

IZA DP No. 6106

## The Effect of School Construction on Test Scores, School Enrollment, and Home Prices

Christopher Neilson  
Seth Zimmerman

November 2011

# The Effect of School Construction on Test Scores, School Enrollment, and Home Prices

**Christopher Neilson**  
*Yale University*

**Seth Zimmerman**  
*Yale University  
and IZA*

Discussion Paper No. 6106  
November 2011

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### The Effect of School Construction on Test Scores, School Enrollment, and Home Prices<sup>\*</sup>

This paper provides new evidence on the effect of school construction projects on home prices, academic achievement, and public school enrollment. Taking advantage of the staggered implementation of a comprehensive school construction project in a poor urban district, we find that, by six years after building occupancy, \$10,000 of per-student investment in school construction raised reading scores for elementary and middle school students by 0.027 standard deviations. For a student receiving the average treatment intensity this corresponds to a 0.21 standard deviation increase. School construction also raised home prices and public school enrollment in zoned neighborhoods.

JEL Classification: I21, I22, H75, R30

Keywords: school construction, test scores, home prices

Corresponding author:

Seth Zimmerman  
Department of Economics  
Yale University  
37 Hillhouse Ave.  
New Haven, CT 06520  
USA  
E-mail: [seth.zimmerman@yale.edu](mailto:seth.zimmerman@yale.edu)

---

<sup>\*</sup> The authors wish to thank the New Haven Public Schools (NHPS) for access to the student data used in this study. We are particularly indebted to William Clark, Laoise King, Catherine McCaslin, and Garth Harries. We thank Daniel Haim of NHPS and Michael DePalma and Robert Lynn of Gilbane Inc. for details on the school construction projects described here, and Kevin Moriarty for additional data assistance. We thank Joseph Altonji, Prashant Bharadwaj, Raj Chetty, Justine Hastings, Amanda Kowalski, Fabian Lange, Richard Mansfield, Alexandre Mas and seminar participants at the Yale University Labor Workshop for their comments and suggestions. All errors are our own. The research reported here was not the result of a for-pay consulting relationship.

# 1 Motivation

## 1.1 Motivation and summary

Investment in school infrastructure is one of the principle ways in which federal, state, and local governments develop physical capital in U.S. communities. In 2008, public expenditures on school construction, land, and building acquisition totaled more than \$58 billion. Of this amount, between \$20 and \$40 billion was spent on the construction of approximately 2,000 new school buildings.<sup>1</sup> This represents not just a large share of total education-related expenditures—roughly ten percent— but a large share of overall infrastructure expenditure: in 2004, public investment in school infrastructure was \$75.9 billion, nearly as much as the \$99.7 billion public investment in all forms of transportation infrastructure, including roads, mass transit, and aviation.<sup>2</sup>

In this paper, we present new evidence on the causal effect of school infrastructure investment on student test scores, neighborhood-specific public school enrollment, and housing prices. We take advantage of a unique natural experiment in which a poor, urban school district embarked upon a comprehensive 15-year, \$1.4 billion school construction program (believed to be the largest per-capita construction program in the nation over the period) to produce estimates that are unbiased by the endogeneity of school construction to school characteristics. Our empirical strategy uses the fact that occupancy dates varied widely across schools, with the first school completed in 1998 and the last slated to be completed in 2014. Specifically, we use a difference-in-differences comparison of test scores, home prices, and public school attendance in neighborhoods on different construction schedules. To the best of our knowledge, this is the first paper to establish the link between school construction and changes in test scores and enrollment patterns.

We find strong evidence that the school construction program led to sustained gains in reading scores for elementary and middle school students. Trends in reading scores are flat in the years leading up to construction, but turn upwards in the year of construction and continue to increase for at least the next six years. By the sixth year following the year of construction, student scores rise by 0.027 standard deviations for each \$10,000 of per-student construction expenditure. For a student receiving the average level of construction

---

<sup>1</sup>Estimates on expenditures vary across sources. See report from the [NCEF](#) - National Clearinghouse for Educational Facilities and Public Education Finances 2008.

<sup>2</sup>Source: [CBO](#) – Congressional Budget Office 2008.

expenditure, this corresponds to a total score gain of 0.21 standard deviations. These gains are large, but not implausibly so; roughly speaking, they are of similar magnitude to those experienced by students who enroll in high-performing charter schools.<sup>3</sup> There is weak evidence of a corresponding increase in math scores.

