) low-income districts changes following SFRs. This result is not robust, and is small relative to the test score impacts that we estimate.

Third, we decompose test scores into two components, and estimate separate SFR effects on each. Specifically, we estimate an individual-level regression of test scores on student demographic characteristics, pooling NAEP data across years for each grade-subject pair and including year fixed effects. We then construct separate achievement-log district income gradients from the fitted values (excluding the fixed effects) for this regression, representing student characteristics that would be affected by SFRs only through changes in sorting, and from the residuals. We find no evidence that reforms affect the demographic component of our test score progressivity measures, supporting our interpretation that our results primarily reflect changes in educational production in low-income school districts (see Appendix Table 2.15).

As a final robustness exercise, we have tested whether the SFR effect on achievement is sensitive to including controls for the presence of a school accountability policy in a state, or whether the SFR effect varies with school accountability. We found evidence for neither.

2.7 Finance Reforms and Statewide Achievement Gaps

The final topic that we investigate is whether finance reforms closed overall test score gaps between high- and low-achieving, minority and white, or low-income and non-low-income students in a state. These are perhaps better measures than our slopes and quintile gaps of the overall effectiveness of a states educational system at delivering equitable, adequate services to disadvantaged students (Card and Krueger (1992b); Krueger and Whitmore (2001)). However, because most inequality is within districts, changes in the distribution of resources across districts may not be well enough targeted to meaningfully close these gaps.

Table 2.8 presents estimates of effects on mean test scores across different subgroups of interest. The first panel shows a DD estimate of the effect on mean (pooled) test scores. The point estimate (not significant) implies a smaller impact per dollar than do our between-district contrasts, though we cannot rule out comparable effect sizes. In any event, our research design is more credible for outcome disparities than for the level of outcomes, as the latter would be confounded by unobserved shocks to average outcomes in a state that are correlated with the timing of school finance reforms ((Hanushek et al. (1996a), Hanushek et al. (1996b)). For example, if SFRs follow negative shocks to mean student achievement, this effect would be downward-biased. Another interpretation is that the marginal productivity of revenues is in fact higher in low-income districts.

The second panel shows impacts on the standard deviation or interquartile range of achievement within states, while the third and fourth panels present results by race and income, respectively. There is no discernible effect on achievement gaps by race or income or on the overall dispersion of test scores. Point estimates are all roughly a full order of magnitude smaller than the earlier estimates for district-level progressivity of mean scores.

Appendix Tables 2.13 and 2.14 resolve the discrepancy. While non-white, low-income, and low-scoring students are more likely than their white, higher-income, and higher-scoring peers to attend school in low-income school districts, the differences are not very large. Roughly one-quarter of non-white and low-scoring students, and one-third of low-income students, live in first-quintile districts, while about 10 percent of each live in fifth-quintile districts (Appendix Table 2.13). This leaves little room for SFRs to substantially affect the relative resources to which the typical minority, low income, or low scoring student is exposed.

To assess this more carefully, we assigned each student the mean revenues for his/her district and estimated event study models for the black-white, income, or test score gap in these imputed revenues. Results, in Appendix Table 2.14, indicate that finance events raise relative per-pupil revenues in the average black students school district by only $195 (S.E. 164), decrease relative per-pupil revenues in the average low-income students district by $33 (S.E. 219), and raise relative per-pupil revenues in the average low-scoring students district by $193 (S.E. 101). Even if funding was much more productive than the average effect implied by our analysis, the funding changes seen here would still not be enough to yield effects on black or low-income students average test scores large enough to detect with our research design. Thus, while reforms aimed at low-income districts appear to have been successful at raising resources and outcomes in these districts, we conclude that within-district changesin the distribution of funding or in other policies that reduce achievement gapswould be necessary to have dramatic impacts on the average low-income, minority, or low-scoring student.

2.8 Discussion

After desegregation, school finance reform is perhaps the most important education policy change in the United States in the last half century. But while the effects of the early reforms on school finance have been well studied, there is little evidence about the finance effects of more recent "adequacy" reforms or about the effects of any of these reforms on student achievement. Our study presents new evidence on each of these questions.

We find that state-level school finance reforms enacted during the adequacy era markedly increased the progressivity of school spending. They did not accomplish this by "leveling down" school funding, but rather by increasing spending across the board, with larger increases in low-income districts. Schools used these additional funds to increase instructional spending, reduce class size, and for capital outlays. Using nationally representative data on student achievement, we find that these reforms were productive: Reforms increased the absolute and relative achievement of students in low-income districts.

Some SFRs were accompanied by other policy changese.g., new curricula, accountability provisions, or new prekindergarten programsthat may have contributed to the achievement effects, though our impression is that for the typical reform the main change was in funding.[25] We thus interpret our estimates as reflecting the productivity of additional resources, though other interpretations cannot be ruled out.

The different time patterns of impacts on resources and on student outcomes, combined with the cumulative nature of the latter, prevents a simple instrumental variables interpretation of the reduced-form coefficients in terms of the achievement effect per dollar spent it is not clear which years revenues are relevant to the accumulated achievement of students tested r years after an event. To assess the magnitude of the impacts we estimate, we focus on estimated effects on student achievement ten years after an event. Because effects on school resources are stable in the years following events, these can be interpreted as the impact of a change in resources for every year of a students career (through 8th grade). Nevertheless, the focus on the $r = 10$ estimate is arbitrary. We would obtain larger estimates

[25]We used our event-study framework to estimate the association of SFRs with changes in state accountability policy, using various measures of accountability rules, and found no relationship. We also investigated specifications that allowed for interactions between finance reform events and the accountability regime, but found no evidence for this either. We are not aware of a systematic classification of other aspects of state policy that might have been affected by SFRs.

of the achievement effect per dollar if we used impacts more than ten years after events, or smaller effects with a shorter window.

Our preferred estimates, based on the gradient of student achievement with respect to district income, indicate that an SFR raises achievement in a district with log average income one point below the state mean, relative to a district at the mean, by 0.1 standard deviations after ten years. Our finance estimates indicate that this district saw an increase in relative state aid of $622 per pupil for each of those ten years, and an increase in total revenues of $424 per pupil.

An increase of $424 per pupil in spending each year from kindergarten through grade 8, discounted to the students kindergarten year using a 3 percent rate, corresponds to a present discounted cost of $3,400. Chetty, John N Friedman, et al. (2011) estimate that a 0.1 standard deviation increase in kindergarten test scores translates into increased earnings in adulthood with present value of $5,350 per pupil. This implies a benefit-cost ratio of 1.5, even when only earnings impacts are counted as benefits.[26]

This ratio is not wholly robust. Our quintile analysis shows larger revenue effects, implying a benefit-cost ratio below one, while the Jackson et al. (2016) study of the effects of earlier finance reforms on students adult outcomes implies much larger benefits per dollar than does our calculation. Thus, although these sorts of calculations are quite imprecise, the evidence appears to indicate that the spending enabled by finance reforms was cost-effective, even without accounting for beneficial distributional effects.

It is important to note that our research design is poorly suited to identifying the optimal allocation of school resources across expenditure categories, or to testing whether actual allocations are close to optimal. It allows us only to say that the average finance reformwhich we interpret to involve roughly unconstrained increases in resources, though in some cases the additional funds were earmarked for particular programs or tied to other reformsled to a productive (though perhaps not maximally productive) use of the funds.

Our results thus show that money can and does matter in education, and complement similar results for the long-run impacts of school finance reforms from Jackson et al. (2016). School finance reforms are blunt tools, and some critics (Hoxby (2001); Hanushek (2006)) have argued that they will be offset by changes in district or voter choices over tax rates or that funds will be spent so inefficiently as to be wasted. Our results do not support these claims. Courts and legislatures can evidently force improvements in school quality for students in low-income districts.

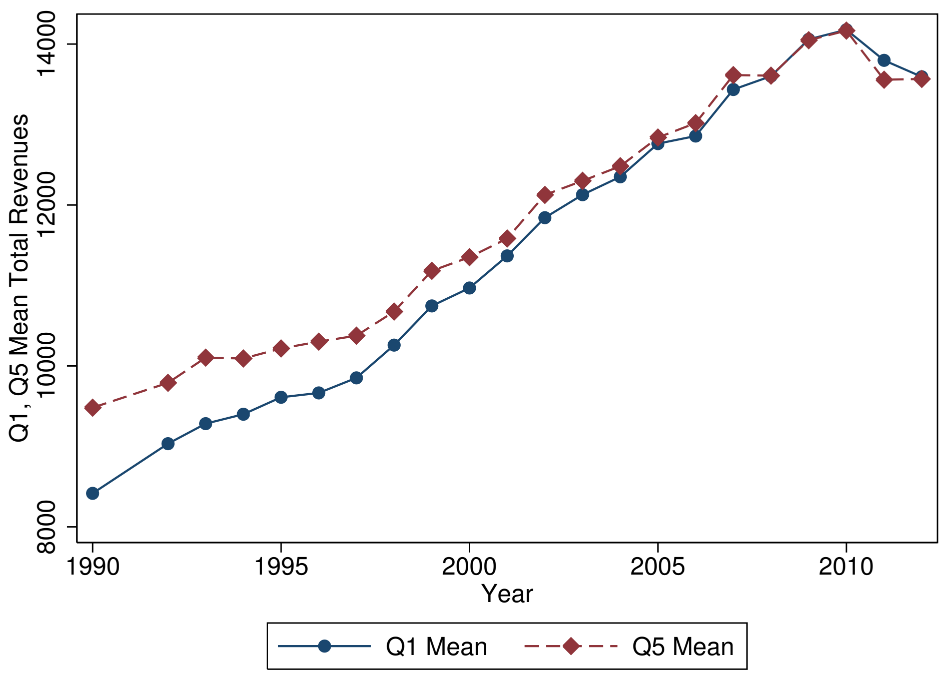
But there is an important caveat to this conclusion. As we discuss in Section 2.7, the average low-income student does not live in a particularly low-income district, so is not well targeted by a transfer of resources to the latter. Thus, we find that finance reforms reduced achievement gaps between high- and low-income school districts but did not have detectable effects on resource or achievement gaps between high- and low-income (or white

[26]The earnings effects of increases in 8th grade test scores are likely larger than those of increases in Kindergarten scores, so using estimates of the latter biases our benefit calculation downward. We do not count the cost of increased spending in grades 9-12, as we have no way to capture its benefits.

and black) students. Attacking these gaps would require policies aimed at the distribution of achievement *within* school districts, something that was generally not a focus of the reforms that we study.

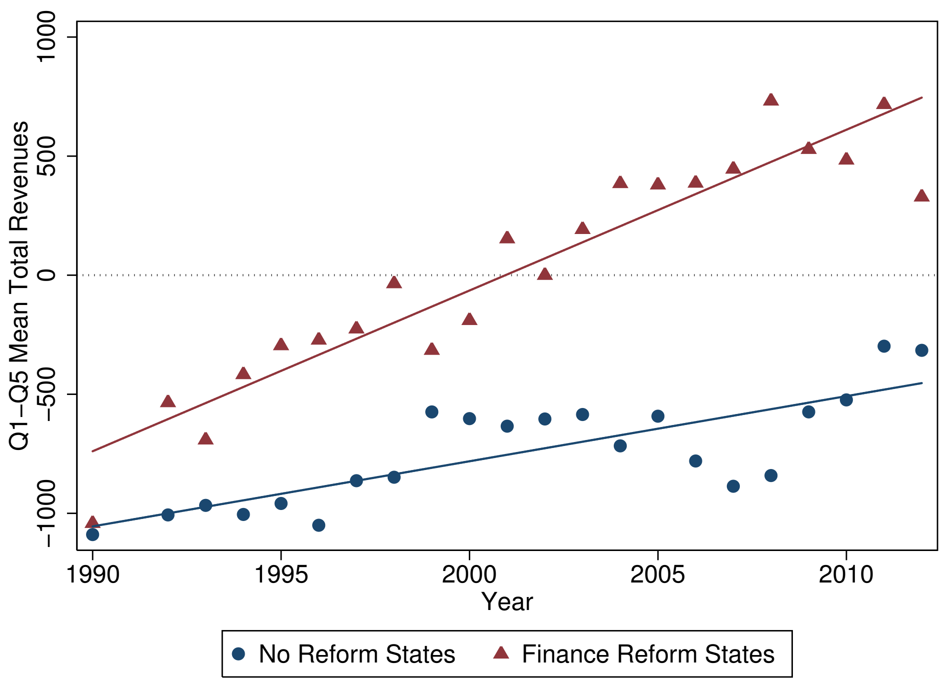
2.9 Figures

Figure 2.1: Mean revenues per pupil for highest and lowest income school districts, 1990-2012



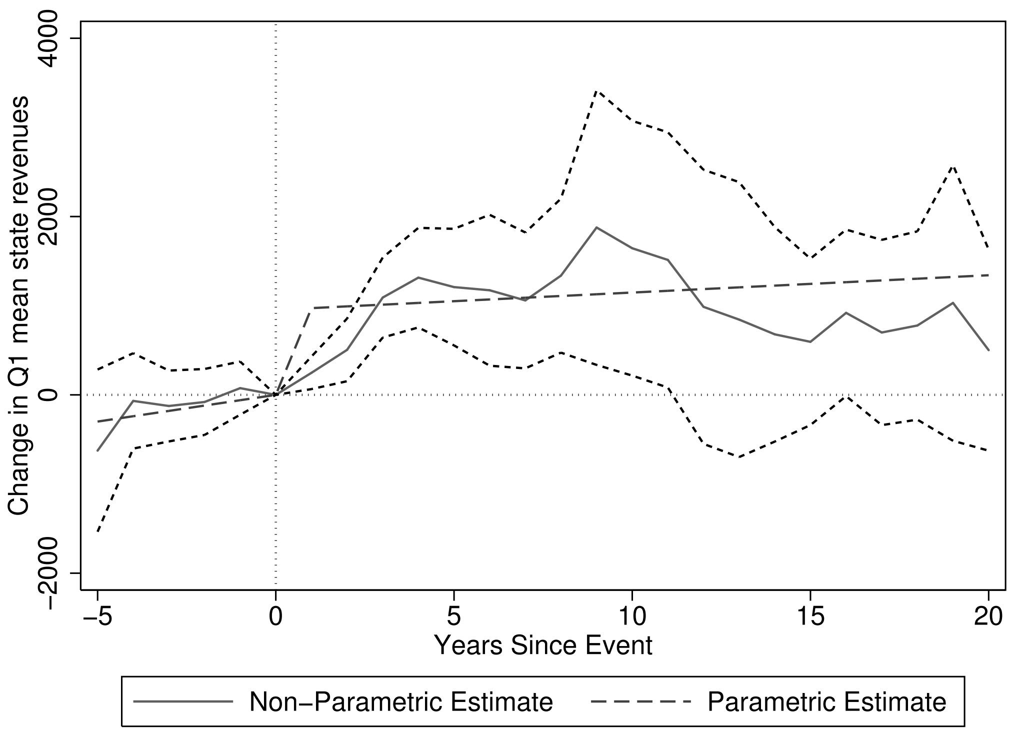
Notes: Highest (lowest) income districts are those in the top (bottom) 20% of their states' district-level distributions of mean household income in 1990, and are labeled as "Q5" and "Q1", respectively. See appendix for details of quintile classifications. Revenues are expressed in real 2013 dollars. Districts are averaged within states, weighing by log district enrollment; states are then averaged without weights. Hawaii and the District of Columbia are excluded.

Figure 2.2: Gap in revenues per pupil between lowest and highest income districts, by state finance reform status, 1990-2012



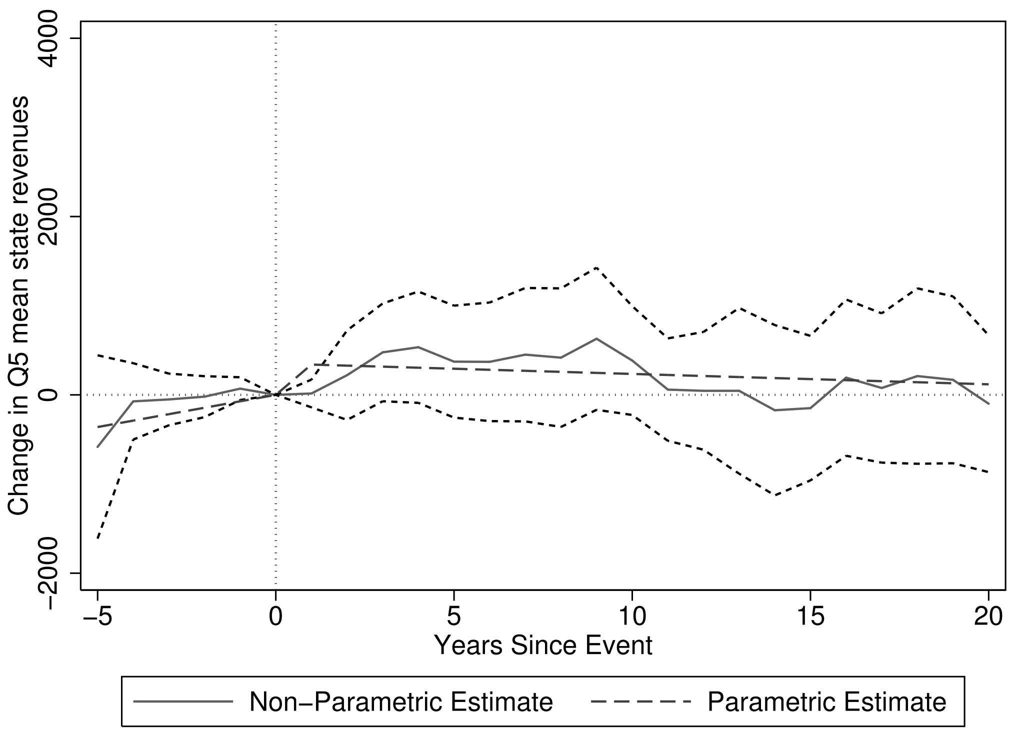
Notes: See notes to Figure 2.1. Finance reform states are those with school finance reforms between 1990 and 2011, as listed in Appendix Table A1. Lines show unweighted best linear fit to time series.

Figure 2.3: Event study estimates of effects of school finance reforms on mean state revenues in lowest income districts



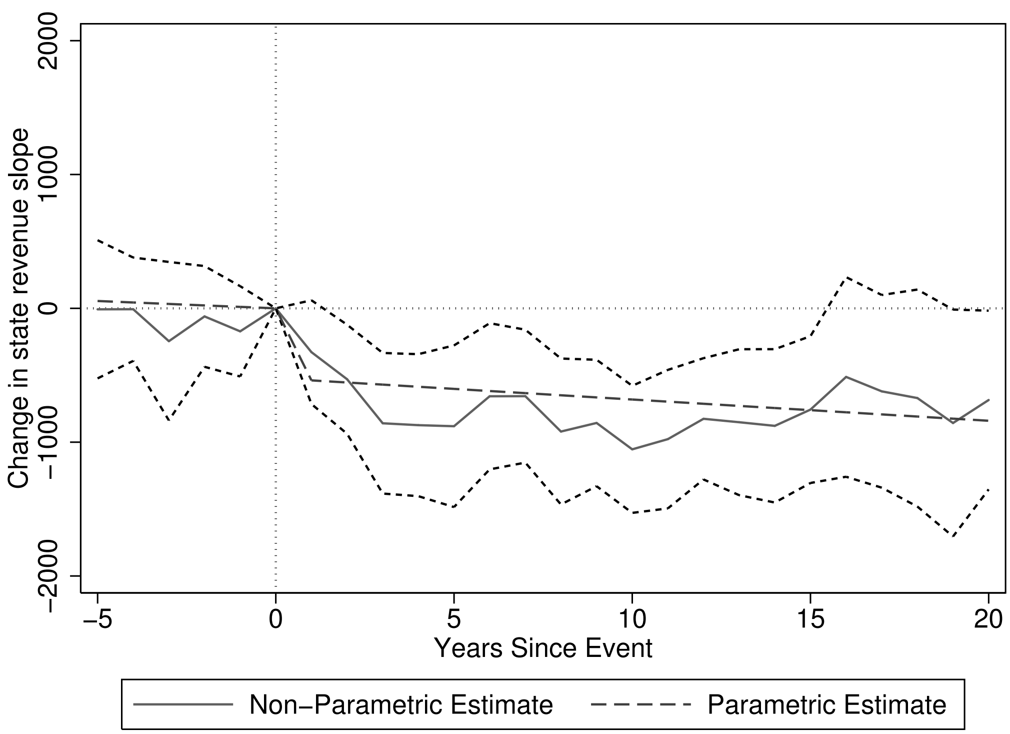
Notes: Figure displays coefficients from event study regressions. Dependent variables are mean state revenues in the lowest income quintile of districts, measured in 2013 dollars per pupil. Dashed lines show the three-parameter parametric model (equation (2.2)). Solid lines shows the non-parametric model (equation (2.3)), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Estimates for the parametric models are reported in Table 2.3, column 2, Panels B and C. The p value for the omnibus hypothesis test of zero pre-event effects in the non-parametric model in is 0.53; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.24; for the test that the post-event jump and change in trend is zero it is 0.01. Standard errors are clustered at the state level.

Figure 2.4: Event study estimates of effects of school finance reforms on mean state revenues in highest income districts



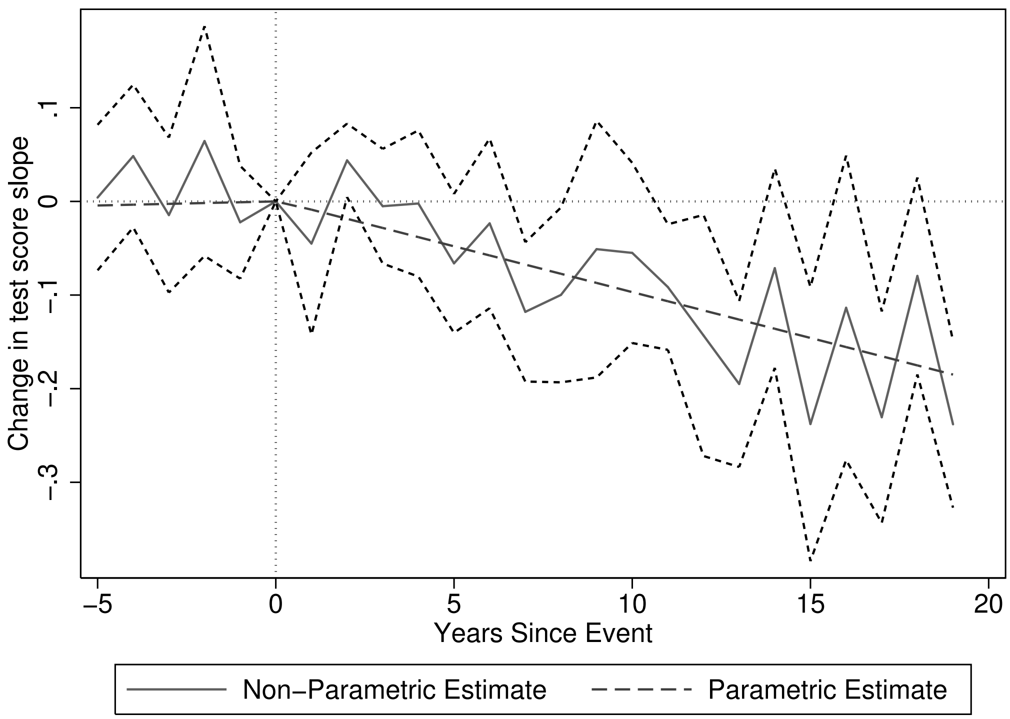
Notes: Figure displays coefficients from event study regressions. Dependent variables are mean state revenues in the highest income quintile of districts, measured in 2013 dollars per pupil. Dashed lines show the three-parameter parametric model (equation (2.2)). Solid lines shows the non-parametric model (equation (2.3)), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Estimates for the parametric models are reported in Table 2.3, column 2, Panels B and C. The p value for the omnibus hypothesis test of zero pre-event effects in the non-parametric model is 0.41; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.21; for the test that the post-event jump and change in trend is zero it is 0.30. Standard errors are clustered at the state level.

Figure 2.5: Event study estimates of effects of school finance reforms on progressivity of state revenues



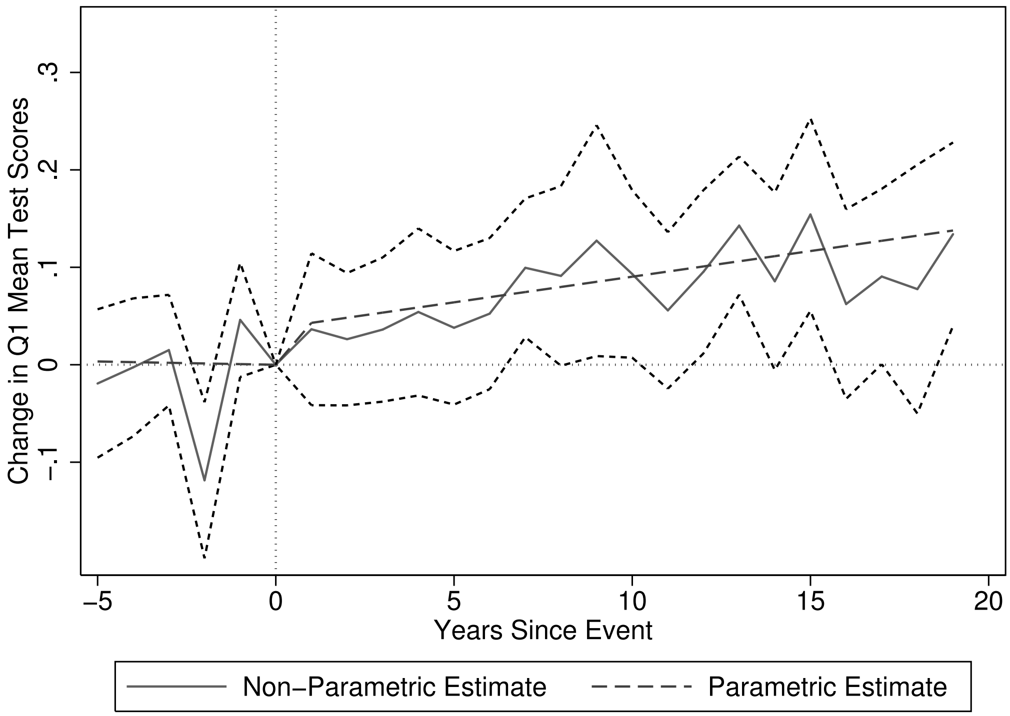
Notes: Figure displays coefficients from event study regressions. Dependent variable is the slope of state per-pupil revenues (in 2013\$) with respect to log mean family income, controlling for log enrollment and district type. Dashed lines show the three-parameter parametric model (equation (2.2)). Solid lines shows the non-parametric model (equation (2.3)), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Estimates for the parametric models are reported in Table 2.3, Panels A and B, columns 5. The p-value for the omnibus hypothesis tests of zero pre-event effects in the non-parametric model is 0.73; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.67; for the test that the post-event jump and change in trend is zero it is 0.05. Standard errors are clustered at the state level.

Figure 2.6: Event study estimates of effects of school finance reforms on progressivity of test scores



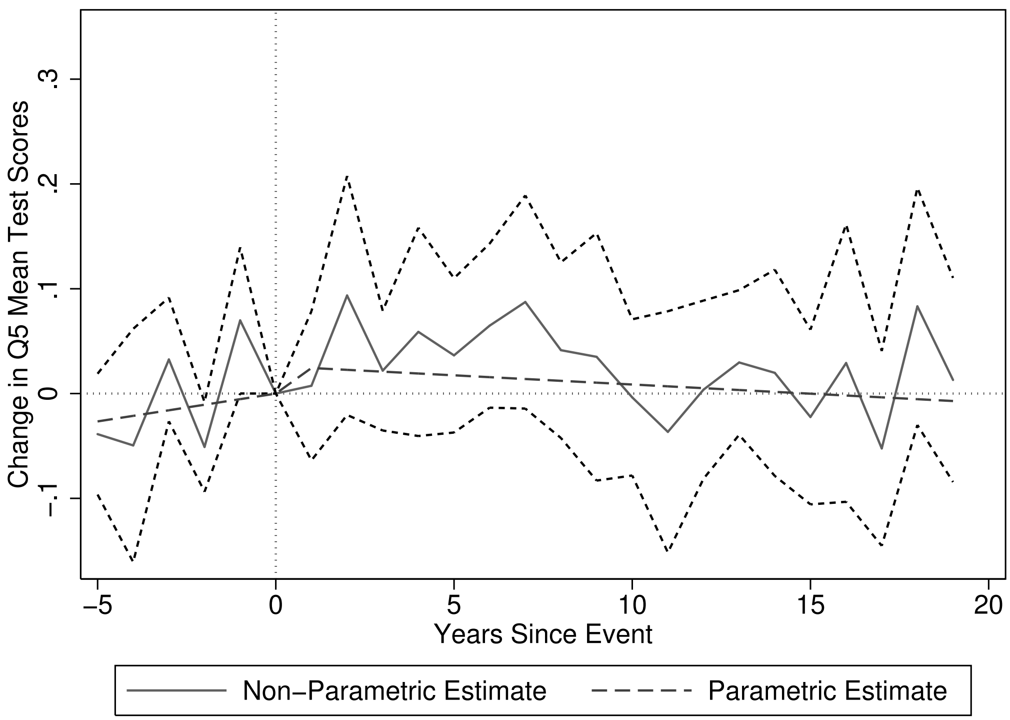
Notes: Figure displays coefficients from event study regressions. Dependent variable is the slope of mean test scores with respect to log mean family income, controlling for log enrollment. Dashed lines show the three-parameter parametric model (equation (2.2)). Solid lines shows the non-parametric model (equation (2.3)), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Both event study regressions include state and subject-grade-year fixed effects. Estimates for the parametric models are reported in Table 2.5, Column 1. The p-value for the omnibus hypothesis test of zero pre-event effects in the non-parametric model is 0.43; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.80; for the test that the post-event jump and change in trend is zero it is 0.02. Standard errors are clustered at the state level.

Figure 2.7: Event study estimates of effects of school finance reforms on mean test scores in lowest income school districts



Notes: Figure displays coefficients from event study regressions. Dependent variables are mean test scores for students at districts in the bottom quintile of the state's distribution of 1990 district mean household incomes. Dashed lines show the three-parameter parametric model (equation (2.2)). Solid lines shows the non-parametric model (equation (2.3)), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Both regressions include state and subject-grade-year fixed effects. The p-value for the omnibus hypothesis test of zero pre-event effects in the non-parametric model is 0.01; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.86; for the test that the post-event jump and change in trend is zero it is 0.01. Standard errors are clustered at the state level.

Figure 2.8: Event study estimates of effects of school finance reforms on mean test scores in highest income school districts



Notes: Figure displays coefficients from event study regressions. Dependent variables are mean test scores for students at districts in the top quintile of the state's distribution of 1990 district mean household incomes. Dashed lines show the three-parameter parametric model (equation (2.2)). Solid lines shows the non-parametric model (equation (2.3)), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Both regressions include state and subject-grade-year fixed effects. The p-value for the omnibus hypothesis test of zero pre-event effects in the non-parametric model is 0.02; the p-value for zero post-event effect is <0.001. In the parametric model, the p-value for the hypothesis that the pre-event trend is zero is 0.15; for the test that the post-event jump and change in trend is zero it is 0.25. Standard errors are clustered at the state level.