Housing prices and neighborhood public school enrollment also respond positively to school construction. Elementary and middle school construction raised home values by 1.3 percent per \$10,000 of per-student expenditure, and the number of school zone residents attending public school by up to 4.4 percent per \$10,000. As with the estimated reading score effects, the timing of these changes coincides with the occupancy of completed buildings. Taken together, our student outcome, home price, and enrollment results suggest that families, and in particular families with children, place a high value on school infrastructure investment. If families only valued infrastructure insofar as it improved education production, this would imply that raising school value added by 0.1 standard deviations would raise neighborhood home prices by 4.7 percent, and enrollment of neighborhood residents in public schools by 16.2 percent. Since school construction also changes neighborhood amenities in other ways, these values should be interpreted as upper bounds on the true elasticities.

Our findings beg the question of *why* school construction raises test scores. Possible pathways include the direct effects of new facilities on pedagogy, effects on student and teacher motivation during school hours, and effects on students' and parents' motivation to invest in academic production at home. Though our empirical work does not allow us to distinguish between these channels, a survey of district principals indicates that student and teacher motivation were at least as important as direct pedagogical effects for improving academic outcomes.

## **1.2 Contributions to the literature**

We build upon and link two distinct strands of literature. The first deals directly with the effects of school infrastructure investment. There are few compelling estimates of how infrastructure expenditures affect student performance. In a review of literature on the

---

<sup>3</sup>Lottery-based evaluations of the effects of attending high-performing KIPP charter schools suggest effects of between 0.12 standard deviations per year of enrollment for reading and 0.35 standard deviations for math (Angrist et al. 2011). Between 6th and 8th grade, students enrolled in Harlem Children's Zone (HCZ) schools experience reading score gains of one quarter to third of a standard deviation and math score gains of 0.8 standard deviations relative to HCZ non-participants (Dobbie and Fryer 2010).

education production function, Hanushek (1997) reports that of 91 correlational studies that examine the relationship between facility quality and student performance, only nine percent found a statistically significant positive relationship, while five percent found a statistically significant negative relationship.<sup>4</sup> Since facility quality is closely associated with other observable and unobservable inputs into education production, the absence of consistent findings is difficult to interpret. Duflo (2001) uses a difference-in-differences approach to obtain plausibly causal estimates of the effects of a large Indonesian school construction program on educational attainment and labor market outcomes. Because her paper focuses on the construction of schools where none had existed before, there is little reason to think results would apply in the US, where the main challenges are those of renovation and rebuilding.

Cellini, Ferreira and Rothstein (2010), henceforth CFR, provide estimates that are both plausibly causal and relevant to US policy. CFR employ a regression discontinuity around the outcomes of school district-level votes on the bond issues used to finance school construction projects to estimate the effects of school construction spending in California on home prices and test scores. They find that a dollar of expenditure on school facilities raises home prices by roughly \$1.50, leading to an average home price increase of between 7 and 10 percent by six years after bond passage. They do not find strong effects for student test scores. CFR's research design has two important limitations. First, CFR use district-level third and fourth grade test scores to measure score effects. Since district-level expenditures are a noisy measure of the expenditures we would expect to improve outcomes for third and fourth graders (e.g., expenditures on elementary schools as opposed to high schools), it is likely that their estimates of test score effects are biased downward towards zero. Second, the cost of identification via regression discontinuity is that estimates cannot necessarily be extrapolated to districts that are not on the electoral margin of bond passage. If electoral outcomes are a function of residents' beliefs about the benefits of school construction, effects in marginal districts will likely differ systematically from effects for districts in which bonds pass or fail by a comfortable amount. In sum, CFR show convincingly that the residents of electorally marginal districts value school construction, but questions remain about test score effects generally, and about home price and test score effects for different types of infra-marginal districts. Of particular interest are poor urban districts, because these districts are frequent targets of policy interventions aimed at improving school qual-

---

<sup>4</sup>23 percent reported statistically insignificant positive relationships, 19 percent reported statistically significant negative relationships, and 44 percent reported relationships of indeterminate sign.

ity and also tend to have the low-quality existing facilities.<sup>5</sup> We establish for the first time a relationship between school construction and test scores, and do so in the context of a poor urban district with baseline facility quality similar to that in other urban districts in the state.

The second strand of literature examines the way housing markets price the quality of local schools. Black (1999) uses discontinuities in the prices of homes on the borders between school districts to estimate the price effects of differences between school-average test scores. Bayer et al. (2007) nests this identification strategy within a model of housing demand and makes the observation that much of the observed price gap is attributable to endogenous socioeconomic segregation along district boundaries, not to test scores *per se*. One implication of the finding that school quality has an independent effect on the housing market is that changes in school quality should set off a dynamic process of residential sorting and changes in home prices. This is how the socioeconomic stratification along zone boundary lines reported in Bayer et al. comes into existence. But because both Black and Bayer et al. estimate static models of housing demand, they cannot observe this process as it unfolds. We use panel data on home prices, public school enrollment, and test scores to document dynamic changes in education production, home prices, and residency patterns in response to the school construction intervention. Our results indicate that changes in home prices, school enrollment among neighborhood residents, and education production all begin at the time of occupancy. Housing markets do not appear to ‘price in’ construction projects in advance of building occupancy. School construction dramatically raises school attendance rates among school zone residents, and students who move to neighborhoods following school construction have observable characteristics associated with higher test score performance. Our findings on sorting into school zones are consistent with the socioeconomic stratification observed in Bayer et al.

We also add to this literature by examining price responses to changes in education production as opposed to aggregate school scores. One limitation of Black and Bayer et al. is that both papers use average scores within school attendance zones as their measure of

---

<sup>5</sup>Filardo (2006) found that rates of investment in new infrastructure were twice as high in rich urban districts as in poor urban districts between 1995 to 2004. Evidence of differential investment can be seen in the heterogeneity of infrastructure quality across schools: 43 percent of schools in which 75 or more percent of students are eligible for free lunch use portable buildings as classrooms, in contrast with 27 percent of schools in which less than 35 percent of students are free-lunch eligible (Source: NCES Digest of Education Statistics : Table 101.) If the marginal returns to investment in infrastructure are decreasing in the quality of existing infrastructure, then poor districts will benefit disproportionately from school construction.

school quality. Zone-level averages represent a mix of the causal effect of zoned schools (i.e., education production) on scores for students living in the neighborhood and the underlying observable and unobservable test score determinants of the students attending school within the zone. Parents trying to optimize over education production will be interested in the causal portion of the average but not the portion that is tied to selection. We use the school construction natural experiment to identify changes in education production at the school zone level, and link these changes to increases in home prices. Because school construction may make neighborhoods more desirable in other ways, our estimates should be interpreted only as upper bounds on the true elasticities of home prices with respect to education production. We believe they nevertheless constitute an important first attempt at pricing educational production.

The paper proceeds as follows. In [section 2](#), we describe the school district and the school construction program. In [section 3](#), we present a theoretical model that we use to guide and interpret our empirical specifications. In [section 4](#), we describe the student and home price data we use to conduct our analysis. Sections 6 through 8 present our empirical findings on the home price, residency, and test score effects of school construction. Section 9 discusses the impact of school construction on student behavior, and describes the results of our survey of school principals. Section 10 concludes.

## **2 The natural experiment**

### **2.1 The school district**

Our project focuses on the public school system in New Haven, Connecticut, which we will refer to as NHPS or the District. New Haven is one of the largest districts in Connecticut and is similar to many urban school districts in the United States. The students mostly come from poor families and overwhelmingly belong to minority groups that have traditionally lagged in educational outcomes such as graduation rates and test scores. The District has an enrollment of approximately 22,000 students, of whom more than 80 percent are eligible for free lunch and approximately 90 percent are either black or Hispanic. One out of four students speaks a language other than English at home. High school dropout rates are triple the state average and test scores are substantially lower than those in the rest of the state. [Table 1](#) summarizes the demographics of the students in the district and contrasts them to state averages in 2009.

In Connecticut, poor urban districts typically have infrastructure of lower quality than other districts. Early in the school construction program, NHPS' facilities were roughly comparable to those of other in-state urban districts. In 2001, public schools in New Haven and Hartford, another urban district in Connecticut, were on average well over 50 years old, and in both cities more than half of schools reported problems related to basic service systems such as heating, air conditioning, plumbing, and lighting, compared to less than one third of schools in Connecticut as a whole.<sup>6</sup>

*Table 1: Student Demographics and Academics in New Haven and Connecticut as a whole*

	NHPS	Connecticut
Hispanic	34.5%	16.5%
Black	51.3%	13.8%
White	12.0%	65.0%
Free Lunch	80.5%	32.9%
English Language Learns	11.9%	5.3%
English Not Spoken at Home	25.5%	13.1%
At Goal or Proficient (CMT)	43.0%	69.0%

Source: Connecticut Department of Education.