2.10 Tables

Table 2.1: NAEP Testing Years

Year	Subjects and grades covered				Number of	Number of
	Math G4	Math G8	Reading G4	Reading G8	States	Students
1990		X			38	97,900
1992	X	X	X		42	321,120
1994			X		41	104,890
1996	X	X			45	228,980
1998			X	X	41	206,810
2000	X	X			42	201,110
2002			X	X	51	270,230
2003	X	X	X	X	51	691,360
2005	X	X	X	X	51	674,420
2007	X	X	X	X	51	711,360
2009	X	X	X	X	51	775,060
2011	X	X	X	X	51	749,250

Notes: In final column, students are cumulated across all tested subjects and grades, and rounded to the nearest 10.

Table 2.2: Summary statistics (district-year panel)

	Overall			Mean by subgroup	
	N	Mean	SD	Q1	Q5
Enrollment	229,386	67,523	181,811	13,537	31,403
Log(mean income, 1990)	223,334	10.53	.2935	10.21	10.9
Total revenue p.p.	229,386	11,087	3,489	10,809	11,871
State	229,386	5,135	2,291	6,371	4,003
Local	229,386	5,094	3,273	3,258	7,349
Federal	229,386	858.2	641.4	1,180	518.4
Expenditures p.p.	229,386	11,264	3,685	10,837	12,116
Instructional	229,386	5,845	1,953	5,659	6,167
Non-instructional	229,386	5,419	2,221	5,178	5,949
NAEP scores	49,867	.2559	.4578	.02925	.5884

Notes: Table reports summary statistics at the district by year level, weighted by district enrollment for the financial variables and by the sum of the student weights for the mean NAEP score.

Table 2.3: Event study estimates of effects of school finance reforms on revenues per pupil

	Q1	Q5	All districts	Q1-Q5 difference	Slope
A: State revenue (1 parameter)					
Post Event	1,225***	527	912**	711**	-622***
	(343)	(378)	(359)	(316)	(223)
B: State revenue (3 parameter)					
Post Event	954***	351	672**	606***	-522**
	(302)	(325)	(320)	(231)	(209)
Trend	60	72	68	-10	-11
	(50)	(56)	(50)	(25)	(25)
Post Event * Yrs Elapsed	-40	-84	-61	42	-5
	(70)	(61)	(60)	(36)	(21)
C: Total revenue (1 parameter)					
Post Event	1,233***	544**	829***	701**	-424
	(370)	(277)	(302)	(309)	(304)
D: Total revenue (3 parameter)					
Post Event	1,164***	471*	839***	696***	-469**
	(287)	(277)	(269)	(243)	(233)
Trend	16	9	9	9	-25
	(39)	(32)	(32)	(24)	(45)
Post Event * Yrs Elapsed	-11	2	-17	-14	53
	(70)	(41)	(52)	(44)	(61)
Observations	1,078	1,076	1,078	1,076	1,078

Notes: Table reports estimates of the parametric event study models, equations (2.1) (panels A and C) and (2.3) (panels B and D). In columns 1-3, dependent variables are mean state (panels A and B) or total (panels C and D) revenues per pupil, weighting districts by their log enrollment; each is computed separately for each state and year. In columns 1 and 2, means are computed over the bottom and top, respectively, quintiles of the states' district 1990 mean household income distributions; in column 3 means are computed over all districts in each state. In column 4, the dependent variable is the gap in state (panels A and B) or total (panels C and D) revenues per pupil between districts in the bottom and top quintiles of the states' district 1990 mean household income distributions. In column 5, the dependent variable is the coefficient from a district-level regression of the state (panels A and B) or total (panels C and D) per-pupil revenue measure on the log of the district's 1990 mean household income, controlling for district log enrollment and district type (elementary / secondary / unified) and weighting by the district's average log enrollment over time. Event study regressions include state and year fixed effects, and are unweighted. Standard errors are clustered at the state level.

Table 2.4: Event study estimates of effects of school finance reforms on components of district finance

	Mean of depvar	Mean	Q1 Mean	Q1-Q5 Mean	Slope
***Revenue Effects*:**					
Total revenue	11,593	829***	1,233***	701**	-424
		(302)	(370)	(309)	(304)
State revenue	5,449	912**	1,225***	711**	-622***
		(359)	(343)	(316)	(223)
Local revenue	5,238	-146	-126	-126	90
		(307)	(233)	(235)	(339)
Federal revenue	907	63	134	116	34
		(83)	(143)	(116)	(33)
***Expenditure Effects*:**					
Total expenditures	11,595	907***	1,377***	753**	-449
		(290)	(367)	(309)	(309)
Current instructional exp.	6,000	443***	604***	243*	-161
		(134)	(155)	(127)	(208)
Teacher salaries + benefits	5,533	339**	449***	143	-103
		(153)	(169)	(117)	(189)
Mean teacher salary	63,321	-30	170	508	-247
		(1,016)	(1,052)	(932)	(1,127)
Pupil teacher ratio	15.50	-0.59***	-0.65***	0.03	0.20
		(0.19)	(0.19)	(0.20)	(0.17)
Non-instructional exp	5,595	464**	773***	511**	-232
		(186)	(257)	(235)	(176)
Student support	3,426	221**	299**	100	-81
		(102)	(119)	(83)	(88)
Total capital outlays	1,076	272**	486***	369**	-87
		(114)	(177)	(181)	(78)
Other current exp.	431.0	7.9	9.2	-2.5	-2.9
		(12.4)	(14.5)	(13.3)	(12.1)

Notes: Each entry in columns 2-5 represents the coefficient from a separate event study regression, using the one-parameter specification in equation (2.1). Dependent variables are constructed from district-level finance summaries indicated by row headings and expressed in per-pupil terms; means across districts are reported in column 1. Specifications correspond to columns 1 and 2 of Table 2.3, panels A (column 2) and B (column 3), and Table 2.4, panels A (column 4) and B (column 5). See notes to Tables 2.3 and 2.4. Standard errors are clustered at the state level.

Table 2.5: Event study estimates of effects of school finance reforms on student achievement

	Slopes		Q1	Q5	Q1-Q5	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Event * Yrs Elapsed	-0.011**	-0.010***	0.007**	-0.001	0.008**	0.013**
	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.006)
Trend	0.001					-0.006
	(0.003)					(0.005)
Post Event	0.001					0.011
	(0.023)					(0.024)
Observations	1498	1498	1509	1506	1504	1504
p, total event effect=0	0.02	0.01	0.02	0.69	0.04	0.07
State FEs	X	X	X	X	X	X
Sub-gr-yr FEs	X	X	X	X	X	X

Notes: Each column represents a separate event study regression, using specification (2.2) and, in columns 2-5, constraining $\beta^{jump} = \beta^{trend} = 0$. Dependent variable in columns 1-2 is the slope of test scores with respect to log mean 1990 income in the district, using NAEP weights and controlling for log district enrollment. In columns 3-4, dependent variable is the weighted mean score in districts in the bottom or top quintile, respectively, of the state district-level income distribution. In columns 5-6, dependent variable is the difference between the bottom and top quintiles. All are computed separately for each state-year-subject-grade cell with available data. All event study specifications include state and subject-grade-year fixed effects, and are weighted by the inverse squared standard error of the dependent variable. p-values for total event effect in columns 1 and 6 test the hypothesis that the β^{jump} and $\beta^{phasein}$ coefficients are both zero; in columns 2-5, the p-value is for the hypothesis that $\beta^{phasein} = 0$, with β^{jump} constrained to zero. Standard errors are clustered at the state level.

Table 2.6: Event study estimates of effects of school finance reforms on student achievement by subject and grade

	Test Score Slope	Q1-Q5 Mean
Pooled	-0.010***	0.008**
	(0.003)	(0.004)
By Subject:		
Math	-0.012***	0.007*
	(0.003)	(0.004)
Reading	-0.006	0.009**
	(0.005)	(0.004)
Difference	-0.006	-0.002
p-value	0.09	0.46
By Grade:		
G4	-0.010**	0.009*
	(0.005)	(0.005)
G8	-0.010**	0.007**
	(0.004)	(0.004)
Difference	0.000	0.001
p-value	0.93	0.72

Notes: First row repeats specifications from Table 2.5, columns 2 and 5. See notes to that table for details. Subsequent models restrict the event study sample to slope and quintile gaps computed in specific subjects or grades. Difference entries report the difference in coefficients between math and reading or grade 4 and grade 8 specifications, with p-values for the hypothesis that the event study coefficient is equal in the two subsamples. Standard errors are clustered at the state level.

Table 2.7: Sensitivity of event study estimates to the treatment of states with multiple events

	Selected Events	All events (stacked)	Initial court events
Panel A: Gradients			
State revenue p.p.	-622***	-479***	-432*
	(223)	(160)	(222)
Total revenue p.p.	-424	-197	-399
	(304)	(269)	(292)
NAEP scores	-0.010***	-0.009***	-0.009***
	(0.003)	(0.003)	(0.003)
Panel B: Q1-Q5 differences			
State revenue p.p.	711**	463**	516
	(316)	(191)	(354)
Total revenue p.p.	701**	448**	584
	(309)	(195)	(398)
NAEP scores	0.008**	0.011***	0.008**
	(0.004)	(0.004)	(0.004)

Notes: Column 1 repeats estimates of the one-parameter parametric event study models from Table 2.4, columns 1 and 3, and Table 2.5, columns 2 and 5. See notes to those tables for details. In column 2, each potential event in each state is included, with a separate copy of the state's finance or test score panel for each event. Event study specification is modified to include state-by-event (-by-grade-by-subject) fixed effects. Column 3 returns to the single-event specification, but uses the first post-1990 court order in each state as its event; states without judicial events are treated as not having finance reforms. Standard errors are clustered at the state level.

Table 2.8: Event study estimates for mean NAEP scores by subgroup

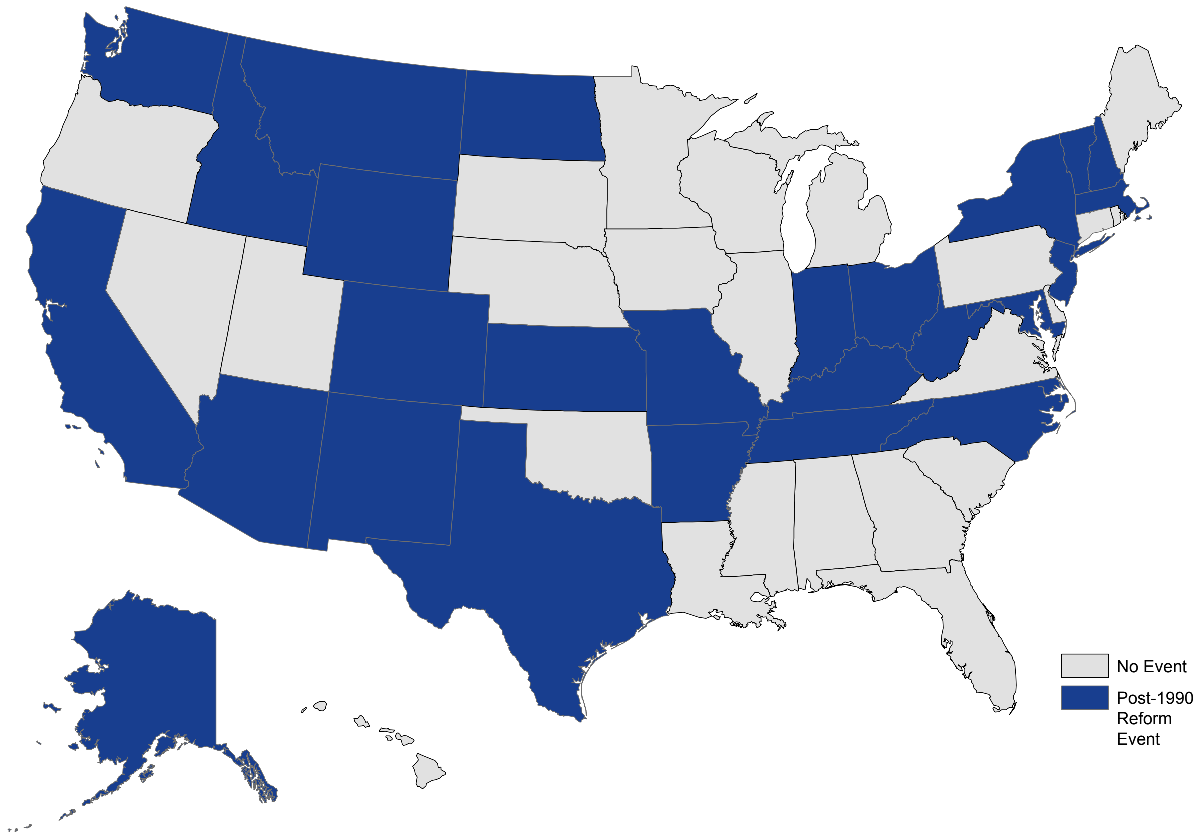
	Post Event * Yrs Elapsed	
Overall mean	0.004	(0.003)
Spread of distribution:		
Std Dev.	-0.000	(0.001)
25th percentile	0.004	(0.003)
75th percentile	0.003	(0.002)
P75 - P25	-0.001	(0.002)
By race:		
Black	0.001	(0.003)
White	0.004*	(0.003)
White - black	0.002	(0.002)
By free lunch status:		
Free lunch	0.001	(0.003)
No free lunch	0.004	(0.003)
No free lunch - free lunch gap	-0.000	(0.002)

Notes: Table reports event study specifications, using equation (2.3) with β^{jump} and $\beta^{phasein}$ constrained to zero. Dependent variables are the indicated summaries of the state-level student achievement distribution: The mean score; the standard deviation of scores; the 25th and 75th percentile scores; the interquartile range; mean scores for black and white students, respectively; the white-black mean score gap; mean scores for free/reduced-price lunch and non-free/reduced-price lunch students; and the gap between these. Standard errors are clustered at the state level.

2.11 Appendix Tables and Figures

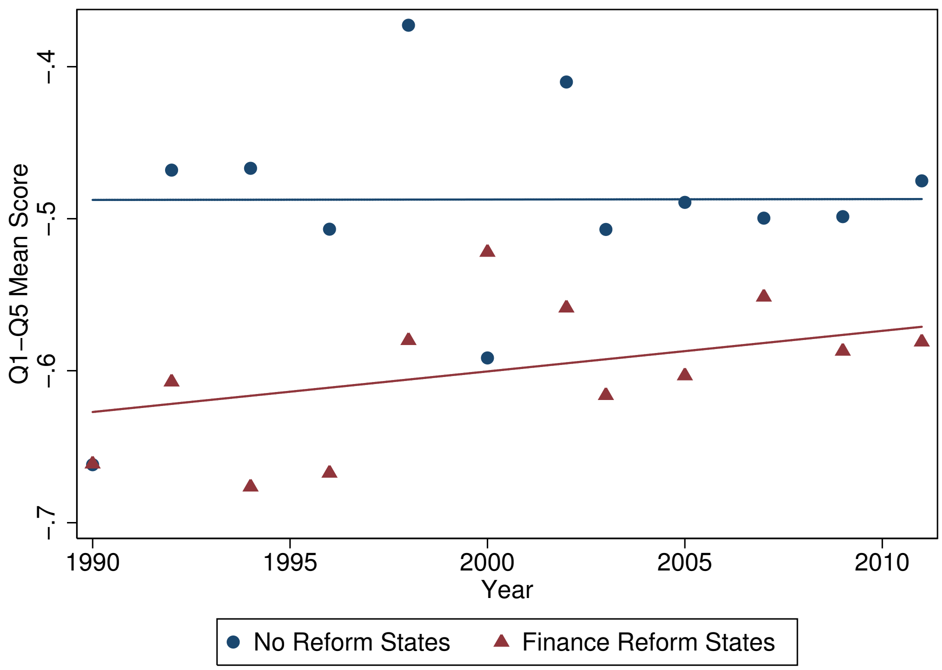
Appendix Figures

Figure 2.9: Geographic distribution of post-1989 school finance events



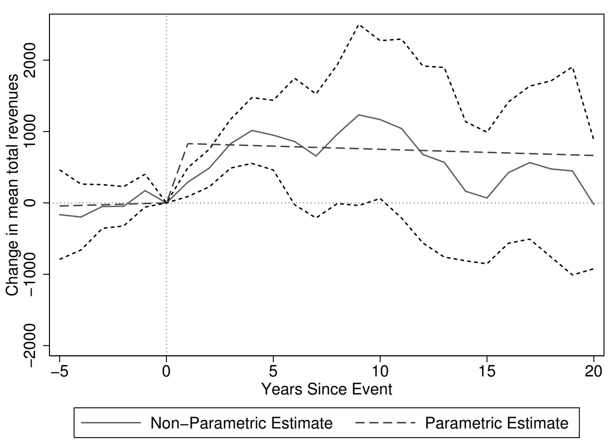
Notes: Map indicates states that had school finance reform events, as listed in Appendix Table 2.9, between 1990 and 2011.

Figure 2.10: Gap in average test scores between lowest and highest income districts, by state finance reform status, 1990-2011

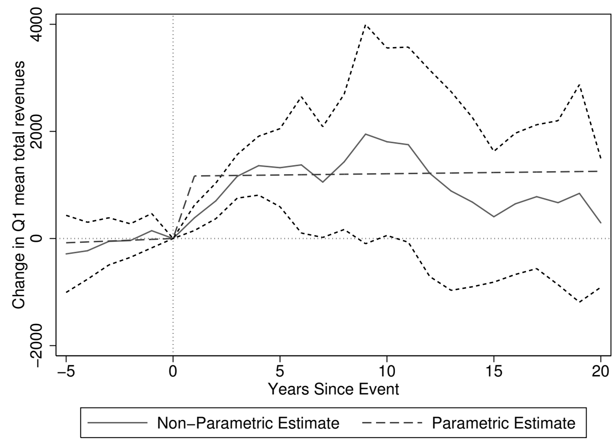


Notes: Lowest (Q1) and highest (Q5) income districts are defined as in Figure 2.1. NAEP observations in districts in each quintile are averaged, using NAEP sampling weights and separately for each grade and subject tested, and the Q1-Q5 difference is computed for each state. State-grade-subject Q1-Q5 differences are averaged separately for each group of states, weighting by the harmonic mean of the sum of the student weights in Q1 and Q5 districts. Lines show best linear fit to the time series.

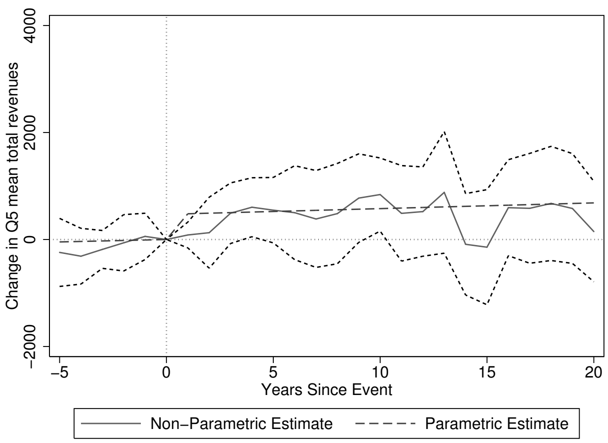
Figure 2.11: Event study estimates of effects of school finance reforms on mean total revenues in lowest and highest income districts



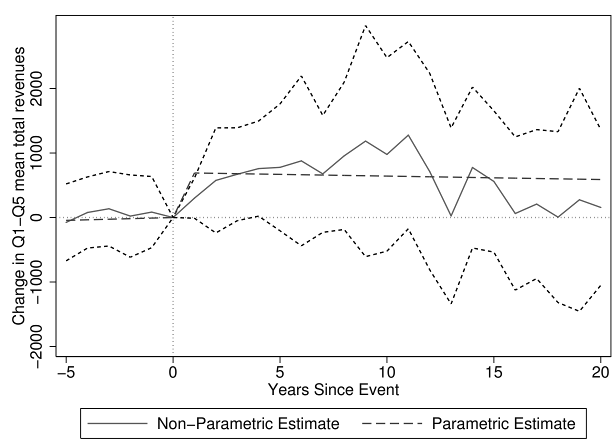
(a) Total revenue, mean



(b) Total revenue, Q1



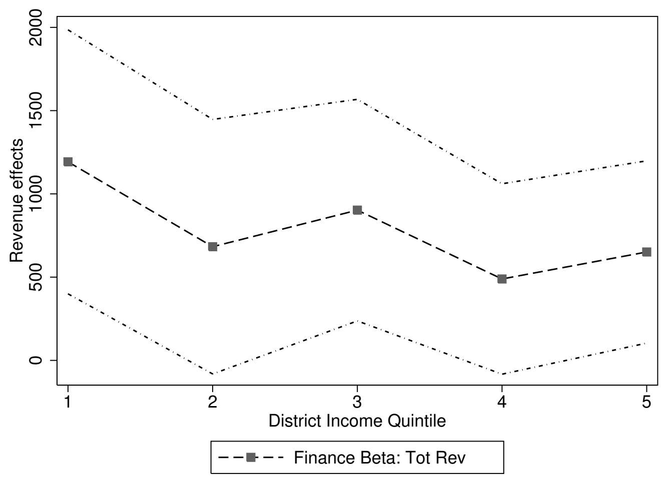
(c) Total revenue, Q5



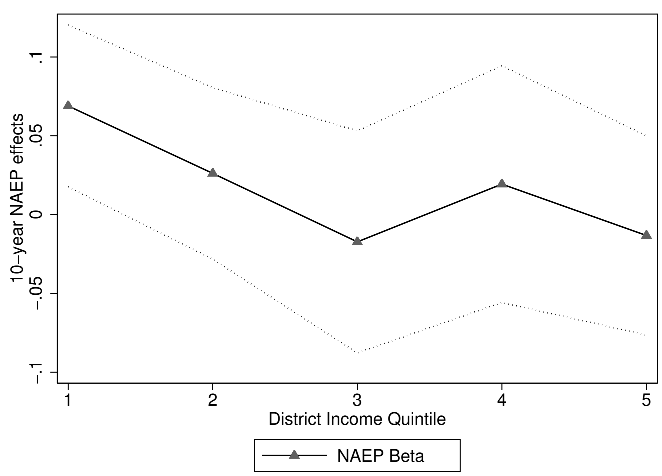
(d) Total revenue, Q1-Q5

Notes: Figure displays coefficients from event study regressions. Dependent variables are mean total revenues per pupil (panel A), mean total revenues per pupil in the lowest income quintile of districts (panel B), mean total revenues per pupil in the highest income quintile of districts (panel C), and the difference in mean total revenues per pupil between districts in the bottom and top income quintile in the state (panel D), all measured in 2013 dollars per pupil. Dashed lines show the three-parameter parametric model (equation (2.2)). Solid lines shows the non-parametric model (equation (2.3)), with the event year (indicated as 0) as the excluded category; dotted lines represent 95% confidence intervals. Estimates for the parametric models are reported in Table 3, panel D, columns 1-4. p values for omnibus hypothesis tests of zero pre-event effects in the non-parametric model in panels A-D are 0.15, 0.40, 0.74, and 0.86, respectively; p-values for zero post-event effects are <0.001 in all panels. In the parametric model, the p-values for the hypothesis that the pre-event trend is zero are 0.79, 0.68, 0.78, and 0.72; for the test that the post-event jump and change in trend is zero they are 0.01, <0.001, 0.22, and 0.01.

Figure 2.12: Event study estimates for total revenues and test scores by district income group



(a) Total revenue



(b) NAEP

Notes: Figure shows event study estimates from one-parameter parametric models for mean revenues and mean test scores in each quintile. Estimates for quintiles 1 and 5 are shown in Table 2.3, panel C, columns 1-2, and Table 2.5, columns 3 and 4. 95% confidence intervals shown by dotted lines. Standard errors are clustered at the state level.

Appendix Tables

Table 2.9: Complete Event List

Appendix Table A1
Complete Event List, 1990-2011

State	Year	Event	Lafortune, Rothstein & Schanzenbach (2016)	Jackson, Johnson & Persico (2016)	Corcoran & Evans (2015)
Alabama	1993	Alabama Coalition for Equity (ACE) v. Hunt; Harper v. Hunt		X	
Alaska	**1999**	Kasayulie v. State of Alaska	Court	X	
Arizona	1994	Roosevelt v. Bishop	Court	X	
	1997	Hull v. Albrecht	Court	X	
	1998	Hull v. Albrecht	Court	X	
	2007	Flores v. Arizona		X	
Arkansas	1994	Lake View v. Arkansas	Court	X	
	1995	Approved Equitable School Finance Plan (Acts 917, 916, and 1194)	Bill		n/a
	2002	Lake View v. Huckabee	Court	X	X
	2005	Lake View v. Huckabee	Court	X	X
	2007	Various acts resulting from Master's Report findings	Bill		n/a
California	1998	Leroy F. Greene School Facilities Act of 1998	Bill		n/a
	2004	Senate Bill 6, Senate Bill 550, Assembly Bill 1550, Assembly Bill 2727, and Assembly Bill 3001	Bill		n/a
Colorado	**2000**	Bill 181; Various Other Acts	Bill		n/a
Connecticut	1995	Sheff v. O'Neill		X	
	2010	Coalition for justice in Education Funding, Inc. v. Rell		X	n/a
Idaho	**1993**	Idaho Schools for Equal Educational Opportunity v. Evans (ISEEO)	Court		
	1994	Senate Bill 1560	Bill		n/a
	1998	Idaho Schools for Equal Educational Opportunity v. State (ISEEO III)		X	
	2005	Idaho Schools for Equal Educational Opportunity v. Evans (ISEEO V)	Court	X	
Indiana	**2011**	HB 1001 (Pl229)	Bill		n/a

(continued)

Appendix Table A1 (continued)

State	Year	Event	LRS (2016)	JJP (2016)	CE (2015)
Kansas	1992	The School District Finance and Quality Performance Act	Bill		n/a
	2005	Montoy v. State; Montoy v. State funding increases	Both	X	X
Kentucky	(1989)	Rose v. Council for Better Education, Inc.	Court	X	X
	1990	Kentucky Education Reform Act (HB 940)	Bill		n/a
Maryland	1996	Bradford v. Maryland State Board of Education	Court		
	2002	Bridge to Excellence in Public Schools Act (BTE) (Senate Bill 856)	Bill		n/a
	2005	Bradford v. Maryland State Board of Education		X	(upheld)
Massachusets	**1993**	McDuffy v. Secretary of the Executive Office of Education; Massachusetts Education Reform Act	Both	X	X
Michigan	1997	Durant v. State of Michigan		X	
Missouri	**1993**	Committee for Educational Equality v. State of Missouri; Outstanding Schools Act (S.B. 380)	Both	X	
	2005	Senate Bill 287	Bill		n/a
Montana	1993	House Bill 667	Bill	X	
	2005	Columbia Falls Elementary School v. State	Court	X	X
	2007	M.C.A. § 20-9-309	Bill		n/a
	2008	Montana Quality Education Coalition v. Montana		X	n/a
New Hampshire	1993	Claremont New Hampshire v. Gregg	Court	X	
	1997	Claremont School District v. Governor	Court	X	X
	1998	Opinion of the Justices--School Financing (Claremont III)			X
	1999	Claremont v. Governor (Claremont III); RSA chapter 193-E	Both	X	X
	2000	Opinion of the Justices--School Financing (Claremont VI)			X
	2002	Claremont School District v. Governor	Court	X	X
	2006	Londonderry School District v. New Hampshire		X	
	2008	SB 539	Bill		n/a

(continued)

Appendix Table A1 (continued)

State	Year	Event	LRS (2016)	JJP (2016)	CE (2015)
New Jersey	1990	The Quality Education Act; Abbot v. Burke	Both	X	X
	1991	Abbott v. Burke		X	
	1994	Abbott v. Burke	Court	X	X
	1996	Comprehensive Educational Improvement and Financing Act of 1996	Bill		n/a
	1997	Special Master's Report; Abbott v. Burke	Bill		X
	1998	Abbott v. Burke	Court		X
	2000	Abbott v. Burke	Court		
	2008	The School Funding Reform Act of 2008	Bill		n/a
New Mexico	1998	Zuni School District v. State		X	
	1999	Zuni School District v. State	Court		
	2001	Deficiencies Corrections Program; Public School Capital Outlay Act	Bill		n/a
New York	2003	Campaign for Fiscal Equity, Inc. v. State	Court	X	X
	2006	Campaign for Fiscal Equity, Inc. v. State	Court	X	
	2007	Education Budget and Reform Act	Bill		n/a
North Carolina	**1997**	Leandro v. State	Court	X	
	2004	Hoke County Board of Education v. State	Court	X	X
North Dakota	**2007**	SB 2200	Bill		n/a
Ohio	**1997**	DeRolph v. Ohio	Court	X	X
	2000	DeRolph v. Ohio; Increased school funding (see 93 Ohio St.3d 309)	Both	X	X
	2001	DeRolph v. Ohio			X
	2002	DeRolph v. Ohio	Court	X	X
Oregon	2009	Pendleton School District 16R v. State		X	n/a
South Carolina	2005	Abbeville County School District v. State		X	

(continued)

Appendix Table A1 (continued)

State	Year	Event	LRS (2016)	JJP (2016)	CE (2015)
Tennessee	1992	The Education Improvement Act	Bill		n/a
	1993	Tennessee Small School Systems v. McWherter	Court	X	X
	1995	Tennessee Small School Systems v. McWherter	Court	X	X
	2002	Tennessee Small School Systems v. McWherter	Court	X	X
Texas	1991	Edgewood Independent School District v. Kirby	Court	X	X
	1992	Carrolton-Farmers Branch ISD v. Edgewood Independent School District	Court	X	X
	1993	Senate Bill 7	Bill		n/a
	2004	West Orange-Cove ISD v. Nelson		X	
	2005	West Orange-Cove Consolidated ISD v. Neeley			X
Vermont	1997	Brigham v. State	Court	X	X
	2003	Revisions to Act 68; H.480	Bill		n/a
Washington	1991	Seattle II		X	
	2007	Federal Way School District v. State		X	
	2010	McCleary v. State	Court	n/a	n/a
West Virginia	**1995**	Tomblin v. Gainer	Court	X	
Wyoming	1995	Campbell County School District v. State	Court	X	X
	1997	The Wyoming Comprehensive Assessment System; The Education Resource Block Grant Model	Bill		n/a
	2001	Campbell II; Recalibration of the MAP model	Bill	X	n/a

Notes: Table lists all events included in any of the Lafortune-Rothstein-Schanzenbach (2016); Jackson-Johnson-Persico (2016); or Corcoran-Evans (2015) event lists, from 1990 onward. Xs indicate events that appear in the relevant event list; n/a indicates events that were out of scope for the relevant list, either because they were too recent or because it included only court cases and not legislative events. In Lafortune et al. column, events are classified as "court," "bill," or "both"; rows without an entry are not included in our event database but are included in one of the comparison samples. Bold years indicate the single event per state selected by our algorithm (see text). Appendix D discusses discrepancies between Lafortune et al. and Jackson et al. lists.