CMT : Connecticut Mastery Test is a standardized test score given to all children in grades 3 to 8.

## 2.2 The school construction project

In contrast to many urban districts, NHPS has had the political and financial backing to enact an ambitious infrastructure investment program. An important contributor to the successful execution of this project was the availability of federal and state financing: the District paid for only 23 percent of the total cost of buildings completed by 2010. The School Construction Project (SCP) had a dramatic affect on primary- and secondary-school infrastructure across the city. The first SCP school was completed in 1998, and the last is

<sup>6</sup>The Connecticut Department of Education collects information on school infrastructure by surveying the school principals. Table A-I shows the frequency with which principals rated service items such as heating, plumbing and air conditioning as either fair or poor in New Haven, Hartford, and Connecticut as a whole in 2001 and 2009.

scheduled to be completed by 2014. Projected total spending is \$1.4 billion, with \$1.1 billion spent on projects that had been completed by 2010. Of 42 school buildings,<sup>7</sup> 12 had been rebuilt completely by 2010, and 18 had been renovated. An additional seven were under construction or under design. The remaining five buildings, all of which house inter-district magnet or small K-1 schools, will not be rebuilt or renovated. School renovations were generally substantial: mean expenditure on renovated schools was \$33 million, compared to a mean expenditure of \$38 million on the rebuilt schools (all dollar values refer to 2005 dollars). The project served students at all educational levels: of nine high schools in the district, five had been rebuilt or renovated and occupied by 2010, with an additional high school in the construction stage. Similarly, of 33 total elementary or middle schools, 25 had been rebuilt or renovated and occupied by 2010, with work on an additional six in the planning or construction phase. The top two panels of **Table A -2** describe the scope of the SCP.

Though the changes made to schools varied depending on the condition of the existing school, SCP administrators did target a number of areas for improvement at all schools. One priority was heating and air conditioning. Prior to the SCP, many schools did not have air conditioning, and some had inadequate heating. A second was classroom technology. Classrooms in new and renovated schools were designed to facilitate the use of computers, and science and media facilities for school-wide use were also improved. A third was community access. SCP administrators designed gyms, playgrounds, and meeting spaces to allow for use by community members as well as students. A fourth was to decrease energy and maintenance costs. A fifth and slightly more abstract goal of the SCP was to make schools more ‘livable’ through subtler changes in design. The design of new school buildings often allowed for more natural light than in the old buildings, and a portion of the budget for each school was allocated to public artwork. For a more detailed description of several of the school construction projects, see the Supplemental Appendix.

School expansion was not among the primary goals of the SCP. In fact, the SCP took place in a time of declining overall demand for classroom space in the District: our data show that District enrollment has been in decline since at least the 2002-2003 school year, before which only four projects were completed. Nor did the SCP seek to change the allocation of students across the district. There were no changes in school zone boundaries over the

---

<sup>7</sup>This count omits charter schools and transitional schools for at-risk youth, and counts each address separately for schools with multiple addresses.

period. One consequence of the neutrality of the SCP with respect to enrollment was that new and renovated buildings typically did not offer much more classroom space than the facilities they replaced. Among the sixteen projects for which we were able to recover pre- and post-completion classroom square footage, the median change was less than seven percent. Classroom space fell by 6.5 percent in the District's best-performing school.

The SCP had a pronounced effect on the quality of the school environment in the District. One way to see this is to track changes in the quality of District schools and compare them to changes at the state level and in other urban districts. Between 2001 and 2009, the percentage of schools reporting failures in basic services like air conditioning and lighting fell from 32 to 18 percent at the state level, and from 54 to 30 percent in Hartford, another poor district in Connecticut. The percentage of NHPS schools reporting such failures fell from 53 to 14 percent. The SCP pushed the quality of NHPS infrastructure from far below the state average to somewhat above it.<sup>8</sup>

### 2.3 Selection of schools in the SCP

Our goal is to identify the effects of school construction on community and student outcomes using a difference-in-differences approach. It is therefore of critical importance for the internal validity of our analysis that school construction not be a *result* of student or community choices. For instance, if schools were chosen for renovation based on projected performance on state tests, our estimates of the effects of construction would likely be biased. Our results would be similarly compromised if the SCP only placed schools in neighborhoods where crime was already dropping.

Fortunately, discussions with SCP administrators and empirical investigation indicate that the process by which schools were chosen was largely exogenous to community-specific factors. First, it is important to note that the comprehensive nature of the SCP rendered the question of *which* schools should be renovated irrelevant; instead, the key question administrators faced was how to choose the order of construction. SCP administrators have stated to the authors that, with the exception of the first few schools, the determinants of construction order were primarily logistical and design hurdles, not community or student characteristics or demands. This claim is consistent with what we observe in the data. As shown in [Figure 1](#), there is no discernible geographic pattern to the order in which schools

---

<sup>8</sup>See [Table A -1](#).

were constructed. Nor do schools built or renovated in the early phases of the SCP differ from schools constructed in the project’s later phases in terms of student demographics. **Table 2** compares students in schools constructed prior to 2006 (the approximate midpoint of the project in terms of completed buildings) to students in schools constructed after 2006. There are no statistically or economically significant differences between the characteristics of students in schools constructed in the first half of the project and the second.<sup>9</sup> In short, the timing of school construction does not appear to have been endogenous to student and community characteristics.

A related question is whether school construction projects coincided with other school- or community-level interventions. If a separate community center were built next to every newly constructed school, it would not be possible to separate the effects of one from the effects of the other using our methodology. It is difficult to conclusively verify the absence of additional interventions, but district officials were not aware of any such policies.

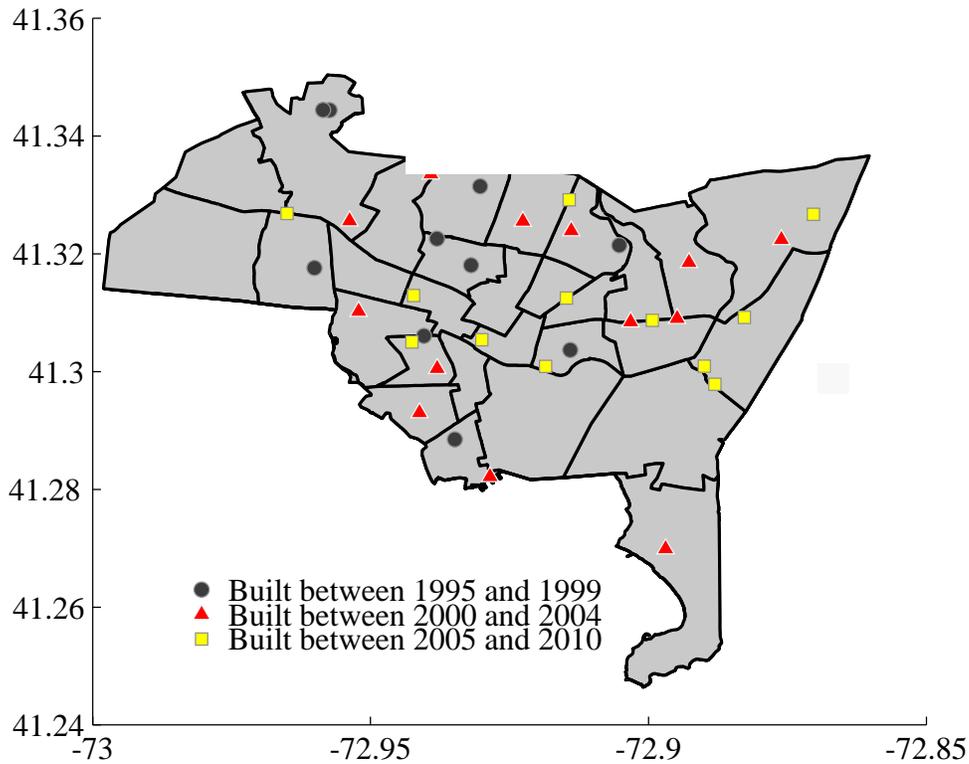
*Table 2: Characteristics of Treated and Untreated Schools in 2006*

	Untreated	Treated	p-value
Count (total)	21	18	
Male	0.51	0.50	0.622
Black	0.53	0.54	0.886
Hispanic	0.32	0.32	0.949
English Language Learners (ELL)	0.14	0.100	0.38
Special Education	0.08	0.08	0.506
Free Lunch	0.80	0.78	0.623

Note: Comparison of student body characteristics of treated and untreated schools at the project midpoint in terms of occupied buildings (2006). The p-value is from the t-test of equality across the two groups. The joint test fails to reject the null hypothesis of no relationship between the two groups at conventional levels.