Table 2.10: Event studies for state budgets

	Per capita	Per pupil
Tax revenues:		
Total revenues	235	2,736
	(258)	(2,044)
Expenditures:		
General expenditures	290**	2,536*
	(138)	(1,505)
Education expenditures	114	1,029
	(70)	(643)
General expenditures (less education)	176**	1,508
	(90)	(977)
Health + welfare expenditures	73	514
	(49)	(457)
General expenditures (less education, health, welfare)	103	993
	(77)	(700)

Notes: Table shows estimates from the one-parameter event study specification (equation (2.1)) for state budgetary aggregates. State and year fixed effects are included. Standard errors are clustered at the state level.

Table 2.11: Comparison to Card-Payne

	State revenues (per capita)			Total revenues (per capita)		
	1977-1992 (CP)	1990-2012 (LRS)		1977-1992 (CP)	1990-2012 (LRS)	
	Long diff	Long diff	Event study	Long diff	Long diff	Event study
Court Ruling:						
Upheld	-0.81			0.20		
	(0.67)			(0.52)		
Unconstitutional	-1.89***			-1.10**		
	(0.62)			(0.48)		
Selected Events:						
Post Event		-2.06	-2.25**		-2.44	-1.61
		(2.24)	(0.89)		(4.73)	(2.38)

Notes: This table shows results using slopes from a regression of per capita state or total funding on district mean household income (note: district mean income here is in *levels*, not logs). Columns 1 and 4 are from table 4 of Card and Payne (2002) and show the long difference from 1977-1992 in the level-level slope coefficient. In columns 2 and 5, we replicate the Card and Payne specification using data from 1990 and 2012. Columns 3 and 6 show estimated effects from the one parameter event study specification (equation (2.1)) where level-level per capita slope coefficients are the dependent variables. Standard errors are clustered at the state level.

Table 2.12: Event study for log income, race, free lunch

(a) Income gradients

	Log mean income		Minority share		Free lunch share	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Event * Yrs Elapsed	-0.0010	0.0008	0.0021	0.0017	0.0058	0.0089
	(0.0029)	(0.0040)	(0.0013)	(0.0015)	(0.0064)	(0.0071)
Trend		-0.0026		0.0008		-0.0023
		(0.0042)		(0.0008)		(0.0031)
Post Event		0.0193		-0.0042		-0.0247
		(0.0368)		(0.0051)		(0.0293)
Observations	147	147	1046	1046	958	958
p(post-event=post-event*trend=0)	0.72	0.87	0.10	0.51	0.37	0.42
State FEs	X	X	X	X	X	X
Yr FEs	X	X	X	X	X	X

(b) Q1-Q5 difference

	Log mean income		Minority share		Free lunch share	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Event * Yrs Elapsed	-0.0017	-0.0008	-0.0012	-0.0016	-0.0035*	-0.0051**
	(0.0029)	(0.0035)	(0.0018)	(0.0020)	(0.0019)	(0.0024)
Trend		-0.0004		0.0003		0.0018
		(0.0035)		(0.0016)		(0.0021)
Post Event		-0.0073		0.0034		-0.0050
		(0.0290)		(0.0085)		(0.0154)
Observations	145	145	1045	1045	962	962
p(post-event=post-event*trend=0)	0.55	0.95	0.49	0.72	0.08	0.07
State FEs	X	X	X	X	X	X
Yr FEs	X	X	X	X	X	X

Notes: Table presents event study specifications where the dependent variable is the slope of the indicated demographic characteristic with respect to the district's 1990 log mean household income (panel A) or the gap between the average for districts in the bottom and top quintiles of the 1990 income distribution (panel B). Minority share and free lunch share are available annually from the Common Core of Data (though missing in some states and some years); log mean income is available from the Census in 1990 and 2000 and from the American Community Survey in 2007-11 (coded as 2011). Standard errors are clustered at the state level.

Table 2.13: Stratification of race, FRL, & achievement, by quintile

	Q1	Q2	Q3	Q4	Q5
Black	0.24	0.24	0.24	0.17	0.11
Black/Hispanic	0.24	0.23	0.24	0.17	0.11
White	0.20	0.20	0.18	0.20	0.22
Free/reduced-price lunch	0.32	0.23	0.20	0.15	0.09
25th pctl or below (NAEP)	0.27	0.21	0.22	0.17	0.13
75th pctl or above (NAEP)	0.14	0.15	0.17	0.22	0.32

Note: Table shows fraction of students of various groups in districts in various quintiles of the state's district income distribution. Each row sums to 1. Racial and free lunch shares are computed using CCD district-level data for the year 1994. The distribution of high- and low-achieving students is based on the 2003 NAEP data, which is the first year of comprehensive data for all grades and subjects.

Table 2.14: Event studies for district-mean resource gaps by race, FRL, & achievement

	Black/White		Free Lunch		25th/75th Pctl (NAEP)	
	St. Rev	Tot. Rev	St. Rev	Tot. Rev	St. Rev	Tot. Rev
Post Event	196	195	-32	-33	143	193*
	(160)	(164)	(193)	(219)	(141)	(101)
Observations	1047	1047	938	938	1509	1509
State FEs	X	X	X	X	X	X
Yr FEs	X	X	X	X		
Sub-gr-yr FEs					X	X

Note: In columns 1 and 2, the dependent variable in event study specifications is the average per-pupil revenue in the district attended by the average black student, less that in the district attended by the average white student in the same state. In columns 3 and 4, analogous revenue gaps are constructed for free/reduced-price lunch and non-free/reduced-price lunch students. In columns 5 and 6, analogous revenue gaps are constructed for students scoring at or below the 25th percentile in the NAEP, and students scoring at or above the 75th percentile in the NAEP. The *Post Event* coefficient shows the estimated event effect from parametric event study model without controlling for prior trends. State and year fixed effects are included in columns 1-4. State and grade-subject-year fixed effects are included in columns 5 and 6. Standard errors are clustered at the state level.

Table 2.15: Impacts of student sorting on student achievement results

	Q1-Q5 difference	Slope
Baseline Estimates	0.008**	-0.010***
	(0.004)	(0.003)
Decomposition 1: Common covariates		
Predicted score	0.003	-0.003
	(0.004)	(0.004)
Residual score	0.005**	-0.007**
	(0.002)	(0.003)
Decomposition 2: Richer covariates		
Predicted score	0.004	-0.004
	(0.004)	(0.003)
Residual score	0.004*	-0.006***
	(0.002)	(0.002)

Notes: First row repeats estimates from Table 2.5, columns 2 and 5. In subsequent rows, dependent variables are modified. We estimate student-level regressions of NAEP scores on student demographic characteristics, with year fixed effects, then compute predicted and residual test scores. We compute separate slopes with respect to district income and quintile gaps for the predicted and residual test scores, and estimate separate event study regressions for each. In decomposition 1, student demographic characteristics are race/ethnicity and gender, along with school means (in the NAEP sample) of each. Decomposition 2 adds indicators for students whose parent is a college graduate and for free or reduced-price lunch receipt, along with indicators for NAEP samples where these variables are unavailable and school means of each. Standard errors are clustered at the state level.

Table 2.16: Multiple events robustness: Monte Carlo simulations

	First event	All events (stacked)
DGP 1: Constant event effect		
Post coefficient	0.789	0.577
DGP 2: Only one event		
Post coefficient	0.788	0.577

Notes: Table reports estimates of average post-event "jump" coefficient from Monte Carlo simulations using the empirical distribution of event dates, in which some states had multiple school finance reform events. Column 1 shows estimates from event study models estimated using only the first event in a state. Column 2 shows estimates using all events in a state, stacking panels and adding a joint state-panel copy fixed effect (see Table 2.7 column 2). In both columns, estimates are from parametric event study models with a single coefficient (equation (2.1)). Row 1 shows estimates from a simulated DGP where every event in a state has a constant effect. Row 2 shows estimates from a DGP where only one event (randomly chosen within state) has an effect. In both DGPs the total event effect over all events within a state is equal to 1. All DGPs include i.i.d. error terms and are simulated 5000 times.

Chapter 3

The Heterogenous Effects of Advanced Math Tracking: Evidence from North Carolina

3.1 Introduction

Numerous studies have demonstrated the importance of strong math skills on individual and societal outcomes (e.g. Hanushek and Woessmann (2015); Ritchie and Bates (2013); Weinberger (2014)). Considerable attention has been devoted to improving the math performance of US students, in particular because American students fare poorly in international comparisons of math skills (Gurria, 2016) and the fraction that go on to purse majors in science, technology, engineering, and math (STEM) (National Science Board, 2010). Still, both educational practitioners and policymakers have often struggled to devise policies that improve students' math preparation and eventual matriculation into STEM fields.

Policies to improve the rigor of mathematical education in U.S. schools often focus on the mathematics course sequences students are exposed to; of particular consequence is the separation of students into distinct math curricular tracks in secondary schools. In nearly all U.S. middle and high schools, assignment of students into mathematics courses relies on tracking: the practice of sorting students into different classrooms based on prior achievement, career intentions, or other objectives. Unlike in many other countries, in which students are tracked into different schools at different ages based on perceived ability or career aspirations, tracking in U.S. public schools is done primarily *within-school*, with students grouped by ability into courses that differ in rigor and/or curriculum.[1]

For most students, the most important divergence in the math curriculum occurs in 8th grade: the highest achieving students are placed in Algebra I, while the rest take another year of math before encountering Algebra I in 9th grade (or sometimes later on in high school). Students who complete Algebra I in 8th grade will generally be on track to complete

[1]See Betts (2011) for an overview of the literature on both within- and between-school tracking regimes.

introductory calculus by the end of high school. For these reasons, 8th grade Algebra is sometimes referred to as a gateway class for selective higher education (e.g. Adelman (2006); Loveless (2013)). Nationally, partitions in the 8th grade math curriculum between students of different ability levels have become ubiquitous: pre-high school exposure to Algebra and other advanced math courses[2] has gone from 16% in 1990 to 47% in 2011. During the same two-decade span, the proportion of students in a basic math course in 8th grade fell from 81% to 48% (Loveless, 2013).

Many individual states and districts have emphasized policies to expand pre-high school exposure to Algebra, with some embracing policies to shift all students into early Algebra, often referred to as the "Algebra for all movement". However, other districts have gone against this trend (e.g. the "detracking" movement of the 1980s; see Burris and Garrity (2008)), often due to concerns over underrepresentation of minority and disadvantaged students in advanced tracks, as well as potential negative effects in lower tracks due to the removal of high-ability peers.[3]

In this paper I contribute new evidence on the heterogenous impacts of early Algebra enrollment for students at different points in the mathematics achievement distribution. Using administrative data from all public K-12 schools in North Carolina, I leverage discontinuities in the propensity to be tracked into 8th grade Algebra around prior year state exam proficiency thresholds to provide credible evidence on the causal impacts of these tracking decisions on student test scores, course enrollments, college entrance exams, and post-high school intentions. Students who score higher on their 7th grade math exams are more likely to enroll in Algebra in 8th grade, and there is evidence of small, precisely estimated discontinuities in this relationship near the state proficiency thresholds. Students crossing the threshold from "level III" to "level IV" (near the 80th percentile in the state grade 7 math distribution) are 5 percentage points (10 percent) more likely to enroll in Algebra in 8th grade rather than later in high school. Regression discontinuity (RD) estimates around this proficiency cutoff indicate that early Algebra is associated with large increases in the mathematics and advanced course-taking, college entrance exam scores, and intentions to major in a STEM field post-graduation. Effects on 4-year college intentions are large, but imprecisely estimated.

For students around the 30th percentile in 7th grade math, there is a smaller 3.5 percentage point (50 percent) discontinuity in the propensity to enroll in 8th grade Algebra crossing the threshold from "level II" to "level III". Notably, this threshold is the cutoff for proficiency in 7th grade math, which is often widely reported and relevant for both state and national educational accountability systems. However, for these students, RD estimates indicate that early Algebra acceleration is associated with losses, rather than gains. Initial

[2]Data are from the NAEP 8th grade math assessment. "Advanced math" courses refer to Algebra I, Algebra II, Geometry, and 2-year Algebra courses. See Loveless (2013) for further detail.

[3]For example, in 2015 San Francisco Unified School District began requiring that all 8th graders enroll in the same math course and delay Algebra I until 9th grade, partially due to equity concerns over the separation students of different achievement or preparation. (`https://ww2.kqed.org/news/2015/07/22/san-francisco-middle-schools-no-longer-teaching-Algebra-1/`)

8th grade test score effects are negative, with estimates of less than −1 SDs for the effect of 8th grade Algebra acceleration on sdtandardized 8th grade math and Algebra I exams. After an initial deleterious effect on achievement, I find no evidence of end-of-high school gains in college entrance exams, mathematics course completion, and 4 year and/or STEM major intentions post-graduation. The heterogeneous RD estimates suggest that early Algebra tracking has vastly different effects based on a student's initial mathematics preparation and achievement. Models estimated using observational variation in the timing of Algebra enrollment closely match RD estimates when the effect of early Algebra tracking is allowed to vary by prior achievement quantiles, suggesting that: (1) estimates of average effects mask important heterogeneity in the effects of secondary school mathematics tracking by students' initial ability and preparation, and (2) selection biases are likely small once effects are allowed to vary by prior ability and controls for prior ability are included.

Importantly, these findings can reconcile differences in the prior literature on the effects of Algebra tracking. To this point, the empirical literature on the student-level effects of early Algebra has been mixed and inconclusive. Much of this literature has focused on either targeted district expansions or average (equilibrium) effects nationally that typically consider students at very different margins in the mathematics achievement distribution. Most studies of district-level expansions consider policies that encourage Algebra enrollment among more academically marginal students, and tend to find negative or null effects on test score outcomes, with only modest effects on later advanced math track persistence (Domina, Penner, et al. (2012), Domina, McEachin, et al. (2015), Clotfelter et al. (2015), Dougherty et al. (2017), McEachin et al. (2017)). On the other hand, studies using observational designs on nationally representative datasets tend to find more positive or null effects (Stein et al. (2011); Aughinbaugh (2012)), where the average student who enrolls in early Algebra is of much higher prior ability. Using multiple discontinuities around prior year test score proficiency levels I am able to show that high ability students see large benefits from early Algebra acceleration, whereas low ability students are adversely affected. I then show that simple models relying on observational variation in Algebra tracking can replicate this pattern of heterogeneity and appear to have minimal selection bias, once the effect of early Algebra tracking is allowed to vary with prior achievement. Taken together, these results suggest that much of the conflicting evidence in the literature on Algebra and secondary school math tracking can be attributed to differences in the mathematics preparation of the students under consideration.

This paper also provides new evidence linking secondary school Algebra tracking decisions to scores on college entrance exams (ACT and SAT) and post-high school college enrollment and major intentions. Prior studies of Algebra tracking using quasi-experimental variation have been unable to examine outcomes to the end of high school or consider effects on post-high school intentions and college readiness. Tracking decisions made in 8th grade affect course placements, rigor, peer quality, and teacher quality throughout high school, and thus one would expect that the effects of such a course diversion would accumulate through the end of high school. I am able to follow several cohorts of students across nearly all public school districts[4] in North Carolina through the end of high school, and find that for high-ability students, 8th grade Algebra acceleration has large effects on STEM major intentions and ACT and SAT exam scores.

The findings in this paper have important implications for mathematics tracking policy. Targeted efforts to expand access to the advanced math curriculum among high-achieving students could lead to large improvements in mathematics skills and college preparedness by the end of high school. This is not an insignificant margin: only about half of students in the upper quartile of the 7th grade math distribution are accelerated into Algebra in 8th grade. Conditional on prior test scores, these discrepancies in advanced math tracking among high-achieving students shows little correlation with race or student income levels (based on free/reduced-price lunch eligibility), but there exist large gaps in propensity to enroll in 8th grade Algebra and advanced high school courses based on parental education levels. Students with more educated parents are considerably more likely to be selected into advanced math tracks in 8th grade and in high school, even among high-achieving students in top of the test score distribution. On the other hand, early placement into Algebra and advanced math courses could harm, rather than help, low-achieving and marginally prepared students. Though few marginally proficient students enroll in advanced 8th grade Algebra (less than 5% in North Carolina), many district-level expansion policies are targeted specifically at this margin.[5] Efforts to increase mathematics skills among this subpopulation of students may be better better targeted towards additional remedial, or "double-dose" mathematics curriculum (Cortes and Goodman (2014); Taylor (2014)).

There are three important caveats worth highlighting. First, in this paper I do not consider the general equilibrium effects of changes in tracking policies. Estimates are specific to marginal students, who are nudged into different 8th grade math tracks under an existing educational policy regime. Targeted policies to expand access among certain students could also have spillover effects on students who are not tracked into advanced courses, perhaps due to changing peer and teacher quality in the both the tracked and non-tracked classrooms.[6] Second, estimates in this paper are specific to cohorts that attended public schools in North Carolina as a part of 7th grade cohorts in 2006 to 2011.[7] 8th grade Algebra participation expanded moderately during this time, especially in North Carolina, where participation had been slightly lower than in other states. Third, the evidence in this paper provides causal estimates of the impact of curricular acceleration in 8th grade, holding constant earlier educational inputs. Policies that aim to accelerate mathematics curriculum, particularly for

[4]A small number of students are excluded from main analyses, mainly those from small schools and districts with only enough student enrollment for one 8th grade classroom. See Section 3.3 for further detail on sample constriction.

[5]For example, the recent targeted expansion in Wake County, NC (Dougherty et al., 2017) set its cutoff around the 20th percentile in the 6th grade math distribution.

[6]Although changing peer and teacher composition between tracks could lead to aggregate effects that differ from those implied by the estimates in this paper, existing evidence suggests that these spillovers are likely to be small (Card and Giuliano, 2016).

[7]Expected/on-time graduation for these cohorts was from 2011 to 2016.

low-achieving students, may well have more positive effects with additional preparation or other alterations in the earlier curriculum (e.g. Dougherty et al. (2017)).

The rest of the paper proceeds as follows. In Section 3.2 I provide brief context on the potential implications 8th grade Algebra acceleration, and review the prior literature examining Algebra tracking and student outcomes. In Section 3.3 I discuss the data sources, sample selection, and test score proficiency levels in North Carolina. Section 3.4 outlines the empirical specification and Regression Discontinuity assumptions. RD estimates are reported in Section 3.5, and Section 3.6 considers the heterogeneity in these estimates, comparing RD LATEs to OLS estimates from more ubiquitous selection on observables models. Finally, in Section 3.7, I discuss the implications of these results for tracking policy decisions and future areas of research.

3.2 Context and Background on 8th Grade Algebra

Early acceleration into Algebra in 8th grade may affect student learning and later course-taking decisions for a number of reasons. For students with a sufficient mastery of prior, lower-order mathematical skills, early access to Algebra allows for more instructional time for higher order topics. In the typical mathematics course sequence in US public high schools, only students who have completed Algebra prior to enrolling in high school are on track to reach calculus by the end of high school (taking only one math course per year). To the extent that skills in Algebra, calculus, and other higher-order mathematics topics are important for later college and labor market outcomes, early Algebra acceleration could be highly beneficial. In North Carolina public schools, students who take Algebra in 8th grade rather than in high school are on average more likely to enroll in Calculus (35% vs 4%), score higher on average on ACT math (23.5 vs 17.7 points) and more likely to intend to enroll in a 4 year college post-graduation (73% vs 34%).[8] Yet students who enroll in 8th grade Algebra are positively selected: the average 8th grade Algebra student scores near the 80th percentile in 7th grade math, while the average student who is not tracked into 8th grade Algebra scores near the 40th percentile in 7th grade math. Figure 3.1 reports Calculus enrollment rates and 4-year college intentions for 8th-grade Algebra takers and non-takers, conditional on prior, grade 7, math test scores. Throughout most of the prior test score distribution, the gap in end of high school outcomes between takers and non-takers is large and remarkably similar. This is suggestive of a positive causal effect, although endogeneity in the selection into 8th grade Algebra may still bias this comparison.

Indeed, prior research using observational variation in Algebra timing in nationally representative survey datasets tends to show strong positive associations between pre-high school exposure to Algebra and academic outcomes (Smith (1996); Gamoran and Hannigan (2000); Stein et al. (2011)) and, more generally, between the rigor of secondary school mathematics curriculum and post-secondary outcomes (Aughinbaugh, 2012). Moreover, besides direct curricular benefits, the timing of Algebra course-taking has important effects on the peers

[8]See Section 3.3 for a description of the data and sample restrictions.

and teachers a student is exposed to both in 8th grade, and throughout high school. Students who are tracked into early Algebra share classrooms with higher ability peers, and Smith (1996) argues that early access to Algebra may also "socialize" students into taking further mathematics classes in high school.

On the other hand, if students are insufficiently prepared for a more rigorous math curriculum in 8th grade, such an acceleration may not be beneficial, and could even be harmful. Positive effects documented in observational studies likely suffer from selection bias, as students who are tracked into early Algebra are on average much higher-achieving and from more socio-economically advantaged backgrounds. Clotfelter et al. (2015) address these selection concerns using variation induced by district-level expansions in 8th grade Algebra access at two large North Carolina school districts. They find large negative effects on performance in Algebra and subsequent math courses, particularly among low-achieving students. While the effects of such large expansions may differ in the short run due to transition costs and may also confound differences in teacher quality, the authors only attribute a small fraction of the negative effects they document to these dimensions. Dougherty et al. (2017) study a similar, but more explicitly *targeted*, district expansion in 8th grade Algebra that affected low-achieving students near the 20th percentile in prior math ability in Wake County, North Carolina. Using an RD design around a test score index cutoff they find evidence of more positive effects, with increased mathematics course-taking and no significant effects on initial test scores. They also find large positive effects on a college readiness exam in 10th grade, and eventual college intentions among students who took the college readiness exam. Similarly, Domina, Penner, et al. (2012) use policy variation in one California school district and find that the expansion was associated with lower academic performance among lower-level students, although the policy change increased the odds that these students enrolled in higher-level math courses.

Clotfelter et al. (2012) argue that much of this discrepancy between the generally positive effects found in observational studies, and the more mixed effects documented in studies of district expansions is due to selection bias: "Once this selection bias is eliminated, the remaining causal effect of accelerating the conventional first course of Algebra into earlier grades, in the absence of other changes in the math curriculum, is for most students decidedly harmful." Yet the existing literature has provided little direct evidence as to the causal effect of the Algebra tracking decision on high-achieving students, who make up the majority of students who are placed into advanced the advanced math sequence in 8th grade.[9] In the paper most similar to this one, McEachin et al. (2017) use an RD design to provide evidence of the effect of 8th grade Algebra at schools with different test score cutoffs, but focus primarily on heterogeneity in the effects of early Algebra by school-level factors, rather than student-level preparedness or ability. The diversion of 8th graders into different math tracks represents the single largest tracking decision in most U.S. public schools, and has implications for the rest of the high school mathematics curriculum. Given the potential

[9]The median 8th grade Algebra student in North Carolina was near the 80th percentile in the 7th grade math distribution from 2006-2011, the years under consideration in this study. See Figure 3.2.

for negative spillover effects and increased educational inequality resulting from tracking decisions, estimates of the total, end-of-high school effect of 8th grade Algebra tracking for students with different levels of prior mathematics skills is crucial for the design of policies that aim to more optimally track students and/or raise mathematics skills among students more generally.

3.3 Data

North Carolina Public School Records

To examine student track assignment and later student outcomes, I use administrative student records from the North Carolina Education Research Data Center. These records contain data for all K-12 students in North Carolina public schools (including public charter schools). The data span the period from 1997-2016 and include transcripts, test scores, and some demographic information for all public school students in North Carolina. Students are linked across both schools and districts, and over time, as long as they are in a North Carolina public school. Students who move out of state or transfer to a private school are not followed in the data. All students in grades 3-8 are tested annually in both math and reading; students in grades 5 and 8 are additionally tested in science. Demographic information is limited to gender, race/ethnicity, free lunch status, disability/exceptionality status, and parental education. These data are available for all years with the exception of the parental education data, which was only collected until 2006. Given the panel nature of the data, this means that I have information on parental education for all students in more recent cohorts who had at least completed grade 3 in a North Carolina public school prior to the 2007 school year.

In North Carolina, there are no comprehensive high school exams. Instead, there are some standardized "End of Course" (EOC) exams on specific subjects in place of final exams. Students who in enroll in Algebra I in any grade - whether in middle or high school - are required to take the state EOC exam for Algebra I. I assign students as having taken Algebra in 8th grade if they have a test score on the Algebra I EOC exam in their 8th grade year. Students who drop out of the class in the first 20 days are exempted from the exam; after this point, the state mandates that every student maintains enrollment in the class and participates in the EOC exam upon completion of the course. Algebra 1 test scores are available from 1997 until 2013. I only use data on the first Algebra I exam taken for each student.

Comprehensive high school transcript data are available in this dataset beginning in 2006, and provide information on course enrollment and grades for all students in grades 9-12. The transcript data also include course classification codes that indicate the level of the course (honors, advanced, AP, etc) in addition to the exact subject. Beginning in the 2012-2013 school year, every 11th grade public school student in North Carolina was required to take the ACT college readiness assessment. Data on ACT composite scores, as well as subject level subscores are available for all 11th grade students from 2013-2015. Similar data on the SAT scores of students are also available beginning in 2009, although only about one-third of North Carolina high school students take the SAT exam. Students who take the SAT also report their intended college major, which I use to determine whether or not a student intends to pursue a STEM major in postsecondary school. Upon graduation, every North Carolina high school student also fills out an exit survey, which includes a description of post-high school college intentions.

Test Score Proficiency Levels

North Carolina "End of Grade" (EOG) exams are administered in both math and reading for grades 3-8, and for science in grades 5 and 8. Scores on these tests are reported in two forms: (1) A "scale" score, which is on a numerical scale with a range of roughly 50 points depending on the exam and year (2) A proficiency test score "level", which ranges from 1 to 4 for years prior to 2013 (in 2013 the state added a fifth proficiency level). the specifics of each score level vary by exam type, year, and grade, but in general, level I corresponds to a "limited command" of the material, level II corresponds to a "partial command", level III corresponds to "solid command", and level IV corresponds to "superior command".[10]

The numerical scale of standardized test scores is not consistent across all years. For the years 2006-2012, the 7th grade math scale scores are on a consistently normed scale from 332 to 383 points. Prior years' scores are scaled differently, and the range and scaling within each proficiency level is neither consistent nor directly comparable with later years. Similarly, in 2013, the exam structure and scoring were changed again and a fifth proficiency level was added, preventing like comparisons with earlier exam years and proficiency cutoffs. 2006 also served as a norming year for the new exam, in which baseline scores were constructed from which later years' scores could be compared. Exam scores in this year were not reported until late in the fall of the following school year, meaning that 2006 test scores were unable to be used in course enrollment decisions in the first semester of the 2006-2007 school year. For these reasons, I include only those students who took the 7th grade math exam for the first time between 2007 and 2012.[11]

For each exam form, a student's score is determined by converting the raw number of correct responses to a developmental scale score that enables direct comparisons across exam forms and years. The distribution of scaled test scores is shown in Figure 3.2. As can be seen in the figure, there are discrete mass points in the test score distribution. Due to rounding in the conversion from raw to scaled scores, two raw scores in the middle of the distribution sometimes correspond to the same scale score on certain exam forms.[12] The aggregation of scores exam forms therefore generates the observed spikes in the scale score distribution.

[10] See Appendix Figure 3.18 for a sample North Carolina EOG student exam report.

[11] The fact that 2006 scores were only reported after course selection had taken place in the following school year generates a useful first stage placebo test, to be discussed in further detail in Section 3.4.

[12] See Appendix Figure 3.17 for an example of the rounding from NC test score technical documentation. Unfortunately, the exam form number is not reported in the student test score data.

Importantly, these are artifacts of rounding error, rather than an indication of systematic sorting of students to different test score thresholds. Concerns over the validity of the RD design with a non-smooth test score distribution will be discussed later in Section 3.4

Sample Selection

The primary sample begins with students who took the 7th grade EOG math exam in the five-year period from 2007-2011. [13] Students with missing test scores in grade 7 are excluded. Prior year test scores are available for the large majority of these students, as long as they had been enrolled in a NC public school and had a valid exam score. Where scores from grades 3 to 6 are used as controls, these scores are standardized within test year, grade, and subject to have a mean 0 and standard deviation of 1. Students who have missing test scores in earlier years are imputed the mean value 0 and assigned a dummy equal to one for missing test scores in that subject-grade combination. Likewise, students with missing demographic data are not excluded, but where demographic controls are used these students are assigned to zero for the relevant category and assigned a dummy equal to one for having missing demographic information.

A small number of students are excluded from the primary estimation sample. First, students who enroll in Algebra prior to 7th grade (only about 2%) are dropped. Next, students with missing or implausible/incorrect school codes (e.g. school codes with only 1 student per year) are excluded. Only schools with enough students to have two classrooms per year (schools with 40 or more students per 8th grade cohort-year on average). Roughly 4.5% of students are at schools this small. Finally, less than 0.5% of students are in schools that enroll no 8th graders in Algebra, and these students are excluded from the main sample. With these restrictions, I am left with a sample of 430,127 students spanning five 7th grade cohorts in 532 North Carolina public middle schools.

Summary statistics for the estimation sample are presented in Table 3.1. 29% of students during this period enroll in Algebra in 8th grade. Of the remaining students who do not take Algebra until high school, most enroll in 9th grade, with less than 2.5% of students taking Algebra in 10th grade or later. Mean standardized test scores in 6th and 7th grades are slightly greater than zero, indicative of a slight positive selection of students into the sample relative to the entire NC population, due to the aforementioned sample restrictions. Students in the sample are predominately white, although 27% are black, and an additional 9% are listed as Hispanic. 46% of students have parents with a high school diploma or less; while about 29% of students have at least one parent who graduated college. The average 8th grade cohort size is 210 students. In 9th grade, roughly 23% of students take an advanced math class of some sort (e.g. Honors, AC, AP), and only 13% of North Carolina students reach calculus by the end of high school. Just under half of high school graduates report the

[13] Recall, the 2006 exam is excluded due to being a norming year in which scores were reported only later in the following year; the 2012 exam year is excluded as this cohort cannot be followed until the end of high school (on-time graduation for students who took the 2012 7th grade EOG exams was in 2017, and the final year in the data sample is 2016).

intention to attend a 4-year college post-high school; about one-quarter of SAT test takers report intending to pursue a STEM major once in college.

3.4 Empirical Strategy

The prior literature has identified two potential issues when attempting to estimate the causal effect of 8th grade Algebra tracking on both short and long-run student outcomes. The first and most notable is selection bias. Students are not randomly assigned into Algebra in 8th grade, and even conditional on prior test scores and available socio-economic background variables, students who are selected into higher level math may differ in unobserved characteristics that may affect outcomes of interest. A priori, one would expect that, on average, such selection issues introduce a positive bias in estimates that rely on observational variation. Second, and more subtle, is the issue of prior mathematics ability and the potential for heterogeneity in the true effect. More able or more high-achieving students are likely better prepared for the demands of an accelerated math curriculum and the more rigorous mathematics courses that follow in high school. One would expect that the effects of 8th grade Algebra acceleration, both initially and in the longer-run, could be very different for students with different baseline levels of mathematics skill.

In this paper I will attempt to address both of these concerns using a regression discontinuity design around prior-year test score proficiency cutoffs. As discussed in Section 3.3, North Carolina EOG test scores are reported as a scaled score, as well as a proficiency level ranging from I to IV. To the extent that test score proficiency levels provide additional information about student achievement or potential for success in future advanced classes, there may be discontinuities in the the propensity to enroll in 8th grade Algebra around these thresholds. These discontinuities provide opportunities for credible identification of causal effects at different points in the student ability distribution.