<sup>9</sup>Informative cross-group comparisons of test scores are not feasible because we do not have test score data prior to 2004, and many schools in the treated category had already been treated by then. Comparisons of 2004 or 2006 scores would reflect the effects of treatment whether or not initial assignment was balanced.

Figure 1: School completion by date and location.



---

Note: Different shaped symbols indicate schools that were completed in different moments of time.

It is also important to ask why district officials chose to pursue such an ambitious infrastructure project. In particular, one might imagine that the district embarked upon the project to compensate for particularly decrepit pre-SCP facilities. If this were the case, it would compromise the generalizability of our results to districts with better baseline levels of infrastructure. As we have discussed above, this is not the case: school buildings in the district were not in observably worse condition than school buildings in similar cities. That said, we caution against applying our findings to schools in wealthier districts with average or above-average levels of existing infrastructure.

### 3 Economic Framework

#### 3.1 A conceptual model

To provide context for our empirical work, we present a simple theoretical framework of neighborhood choice and school performance in a dynamic setting. We use the framework to explore the relationship between home prices, neighborhood choice, and school performance in a way that clarifies the assumptions necessary for unbiased estimation of the causal effects of school construction on home prices and test scores using variation in the timing of construction projects.

Let a measure of families reside in the district. Families are differentiated in terms of preferences for local amenities  $\theta_i$  and resources  $\omega_i$ . School quality is one important local amenity. Each period, an exogenously selected proportion  $\gamma$  of individuals leave the district, and a demographically equivalent proportion enter, so that the distribution of  $\theta_i$  and  $\omega_i$  within the district remains constant over time. This assumption simplifies discussion and will not affect our conclusions so long as school construction does not have a large effect on cross-district migration.

All district residents must choose a home  $h \in H$  in which to live. Each home is located within a school zone  $z \in Z$ , and the supply of homes within each zone is fixed. District residents who lived in the district in the previous period may choose either to remain in their current home, or to pay a fixed cost  $C$  to relocate within the district. If they choose to relocate within the district in period  $t$ , their indirect utility can be written as

$$U(\theta_i, \omega_i) = \max_{z \in Z, h \in H} \sum_{s=t}^T \beta^{(s-t)} E \left( u(\kappa_z, \chi_z, X^h, p_{zt}^h; \theta_i, \omega_i) \middle| \Omega_{it} \right) - C \quad (1)$$

Here,  $\kappa_z$  is the quality of the school for zone  $z$ ,  $\chi_z$  are the local amenities,  $X^h$  are the characteristics of home  $h$ , and  $p_{zt}^h$  is the price of purchasing home  $h$  in zone  $z$  at time  $t$ .  $0 < \beta < 1$  is a discount factor. The expectation is taken over an information set given by  $\Omega_{it}$  which can include beliefs regarding the timing of specific school construction projects. New district arrivals face the same problem but do not pay the fixed relocation cost.

This model captures in a general way the dynamic considerations that underly the interrela-

tionships between migration, school enrollment, and school quality. Models of this type are not analytically tractable, and present substantial computational challenges as well (Meghir and Rivkin (2010), Kennan and Walker (2011)). These challenges are beyond the scope of this paper. Instead, we employ the model as a conceptual guide for predicting the effects of school construction, and assessing the biases that may arise in reduced-form estimates of these effects.

We focus in particular on three intuitive implications of this model:

- i. School construction projects that raise neighborhood amenities  $\chi_z$  or school quality  $\kappa_z$  will tend to raise home prices in that neighborhood. Excess demand for that location, given the original prices, will come from existing families who want to readjust and also newly arrived families who will choose that location with a higher probability.
- ii. If families with school age children value school quality more than families without children, raising  $\kappa_z$  will tend to raise the share of families with children in neighborhood  $z$ . Further, if families with some values of  $\theta_i$  and  $\omega_i$  value school quality or amenities more than others, school construction will change the distribution of preferences and resources across neighborhoods.
- iii. In the short run, the size of price and demographic responses to school construction will rise with the size of associated changes in school and amenity quality, decline with fixed cost  $C$ , and rise with migration rate  $\gamma$ . The larger the positive impact of school construction and the easier it is for residents to respond, the larger price changes and demographic consequences will be.