Figure 3.3 shows the mean proportion enrolled in Algebra in 8th grade, by prior-year, 7th grade math test scores, for students who score close to the level IV (panel (a)) and level III (panel (b)) thresholds.[14] Dashed vertical lines indicate proficiency threshold cutoffs. Panel (a) zooms into a 10-point window around the level IV proficiency threshold, which corresponds to near the 80th percentile in the 7th grade math test score distribution. A simple linear trend is fit separately on both sides of the cutoff, indicated by the solid line. The figure seems to suggest a small, but notable 5 percentage point discontinuity in the propensity to enroll in 8th grade Algebra around the level IV proficiency threshold.[15] Panel (b) displays the analogous means by prior test score and estimated linear fit for the 10-point window

[14]See Panel (a) of Appendix Figure 3.7 for the mean 8th grade algebra enrollment by 7th grade test scores over the entire distribution of test scores.

[15]The strong upward slope in the relationship between 8th grade Algebra enrollment and 7th grade test scores makes the (small) discontinuities more difficult to see visually. Appendix Figure 3.16 reports analogous figures where the propensity to enroll in 8th grade Algebra is de-trended with respect to prior-year test scores; discontinuities are more visually apparent in these more zoomed-in figures.

around both sides of the level III proficiency threshold. Around the level III proficiency threshold (roughly the 30th percentile in 7th grade math test score distribution), there is a smaller, approximately 3.5 percentage point discontinuity in 8th grade Algebra enrollment. Note that in Panel (b), the sample is restricted to only include the half of students at schools with above median 8th grade Algebra enrollment. At schools with below-median enrollment of students into Algebra I the proportion of 8th grade students enrolled in Algebra I is nearly zero. For this reason all analyses around the level III proficiency threshold will include only those students at schools with above median total enrollment in 8th grade Algebra.[16]

Discontinuities in the probability of enrollment in 8th grade Algebra around these proficiency levels imply that students and/or schools use information from these proficiency levels *independently from the underlying scale scores* to determine enrollment in 8th grade Algebra. This could be due to explicit school policies, differences in course recommendations by school counselors based on test score levels, or perhaps differences in student and parent beliefs about a students ability inferred from test score levels. Anecdotal evidence suggests that many districts have proficiency level requirements for enrollment in certain advanced classes and give recommendations to students and parents based on the proficiency levels the student achieved on prior year exams.[17] This is not necessarily the case for *all* schools and districts in the state, as assignment rules vary significantly by district, school, and even year. Some schools have no formal requirements, while others may rely on earlier test scores, especially those that begin math curriculum tracking before grade 8. For example, the Wake County School District assigns students to Algebra I using a predicted probability of success derived from multiple years of prior grades and test scores before grade 7 (Dougherty et al., 2017).[18] Rather than attempting to identify a specific subset of districts and/or schools that use test score proficiency levels in Algebra assignment, I remain agnostic about the specific policies in any given school and instead identify the average effect over all schools and districts that meet the sampling criteria outlined in Section 3.3.

Scoring on either side of the threshold is a matter of getting one more or one fewer question correct on the 7th grade math exam. For students sufficiently close to the test score level threshold, achievement of a given proficiency level is effectively randomly assigned. Treating these proficiency levels as assignment rules for 8th grade Algebra with imperfect

[16]Panel (b) of Appendix Figure 3.7 compares 8th grade Algebra propensity near the level III proficiency thresholds for schools with above and below median total 8th grade Algebra enrollment. As can be seen in the figure, schools below the median in Algebra enrollment enroll only about 1% of students near the level 3 threshold in 8th grade Algebra, with no discontinuous relationship, compared to 7% enrolled at schools with above median participation.

[17]Conversations with staff members from several school districts in North Carolina indicated that prior test scores and specifically prior proficiency levels are sometimes used in 8th grade math curriculum assignment. These are typically not strict cutoffs, as there are other assignment criteria besides prior year test scores. Notably, some districts made explicit mention of the importance of "parent advocates" in the assignment process.

[18]Reassuringly, in Wake County School District there exist no discontinuities in 8th grade Algebra enrollment around prior proficiency cutoffs, as expected given that their algorithm relies on the scale scores and not proficiency levels.

compliance motivates the use of a (fuzzy) regression discontinuity (RD) design. This design will identify the effect of early enrollment in Algebra for students around each proficiency level threshold, when there are discontinuities in the relationship between 7th grade math scores and 8th grade Algebra enrollment. For this design to yield unbiased causal estimates of the local average treatment effect (LATE), students must not be able to systematically sort or alter their test score to land on either side of the proficiency threshold. As will be further discussed in Section 3.4, exams are scored and scaled by statewide committee separately for each subject, grade, and year, yielding no scope for precise manipulation by any student. Furthermore, dedicated parents need not rely on high test scores to get their children to enroll in 8th grade Algebra: nearly half of students below the cutoff still enroll in Algebra I.[19]

First Stage Estimates

To estimate the potential discontinuities in Algebra assignment around prior-year proficiency thresholds I use local linear regressions that allow for a different slope with respect to test scores on both sides of the proficiency cutoffs:

$$A_i = \beta Z_i + \delta_0 \theta_i + \delta_1 \theta_i * Z_i + \epsilon_i \tag{3.1}$$

Here, i denotes student, and θ_i denotes a student's grade 7 math test score, in the reported scaled units. $Z_i = 1[\theta_i \geq p]$ is an indicator for crossing the relevant proficiency threshold, p. The dependent variable, A_i is an indicator for enrollment in Algebra in 8th grade. Estimates of equation (3.1) are shown in column (1) of Table 3.2, panels (a) and (b), which correspond to the first stage estimates displayed in Figure 3.3, respectively. Both regressions are estimated using data only within a small bandwidth of 5 scale score points on either side of the threshold, for which the linear approximation appears to fit the data quite well.[20] Estimates of equation (3.1) show that there is a marginally significant discontinuity of 5.0 (SE 2.7) percentage points at the level IV proficiency cutoff. Even at the lower level III threshold, where less than 10 percent track into Algebra in 8th grade, gaining an additional point and crossing into the 3rd proficiency level is associated with a statistically significant 3.4 (SE 1.5) percentage point greater propensity to enroll in Algebra in 8th grade.

These estimates pool data across cohorts and schools, and do not adjust for differences in student demographics or prior student achievement. Columns (2) and (3) of Table 3.2 add in fixed effects for cohort (α_c) and school (ϕ_s), respectively:

$$A_{ics} = \alpha_c + \phi_s + \beta Z_{ics} + \delta_0 \theta_{ics} + \delta_1 \theta_{ics} * Z_{is} + \epsilon_{ics} \tag{3.2}$$

[19]Within-school estimates of discontinuities across proficiency thresholds yield only a handful of small (mostly charter) schools in which state proficiency thresholds appear to be binding for course selection.

[20]For small bandwidths, estimated discontinuities are quite similar. However, nonlinearities in the overall distribution lead to increasing bias from using larger bandwidths with a linear fit. See Appendix Figure 3.13 for a comparison of first stage estimates by bandwidth.

Point estimates are largely unchanged across these different specifications at both proficiency thresholds, but standard errors decrease substantially with the addition of fixed effects for cohort and school. In particular, the addition of cohort fixed effects markedly improves the precision of the estimates of β. Why is this the case? Overall participation in 8th grade Algebra had been increasing both nationally and in North Carolina during this time period, meaning that the propensity for a given student to be tracked into early Algebra differed between different cohorts, even conditional on prior test scores. For this reason, allowing for a different intercept within each cohort reduces noise and improves the estimation of these discontinuities. In column (4) controls for prior test scores and for student-level demographic characteristics are added. Reassuringly, first stage estimates for both discontinuities are very robust to the inclusion of these controls.

Potential Threats to Identification

Estimates in Table 3.2 imply that students, parents, and/or school staff use the information in proficiency thresholds in the assignment process for Algebra in 8th grade. Importantly, the validity of the RD approach relies on the (untestable) assumption of local randomization around these thresholds (e.g. Lee and Lemieux 2010). Put differently, we must have that the “treatment” of early Algebra enrollment is as good as randomly assigned around the threshold in order for a causal interpretation of RD estimates. There a two notable potential threats to identification inherent to this particular setting: (1) systematic sorting of students around the thresholds, and (2) spurious discontinuities, perhaps due to functional form assumptions.

If students are able to manipulate test scores in such a manner as to systematically sort around the proficiency thresholds, the assumption of local randomization would be violated. Figure 3.2 plots the distribution of 7th grade test scores, and there are notable mass points in the middle of the distribution, between the level III and level IV proficiency thresholds. Importantly, these mass points are not directly to the right of either threshold, indicating that this bunching is not the result of systematic sorting to end up on the “right side” of the cutoffs; however, there is one notable mass point just *before* the level IV cutoff. As mentioned earlier, these mass points are are artifacts of rounding error in the scaling process in which raw test scores (the total number correct) are converted into the reported and student-normed scale.[21] The exact mappings are determined each year by a statewide committee as exams are scored, after they have been taken, and vary between exam forms and years. Thus, despite the non-smooth distribution in prior scaled test scores, any systematic sorting of students around any particular score threshold would be nearly impossible in practice. Moreover, student selection into advanced math courses is determined by more than test scores, and proficiency cutoffs are not binding in the assignment process: to the left of the level IV cutoff, 40% of students still enroll in 8th grade Algebra (Figure 3.3). To the extent that students, parents, or school staff want to influence assignment into advanced courses,

[21]See Appendix Figure 3.17.

precise test score manipulation would not only be essentially impossible, but it would be unnecessary.

Appendix Table 3.9 provides an additional test for sorting and violations of local randomization. Under local randomization, there should be no discontinuous differences in baseline, predetermined student characteristics across either threshold. In both panels of Table 3.9, estimates of β from equation (3.2) are shown where the dependent variables are predetermined student characteristics. In the first row, an index of the predicted probability of 8th grade Algebra enrollment is included; this measure is constructed from the predicted values of a regression of A_{ics} on all available student demographic and prior-test score characteristics. Reassuringly, the predicted index shows no significant difference across either threshold. While most individual characteristics show no significant discontinuities across the threshold, there are a small number that are, in fact, statistically significant. Still, in every instance these estimated discontinuities are of an incredibly small magnitude, and often of a different sign. While these differences do not appear to indicate any notable violations of local randomization, I nonetheless include controls for each of these characteristics in all baseline RD estimates.[22]

Other potential threats to the validity of the RD design stem from the inherent functional form assumptions. One might be concerned that the discontinuity in 8th grade Algebra enrollment around the level III / level IV proficiency threshold is spuriously driven by the spike in prior test score distribution right before the level IV cutoff, if the discrete change in true "ability" is not constant with each scale score point increment. However, if this were the case, one would expect to find similar discontinuities around the other mass points elsewhere in the prior score distribution. Appendix Figure 3.8 plots estimated discontinuities at each scale score within 5 points above or below the level III and level IV proficiency thresholds. As can be seen in the figures, there are no sign of discontinuities near these other mass points. Moreover, the estimated discontinuities for both thresholds are only large and significant at the actual proficiency thresholds; if the discontinuities were driven by bias from a misspecified functional form, we would expect other spurious discontinuities at other test scores, which we see no evidence of.[23]

Finally, Appendix Figure 3.9 provides additional evidence that the discontinuities are indeed non-spurious, and not artifacts of functional form or other assumptions. The figure plots estimated coefficients for the level III and level IV cutoff separately by year. In panel (a), for the level IV cutoff, the coefficients are all roughly 4-5 percentage points and do not vary significantly for all years after 2006. Notably, there is no discontinuity in 2006, as the point estimate is close to zero and insignificant. This is not a blip: as mentioned earlier,

[22] As would be expected given the estimates in Table 3.9, estimates are robust to the inclusion or exclusion of controls. See Appendix Tables 3.12, 3.13, 3.14, and 3.15, which report results from models estimated without controls.

[23] Estimates in Appendix Figure 3.8 are still statistically significant when the cutoff is erroneously set just above or below the true proficiency threshold. This is to be expected: misspecification of the true location of a discontinuity using a linear functional form will result in attenuation of the discontinuity, as the specified discontinuity point gets further from the "true" discontinuity.

state exams were updated in 2006 and the 2006 population served as a norming reference for students in all subsequent years. Exam scores were not released to students or schools until October, well after the beginning of the academic year and after course assignments had already been determined. Prior exam scores for this cohort were thus not known at the time of 8th grade math assignment; only in those exam years where scores were available prior to course assignment does there exist a discontinuity in 8th grade Algebra enrollment around the level IV proficiency threshold. The pattern is indeed similar for estimates near the level III cutoff in panel (b). Here, the estimate is again insignificant and near zero in 2006, and positive and similar in magnitude in all other years, with the exception of the 2009 grade 7 cohort. To my knowledge, there is no reason that the 2009 exam year differed with respect for 8th grade Algebra assignment for students around the level III but not the level IV threshold, so this is possibly just a blip in the data. For this reason, the 2009 grade 7 cohort is still included in all baseline estimates.

Baseline Specification

To estimate the effect of 8th grade Algebra enrollment on later outcomes I build on the specification in equation (3.2). Local linear regressions are estimated within a narrow window around both proficiency thresholds, separately for the level III and level IV thresholds. The first stage and reduced form equations give the raw discontinuities in 8th grade Algebra enrollment and the outcome of interest, respectively, from crossing a given proficiency threshold:

$$A_{ics} = \alpha_c^{\text{FS}} + \phi_s^{\text{FS}} + \beta^{\text{FS}} Z_{ics} + \delta_0^{\text{FS}} \theta_{ics} + \delta_1^{\text{FS}} \theta_{ics} * Z_{ics} + X'_{ics} \Gamma^{\text{FS}} + \epsilon_{ics}^{\text{FS}} \tag{3.3}$$
$$Y_{ics} = \alpha_c^{\text{RF}} + \phi_s^{\text{RF}} + \beta^{\text{RF}} Z_{ics} + \delta_0^{\text{RF}} \theta_{ics} + \delta_1^{\text{RF}} \theta_{ics} * Z_{ics} + X'_{ics} \Gamma^{\text{RF}} + \epsilon_{ics}^{\text{RF}} \tag{3.4}$$

Here i denotes student, c denotes cohort, and s denotes school. Again, the running variable, θ_{ics}, is the student's 7th grade math score. Enrollment in Algebra in 8th grade is denoted by A_{ics}. As mentioned in Section 3.4, controls are included in all baseline estimates. X_{isc} includes controls for prior test scores (grade 3 to grade 6), race, gender, free lunch status, limited english proficiency (LEP) status, and parental education. To account for any differences in the relationship between 7th grade math scores and 8th grade Algebra tracking across schools and cohorts, school and cohort fixed effects are included. Following the convention of Lee and Lemieux (2010), the "Local Average Treatment Effect" (LATE) (Angrist et al., 1996) of tracking into 8th grade Algebra can be estimated via local linear regressions in the following Two-Stage-Least-Squares (2SLS) system, using an indicator for reaching a particular 7th grade math proficiency threshold p, $Z_{is} = 1[\theta_{ics} \geq p]$, as an instrument for early enrollment in Algebra, A_{ics}:

$$A_{ics} = \alpha_c^{\text{FS}} + \phi_s^{\text{FS}} + \beta^{\text{FS}} Z_{ics} + \delta_0^{\text{FS}} \theta_{ics} + \delta_1^{\text{FS}} \theta_{ics} * Z_{ics} + X'_{ics} \Gamma^{\text{FS}} + \epsilon_{ics}^{\text{FS}} \tag{(3.3)}$$
$$Y_{ics} = \alpha_c + \phi_s + \beta A_{ics} + \delta_0 \theta_{ics} + \delta_1 \theta_{ics} * Z_{ics} + X'_{ics} \Gamma + \epsilon_{ics} \tag{3.5}$$

As with the first stage estimates reported in Table 3.2, estimated LATEs from this 2SLS system are estimated within a small window of 5 points on either side of the proficiency cutoff. I report estimates from local linear regressions in a tight bandwidth around the proficiency thresholds, where the linear function fits the data quite well (Figure 3.3). Results are generally robust to alternative choices of bandwidth or the use of IK "optimal" bandwidths (Imbens and Kalyanaraman, 2012).[24] As the 7th grade scale scores are discrete and there are differences in scaling across years, I follow the advice of Lee and Card (2008) and two-way cluster standard errors by scale score-by-exam year, and by school.

3.5 Results

Short-Run Test Score Outcomes

Much of the prior literature on Algebra acceleration has found negative impacts on initial performance in Algebra and on standardized 8th grade test scores. Table 3.3 presents estimates of the effect of 8th grade Algebra enrollment on standardized 8th grade test scores, as well as state-standardized EOC test scores in Algebra and English. Columns (1) and (2) report the mean of the dependent variable, just before the level III and level IV thresholds, respectively. In column (3) reports the mean (OLS) effect, from a regression using observational variation in 8th grade Algebra participation and controls from all prior test score from grade 3 to grade 7 in math, reading, and science, as well as controls for demographics. Columns (4) and (5) report the reduced form (equation (3.4)) and 2SLS estimates (equation (3.5)) for the level III threshold; Columns (5) and (6) repeat the same for the level IV threshold. OLS results suggest positive effects of 8th grade Algebra on 8th grade test scores, 9th grade English 1 test scores, but no effect on Algebra test scores.[25] However, these estimates are likely biased by positive selection, even conditional on prior test scores, and mask any potential heterogeneity in the causal effect for students of differing ability.

Indeed, the 2SLS estimates at both discontinuities differ from the OLS estimates. At the level III discontinuity, the effect of early Algebra acceleration is extremely negative. Students score 1.23 SDs (SE 0.47) *lower* in Algebra[26], and 1.18 SDs (SE 0.31) *lower* in 8th grade math. Point estimates for reading in 8th grade, and English in 9th grade are also large and negative, but imprecisely estimated. Students who cross the level III threshold are only marginally

[24]Baseline estimates are shown for a bandwidth of 5 scale score points, although results are robust to different choices of bandwidth. See Appendix Figures 3.13, 3.14, 3.15. Including higher order polynomials in the running variable and using a larger bandwidth produces estimates that are qualitatively similar, but very imprecise.

[25]The Algebra I EOC is taken at the end of the year by all students who enrolled in Algebra, in any grade. For students who took the exam multiple times (typically if the course was failed initially and then retaken), only the first exam score is used. For students who enroll in 8th grade Algebra, both the Algebra I EOC exam and the G8 Math EOG exams are taken at the end of the year.

[26]Reduced form figures for Algebra I exam scores near the level III threshold are shown in panel (a) of Appendix Figures 3.10 and 3.11.

proficient in 7th grade math, and near the 30th percentile overall within the state. That the effect for these students is highly negative is therefore not surprising and entirely consistent with the aforementioned literature that has examined district-level expansions among low-achieving students: students who are only marginally proficient at their grade level are ill-prepared for early acceleration. Moreover, these students are generally very near or at the bottom of the achievement distribution within their class, which could be detrimental relative to the counterfactual of being towards the middle or top of the distribution in a less-advanced classroom.

For students at the level IV margin, the effects are quite different. Panel (a) of Figure 3.4 reports the reduced form relationship between grade 7 math scores and Algebra test scores, and shows little evidence of any sizable discontinuity in Algebra test scores across the level IV threshold.[27] Indeed, 2SLS estimates for Algebra, 8th grade math, and 8th grade reading standardized exams are small and insignificant. Importantly, this is not a "null" effect; students are getting essentially the same achievement a year earlier, despite the curricular intensification. In addition, there is a large and significant effect on 9th grade English test scores (0.37 SDs, SE 0.18). This effect seems a bit puzzling, given that the 8th grade Algebra acceleration only initially affects the math curriculum. However, advanced-track course taking at many schools is highly correlated across subjects, as many students who take advanced courses in one subject are tracked into advanced courses in other subjects as well. Indeed, evidence in Table 3.4 appears to reconcile this somewhat: not only does 8th grade Algebra enrollment increase the likelihood of enrolling in an advanced 9th grade math course, but it also is associated with a 13 percentage point (SE 5.9) enrollment in an advanced English course.

Tracking Persistence and Course Outcomes

For high-ability students, early Algebra has no discernible positive or negative short-run effect on test score outcomes, while for low-ability students, early acceleration appears to be especially detrimental to standardized math test scores. However, the goal of 8th grade math acceleration is not to maximize 8th and 9th grade test scores, but rather to increase overall mathematics skills and preparation through the end of high school. Table 3.4 presents evidence on the effects of 8th grade Algebra enrollment on subsequent course trajectories in high school. Students at the level III threshold are 31 percentage points (SE 8) more likely to enroll in an advanced math course in 9th grade, relative to a baseline mean of 2% just below the threshold. On the other hand, higher-achieving students near the level IV cutoff are 64 percentage points (SE 6) more likely to enroll in an advanced math class, and 13 percentage points (SE 5.9) more likely to enroll in an advanced English course in 9th grade.

[27] As the strong upward slopes in the relationships between the outcomes of interest and 7th grade test scores make small discontinuities difficult to see visually, reduced form figures in Figure 3.4 are "zoomed-in" by de-trending the outcomes with respect to prior-year test scores. "Raw" reduced form figures for the level IV threshold are reported in Appendix Figure 3.12.

Upon completion of Algebra, the typical high school course trajectory is Geometry, Algebra II, Pre-Calculus (sometimes referred to as Analysis) and finally Calculus, although the order of Geometry and Algebra II is sometimes reversed. Thus, for students who enroll in Algebra in 8th grade and stay "on track", a student will have completed both Geometry and Algebra II by grade 10, Pre-Calculus by grade 11, and Calculus by grade 12. Columns 3-5 provide estimates of this track persistence. For low-achieving students, there appears to be only modest persistence in the advanced mathematics track in 9th and 10th grade, with no significant effects by 11th and 12th grade. They are 39 percentage points (SE 18) more likely to have completed both Geometry and Algebra II by grade 10, but effects on Pre-Calculus and Calculus in grades 11 and 12 are small and insignificant. Conversely, for high-achieving students, the mathematics course pipeline appears to have much less leakage. Students are 61 percentage points (SE 8) more likely to make it through Geometry and Algebra II by grade 10 and 54 percentage points (SE 8) more likely to have made it to Pre-Calculus. By the end of high school, 8th grade Algebra enrollment is associated with a 41 percentage point (SE 9) increase in Calculus enrollment, relative to a baseline of only 15% just below the threshold. These effects for high-achieving students are illustrated graphically in Figure 3.5, panel (a). High-achieving students also take more difficult versions of these courses: the effect on total number of AP courses taken is positive and significant, although the effect is less than one (0.35 additional AP math courses, SE 0.14). Effects on non-math AP courses are similar in magnitude, but imprecise and insignificant. Overall, for high-achieving students near the 80th percentile, early acceleration into Algebra appears to be highly beneficial for increasing both the rigor and the overall mathematics preparation through the entirety of high school. Low-achieving students see a slight initial increase in mathematics enrollment in 9th and 10th grades in Geometry and Algebra II, but by 11th grade, low-achieving students who were tracked into early Algebra were no more likely to remain on the advanced math track than their similarly low-scoring peers who were not accelerated in 8th grade.

College Entrance Exams

Table 3.5 and panel (b) of Figure 3.5 report estimates for ACT and SAT exam scores. Beginning in 2013, all North Carolina 11th graders in public schools took the ACT as a part of the standard course of study. Conversely, participation in the SAT is selected, and only about one-third of North Carolina high school students took the SAT during the sample window, with most taking the exam in 11th or 12th grade. While this creates the potential for selection bias, I find no differences in the propensity to take the SAT from 8th grade Algebra acceleration for students near either proficiency margin in 2SLS estimates. Moreover, for high-achieving students near the level IV threshold results on both the ACT and SAT are very similar. For these students early Algebra acceleration is associated with very large gains on both college entrance exams, particularly on the math subsections, on the order of roughly 0.4 to 0.5 SDs (Figure 3.5, panel b).

I find little evidence of any positive effects on either ACT or SAT scores for low-achieving students near the level III proficiency threshold. With the exception of ACT Science, for which the estimated effect is quite large and marginally significant, estimated effects are statistically insignificant and very imprecisely estimated. For most outcomes I cannot rule out large effects, although the lack of a long-run college entrance exam effect for low-achieving students from early Algebra acceleration is perhaps to be expected: as shown in Tables 3.3 and 3.4, these students experienced large initial declines in math performance, and saw only a modest increase in their likelihood to remain on the advanced math track through only 9th and 10th grade.

For students at the level IV threshold, the estimates indicate that college entrance exam performance was significantly improved by early Algebra. Early Algebra is associated with gain of 2.97 points on the ACT Math subsection (SE 0.71), and a total gain 2.53 points (SE 0.89) on the overall composite. The ACT is scored on a scale from 1 to 36, and the typical standard deviation is around 6 points, meaning that this increase corresponds to roughly a 0.5 SD effect on math, and a 0.4 SD effect overall. The overall composite score gain comes no only through the increase in math, but from marginally significant increases in ACT English/Writing and Reading, and a smaller positive but insignificant increase in ACT Science. On the SAT, math scores increase by 48.7 points (SE 14.9), verbal scores by 42.8 points (SE 17.4), and writing scores by 75.8 points (SE 7.8). These effects are consistent with ACT estimates and indicate that early Algebra is associated with roughly 0.5 SD increases on Math subsections of college entrance exams, with large gains onto other non-math sections as well. Why would early Algebra be associated with gains on non-math subsections? Recall in Table 3.4 that students were also more likely to be enrolled in advanced English in 9th grade, and saw large gains on 9th grade English exams as well. In addition to overall curricular rigor increasing, peer effects may also be important, as students who are consistently enrolled in advanced courses in high school are surrounded by peers of overall higher quality.

End-of-High School Outcomes

Table 3.6 reports estimates for end-of-high school outcomes, including graduation, high school GPA, and post-high school college enrollment and major intentions. A student's GPA is measured at the end of high school, and is scaled on the standard 0 to 4 scale.[28] College enrollment intentions are collected from a student survey at the time of graduation. STEM major intentions are computed using information from a survey collected at the time of SAT taking, only available for the subsample of students who took the SAT during high school.

For students at both proficiency thresholds, the overall effect on the probability of high school graduation is small and insignificant. While passing Algebra is typically a graduation requirement, the timing of Algebra enrollment, whether in 8th or 9th grade, or even later, appears to have little impact on this margin of high school completion. Low-achieving students who were accelerated in 8th grade see no higher or lower GPAs by the end of high school, and report no greater intentions to enroll in either 2 or 4 year postsecondary institutions, nor

[28]Some high schools assign higher GPA marks for equivalent grades in advanced or honors versions of certain courses. Here I use an "unweighted" GPA where the mapping between course grade and GPA marks is consistent between courses of any given level.

to pursue a STEM major post graduation. Similarly, high-achieving students accelerated in 8th grade see no effect on GPA, but this is not necessarily a null effect: equivalent grades in higher level, more rigorous courses is indicative of positive gains in both mathematics and overall academic skills. Recall that these students generally persisted on the advanced math track, with over 40% reaching calculus by the end of high school, and many enrolling in AP versions of math courses.

Estimates for students near the level IV threshold on both 2 and 4 year college intentions variables are large but insignificant. Point estimates appear to indicate a substitution between 2 and 4-year college, though effects are imprecise: the point estimate on 4-year college intentions is 17 percentage points (SE 12), with the estimated effect on 2-year college intentions nearly equal and opposite, -15 percentage points (SE 13). Notably, the margin for improvement in college enrollment for high-achieving math students is somewhat small. Approximately 84% of these students intended to enroll in college post graduation, with two-thirds stating the intention to pursue a 4-year postsecondary degree. On the other hand, relatively few students at this margin plan on pursuing STEM majors. The effect of early Algebra on STEM intentions is substantial. Estimates indicate a 33 percentage point increase (SE 13) in postsecondary STEM intentions, relative to a baseline mean of 23% just below the level IV threshold.

Heterogeneity: Student-level

Overall, 8th grade Algebra acceleration leads to large gains in mathematics course-taking and college entrance exams among high-achieving students, with little indication of any negative short-run test score effects. Low-achieving students, conversely, experience large initial test score losses and limited mathematics track persistence, with little evidence of any end of high school gains (or losses). How did these effects at both prior test score margins vary by student characteristics?[29] Tables 3.7 and 3.8 report heterogeneity in 2SLS estimates at the level IV and level III margins, respectively. Each column reports 2SLS estimates from equation (3.5) where A_{ics} is interacted with student characteristics, and is instrumented with Z_{ics} interacted with student characteristics.[30]

Table 3.7 shows some indication heterogeneity in effects for high-achieving students by gender, race, free-lunch eligibility, and parental education, although most differences are small in magnitude. ACT Math and STEM major effects are larger for boys than girls, although advanced math effects are slightly larger for girls. Black and Hispanic students saw larger gains in 9th grade advanced math propensity, ACT math scores, and in 4 year

[29] Additional estimates of heterogeneity in 2SLS effects by school and district characteristics are reported in Appendix Tables 3.10 and 3.11.

[30] Here, estimates are produced from interactions in the pooled model, which restricts the coefficients on the trends in the running variable (δ_1, δ_2) and the included covariates (Γ) to be equal for all students, while allowing for β to vary. Heterogeneous effects could also be estimated by running models separately for each category, which would allow for all coefficients to vary by student demographic subcategory. Results are very similar from either approach.

college intentions. When split more directly by student socioeconomic status, estimates indicate that more advantaged students generally see larger gains, with larger effects on most outcomes for non-free lunch students and students of college graduates, particularly in course-taking (both advanced in 9th grade and calculus by 12th grade).

For low-achieving students, heterogeneity by student demographic subgroup is somewhat consistent with that for high-achievers. Table 3.8 shows that persistence into advanced math in 9th grade is larger for girls than boys, and is larger among non-free lunch students and students with college educated parents. Similarly, the negative effects on initial Algebra test scores are worse for Black and Hispanic students, and for students with non-college educated parents, although that difference is not statistically significant ($p = 0.11$). Still, the magnitude of these differences is only small overall; the evidence across subgroups consistently indicates that accelerated students near the level III threshold fared poorly on initial test scores and saw little to no improvements in course-taking, college entrance exam scores, or postsecondary intentions.

The final panel of both Tables 3.7 and 3.8 shows results split by *6th grade* math achievement level. In Table 3.7, the results for high-achieving 7th graders are split by whether a student was high-achieving (level IV) in grade 6 math or not. In Table 3.8, results are split by whether a student was middle or high-achieving (levels 3 or 4) in grade 6 or not. Given that test scores are noisy measures of student ability, heterogeneity by earlier achievement can provide an additional test of heterogeneity in the effect of early Algebra by mathematics ability. Results in both tables reflect the general pattern of results reported thus far. Among high-achieving students in 7th grade, those who were also high-achieving in grade 6 math saw even larger gains in course taking and ACT math scores; among low-achieving students in 7th grade, those who were also low-achieving in grade 6 math experienced larger initial losses in test scores, and slightly worse persistence, although these differences are not statistically significant. Still, the magnitude of these differences is small (and often statistically insignificant) relative to the difference between pooled estimates of the effects at the level III and level IV thresholds.