The first two points indicate that we can interpret increases in neighborhood-specific property values and populations of families with children as evidence that school construction had an impact on school effects and/or amenities. The second and third points indicate that selection of students into schools and neighborhoods will be an important challenge to deal with in the estimation procedure, and provide some insight into when selection problems will be more or less severe. The next section describes our procedure for estimating reduced-form effects of school construction on home prices, residential choice, and test scores.

### 3.1.1 Empirical Model for Home Prices

Home prices are a function of neighborhood amenities  $\chi_z$ , school quality  $\kappa_z$ , and specific home characteristics  $X^h$ . This function may vary over time as the characteristics of neighborhoods and neighborhood residents change. It also depends on the school construction timeline: families will price in school construction as information about the project enters their information set  $\Omega_{it}$ .

To estimate this function, we assume linearity, so that

$$p_{zt}^h = X_t^h \beta + \alpha_t + \alpha_z + \gamma_z \cdot t + g_{t,\{\ell\}}(S_z) + \epsilon_z(t)^h \quad (2)$$

We approximate the variation in prices that is unrelated to school construction using a set of time-varying property-level components ( $X_t^h \beta$ ), neighborhood effects ( $\alpha_z$ ), district-wide time effects ( $\alpha_t$ ), and neighborhood level time trends ( $\gamma_z t$ ). School construction enters the price equation through a function  $g_{t,\{\ell\}}(S_z)$  that maps per-capita investment in school construction  $S_z$  to prices, and  $\epsilon_z(t)^h$  is an i.i.d noise component.

We choose the form of  $g_{t,\{\ell\}}$  to balance our preference for simplicity with the need to reflect the fact that school construction projects will be priced in gradually. Let  $\{\ell\} = \{\ell_{fz}, \ell_{cz}, \ell_{oz}\}$ , where  $\ell_{fz}$  is the date the project in neighborhood  $z$  was announced (the ‘filing date’),  $\ell_{cz}$  is the date construction began in neighborhood  $z$ , and  $\ell_{oz}$  is the occupancy date for the project in neighborhood  $z$ . Then

$$g_{t,\{\ell\}}(S_z) = S_z \times (\delta_f 1[t \geq \ell_{fz}] + \delta_c 1[t \geq \ell_{cz}] + \delta_o 1[t \geq \ell_{oz}]). \quad (3)$$

Home prices jump discontinuously as the project enters each new phase. The idea is that home prices reflect the expected future value of the school to the neighborhood, and therefore change as uncertainty regarding the project is resolved. Effects are scaled by the per capita dollar value of the project, computed by dividing total cost by total occupancy.

Substituting equation [Equation 3](#) into equation [Equation 2](#) leads to a difference-in-differences style specification with three phases of treatment. That is, we estimate the effect of school construction on home prices by comparing changes in home prices in neighborhoods in which projects were announced, constructed, or occupied to simultaneous changes in home

prices in neighborhoods where those events did not occur or occurred at different times. For this model to yield unbiased estimates of construction effects, a) treatment dummies must not be correlated with changes in the unobservable price determinants of transacted homes, and b) treatments cannot coincide with other discontinuities in neighborhood-specific trends. Assumption a) will be violated if families with the resources and tastes to select into neighborhoods with new schools prefer homes that are unobservably more expensive than other families. Fortunately, we can use assessor estimates of ‘unobserved’ home quality to address this issue directly. We discuss this in more detail in the section 5. As mentioned above, discussions with district officials do not indicate that assumption b) is a major concern.

### 3.2 Academic Achievement

Let the academic outcomes of child  $i$  at the school for zone  $z$  in grade  $g$  be given by

$$T_{iz}^g = \kappa_z + \eta_i^g + X_i\beta + \epsilon_{iz}^g \quad (4)$$

where  $\kappa_z$  is the school’s value added to the student’s test score,  $X_i$  are observed student demographics, and  $\epsilon_{iz}^g$  is an i.i.d. shock. The  $\eta_i^g$  are unobserved factors that determine test scores, such as parental investments or student effort. In general, the  $\eta_i^g$  will be correlated with family resources  $\omega_i$  and preferences  $\theta_i$ ; we posit a specific relationship below.