3.6 Comparing RD and Observational LATEs

Estimates of the causal effect of 8th grade Algebra acceleration for students at different 7th grade test score proficiency threshold indicate substantial heterogeneity in the effect by prior achievement. Moreover, as indicated in Tables 3.3, 3.4, 3.5, and 3.6, estimates also tend to differ substantially from OLS estimates using observational variation. For low-achieving students, RD estimates around the level III threshold are consistent with the idea that observational estimates of 8th grade Algebra acceleration suffer from positive selection bias. For high-achieving students, however, the pattern is reversed: RD estimates at the level IV threshold are generally much larger than the observational estimates, most notably for effects on ACT and SAT scores, pre-calculus and calculus enrollment, and STEM major intentions. These observational estimates, as with most reported in the prior literature, are estimates of effects pooled over the entire distribution of student ability among 8th grade Algebra takers. Even in the absence of selection bias, pooled observational estimates mask any heterogeneity in the effects, and differences between observational and quasi-experimental estimates may derive from differences in prior student ability among the population studied.

Figure 3.6 plots estimated LATEs from Tables 3.3, 3.4, 3.5, and 3.6 at both the level III and level IV proficiency thresholds, as well as estimates using observational variation, allowing the effect to vary non-parametrically by decile of 7th grade math achievement. Specifically, I estimate:

$$Y_{ics} = \alpha_c + \phi_s + \sum_{d=1}^{10} \beta_d A_{ics} * \mathbb{1}[\theta^d_{ics} = d] + \sum_{d=1}^{10} \delta_d \mathbb{1}[\theta^d_{ics} = d] + X'_{ics}\Gamma + \epsilon_{ics} \tag{3.6}$$

Here, θ^d_{ics} denotes a student's 7th grade math score decile.[31] X_{ics} includes controls for all other test scores from grades 3-6 in math, science, and reading, as well as grade 7 reading scores and demographic background variables. As in equation (3.5), cohort (α_c) and school (ϕ_s) fixed effects are included. Figure 3.6 plots each of the β_d's, along with the RD estimates for the level III and level IV thresholds, which are nearest the 30th and 80th percentiles, respectively.

Results presented in Figure 3.6 indicate that there is substantial heterogeneity in observational estimates of the effect of 8th grade Algebra by prior achievement decile. For most outcomes, RD estimates at both discontinuities are similar to the analogous β_d estimated in equation (3.6). Early acceleration into Algebra in 8th grade is beneficial across a number of test score, course-taking, and post-high school outcomes for students in the upper half of the distribution, and in the top deciles in particular. For students below 50th percentile, however, early Algebra appears to be less beneficial. Observational estimates indicate negative impacts for these students on initial Algebra test scores, and the RD estimate from the level III threshold is even more negative than the corresponding OLS estimate. The positive benefits of early Algebra on ACT scores and calculus enrollment only appear among students in the top half of the distribution, and in both of these cases RD estimates at the 80th percentile are actually larger than the corresponding OLS estimate. Overall, both observational and RD estimates in this paper indicate that the effects of 8th grade Algebra depend crucially on the mathematics ability and level of preparation of the student.

3.7 Conclusion

The timing of a student's first exposure to Algebra remains the most important and consequential mathematics curricular divergence in US public K-12 education. Students who are accelerated into Algebra in 8th grade generally experience a more rigorous math curriculum, are on track to reach calculus by the end of high school, and are in classrooms with higher ability peers on average. Often referred to as a "gateway" into selective high education,

[31] Relative to all North Carolina 7th graders tested during the sample period.

efforts to increase mathematics skills and college preparation in secondary schools often focus on the timing of a student's enrollment in Algebra. Despite this importance, the prior literature on the effects of 8th grade Algebra on student outcomes has been mixed and inconclusive. The best available evidence to date comes from studies using quasi-experimental variation, but most only examine one or a small number of districts, and to date have been unable to follow students through the end of high school or examine effects on college readiness and post-high school college and major decisions. In this paper I provide credible new estimates of heterogeneous causal effects of 8th grade Algebra acceleration using administrative data from nearly all North Carolina public school students from the 2007-2011 7th grade cohorts.

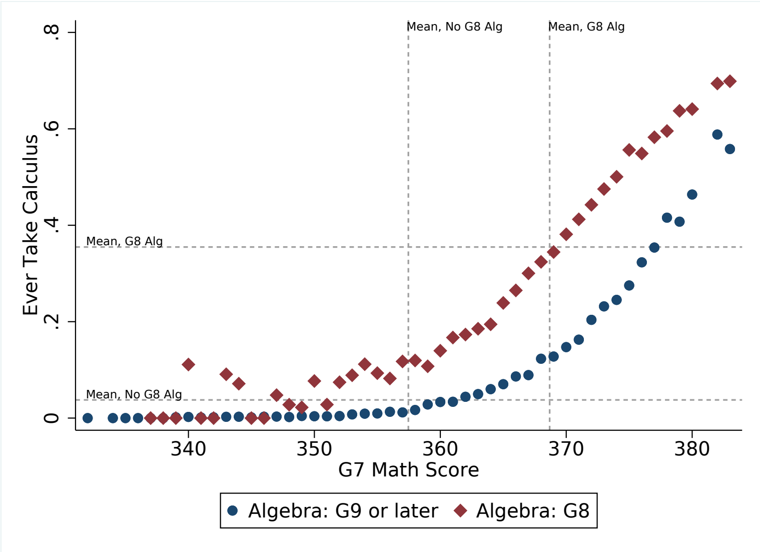
RD estimates from two discontinuities near the 30th and 80th percentile in the 7th grade math test score distribution provide evidence of large positive impacts for high-achieving students, but large negative effects for low-achieving students. Low-achieving students see large initial math test score losses, with little persistence on the advanced math track in high school, and no evidence of end-of-high school gains. For high-achieving students, early Algebra causes no detrimental achievement effect initially, and leads to large gains in mathematics course taking and rigor through high school. By the end of high school, high-achieving 7th graders with earlier exposure to Algebra see improvements in college entrance exam scores and intentions to pursue a STEM major post-high school. Evidence on the college enrollment margin is imprecise and less conclusive, but suggestive of substitution away from 2-year and into 4-year postsecondary institutions. The estimated heterogeneity in the impacts of Algebra acceleration in this paper can reconcile much of the conflicting evidence in the literature, which can be attributed to differences in the mathematics preparation of the students under consideration in a particular study.

Access to early Algebra among high-achieving students is surprisingly low: only about half of students in the top quartile of 7th grade math achievement enroll in Algebra in 8th grade. Efforts to expand access to Algebra among such well-prepared, high-achieving students could yield large improvements in mathematics skills, college preparedness, and eventual STEM major enrollment. However, the findings of this paper suggest caution if these efforts are targeted to students with lower mathematics ability and preparation: for such students, early Algebra will be less beneficial and may even be counterproductive.

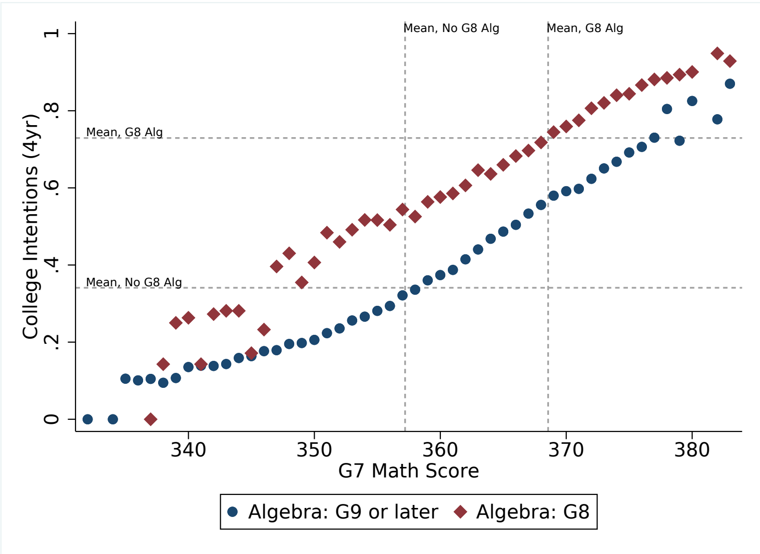
3.8 Figures

Figure 3.1: End of HS Outcomes by G7 Math Score and G8 Algebra

(a) Whether take calculus



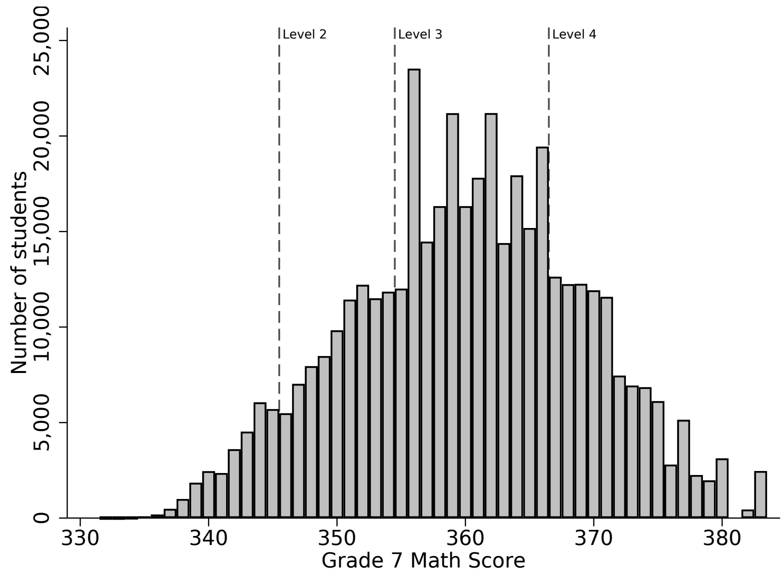
(b) College Intentions: 4yr



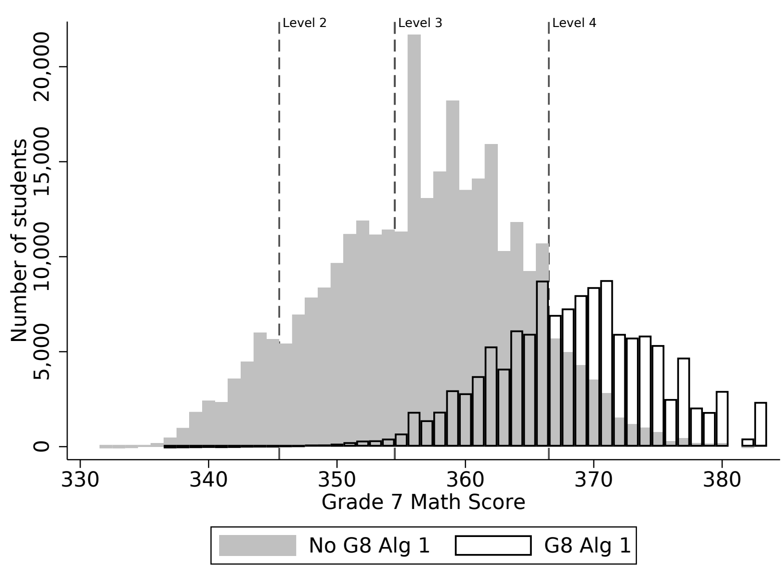
Notes: Figure shows the proportion of students who take Calculus in high school (panel a) or intend to go to a 4-year college (panel b) for a given grade 7 math test score. Blue circles (red diamonds) indicate the proportion at a given test score conditional on taking Algebra in 9th grade or later (in 8th grade). Horizontal dashed lines report the mean of the outcome for all students, by Algebra enrollment timing. Vertical dashed lines report the mean 7th grade test score for all students, by Algebra enrollment timing.

Figure 3.2: G7 Math Score Distribution

(a) Aggregate



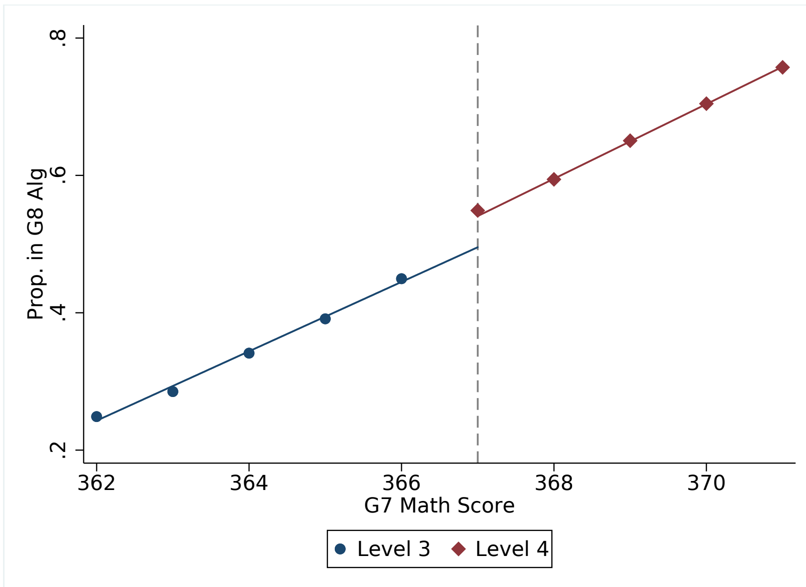
(b) By 8th Grade Algebra



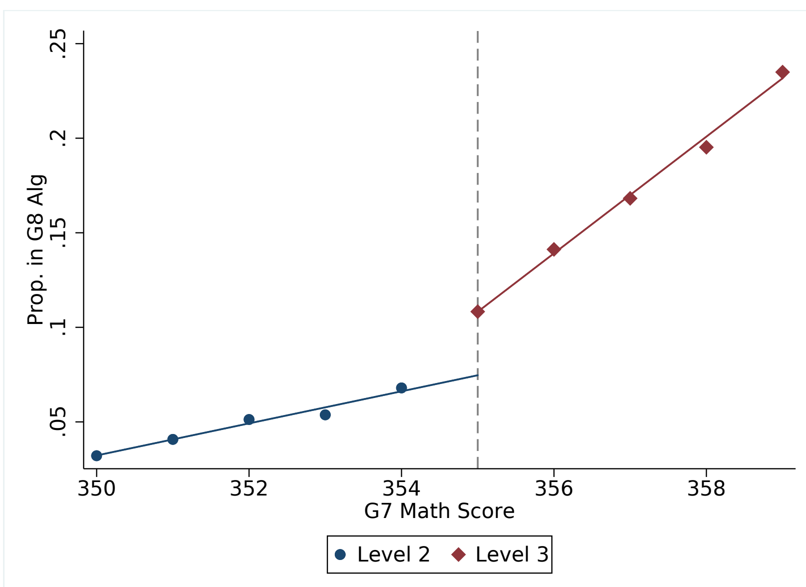
Notes: Panel (a) reports the distribution of 7th grade math test scores. Panel (b) reports the distribution of 7th grade math scores for students who enroll in 8th grade Algebra (black outline) and for students who enroll in Algebra in grade 9 or later (solid gray). Dashed vertical lines indicate state proficiency thresholds.

Figure 3.3: First Stage Figures

(a) Level IV Cutoff



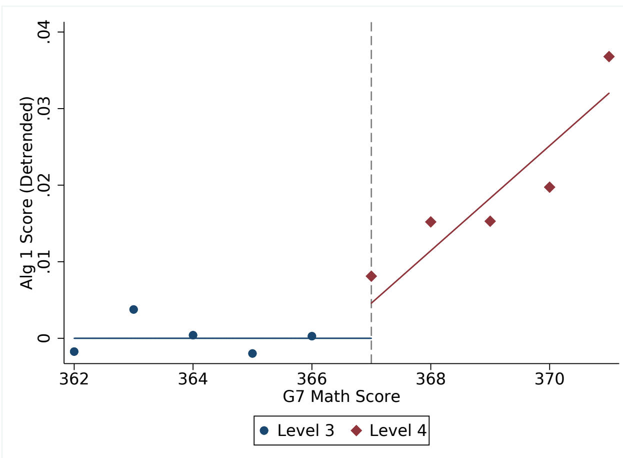
(b) Level III Cutoff



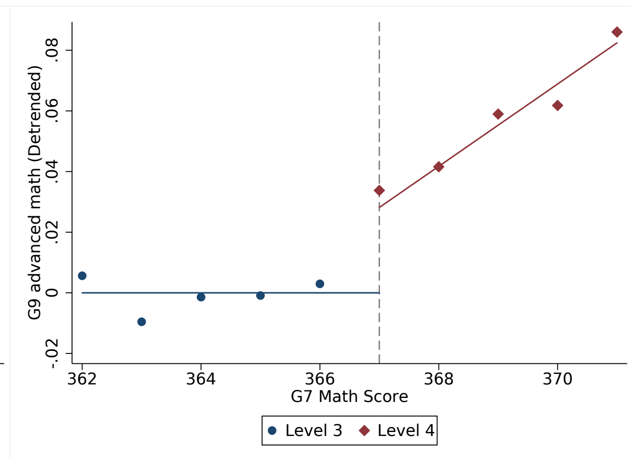
Notes: Figure shows the proportion of students who enroll in 8th grade algebra by grade 7 test score. Panel (a) restricts to 7th grade math test scores within a 5-point window on either side of the level IV proficiency threshold; panel (b) restricts to 7th grade scores within a 5-point window on either side of the level III proficiency threshold. Solid lines report the linear fit to the data on either side of the thresholds in both panels.

Figure 3.4: Reduced Form: Level IV Threshold (De-trended)

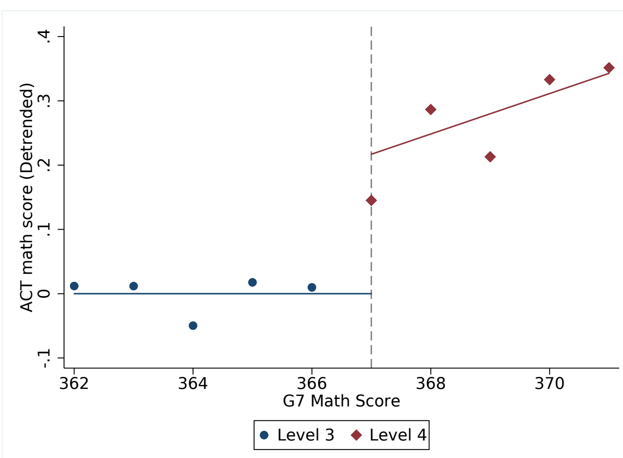
(a) Algebra 1 Score



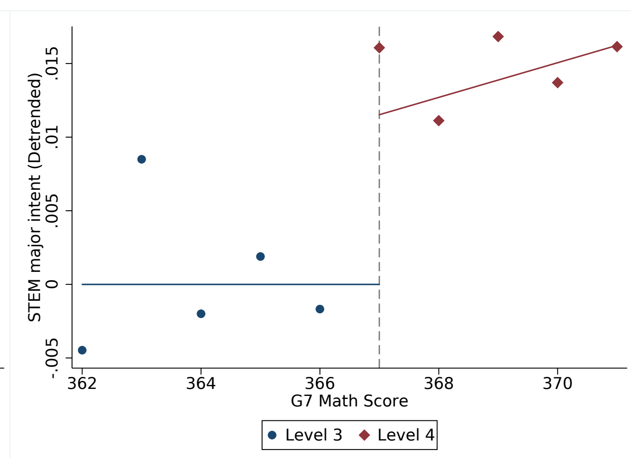
(b) G9 Advanced Math



(c) ACT Math



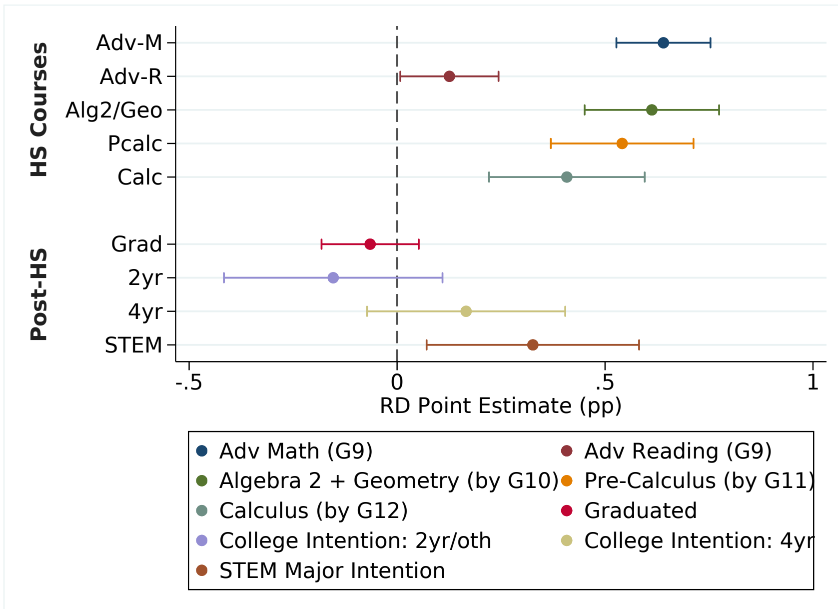
(d) STEM Major Intent



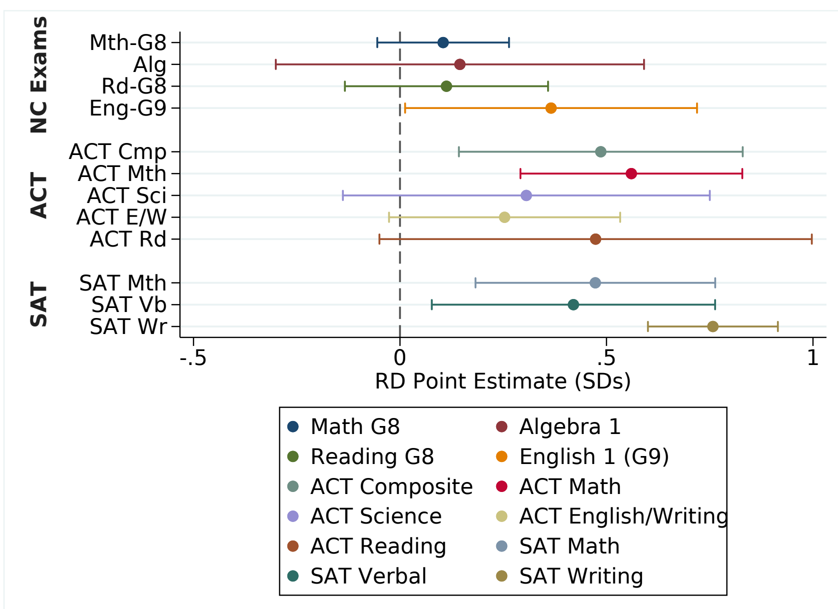
Notes: Figure reports reduced form relationship between grade 7 test scores and Algebra 1 test score (panel a), 9th grade advanced math (panel b), ACT math scores (panel c), and STEM major intention (panel d) within a 5-point window on either side of the level IV proficiency threshold. Variables on the y-axis are de-trended with respect to a linear trend in 7th grade test scores below the level IV threshold. Solid lines report the linear fit to the data on either side of the threshold in all panels.

Figure 3.5: RD Estimates: Level IV Threshold

(a) Course Pipeline / End of HS



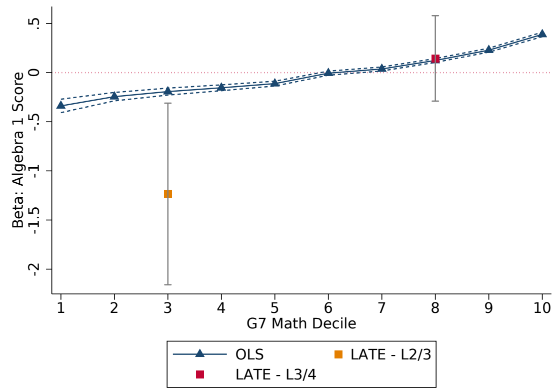
(b) Test Scores



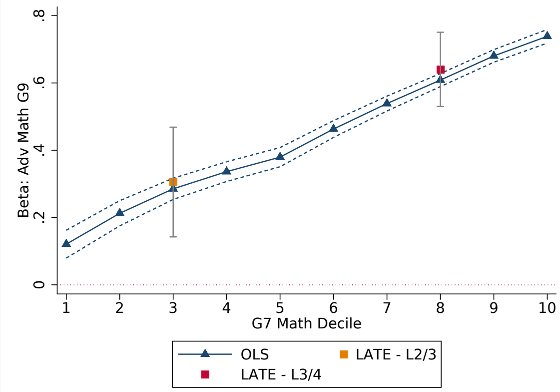
Notes: Figure reports estimates of β from equation (3.5) at the level IV threshold for course-pipeline and end-of-high school outcomes (panel a), and test score outcomes (panel b). Outcomes in panel a are reported in percentage point units; outcomes in panel b are reported in test score standard deviation units. See Section 3.5 in the text for greater detail.

Figure 3.6: LATEs by Prior Test Score Decile

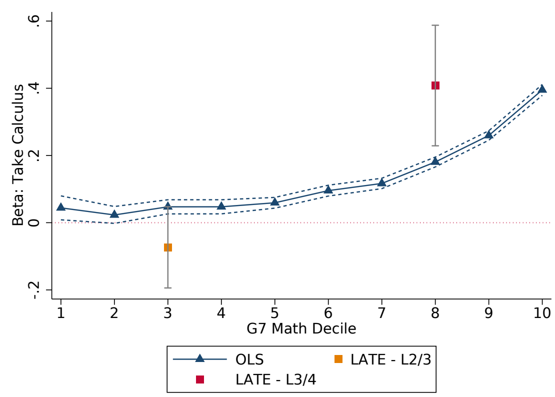
(a) Algebra 1 Score



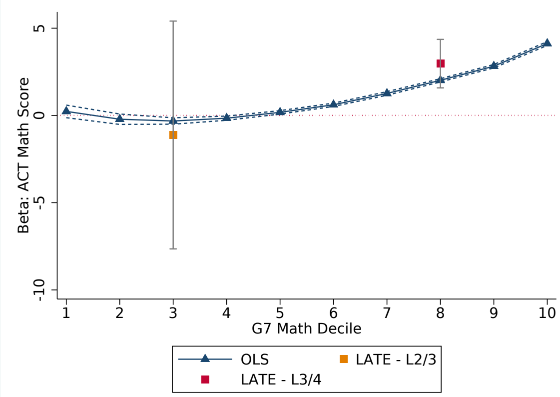
(b) G9 Advanced Math



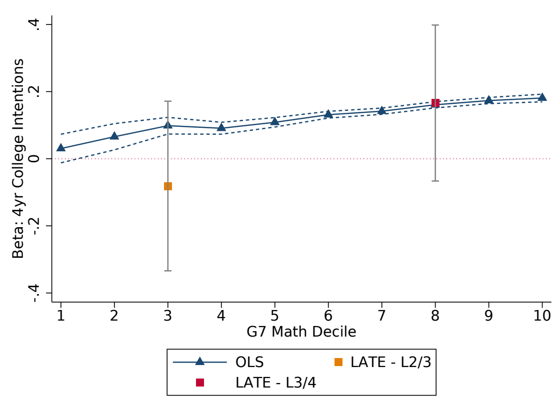
(c) Calculus



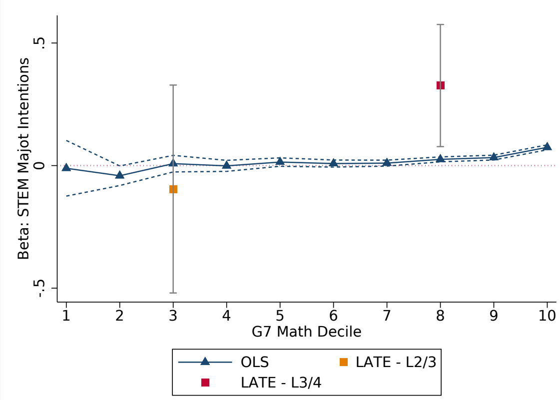
(d) ACT Math



(e) College Intention: 4yr



(f) STEM Major Intention



Notes: Figure plots estimates of β from equation (3.5) at both the level III and level IV thresholds, and estimates of β_d from equation (3.6), for Algebra 1 test score (panel a), 9th grade advanced math (panel b), Calculus enrollment (panel c), ACT math scores (panel d), 4-year college intentions (panel e) and STEM major intentions (panel f). See Section 3.6 in the text for greater detail.

3.9 Tables

Table 3.1: Summary Statistics

	mean	sd
G8 Algebra I	0.29	(0.46)
Prior Scores:		
G7 math score (scale)	360.14	(9.13)
G6 math score (std)	0.06	(0.95)
G7 read score (std)	0.07	(0.95)
Demographics:		
Black	0.27	(0.44)
Hispanic	0.09	(0.29)
Asian	0.02	(0.15)
Female	0.51	(0.50)
Free/reduced lunch	0.38	(0.49)
Parent HS dropout	0.08	(0.28)
Parent HS grad	0.38	(0.49)
Parent some college	0.24	(0.43)
Parent college grad	0.29	(0.46)
School/District:		
G8 cohort size	209.50	(81.94)
School fraction G8 Alg1	0.29	(0.15)
District fraction G8 Alg1	0.29	(0.13)
Selected Outcomes:		
Alg 1 Score	0.30	(0.93)
Adv Math G9	0.23	(0.42)
Take Calculus	0.13	(0.34)
College Intent: 4yr	0.46	(0.50)
STEM Major Intent	0.23	(0.42)
ACT Math	19.47	(4.61)
Observations	430127	

Notes: Table reports summary statistics for the main estimation sample, at the student-level. See Section 3.4 in the text for greater detail.

Table 3.2: First Stage Results

(a) Level IV Cutoff

	(1)	(2)	(3)	(4)
Above L3 Threshold	0.050*	0.049***	0.047***	0.044***
	(0.027)	(0.006)	(0.002)	(0.003)
N	148795	148795	148795	148795
F-statistic	3.36	79.04	581.87	168.32
Sch FEs			X	X
Cohort FEs		X	X	X
Controls				X
Clust by Score	X	X	X	X
Clust by Sch			X	X

(b) Level III Cutoff

	(1)	(2)	(3)	(4)
Above L2 Threshold	0.034**	0.034***	0.033***	0.035***
	(0.015)	(0.010)	(0.010)	(0.011)
N	65884	65884	65884	65884
F-statistic	5.13	12.41	10.43	11.16
Sch FEs			X	X
Cohort FEs		X	X	X
Controls				X
Clust by Score	X	X	X	X
Clust by Sch			X	X

Notes: Table reports first stage estimates of β from equations (3.1), (3.2), and (3.3). Panel (a) reports estimates of the level IV first stage discontinuity, while panel (b) reports estimates of the level III first stage discontinuity. Column (1) corresponds to equation (3.1), and column (2) adds cohort fixed effects, column (3) adds school fixed effects (equation (3.2)), and column 4 adds controls (equation (3.3)). Standard errors are two-way clustered by school and test score-by-exam year. See Section 3.4 in the text for greater detail.

Table 3.3: Results: "Short-run" Outcomes

	Mean of Depvar		OLS	Level 2/3 Cutoff		Level 3/4 Cutoff	
	L 2/3	L 3/4		RF	2SLS	RF	2SLS
Math:							
G8 Math Score	-0.53	0.53	0.08***	-0.04***	-1.18***	0.00	0.10
			(0.01)	(0.01)	(0.31)	(0.00)	(0.08)
Algebra 1 Score	-0.26	0.70	-0.01	-0.04***	-1.23***	0.01	0.15
			(0.01)	(0.01)	(0.47)	(0.01)	(0.22)
English/Reading:							
G8 Reading Score	-0.46	0.46	0.03***	-0.01	-0.35	0.00	0.11
			(0.00)	(0.01)	(0.24)	(0.01)	(0.12)
G9 English Score	-0.12	0.69	0.05***	-0.01	-0.48	0.02**	0.37**
			(0.00)	(0.01)	(0.37)	(0.01)	(0.18)

Notes: Table reports estimates of the effect of 8th grade Algebra on 8th and 9th grade test score outcomes. Dependent variables are test scores in 8th grade math (row 1), Algebra I (row 2), 8th grade reading (row 3), and 9th grade English (row 4). Columns (1) and (2) report the mean of the dependent variable within 2 scale score points below the 7th grade level III and level IV proficiency levels, respectively. Column (3) reports estimates of the effect using observational variation in 8th grade Algebra enrollment using the entire sample. Columns (4) and (6) report reduced form estimates of β^{RF} from local linear regressions corresponding to equation (3.4). Columns (5) and (7) report RD estimates of β using 2SLS, corresponding to equation (3.5). Columns (4) and (5) correspond to the level III threshold; columns (6) and (7) correspond to the level IV threshold. All RF and 2SLS effects are estimated within a bandwidth of 5 scale score points around the relevant proficiency threshold. Standard errors are two-way clustered by school and test score-by-exam year. See Sections 3.4 and 3.5 in the text for greater detail.

Table 3.4: Results: Course Progression Outcomes

	Mean of Depvar		OLS	Level 2/3 Cutoff		Level 3/4 Cutoff	
	L 2/3	L 3/4		RF	2SLS	RF	2SLS
Advanced, G9:							
Adv Math G9	0.02	0.33	0.56***	0.01**	0.31***	0.03***	0.64***
			(0.01)	(0.01)	(0.08)	(0.00)	(0.06)
Adv Reading G9	0.17	0.60	0.20***	0.00	0.14	0.01*	0.13**
			(0.01)	(0.01)	(0.15)	(0.00)	(0.06)
Mathematics Pipeline:							
Alg 2/Geom by G10	0.07	0.47	0.63***	0.01	0.39**	0.03***	0.61***
			(0.01)	(0.01)	(0.18)	(0.00)	(0.08)
Pre-Calc by G11	0.03	0.31	0.35***	0.00	0.11	0.02***	0.54***
			(0.01)	(0.00)	(0.12)	(0.00)	(0.08)
Calculus by G12	0.01	0.15	0.20***	-0.00	-0.07	0.02***	0.41***
			(0.01)	(0.00)	(0.06)	(0.01)	(0.09)
AP Course-Taking:							
Number AP (Math)	0.01	0.21	0.28***	-0.00	-0.01	0.02**	0.35**
			(0.01)	(0.00)	(0.06)	(0.01)	(0.14)
Number AP (Non-Math)	0.18	0.82	0.41***	0.03**	0.74	0.02	0.46
			(0.02)	(0.01)	(0.54)	(0.02)	(0.41)

Notes: Table reports estimates of the effect of 8th grade Algebra on high school course-taking outcomes. Dependent variables are enrollment in 9th grade advanced math (row 1), advanced reading (row 2), Algebra 2 and Geometry by grade 10 (row 3), Pre-Calculus by grade 11 (row 4), Calculus by grade 12 (row 5), total number of AP math classes (row 6), and total number of non-math AP classes (row 7). Columns (1) and (2) report the mean of the dependent variable within 2 scale score points below the 7th grade level III and level IV proficiency levels, respectively. Column (3) reports estimates of the effect using observational variation in 8th grade Algebra enrollment using the entire sample. Columns (4) and (6) report reduced form estimates of β^{RF} from local linear regressions corresponding to equation (3.4). Columns (5) and (7) report RD estimates of β using 2SLS, corresponding to equation (3.5). Columns (4) and (5) correspond to the level III threshold; columns (6) and (7) correspond to the level IV threshold. All RF and 2SLS effects are estimated within a bandwidth of 5 scale score points around the relevant proficiency threshold. Standard errors are two-way clustered by school and test score-by-exam year. See Sections 3.4 and 3.5 in the text for greater detail.

Table 3.5: Results: ACT and SAT

(a) Short-Run

	Mean of Depvar		OLS	Level 2/3 Cutoff		Level 3/4 Cutoff	
	L 2/3	L 3/4		RF	2SLS	RF	2SLS
ACT:							
ACT Composite	15.10	20.20	1.39***	0.05	2.89	0.11***	2.53***
			(0.04)	(0.07)	(3.72)	(0.04)	(0.89)
ACT Math	16.10	20.80	1.62***	-0.02	-1.12	0.13***	2.97***
			(0.04)	(0.06)	(3.33)	(0.03)	(0.71)
ACT Science	15.50	20.10	1.24***	0.15*	8.43*	0.07	1.65
			(0.04)	(0.08)	(4.21)	(0.05)	(1.19)
ACT Eng/Writing	12.90	18.20	1.26***	0.03	1.92	0.07*	1.67*
			(0.03)	(0.08)	(4.30)	(0.04)	(0.92)
ACT Reading	15.10	20.30	1.33***	0.01	0.68	0.13*	3.03*
			(0.04)	(0.10)	(5.29)	(0.07)	(1.67)
SAT:							
Took SAT	0.19	0.35	0.04***	-0.00	-0.06	-0.00	-0.00
			(0.00)	(0.01)	(0.37)	(0.01)	(0.27)
SAT Math	400.00	514.00	15.52***	0.92	15.94	2.04***	48.72***
			(0.65)	(1.87)	(31.71)	(0.72)	(14.88)
SAT Verbal	402.00	497.00	8.71***	-2.72	-47.07	1.80**	42.84**
			(0.59)	(1.91)	(36.95)	(0.91)	(17.41)
SAT Writing	386.00	472.00	13.39***	-0.41	-7.17	3.18***	75.77***
			(0.67)	(1.23)	(15.40)	(0.62)	(7.83)

Notes: Table reports estimates of the effect of 8th grade Algebra on college entrance exam scores. Dependent variables are test scores on the ACT composite (row 1), ACT math (row 2), ACT science (row 3), ACT English/writing (row 4), ACT reading (row 5), SAT math (row 7), SAT verbal (row 8), SAT writing (row 9), and an indicator for taking the SAT (row 6). Columns (1) and (2) report the mean of the dependent variable within 2 scale score points below the 7th grade level III and level IV proficiency levels, respectively. Column (3) reports estimates of the effect using observational variation in 8th grade Algebra enrollment using the entire sample. Columns (4) and (6) report reduced form estimates of β^{RF} from local linear regressions corresponding to equation (3.4). Columns (5) and (7) report RD estimates of β using 2SLS, corresponding to equation (3.5). Columns (4) and (5) correspond to the level III threshold; columns (6) and (7) correspond to the level IV threshold. All RF and 2SLS effects are estimated within a bandwidth of 5 scale score points around the relevant proficiency threshold. Standard errors are two-way clustered by school and test score-by-exam year. See Sections 3.4 and 3.5 in the text for greater detail.

Table 3.6: Results: End of High School

	Mean of Depvar		OLS	Level 2/3 Cutoff		Level 3/4 Cutoff	
	L 2/3	L 3/4		RF	2SLS	RF	2SLS
End of High School:							
Graduated	0.82	0.91	0.00	0.00	0.05	-0.00	-0.06
			(0.00)	(0.01)	(0.24)	(0.00)	(0.06)
GPA (unweighted)	2.40	3.10	0.13^{***}	0.00	0.12	-0.00	-0.02
			(0.00)	(0.01)	(0.22)	(0.01)	(0.13)
Post-High School:							
College Intent: 2yr	0.53	0.33	-0.10^{***}	0.00	0.13	-0.01	-0.15
			(0.00)	(0.01)	(0.25)	(0.01)	(0.13)
College Intent: 4yr	0.27	0.57	0.12^{***}	-0.00	-0.08	0.01	0.17
			(0.00)	(0.00)	(0.13)	(0.01)	(0.12)
STEM Major Intent	0.17	0.23	0.02^{***}	-0.01	-0.10	0.01^{***}	0.33^{**}
			(0.00)	(0.01)	(0.22)	(0.01)	(0.13)

Notes: Table reports estimates of the effect of 8th grade Algebra on end-of-high school outcomes. Dependent variables are indicators for graduation (row 1), 2-year college intentions (row 3), 4-year college intentions (row 4), STEM major intentions (row 5), and cumulative, unweighted high school GPA (row 2). Columns (1) and (2) report the mean of the dependent variable within 2 scale score points below the 7th grade level III and level IV proficiency levels, respectively. Column (3) reports estimates of the effect using observational variation in 8th grade Algebra enrollment using the entire sample. Columns (4) and (6) report reduced form estimates of β^{RF} from local linear regressions corresponding to equation (3.4). Columns (5) and (7) report RD estimates of β using 2SLS, corresponding to equation (3.5). Columns (4) and (5) correspond to the level III threshold; columns (6) and (7) correspond to the level IV threshold. All RF and 2SLS effects are estimated within a bandwidth of 5 scale score points around the relevant proficiency threshold. Standard errors are two-way clustered by school and test score-by-exam year. See Sections 3.4 and 3.5 in the text for greater detail.

Table 3.7: Heterogeneity by Student Characteristics (Level IV)

	Alg 1 Score	Adv Math G9	Calculus	ACT Math	4yr Coll Intent	STEM Intent
By gender:						
Boys	0.13	0.62^{***}	0.42^{***}	3.25^{***}	0.18	0.37^{***}
	(0.23)	(0.05)	(0.09)	(0.73)	(0.12)	(0.13)
Girls	0.16	0.66^{***}	0.39^{***}	2.63^{***}	0.15	0.27^{**}
	(0.22)	(0.06)	(0.10)	(0.70)	(0.12)	(0.12)
P-value	0.26	0.01	0.33	0.00	0.09	0.00
By race:						
Black/Hispanic	0.15	0.65^{***}	0.41^{***}	3.08^{***}	0.19	0.33^{**}
	(0.22)	(0.06)	(0.09)	(0.70)	(0.12)	(0.13)
White/Asian/Other	0.10	0.59^{***}	0.41^{***}	2.30^{***}	0.01	0.33^{**}
	(0.22)	(0.06)	(0.09)	(0.75)	(0.12)	(0.14)
P-value	0.13	0.01	0.99	0.00	0.00	0.90
By FRL status:						
FRL	0.09	0.56^{***}	0.19	2.55^{***}	0.12	0.29^{**}
	(0.25)	(0.07)	(0.14)	(0.78)	(0.12)	(0.11)
No FRL	0.05	0.69^{***}	0.27^{**}	3.10^{***}	0.21^{*}	0.24^{**}
	(0.25)	(0.07)	(0.12)	(0.66)	(0.12)	(0.10)
P-value	0.32	0.00	0.06	0.03	0.01	0.34
By parental education:						
Parent: No college	0.21	0.60^{***}	0.36^{***}	3.26^{***}	0.20	0.26^{*}
	(0.15)	(0.08)	(0.10)	(0.79)	(0.13)	(0.14)
Parent: college grad	0.25^{*}	0.69^{***}	0.44^{***}	3.51^{***}	0.20	0.24^{*}
	(0.15)	(0.08)	(0.11)	(0.74)	(0.13)	(0.15)
P-value	0.12	0.00	0.00	0.26	0.93	0.73
By G6 Math Level:						
Low/Mid G6 Math	0.15	0.57^{***}	0.31^{***}	2.47^{***}	0.14	0.35^{**}
	(0.24)	(0.06)	(0.11)	(0.65)	(0.12)	(0.14)
High G6 Math	0.18	0.63^{***}	0.38^{***}	2.83^{***}	0.16	0.33^{**}
	(0.23)	(0.06)	(0.11)	(0.63)	(0.11)	(0.13)
P-value	0.10	0.00	0.00	0.00	0.15	0.37

Notes: Table reports heterogeneity in level IV 2SLS effects for selected outcomes from Tables 3.3, 3.4, 3.5, and 3.6. Each panel reports estimates of β from equation (3.5) interacted with student characteristics. The p-value on the difference between the coefficients is reported in the final row of each panel. Standard errors are two-way clustered by school and test score-by-exam year. See Section 3.5 in the text for greater detail.

Table 3.8: Heterogeneity by Student Characteristics (Level III)

	Alg 1 Score	Adv Math G9	Calculus	ACT Math	4yr Coll Intent	STEM Intent
By gender:						
Boys	-1.28**	0.25***	-0.10	-0.97	-0.24*	-0.04
	(0.54)	(0.09)	(0.08)	(4.14)	(0.13)	(0.24)
Girls	-1.21***	0.34***	-0.06	-1.16	0.00	-0.13
	(0.43)	(0.08)	(0.06)	(3.10)	(0.12)	(0.21)
P-value	0.59	0.00	0.29	0.87	0.00	0.36
By race:						
Black/Hispanic	-1.38***	0.31***	-0.12	0.43	-0.11	-0.12
	(0.53)	(0.10)	(0.08)	(3.98)	(0.16)	(0.26)
White/Asian/Other	-1.12**	0.30***	-0.05	-2.21	-0.06	-0.08
	(0.44)	(0.08)	(0.06)	(3.35)	(0.11)	(0.18)
P-value	0.09	0.76	0.07	0.00	0.49	0.74
By FRL status:						
FRL	-1.81	0.24	-0.86**	-6.06	0.02	1.20
	(1.44)	(0.22)	(0.41)	(6.44)	(0.36)	(1.08)
No FRL	-1.70	0.30*	-0.61	-2.29	0.27	0.75
	(1.14)	(0.18)	(0.40)	(5.29)	(0.31)	(0.87)
P-value	0.76	0.24	0.04	0.01	0.00	0.25
By parental education:						
Parent: No college	-1.54***	0.23**	-0.10	-2.59	-0.42	-0.22
	(0.48)	(0.10)	(0.07)	(3.40)	(0.26)	(0.24)
Parent: college grad	-1.17***	0.37***	0.05	0.18	0.05	-0.17
	(0.31)	(0.08)	(0.05)	(2.63)	(0.21)	(0.18)
P-value	0.11	0.01	0.00	0.01	0.00	0.56
By G6 Math Level:						
Low G6 Math	-1.48***	0.28***	-0.08	-3.12	-0.22	-0.04
	(0.55)	(0.10)	(0.07)	(3.39)	(0.18)	(0.23)
Mid/High G6 Math	-1.15***	0.32***	-0.06	-1.03	-0.05	-0.02
	(0.40)	(0.07)	(0.07)	(2.63)	(0.14)	(0.23)
P-value	0.10	0.31	0.34	0.06	0.02	0.73

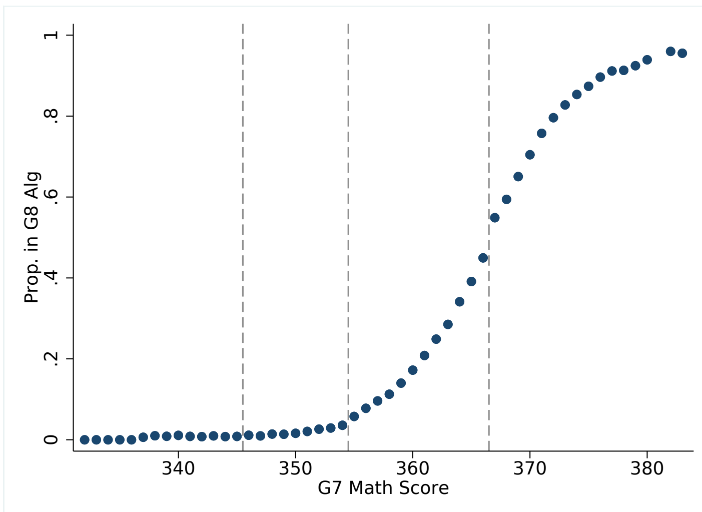
Notes: Table reports heterogeneity in level III 2SLS effects for selected outcomes from Tables 3.3, 3.4, 3.5, and 3.6. Each panel reports estimates of β from equation (3.5) interacted with student characteristics. The p-value on the difference between the coefficients is reported in the final row of each panel. Standard errors are two-way clustered by school and test score-by-exam year. See Section 3.5 in the text for greater detail.

3.10 Appendix Tables and Figures

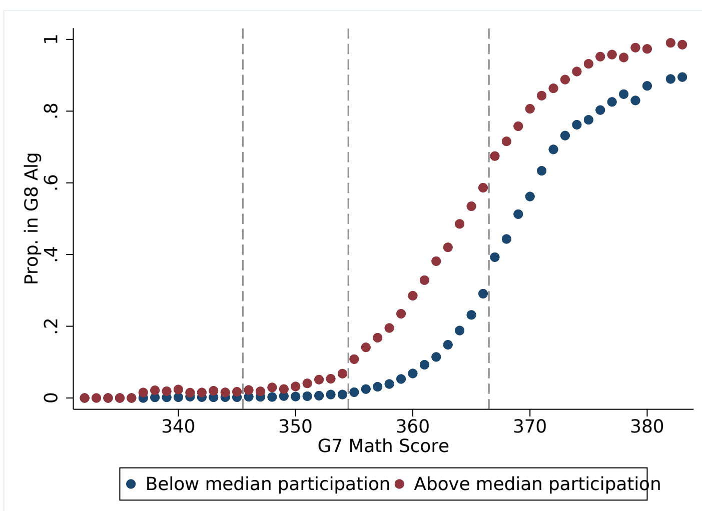
Appendix Figures

Figure 3.7: G7 Math Scores and G8 Algebra

(a) Overall



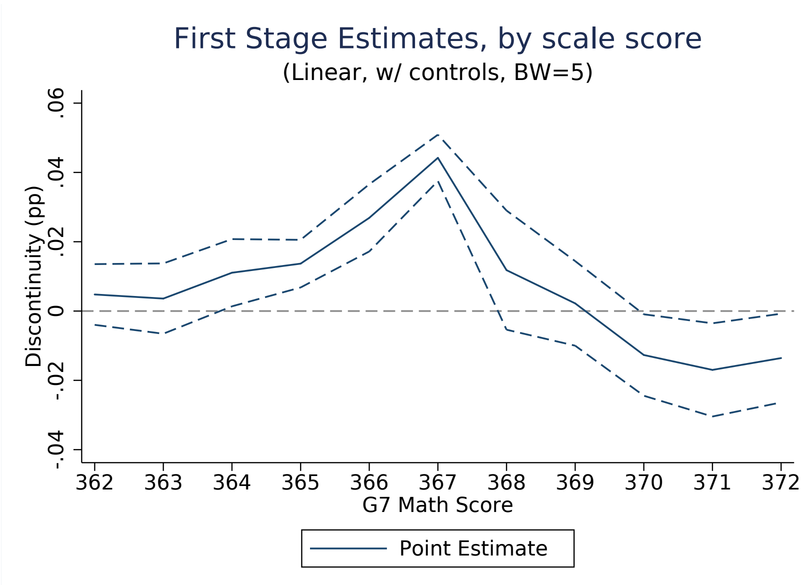
(b) By school proportion enrolled in G8 algebra



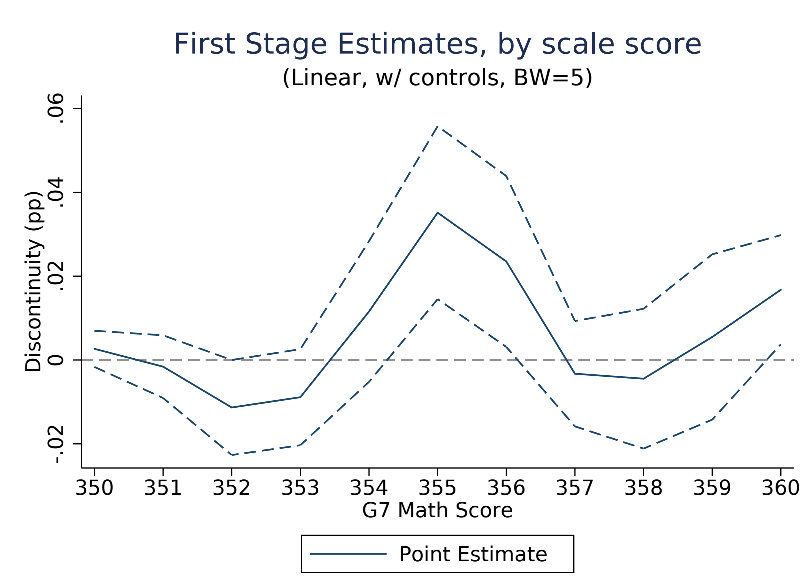
Notes: Figure shows the proportion of students who enroll in 8th grade algebra by grade 7 test score. Panel (a) shows proportions for all 7th grade math test score values. Panel (b) reports these proportions, split by whether a student's school was above (red dot) or below (blue dot) the statewide school-level median proportion of students enrolled in 8th grade Algebra. Dashed vertical lines indicate state proficiency cutoffs.

Figure 3.8: First Stage, by Cutoff Location (Placebo)

(a) Level IV Cutoff



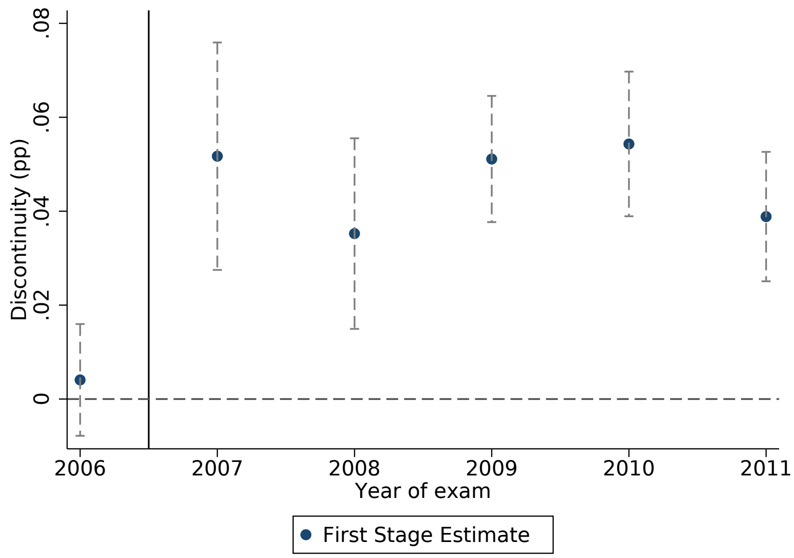
(b) Level III Cutoff



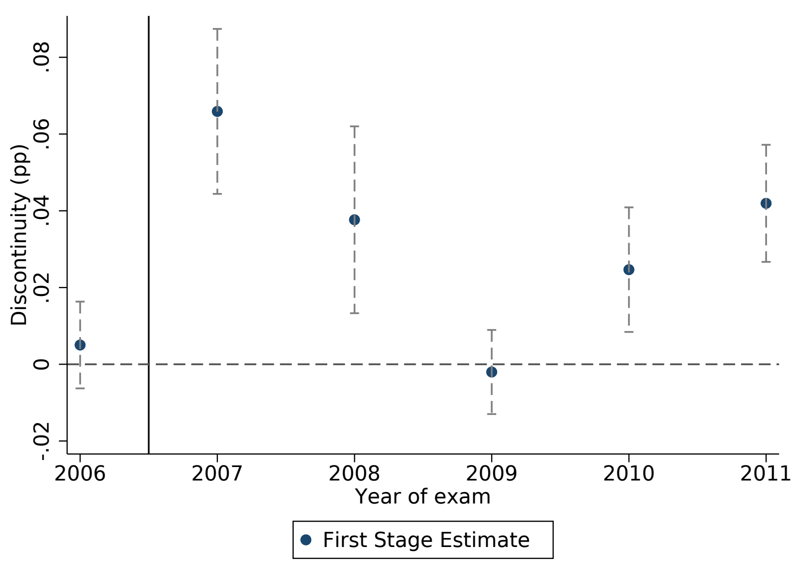
Notes: Figure reports estimates of β from equation (3.3) where the proficiency cutoff is incorrectly placed at different scale score values. Panel (a) reports estimates near the level IV threshold; the estimate at a score of 367 corresponds to first stage estimate from column (4) of Table 3, panel (a). Panel (b) reports estimates near the level III threshold; the estimate at a score of 355 corresponds to first stage estimate from column (4) of Table 3, panel (b). Standard errors are two-way clustered by school and test score-by-exam year.

Figure 3.9: First Stage, By Exam Year

(a) Level 3/4 Cutoff

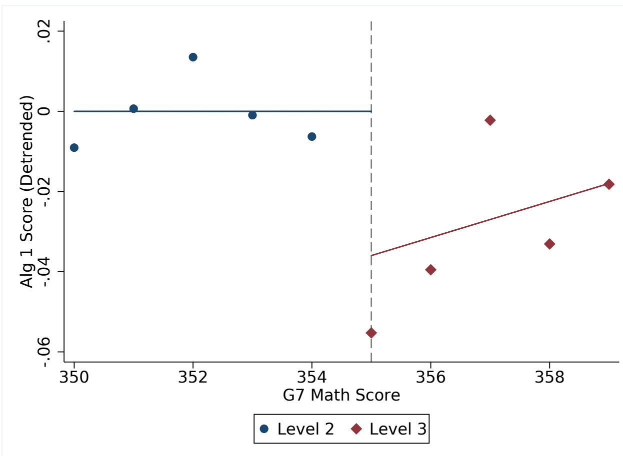


(b) Level 2/3 Cutoff

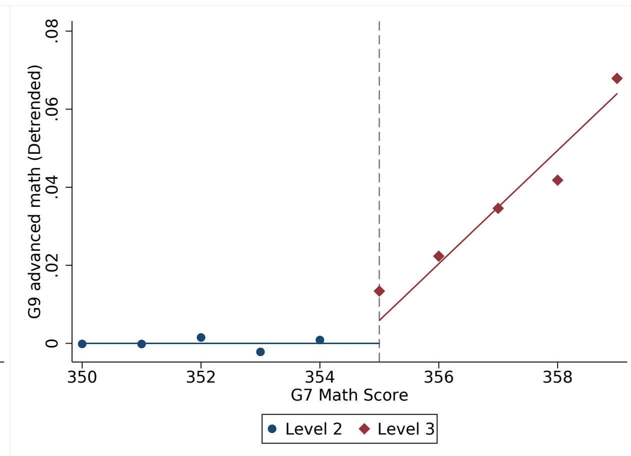


Notes: Figure reports estimates of β from equation (3.3) separately by exam year. Panel (a) shows estimates for the level IV threshold, and panel (b) shows estimates for the level III threshold. Standard errors are two-way clustered by school and test score-by-exam year. See Section 3.4 and 3.4 in the text for greater detail.

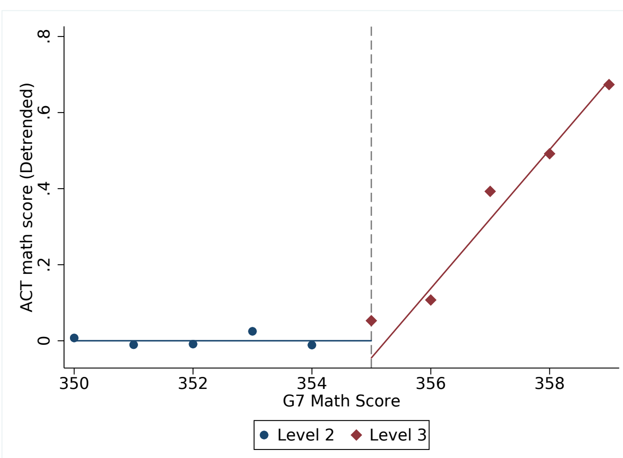
Figure 3.10: Reduced Form: Level III (De-trended)



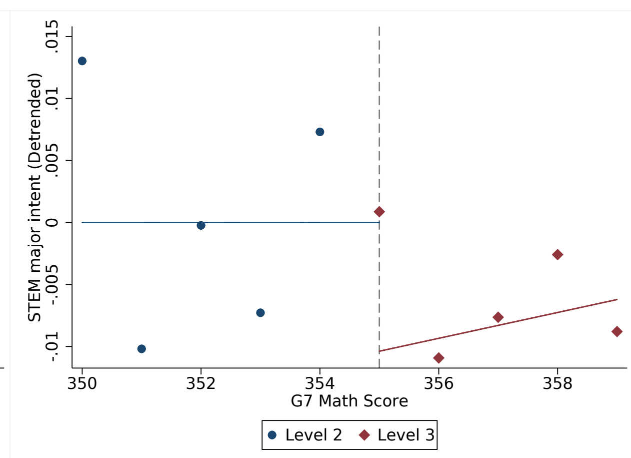
(a) Algebra 1 Test Score



(b) G9 Advanced Math



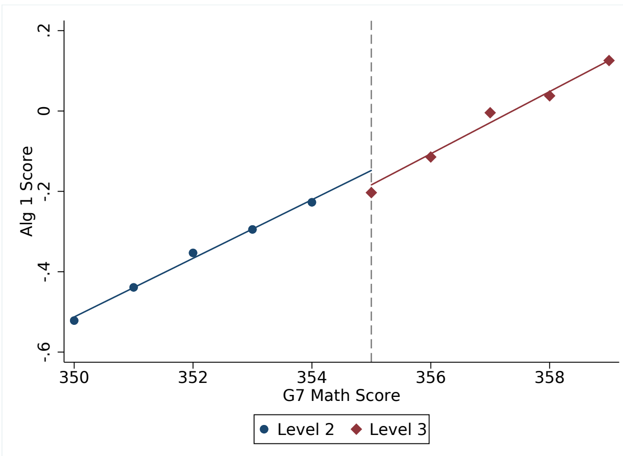
(c) ACT Math



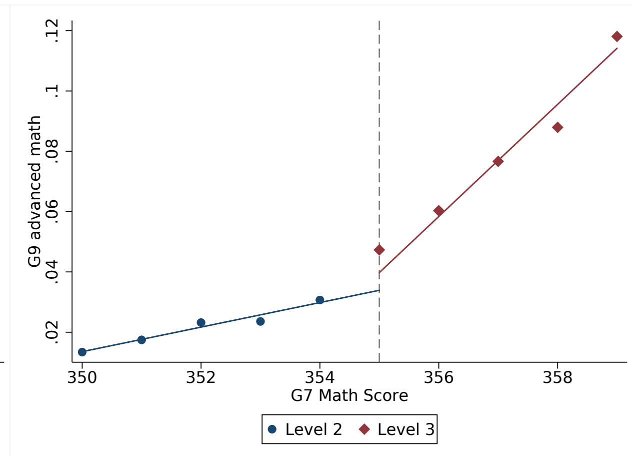
(d) STEM Intentions: 4yr

Notes: Figure reports reduced form relationship between grade 7 test scores and Algebra 1 test score (panel a), 9th grade advanced math (panel b), ACT math scores (panel c), and STEM major intention (panel d) within a 5-point window on either side of the level III proficiency threshold. Variables on the y-axis are de-trended with respect to a linear trend in 7th grade test scores below the level III threshold. Solid lines report the linear fit to the data on either side of the threshold in all panels.

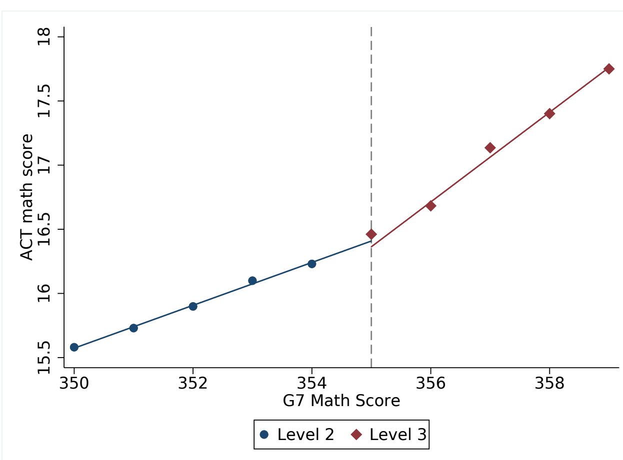
Figure 3.11: Reduced Form: Level III (Raw)



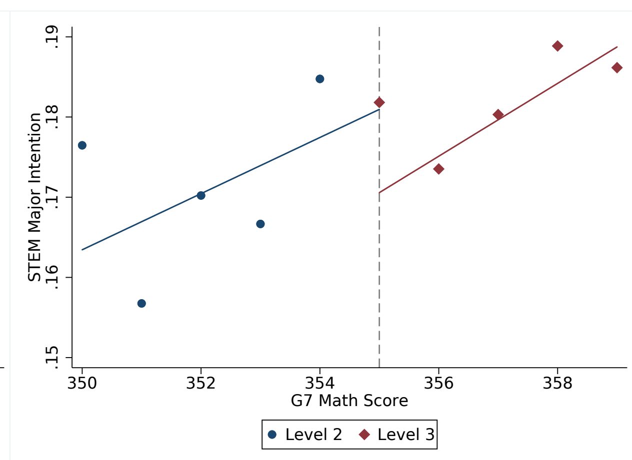
(a) Algebra 1 Test Score



(b) G9 Advanced Math



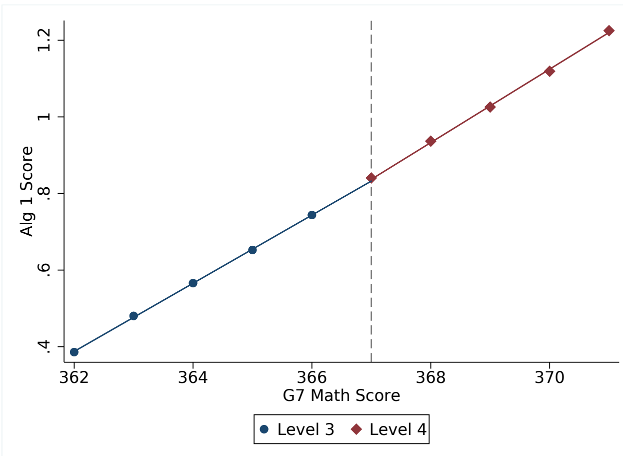
(c) ACT Math



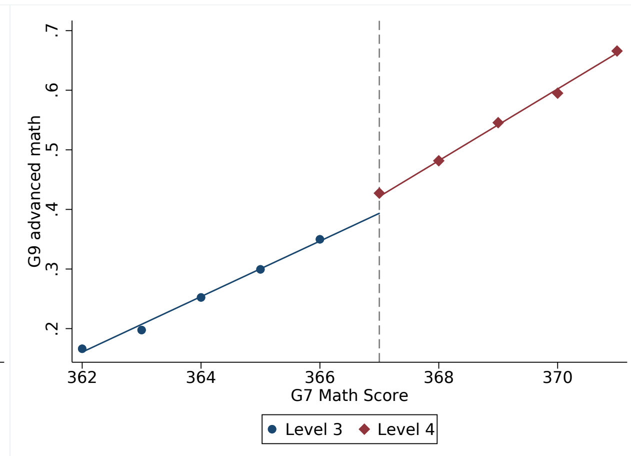
(d) STEM Intentions: 4yr

Notes: Figure reports reduced form relationship between grade 7 test scores and Algebra 1 test score (panel a), 9th grade advanced math (panel b), ACT math scores (panel c), and STEM major intention (panel d) within a 5-point window on either side of the level III proficiency threshold. Solid lines report the linear fit to the data on either side of the threshold in all panels.

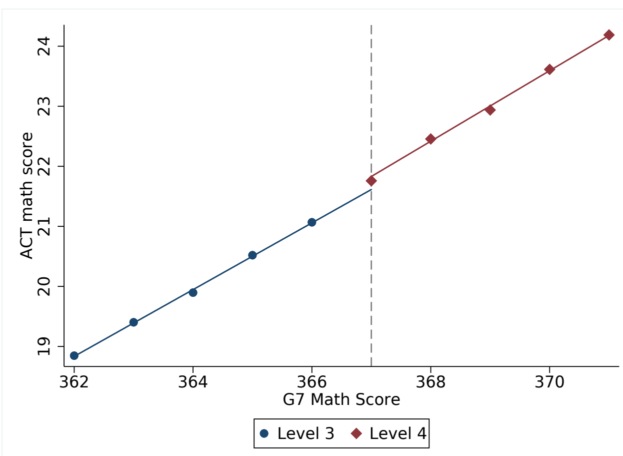
Figure 3.12: Reduced Form: Level IV (Raw)



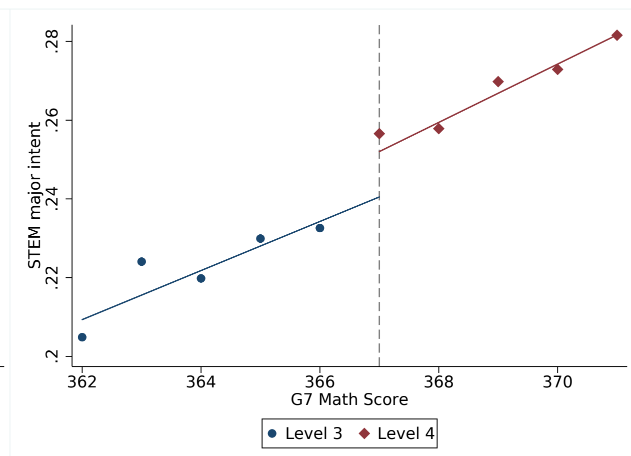
(a) Algebra 1 Score



(b) G9 Advanced Math



(c) ACT Math

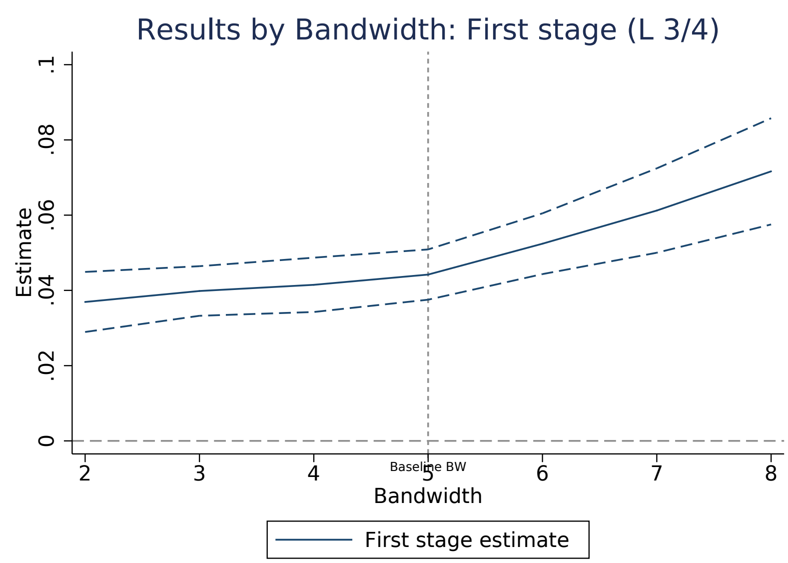


(d) STEM Major Intent

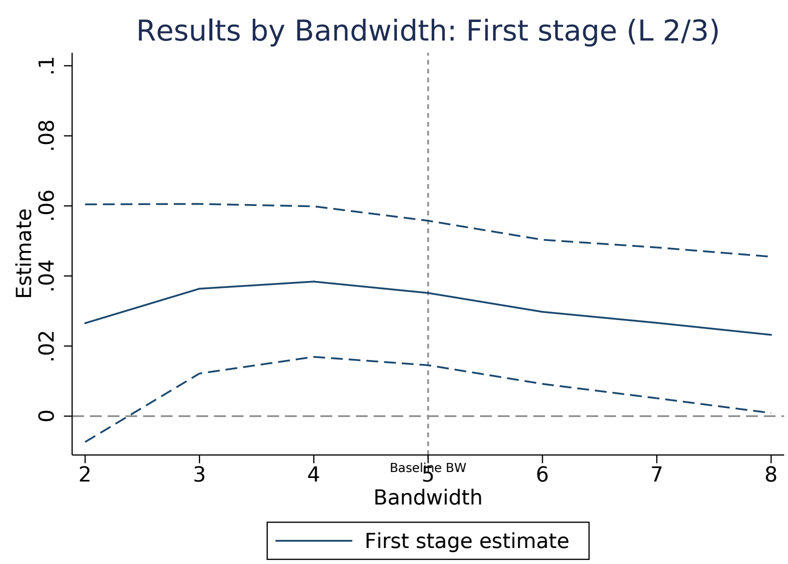
Notes: Figure reports reduced form relationship between grade 7 test scores and Algebra 1 test score (panel a), 9th grade advanced math (panel b), ACT math scores (panel c), and STEM major intention (panel d) within a 5-point window on either side of the level IV proficiency threshold. Solid lines report the linear fit to the data on either side of the threshold in all panels.

Figure 3.13: First Stage, By BW

(a) Level IV Cutoff

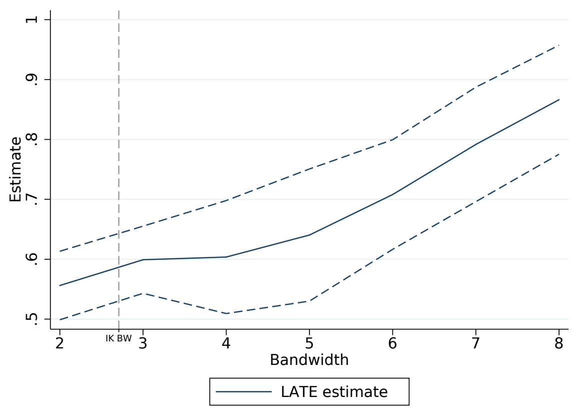


(b) Level III Cutoff

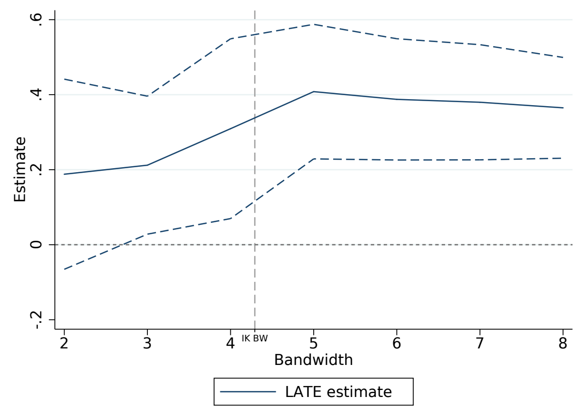


Notes: Figure reports first stage estimates of β from equation (3.3) for different choices of bandwidth. Panel (a) reports estimates for the level IV proficiency threshold; panel (b) reports estimates for the level III proficiency threshold. The dashed vertical line indicates the baseline bandwidth used in the main estimates. Standard errors are two-way clustered by school and test score-by-exam year. See Section 3.4 in the text for greater detail.

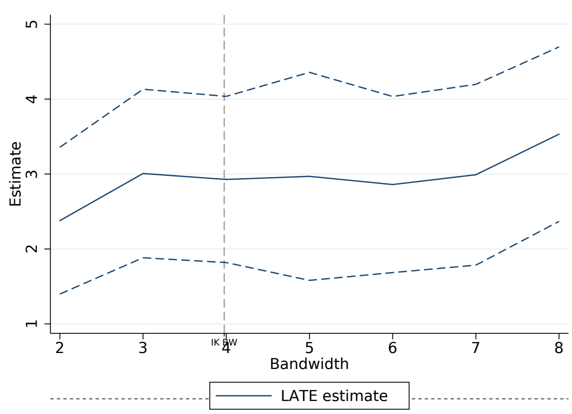
Figure 3.14: Results by BW: Level IV



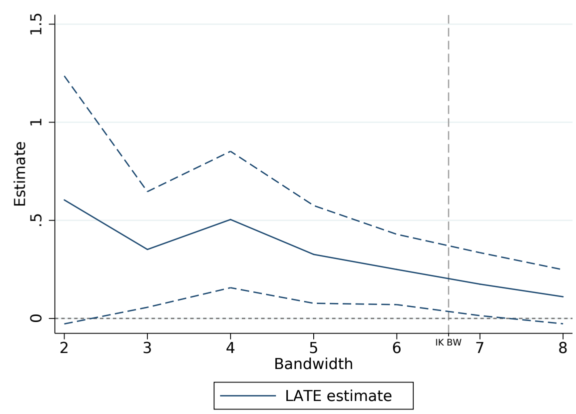
(a) Algebra 1 Score



(b) G9 Advanced Math



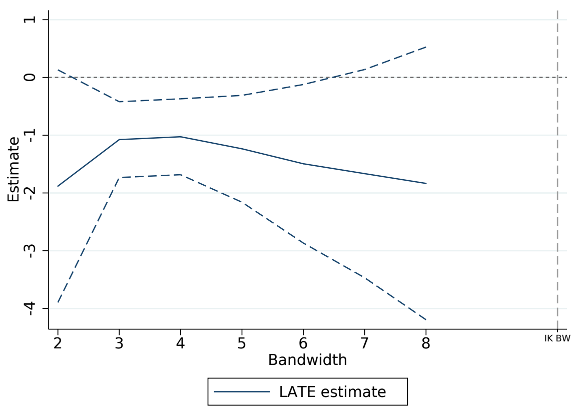
(c) ACT Math



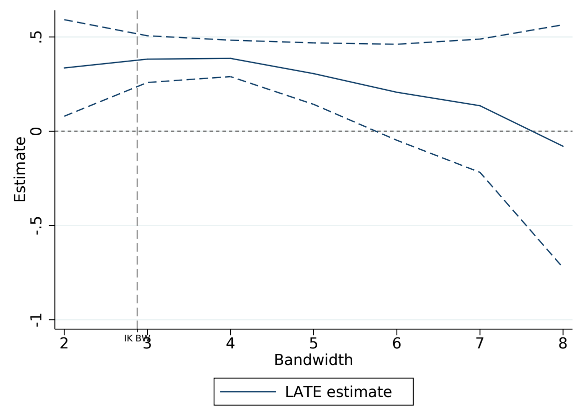
(d) STEM Major Intention

Notes: Figure reports 2SLS estimates of β from equation (3.5) for different choices of bandwidth around the level IV threshold.. Dependent variables are Algebra 1 test score (panel a), 9th grade advanced math (panel b), ACT math scores (panel c), and STEM major intention (panel d). The dashed vertical line indicates the optimal IK bandwidth. Standard errors are two-way clustered by school and test score-by-exam year. See Section 3.4 in the text for greater detail.

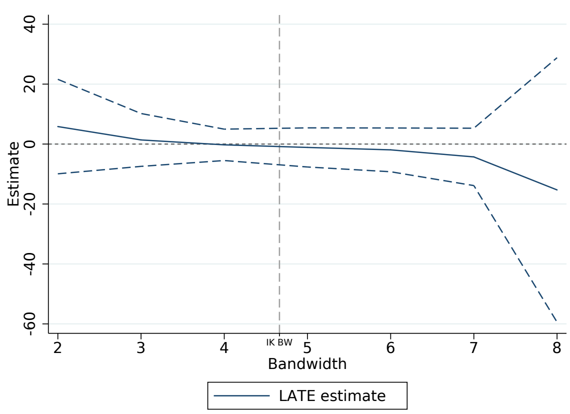
Figure 3.15: Results by BW: Level III



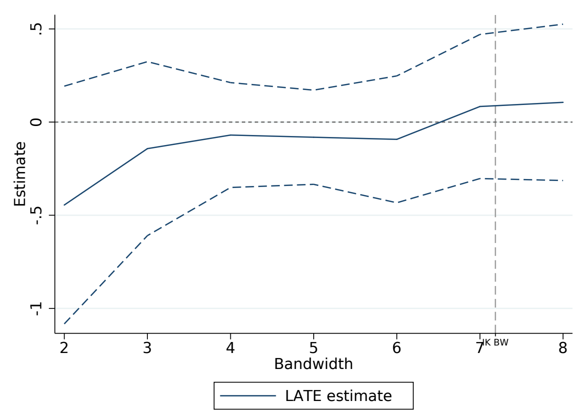
(a) Algebra 1 Score



(b) G9 Advanced Math



(c) ACT Math

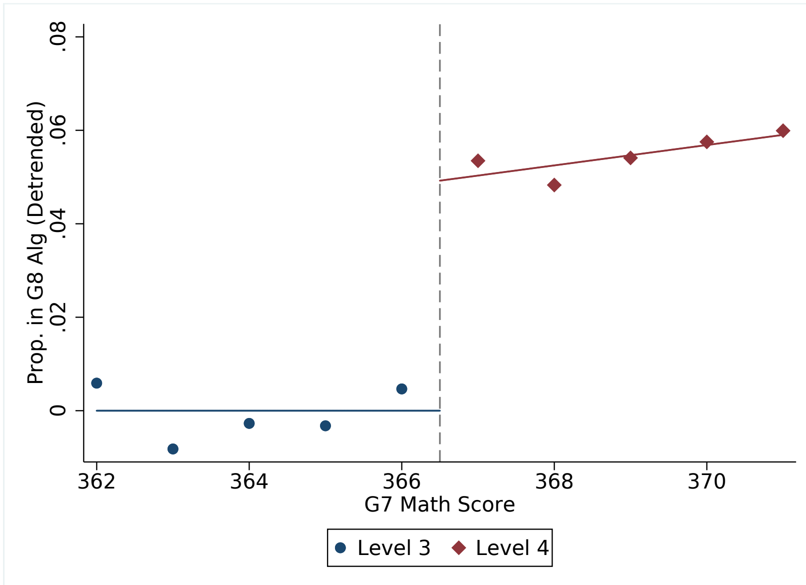


(d) College Intentions: 4yr

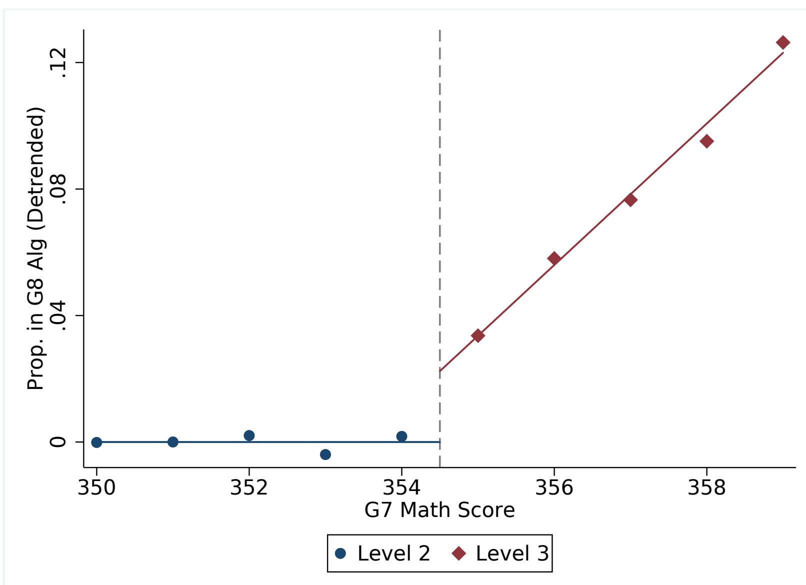
Notes: Figure reports 2SLS estimates of β from equation (3.5) for different choices of bandwidth around the level III threshold.. Dependent variables are Algebra 1 test score (panel a), 9th grade advanced math (panel b), ACT math scores (panel c), and STEM major intention (panel d). The dashed vertical line indicates the optimal IK bandwidth. Standard errors are two-way clustered by school and test score-by-exam year. See Section 3.4 in the text for greater detail.

Figure 3.16: First Stage Figures - Detrended

(a) Level IV Cutoff



(b) Level III Cutoff



Notes: Figure shows the (de-trended) proportion of students who enroll in 8th grade algebra by grade 7 test score. Panel (a) restricts to 7th grade math test scores within a 5-point window on either side of the level IV proficiency threshold; panel (b) restricts to 7th grade scores within a 5-point window on either side of the level III proficiency threshold. 8th grade Algebra enrollment (y-axis) is de-trended with respect to a linear trend in 7th grade test scores below the level IV threshold (panel a) or level III threshold (panel b). Solid lines report the linear fit to the data on either side of the thresholds in both panels.

Figure 3.17: North Carolina Test Score Scaling and Rounding

Grade 7 Form K

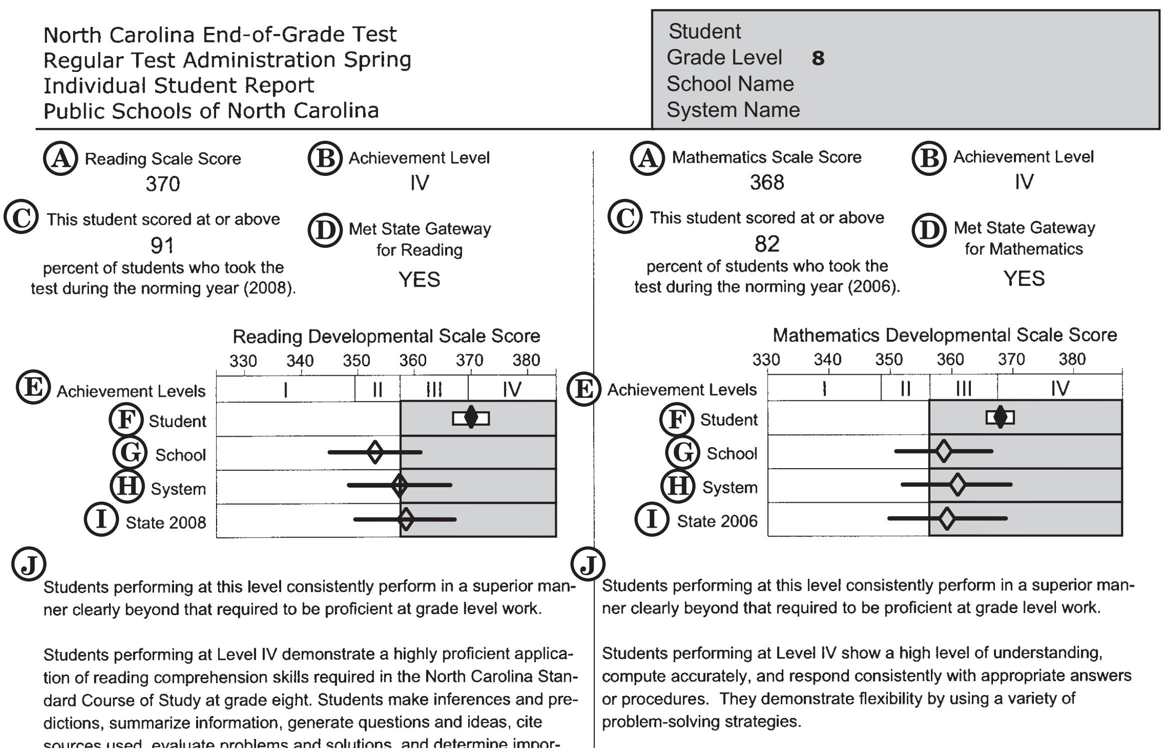
Score	EAP	SD	Score	SE
0	-2.10	0.52	332	5
1	-2.06	0.53	333	5
2	-2.01	0.53	333	5
3	-1.95	0.54	334	5
4	-1.90	0.54	334	6
5	-1.83	0.55	335	6
6	-1.76	0.55	336	6
7	-1.68	0.55	337	6
8	-1.60	0.55	337	6
9	-1.51	0.55	338	6
10	-1.41	0.55	339	6
11	-1.30	0.54	340	6
12	-1.18	0.53	342	5
13	-1.06	0.51	343	5
14	-0.94	0.49	344	5
15	-0.82	0.47	345	5
16	-0.70	0.44	347	4
17	-0.58	0.41	348	4
18	-0.46	0.38	349	4
19	-0.35	0.36	350	4
20	-0.25	0.33	351	3
21	-0.16	0.31	352	3
22	-0.06	0.30	353	3
23	0.02	0.28	354	3
24	0.11	0.27	355	3
25	0.19	0.26	356	3
26	0.27	0.25	356	3
27	0.34	0.25	357	3

Grade 7 Form L

Score	EAP	SD	Score	SE
0	-2.17	0.53	332	5
1	-2.12	0.53	332	5
2	-2.06	0.54	333	5
3	-2.00	0.54	333	5
4	-1.94	0.55	334	6
5	-1.86	0.55	335	6
6	-1.78	0.55	336	6
7	-1.69	0.55	336	6
8	-1.60	0.55	337	6
9	-1.50	0.55	338	6
10	-1.39	0.54	340	6
11	-1.27	0.53	341	5
12	-1.15	0.52	342	5
13	-1.03	0.50	343	5
14	-0.91	0.48	344	5
15	-0.78	0.45	346	5
16	-0.66	0.42	347	4
17	-0.55	0.40	348	4
18	-0.44	0.37	349	4
19	-0.33	0.35	350	4
20	-0.24	0.33	351	3
21	-0.14	0.31	352	3
22	-0.05	0.29	353	3
23	0.03	0.28	354	3
24	0.11	0.27	355	3
25	0.19	0.26	356	3
26	0.27	0.25	356	3
27	0.34	0.24	357	2

Notes: Figure shows the mapping between raw number correct ("score" in column 1) and scale score ("score" in column 4) for the 7th grade math exam in 2008, forms K (left panel) and L (right panel). Solid red boxes outline examples where rounding in scaling process converts two raw scores to the same scale score value. Source: North Carolina Mathematics Tests: Edition 3 Technical Report, June 2008.

Figure 3.18: North Carolina Test Score Report Sample



Notes: Figure shows a sample score report for a hypothetical 8th grade student in North Carolina.

Appendix Tables

Table 3.9: Covariate Balance

(a) Level IV Cutoff

	Above L3	P-Value
Predicted Pr(G8 Alg)	0.003 (0.002)	0.108
Black	-0.001 (0.003)	0.638
Hispanic	-0.001 (0.002)	0.695
Asian	0.001 (0.001)	0.703
Female	-0.002 (0.005)	0.716
Free/Reduced Lunch	-0.003 (0.005)	0.468
Parent: College Grad	0.013*** (0.004)	0.002
Parent: Some College	-0.007 (0.006)	0.226
Parent: HS Grad	-0.002 (0.005)	0.657
Parent: HS Dropout	-0.003 (0.002)	0.119
G7: Reading Score	0.001 (0.006)	0.833
G6: Math Score	0.012** (0.005)	0.017
G6: Reading Score	0.005 (0.006)	0.395
G5: Math Score	0.006 (0.005)	0.193
G5: Reading Score	-0.004 (0.007)	0.567
G5: Science Score	-0.017 (0.025)	0.506
G4: Math Score	0.003 (0.006)	0.688
G4: Reading Score	0.005 (0.007)	0.526
G3: Math Score	0.006 (0.004)	0.153
G3: Reading Score	0.004 (0.007)	0.550
Limited English Proficient	-0.000 (0.001)	0.885

(b) Level III Cutoff

	Above L2	P-Value
Predicted Pr(G8 Alg)	-0.003 (0.002)	0.269
Black	0.008 (0.006)	0.210
Hispanic	0.006 (0.006)	0.300
Asian	0.002 (0.002)	0.364
Female	0.001 (0.008)	0.951
Free/Reduced Lunch	0.023*** (0.007)	0.003
Parent: College Grad	0.005 (0.006)	0.469
Parent: Some College	-0.015* (0.008)	0.063
Parent: HS Grad	0.010 (0.007)	0.160
Parent: HS Dropout	0.000 (0.004)	0.912
G7: Reading Score	-0.003 (0.011)	0.772
G6: Math Score	-0.014 (0.010)	0.171
G6: Reading Score	-0.026* (0.013)	0.055
G5: Math Score	-0.016 (0.010)	0.137
G5: Reading Score	-0.023 (0.016)	0.148
G5: Science Score	-0.024 (0.037)	0.522
G4: Math Score	0.009 (0.014)	0.534
G4: Reading Score	0.002 (0.015)	0.907
G3: Math Score	-0.021 (0.015)	0.187
G3: Reading Score	-0.031* (0.018)	0.087
Limited English Proficient	0.004 (0.004)	0.404

Notes: Table reports estimated discontinuities in selected prior test score and demographic outcomes at the level IV (panel a) and level III (panel b) thresholds. Estimates are from regressions analogous to equation (3.2), with pre-treatment variables as the dependent variable. Each row an presents estimate from a single regression. In the first row in both panels the dependent variables are predicted values from a regression of 8th grade Algebra enrollment on all available pre-treatment variables. Standard errors are two-way clustered by school and test score-by-exam year. See Section 3.4 in the text for greater detail.

Table 3.10: Heterogeneity: School/District (Level IV)

	Alg 1	Adv Math G9	Calculus	ACT Math	4yr Coll Intent	STEM Intent
By School Poverty:						
Sch: High Pov	0.15	0.63^{***}	0.40^{***}	3.04^{***}	0.17	0.30^{***}
	(0.22)	(0.06)	(0.09)	(0.69)	(0.12)	(0.12)
Sch: Low Pov	0.14	0.66^{***}	0.43^{***}	2.78^{***}	0.16	0.37^{***}
	(0.23)	(0.06)	(0.10)	(0.78)	(0.12)	(0.14)
P-value	0.78	0.36	0.20	0.26	0.47	0.08
By Sch Achievement:						
Sch: Low Ach	0.08	0.65^{***}	0.45^{***}	2.66^{***}	0.13	0.37^{***}
	(0.21)	(0.05)	(0.09)	(0.69)	(0.11)	(0.13)
Sch: High Ach	0.18	0.63^{***}	0.39^{***}	3.14^{***}	0.18	0.31^{**}
	(0.23)	(0.06)	(0.10)	(0.73)	(0.12)	(0.13)
P-value	0.01	0.49	0.05	0.03	0.01	0.07
By School G8 Algebra Prop:						
Sch: Low Prop Alg	0.10	0.63^{***}	0.40^{***}	2.75^{***}	0.17	0.31^{**}
	(0.21)	(0.06)	(0.09)	(0.70)	(0.11)	(0.12)
Sch: High Prop Alg	0.22	0.67^{***}	0.42^{***}	3.21^{***}	0.16	0.35^{**}
	(0.24)	(0.06)	(0.09)	(0.71)	(0.13)	(0.14)
P-value	0.01	0.15	0.63	0.03	0.51	0.20
By District Size:						
Dist: Small	0.14	0.63^{***}	0.42^{***}	2.88^{***}	0.18	0.35^{***}
	(0.22)	(0.06)	(0.10)	(0.72)	(0.12)	(0.13)
Dist: Large	0.16	0.66^{***}	0.39^{***}	3.07^{***}	0.15	0.29^{**}
	(0.22)	(0.06)	(0.09)	(0.70)	(0.12)	(0.12)
P-value	0.51	0.19	0.49	0.33	0.08	0.04

Notes: Table reports heterogeneity in level IV 2SLS effects for selected outcomes from Tables 3.3, 3.4, 3.5, and 3.6. Each panel reports estimates of β from equation (3.5) interacted with school and district characteristics. The p-value on the difference between the coefficients is reported in the final row of each panel. Standard errors are two-way clustered by school and test score-by-exam year.

Table 3.11: Heterogeneity: School/District (Level III)

	Alg 1	Adv Math G9	Calculus	ACT Math	4yr Coll Intent	STEM Intent
By School Poverty:						
Sch: High Pov	-1.41**	0.24**	-0.12	0.70	-0.00	-0.09
	(0.60)	(0.11)	(0.08)	(5.17)	(0.17)	(0.32)
Sch: Low Pov	-1.14***	0.34***	-0.05	-1.66	-0.13	-0.10
	(0.42)	(0.08)	(0.06)	(2.93)	(0.11)	(0.16)
P-value	0.25	0.05	0.18	0.31	0.20	0.94
By Sch Achievement:						
Sch: Low Ach	-1.14**	0.34***	-0.05	-1.02	-0.10	-0.09
	(0.45)	(0.08)	(0.06)	(3.60)	(0.11)	(0.17)
Sch: High Ach	-1.62**	0.17	-0.20	0.55	-0.00	-0.13
	(0.76)	(0.14)	(0.12)	(8.23)	(0.20)	(0.40)
P-value	0.18	0.02	0.07	0.74	0.40	0.86
By District Size:						
Dist: Small	-1.43**	0.30***	-0.26**	-1.37	-0.18	-0.01
	(0.63)	(0.10)	(0.12)	(3.63)	(0.18)	(0.35)
Dist: Large	-1.18***	0.31***	-0.07	-1.09	-0.06	-0.09
	(0.44)	(0.08)	(0.07)	(3.28)	(0.12)	(0.22)
P-value	0.37	0.92	0.01	0.62	0.20	0.61

Notes: Table reports heterogeneity in level III 2SLS effects for selected outcomes from Tables 3.3, 3.4, 3.5, and 3.6. Each panel reports estimates of β from equation (3.5) interacted with school and district characteristics. The p-value on the difference between the coefficients is reported in the final row of each panel. Standard errors are two-way clustered by school and test score-by-exam year.

Table 3.12: Results: "Short-run" Outcomes (No Controls)

	Mean of Depvar		OLS	Level 2/3 Cutoff		Level 3/4 Cutoff	
	L 2/3	L 3/4		RF	2SLS	RF	2SLS
Math Scores:							
G8 Math Score	-0.53	0.53	0.08***	-0.05***	-1.37***	0.01**	0.19**
			(0.01)	(0.01)	(0.41)	(0.00)	(0.09)
Algebra 1 Score	-0.26	0.70	-0.01	-0.05***	-1.39**	0.01	0.22
			(0.01)	(0.01)	(0.55)	(0.01)	(0.22)
English/Reading Scores:							
G8 Reading Score	-0.46	0.46	0.03***	-0.02	-0.69	0.01	0.14
			(0.00)	(0.01)	(0.46)	(0.01)	(0.13)
G9 English Score	-0.12	0.69	0.05***	-0.02	-0.56	0.02**	0.36**
			(0.00)	(0.01)	(0.44)	(0.01)	(0.18)

Notes: Table reports estimates of the effect of 8th grade Algebra on 8th and 9th grade test score outcomes, without controls. Dependent variables are test scores in 8th grade math (row 1), Algebra I (row 2), 8th grade reading (row 3), and 9th grade English (row 4). Columns (1) and (2) report the mean of the dependent variable within 2 scale score points below the 7th grade level III and level IV proficiency levels, respectively. Column (3) reports estimates of the effect using observational variation in 8th grade Algebra enrollment using the entire sample. Columns (4) and (6) report reduced form estimates of β^{RF} from local linear regressions corresponding to equation (3.4) (excluding control variables). Columns (5) and (7) report RD estimates of β using 2SLS, corresponding to equation (3.5) (excluding control variables). Columns (4) and (5) correspond to the level III threshold; columns (6) and (7) correspond to the level IV threshold. All RF and 2SLS effects are estimated within a bandwidth of 5 scale score points around the relevant proficiency threshold. Standard errors are two-way clustered by school and test score-by-exam year. See Sections 3.4 and 3.5 in the text for greater detail.

Table 3.13: Results: Course Progression (No Controls)

	Mean of Depvar		OLS	Level 2/3 Cutoff		Level 3/4 Cutoff	
	L 2/3	L 3/4		RF	2SLS	RF	2SLS
Advanced, G9:							
Adv Math G9	0.02	0.33	0.56***	0.01*	0.30***	0.03***	0.67***
			(0.01)	(0.01)	(0.08)	(0.00)	(0.05)
Adv Reading G9	0.17	0.60	0.20***	0.00	0.06	0.01**	0.15**
			(0.01)	(0.01)	(0.16)	(0.00)	(0.07)
Mathematics Pipeline:							
Alg 2/Geom by G10	0.07	0.47	0.63***	0.01	0.39**	0.03***	0.63***
			(0.01)	(0.01)	(0.19)	(0.00)	(0.09)
Pre-Calc by G11	0.03	0.31	0.35***	0.00	0.11	0.03***	0.56***
			(0.01)	(0.00)	(0.11)	(0.00)	(0.09)
Calculus by G12	0.01	0.15	0.20***	-0.00*	-0.07	0.02***	0.42***
			(0.01)	(0.00)	(0.06)	(0.01)	(0.08)
AP Course-Taking:							
Number AP (Math)	0.01	0.21	0.28***	-0.00	-0.01	0.02**	0.39***
			(0.01)	(0.00)	(0.06)	(0.01)	(0.13)
Number AP (Non-Math)	0.18	0.82	0.41***	0.03***	0.82	0.03	0.64
			(0.02)	(0.01)	(0.56)	(0.02)	(0.41)

Notes: Table reports estimates of the effect of 8th grade Algebra on high school course-taking outcomes, without controls. Dependent variables are enrollment in 9th grade advanced math (row 1), advanced reading (row 2), Algebra 2 and Geometry by grade 10 (row 3), Pre-Calculus by grade 11 (row 4), Calculus by grade 12 (row 5), total number of AP math classes (row 6), and total number of non-math AP classes (row 7). Columns (1) and (2) report the mean of the dependent variable within 2 scale score points below the 7th grade level III and level IV proficiency levels, respectively. Column (3) reports estimates of the effect using observational variation in 8th grade Algebra enrollment using the entire sample. Columns (4) and (6) report reduced form estimates of β^{RF} from local linear regressions corresponding to equation (3.4) (excluding control variables). Columns (5) and (7) report RD estimates of β using 2SLS, corresponding to equation (3.5) (excluding control variables). Columns (4) and (5) correspond to the level III threshold; columns (6) and (7) correspond to the level IV threshold. All RF and 2SLS effects are estimated within a bandwidth of 5 scale score points around the relevant proficiency threshold. Standard errors are two-way clustered by school and test score-by-exam year. See Sections 3.4 and 3.5 in the text for greater detail.

Table 3.14: Results: ACT and SAT (No Controls)

	Mean of Depvar		OLS	Level 2/3 Cutoff		Level 3/4 Cutoff	
	L 2/3	L 3/4		RF	2SLS	RF	2SLS
***ACT*:**							
ACT Composite	15.10	20.20	1.39***	0.09	4.59	0.14***	3.11***
			(0.04)	(0.07)	(3.45)	(0.04)	(0.98)
ACT Math	16.10	20.80	1.62***	-0.01	-0.37	0.16***	3.42***
			(0.04)	(0.05)	(2.80)	(0.04)	(0.79)
ACT Science	15.50	20.10	1.24***	0.18**	9.32**	0.10*	2.13*
			(0.04)	(0.08)	(4.08)	(0.05)	(1.16)
ACT Eng/Writing	12.90	18.20	1.26***	0.08	3.97	0.12***	2.44**
			(0.03)	(0.07)	(3.91)	(0.04)	(1.03)
ACT Reading	15.10	20.30	1.33***	0.07	3.71	0.17**	3.67**
			(0.04)	(0.09)	(4.93)	(0.07)	(1.76)
***SAT*:**							
Took SAT	0.19	0.35	0.04***	-0.00	-0.10	0.00	0.03
			(0.00)	(0.01)	(0.45)	(0.01)	(0.31)
SAT Math	400.00	514.00	15.52***	1.01	17.87	2.29***	52.37***
			(0.65)	(2.10)	(36.73)	(0.87)	(16.09)
SAT Verbal	402.00	497.00	8.71***	-2.59	-45.68	2.12*	48.67**
			(0.59)	(1.89)	(39.45)	(1.11)	(22.05)
SAT Writing	386.00	472.00	13.39***	-0.64	-11.25	3.29***	75.36***
			(0.67)	(1.79)	(31.54)	(0.78)	(15.11)

Notes: Table reports estimates of the effect of 8th grade Algebra on college entrance exam scores, without controls. Dependent variables are test scores on the ACT composite (row 1), ACT math (row 2), ACT science (row 3), ACT English/writing (row 4), ACT reading (row 5), SAT math (row 7), SAT verbal (row 8), SAT writing (row 9), and an indicator for taking the SAT (row 6). Columns (1) and (2) report the mean of the dependent variable within 2 scale score points below the 7th grade level III and level IV proficiency levels, respectively. Column (3) reports estimates of the effect using observational variation in 8th grade Algebra enrollment using the entire sample. Columns (4) and (6) report reduced form estimates of β^{RF} from local linear regressions corresponding to equation (3.4) (excluding control variables). Columns (5) and (7) report RD estimates of β using 2SLS, corresponding to equation (3.5) (excluding control variables). Columns (4) and (5) correspond to the level III threshold; columns (6) and (7) correspond to the level IV threshold. All RF and 2SLS effects are estimated within a bandwidth of 5 scale score points around the relevant proficiency threshold. Standard errors are two-way clustered by school and test score-by-exam year. See Sections 3.4 and 3.5 in the text for greater detail.

Table 3.15: Results: End of High School (No Controls)

	Mean of Depvar		OLS	Level 2/3 Cutoff		Level 3/4 Cutoff	
	L 2/3	L 3/4		RF	2SLS	RF	2SLS
End of High School:							
Graduated	0.82	0.91	0.00	-0.00	-0.00	-0.00	-0.05
			(0.00)	(0.01)	(0.23)	(0.00)	(0.06)
GPA (unweighted)	2.40	3.10	0.13***	0.00	0.13	0.00	0.07
			(0.00)	(0.01)	(0.28)	(0.01)	(0.12)
Post-High School:							
College Intent: 2yr	0.53	0.33	-0.10***	0.01	0.17	-0.01	-0.18
			(0.00)	(0.01)	(0.25)	(0.01)	(0.12)
College Intent: 4yr	0.27	0.57	0.12***	-0.00	-0.13	0.01*	0.20*
			(0.00)	(0.00)	(0.12)	(0.01)	(0.12)
STEM Major Intent	0.17	0.23	0.02***	-0.01	-0.11	0.01***	0.33***
			(0.00)	(0.01)	(0.22)	(0.01)	(0.13)

Notes: Table reports estimates of the effect of 8th grade Algebra on end-of-high school outcomes, without controls. Dependent variables are indicators for graduation (row 1), 2-year college intentions (row 3), 4-year college intentions (row 4), STEM major intentions (row 5), and cumulative, unweighted high school GPA (row 2). Columns (1) and (2) report the mean of the dependent variable within 2 scale score points below the 7th grade level III and level IV proficiency levels, respectively. Column (3) reports estimates of the effect using observational variation in 8th grade Algebra enrollment using the entire sample. Columns (4) and (6) report reduced form estimates of β^{RF} from local linear regressions corresponding to equation (3.4) (excluding control variables). Columns (5) and (7) report RD estimates of β using 2SLS, corresponding to equation (3.5) (excluding control variables). Columns (4) and (5) correspond to the level III threshold; columns (6) and (7) correspond to the level IV threshold. All RF and 2SLS effects are estimated within a bandwidth of 5 scale score points around the relevant proficiency threshold. Standard errors are two-way clustered by school and test score-by-exam year. See Sections 3.4 and 3.5 in the text for greater detail.

Bibliography

[1] Atila Abdulkadiroğlu, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters. “Do Parents Value School Effectiveness?” *Working Paper, National Bureau of Economic Research* (2017).

[2] Clifford Adelman. “The toolbox revisited: Paths to degree completion from high school through college.” *US Department of Education* (2006).

[3] Debbie Alexander and Laurie Lewis. “Condition of America’s Public School Facilities: 2012-13. First Look. NCES 2014-022.” *National Center for Education Statistics* (2014).

[4] Joshua D Angrist, Guido W Imbens, and Donald B Rubin. “Identification of causal effects using instrumental variables”. *Journal of the American statistical Association* 91.434 (1996), pp. 444–455.

[5] David Arsen and Thomas Davis. “Taj Mahals of decaying shacks: Patterns in local school capital stock and unmet capital need”. *Peabody Journal of Education* 81.4 (2006), pp. 1–22.

[6] Alison Aughinbaugh. “The effects of high school math curriculum on college attendance: Evidence from the NLSY97”. *Economics of Education Review* 31.6 (2012), pp. 861–870.

[7] Jushan Bai. “Estimation of a change point in multiple regression models”. *Review of Economics and Statistics* 79.4 (1997), pp. 551–563.

[8] Lisa Barrow. “School choice through relocation: evidence from the Washington, DC area”. *Journal of Public Economics* 86.2 (2002), pp. 155–189.

[9] Lisa Barrow and Cecilia Elena Rouse. “Using market valuation to assess public school spending”. *Journal of Public Economics* 88.9 (2004), pp. 1747–1769.

[10] Patrick Bayer, Fernando Ferreira, and Robert McMillan. “A unified framework for measuring preferences for schools and neighborhoods”. *Journal of Political Economy* 115.4 (2007), pp. 588–638.

[11] Christopher Berry and Charles Wysong. “Making Courts Matter: Politics and the Implementation of State Supreme Court Decisions”. *The University of Chicago Law Review* 79.1 (2012), pp. 1–29.

[12] Julian R Betts. "The economics of tracking in education". *Handbook of the Economics of Education*. Vol. 3. Elsevier, 2011, pp. 341–381.

[13] Barbara Biasi. "School Finance Equalization and Intergenerational Mobility: A Simulated Instruments Approach" (2017).

[14] Sandra E Black. "Do better schools matter? Parental valuation of elementary education". *The Quarterly Journal of Economics* 114.2 (1999), pp. 577–599.

[15] Alex J Bowers and Angela Urick. "Does high school facility quality affect student achievement? A two-level hierarchical linear model". *Journal of Education Finance* 37.1 (2011), pp. 72–94.

[16] Jan K Brueckner. "Property values, local public expenditure and economic efficiency". *Journal of Public Economics* 11.2 (1979), pp. 223–245.

[17] Jack Buckley, Mark Schneider, and Yi Shang. "The Effects of School Facility Quality on Teacher Retention in Urban School Districts." *National Clearinghouse for Educational Facilities* (2004).

[18] Carol Corbett Burris and Delia T. Garrity. *Detracking for excellence and equity*. Association for Supervision and Curriculum Development, 2008.

[19] Gary Burtless. *Does money matter? The effect of school resources on student achievement and adult success*. Brookings Institution Press, 1996.

[20] Christopher A Candelaria and Kenneth A Shores. "The Sensitivity of Causal Estimates from Court-Ordered Finance Reform on Spending and Graduation Rates". *Center for Education Policy Analysis Working Paper* 16-05 (2015).

[21] David Card and Laura Giuliano. "Can tracking raise the test scores of high-ability minority students?" *American Economic Review* 106.10 (2016), pp. 2783–2816.

[22] David Card and Alan B Krueger. "Does school quality matter? Returns to education and the characteristics of public schools in the United States". *Journal of Political Economy* 100.1 (1992), pp. 1–40.

[23] David Card and Alan B Krueger. "School quality and black-white relative earnings: A direct assessment". *The Quarterly Journal of Economics* 107.1 (1992), pp. 151–200.

[24] David Card, Alexandre Mas, and Jesse Rothstein. "Tipping and the Dynamics of Segregation". *The Quarterly Journal of Economics* 123.1 (2008), pp. 177–218.

[25] David Card and A Abigail Payne. "School finance reform, the distribution of school spending, and the distribution of student test scores". *Journal of public economics* 83.1 (2002), pp. 49–82.

[26] Elizabeth U Cascio, Nora Gordon, and Sarah Reber. "Local responses to federal grants: Evidence from the introduction of title I in the South". *American Economic Journal: Economic Policy* 5.3 (2013), pp. 126–59.

[27] Elizabeth U Cascio and Sarah Reber. “The poverty gap in school spending following the introduction of Title I”. *American Economic Review* 103.3 (2013), pp. 423–27.

[28] Stephanie Riegg Cellini, Fernando Ferreira, and Jesse Rothstein. “The value of school facility investments: Evidence from a dynamic regression discontinuity design”. *The Quarterly Journal of Economics* 125.1 (2010), pp. 215–261.

[29] Raj Chetty, John N. Friedman, and Jonah E. Rockoff. “Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates”. *American Economic Review* 104.9 (Sept. 2014), pp. 2593–2632.

[30] Raj Chetty, John N Friedman, et al. “How does your kindergarten classroom affect your earnings? Evidence from Project STAR”. *The Quarterly Journal of Economics* 126.4 (2011), pp. 1593–1660.

[31] Melissa A Clark. *Education reform, redistribution, and student achievement: Evidence from the Kentucky Education Reform Act*. Mathematica Policy Research, NJ, 2003.

[32] Charles T Clotfelter, Helen F Ladd, and Jacob L Vigdor. *Algebra for 8th graders: Evidence on its effects from 10 North Carolina districts*. Tech. rep. National Bureau of Economic Research, 2012.

[33] Charles T Clotfelter, Helen F Ladd, and Jacob L Vigdor. “The aftermath of accelerating algebra evidence from district policy initiatives”. *Journal of Human Resources* 50.1 (2015), pp. 159–188.

[34] James S Coleman et al. “Equality of Education Opportunity Study, Washington, DC: US Department of Health, Education, and Welfare, Office of Education”. *National Center for Education Statistics* (1966).

[35] Courtney A Collins and Erin K Kaplan. “Capitalization of School Quality in Housing Prices: Evidence from Boundary Changes in Shelby County, Tennessee”. *American Economic Review: Papers and Proceedings* 107.5 (2017), pp. 628–632.

[36] Michael Conlin and Paul N Thompson. “Impacts of New School Facility Construction: An Analysis of a State-Financed Capital Subsidy Program in Ohio”. *Economics of Education Review* (2017).

[37] John E Coons, William H Clune, and Stephen D Sugarman. “Private wealth and public education” (1970).

[38] Sean P Corcoran and William N Evans. “Equity, adequacy, and the evolving state role in education finance”. *Handbook of research in education finance and policy* 332 (2015).

[39] Kalena E Cortes and Joshua S Goodman. “Ability-tracking, instructional time, and better pedagogy: The effect of double-dose algebra on student achievement”. *American Economic Review: Papers & Proceedings* 104.5 (2014), pp. 400–405.

[40] Faith E Crampton, David C Thompson, and Janis M Hagey. "Creating and sustaining school capacity in the twenty-first century: Funding a physical environment conducive to student learning". *Journal of Education Finance* 27.2 (2001), pp. 633–652.

[41] Thurston Domina, Andrew McEachin, Andrew Penner, and Emily Penner. "Aiming high and falling short: California's eighth-grade algebra-for-all effort". *Educational Evaluation and Policy Analysis* 37.3 (2015), pp. 275–295.

[42] Thurston Domina, Andrew M Penner, Emily K Penner, and AnneMarie Conley. "Does Detracking Work? Evidence from a Mathematics Curricular Reform." *Society for Research on Educational Effectiveness* (2012).

[43] Shaun M Dougherty, Joshua S Goodman, Darryl V Hill, Erica G Litke, and Lindsay C Page. "Objective course placement and college readiness: Evidence from targeted middle school math acceleration". *Economics of Education Review* 58 (2017), pp. 141–161.

[44] Thomas Downes and Leanna Stiefel. "Measuring equity and adequacy in school finance". *Handbook of Research in Education Finance and Policy* (2015), pp. 222–236.

[45] Esther Duflo. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment". *American Economic Review* 91.4 (Sept. 2001), pp. 795–813.

[46] Esther Duflo. "The medium run effects of educational expansion: Evidence from a large school construction program in Indonesia". *Journal of Development Economics* 74.1 (2004), pp. 163–197.

[47] William Duncombe and John Yinger. "Measurement of cost differentials". *Handbook of Research in Education Finance and Policy* (2015).

[48] Susan Dynarski, Joshua Hyman, and Diane Whitmore Schanzenbach. "Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion". *Journal of Policy Analysis and Management* 32.4 (2013), pp. 692–717.

[49] David N Figlio and Maurice E Lucas. "What's in a grade? School report cards and the housing market". *The American Economic Review* 94.3 (2004), pp. 591–604.

[50] Mary Filardo. "State of Our Schools: America's K–12 Facilities 2016". *Washington, DC: 21st Century School Fund* (2016).

[51] Mary W Filardo, Jeffrey M Vincent, Ping Sung, and Travis Stein. "Growth and Disparity: A Decade of US Public School Construction." *21st Century School Fund* (2006).

[52] Ann Flanagan and Sheila Murray. "A decade of reform: The impact of school reform in Kentucky". *Helping children left behind: State aid and the pursuit of educational equity* (2004), pp. 165–213.

[53] Bruce Fuller et al. “Building schools, rethinking quality? Early lessons from Los Angeles”. *Journal of Educational Administration* 47.3 (2009), pp. 336–349.

[54] Adam Gamoran and Eileen C Hannigan. “Algebra for everyone? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school”. *Educational Evaluation and Policy Analysis* 22.3 (2000), pp. 241–254.

[55] Felipe Goncalves. “The Effects of School Construction on Student and District Outcomes: Evidence from a State-Funded Program in Ohio” (2015).

[56] Angel Gurria. “Pisa 2015 results in focus”. *PISA in Focus* 67 (2016), p. 1.

[57] Jonathan Guryan. *Does money matter? Regression-discontinuity estimates from education finance reform in Massachusetts*. Tech. rep. National Bureau of Economic Research, 2001.

[58] Eric A Hanushek. “Assessing the effects of school resources on student performance: An update”. *Educational Evaluation and Policy Analysis* 19.2 (1997), pp. 141–164.

[59] Eric A Hanushek. “School resources”. *Handbook of the Economics of Education* 2 (2006), pp. 865–908.

[60] Eric A Hanushek. “The economics of schooling: Production and efficiency in public schools”. *Journal of Economic Literature* 24.3 (1986), pp. 1141–1177.

[61] Eric A Hanushek. “The failure of input-based schooling policies”. *The Economic Journal* 113.485 (2003).

[62] Eric A Hanushek, John F Kain, and Steven G Rivkin. “Disruption versus Tiebout improvement: The costs and benefits of switching schools”. *Journal of Public Economics* 88.9 (2004), pp. 1721–1746.

[63] Eric A Hanushek and Alfred A Lindseth. *Schoolhouses, courthouses, and statehouses: Solving the funding-achievement puzzle in America’s public schools*. Princeton University Press, 2009.

[64] Eric A Hanushek, Steven G Rivkin, and Lori L Taylor. *Aggregation and the estimated effects of school resources*. Tech. rep. National Bureau of Economic Research, 1996.

[65] Eric A Hanushek, Steven G Rivkin, and Lori L Taylor. “The identification of school resource effects”. *Education Economics* 4.2 (1996), pp. 105–125.

[66] Eric A Hanushek and Ludger Woessmann. *The knowledge capital of nations: Education and the economics of growth*. MIT Press, 2015.

[67] Kai Hong and Ron Zimmer. “Does Investing in School Capital Infrastructure Improve Student Achievement?” *Economics of Education Review* 53 (2016), pp. 143–158.

[68] Richard Hornbeck and Daniel Keniston. “Creative Destruction: Barriers to Urban Growth and the Great Boston Fire of 1872”. *American Economic Review* 107.6 (June 2017), pp. 1365–98.

[69] Harold W Horowitz. "Unseparate but unequal:The emerging fourteenth amendment issue in public school education". *UCLA Law Review* 13 (1965), p. 1147.

[70] Caroline M Hoxby. "All school finance equalizations are not created equal". *The Quarterly Journal of Economics* 116.4 (2001), pp. 1189–1231.

[71] Joshua Hyman. "Does Money Matter in the Long Run? Effects of School Spending on Educational Attainment". *American Economic Journal: Economic Policy* (Forthcoming).

[72] Guido Imbens and Karthik Kalyanaraman. "Optimal bandwidth choice for the regression discontinuity estimator". *The Review of economic studies* 79.3 (2012), pp. 933–959.

[73] Scott A Imberman and Michael F Lovenheim. "Does the market value value-added? Evidence from housing prices after a public release of school and teacher value-added". *Journal of Urban Economics* 91 (2016), pp. 104–121.

[74] C Kirabo Jackson, Rucker C Johnson, and Claudia Persico. "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms". *The Quarterly Journal of Economics* 131.1 (2016), pp. 157–218.

[75] Brian Jacob and Jesse Rothstein. "The Measurement of Student Ability in Modern Assessment Systems". *Journal of Economic Perspectives* 30.3 (Sept. 2016), pp. 85–108.

[76] Thomas J Kane and Douglas O Staiger. *Estimating teacher impacts on student achievement: An experimental evaluation*. Tech. rep. National Bureau of Economic Research, 2008.

[77] David Kirp. "The poor, the schools, and equal protection". *Harvard Educational Review* 38.4 (1968), pp. 635–668.

[78] William S Koski and Jesse Hahnel. "The past, present, and possible futures of educational finance reform litigation". *Handbook of Research in Education Finance and Policy* (2015).

[79] Alan B Krueger. "Economic considerations and class size". *The Economic Journal* 113.485 (2003).

[80] Alan B Krueger. "Experimental estimates of education production functions". *The Quarterly Journal of Economics* 114.2 (1999), pp. 497–532.

[81] Alan B Krueger and Diane M Whitmore. *Rich Schools, Poor Schools: The Promise of Equal Educational Opportunity*. University of Chicago Press, Chicago IL, 1967.

[82] Alan B Krueger and Diane M Whitmore. *Would smaller classes help close the black-white achievement gap?* Vol. 451. Industrial Relations Section, Princeton University, 2001.

[83] Helen F Ladd and Edward B Fiske. *Handbook of research in education finance and policy*. Routledge, 2015.

[84] Julien Lafortune, Jesse Rothstein, and Diane Whitmore Schanzenbach. *School Finance Reform and the Distribution of Student Achievement*. Tech. rep. National Bureau of Economic Research, Inc, 2016.

[85] Julien Lafortune, Jesse Rothstein, and Diane Whitmore Schanzenbach. "School finance reform and the distribution of student achievement". *American Economic Journal: Applied Economics* (Forthcoming).

[86] Edward P. Lazear. "Educational Production". *The Quarterly Journal of Economics* 116.3 (2001), pp. 777–803.

[87] David S Lee and David Card. "Regression discontinuity inference with specification error". *Journal of Econometrics* 142.2 (2008), pp. 655–674.

[88] David S Lee and Thomas Lemieux. "Regression discontinuity designs in economics". *Journal of economic literature* 48.2 (2010), pp. 281–355.

[89] Tom Loveless. *The 2013 Brown Center Report on American Education: How Well are American Students Learning?* Vol. 2. 3. Brown Center on Education Policy at The Brookings Institution, 2013.

[90] Paco Martorell, Kevin Stange, and Isaac McFarlin. "Investing in schools: capital spending, facility conditions, and student achievement". *Journal of Public Economics* 140 (2016), pp. 13–29.

[91] Andrew McEachin, Thurston Domina, and Andrew M Penner. *Understanding the Effects of Middle School Algebra: A Regression Discontinuity Approach*. RAND Corporation, 2017.

[92] Joel McFarland et al. "The Condition of Education 2017. NCES 2017-144." *National Center for Education Statistics* (2017).

[93] Steven C McMullen and Kathryn E Rouse. "School crowding, year-round schooling, and mobile classroom use: Evidence from North Carolina". *Economics of Education Review* 31.5 (2012), pp. 812–823.

[94] Sheila E Murray, William N Evans, and Robert M Schwab. "Education-finance reform and the distribution of education resources". *American Economic Review* (1998), pp. 789–812.

[95] National Science Board. "Preparing the next generation of stem innovators: Identifying and developing our nation's human capital". *National Science Foundation* (2010).

[96] Christopher A Neilson and Seth D Zimmerman. "The effect of school construction on test scores, school enrollment, and home prices". *Journal of Public Economics* 120 (2014), pp. 18–31.

[97] John Ries and Tsur Somerville. "School quality and residential property values: evidence from Vancouver rezoning". *The Review of Economics and Statistics* 92.4 (2010), pp. 928–944.

[98] Stuart J Ritchie and Timothy C Bates. “Enduring links from childhood mathematics and reading achievement to adult socioeconomic status”. *Psychological science* 24.7 (2013), pp. 1301–1308.

[99] J Rogers et al. *California educational opportunity report: Listening to public school parents.* 2009.

[100] Paul A Samuelson. “The pure theory of public expenditure”. *The Review of Economics and Statistics* 36.4 (1954), pp. 387–389.

[101] Melissa Sands. “The Distributive Politics of Education Policy: Party Control of State Government and Transfers to Localities” (2015).

[102] David P Sims. “Lifting all boats? Finance litigation, education resources, and student needs in the post-Rose era”. *Education Finance and Policy* 6.4 (2011), pp. 455–485.

[103] David P Sims. “Suing for your supper? Resource allocation, teacher compensation and finance lawsuits”. *Economics of Education Review* 30.5 (2011), pp. 1034–1044.

[104] Julia B Smith. “Does an extra year make any difference? The impact of early access to algebra on long-term gains in mathematics attainment”. *Educational Evaluation and Policy Analysis* 18.2 (1996), pp. 141–153.

[105] Mary Kay Stein, Julia Heath Kaufman, Milan Sherman, and Amy F Hillen. “Algebra: A challenge at the crossroads of policy and practice”. *Review of Educational Research* 81.4 (2011), pp. 453–492.

[106] Eric Taylor. “Spending more of the school day in math class: Evidence from a regression discontinuity in middle school”. *Journal of Public Economics* 117 (2014), pp. 162–181.

[107] Richard R. Terzian. *Recommendations for Improving the School Facility Program in Los Angeles Unified School District.* Little Hoover Commission, 1999.

[108] Cynthia Uline and Megan Tschannen-Moran. “The walls speak: The interplay of quality facilities, school climate, and student achievement”. *Journal of Educational Administration* 46.1 (2008), pp. 55–73.

[109] Catherine J Weinberger. “The increasing complementarity between cognitive and social skills”. *Review of Economics and Statistics* 96.4 (2014), pp. 849–861.

[110] William Welsh, Erin Coghlan, Bruce Fuller, and Luke Dauter. “New Schools, Overcrowding Relief, and Achievement Gains in Los Angeles–Strong Returns from a $19.5 Billion Investment. Policy Brief 12-2.” *Policy Analysis for California Education, PACE (NJ1)* (2012).