The school value added  $\kappa_z$  can be interpreted as a function of a variety of inputs, such as the quality of teachers and administrators, the school culture, and school infrastructure. Given that families choose their neighborhood through [Equation 1](#), the assignment of students to schools is not orthogonal to unobservable student characteristics  $\eta_i^g$ . This means that it is not possible to directly estimate [Equation 4](#) and recover  $\kappa_z$ , even when conditioning on  $X_i$ . This is a classic problem of selection.

As with housing prices, we allow school construction projects to shift school quality  $\kappa_z$  through a step function which varies with time relative to occupancy of the new school building. We restrict the effects of school construction to be homogeneous across individuals at each time relative to treatment. That is,

$$\kappa_z = \kappa_z(\ell) = \bar{\kappa}_z + \delta_\ell^s S_z \quad (5)$$

where  $\ell$  denotes time relative to occupancy (in this case measured in academic years), and  $\delta_\ell^s$  is the effect on school value added at time relative to treatment  $\ell$  for each dollar or per-capita neighborhood-specific expenditure  $S_z$ . We interpret  $\delta_\ell^s$  as the cumulative effect of current and lagged school construction treatments; i.e., we do not attempt to distinguish between the effects of having a new school this year on this year's score from the effects of having a new school last year on this year's score.

We also allow the construction project to affect the unobserved student and family contribution to test scores by changing the effort parents and students put into schoolwork. This is expressed through the unobserved component of student test scores  $\eta_i^g$ . We model  $\eta_i^g$  as a function of all current and prior investments that are not contingent on construction,  $\bar{\eta}(\theta_i, \omega_i)$ , plus a component that changes with the time relative to occupancy of the new school building and per capita school building expenditures:

$$\eta_i^g = \eta_i^g(\ell) = \sum_{s=0}^g \rho_{gs} [I_s(\theta_i, \omega_i)] + \delta_\ell^f S_z = \bar{\eta}_i^g(\theta_i, \omega_i) + \delta_\ell^f S_z \quad (6)$$

The  $\rho_{gs}$  are discount factors that capture the effects of year  $s$  inputs on grade  $g$  outcomes. Like the  $\delta_\ell^s$ , the  $\delta_\ell^f$  are restricted to be the the same for all students and are interpreted as the cumulative impact of construction to date.

Substituting [Equation 5](#) and [Equation 6](#) into [Equation 4](#) we obtain the following expression for student test scores:

$$T_{igz} = \bar{\kappa}_z + \delta^s D(\ell) S_z + \bar{\eta}_i^g(\theta_i, \omega_i) + \delta^f D(\ell) S_z + X_i \beta + \epsilon_{iz} \quad (7)$$

$$= \bar{\kappa}_z + \bar{\eta}_i^g(\theta_i, \omega_i) + \Delta D(\ell) S_z + X_i \beta + \epsilon_{iz} \quad (8)$$

$D(\ell)$  is a row vector of indicator variables which signal time relative to occupancy of the new school building for  $\ell \in \{\ell_{min}, \dots, \ell_{max}\}$ .  $\delta^f = [\delta_{\ell_{min}}^f, \dots, \delta_{\ell_{max}}^f]'$ ,  $\delta^s$  is defined analogously, and  $\Delta = \delta^s + \delta^f$  is a vector of the added family, student and school effects of the school construction project. Separately identifying  $\delta^f$  and  $\delta^s$  would require data on student and school investment that cannot be obtained in practice. Our goal is therefore to estimate  $\Delta$ , the total effect of school construction on academic achievement.

### 3.2.1 Post-randomization readjustment and effect identification

Bias in OLS estimates of  $\Delta$  in Equation 7 stems from correlation between  $D(\ell)$  and  $\bar{\kappa}_z$  or  $\bar{\eta}_i^g$ . We will control directly for neighborhood effects in each of our empirical specifications, and in some specifications control for neighborhood trends as well. Coupled with the fact that school construction was not systematically associated with other school policy changes, this mitigates our concern about bias arising through correlation with  $\bar{\kappa}_z$ . Further, because the initial assignment of school construction projects is unrelated to student characteristics,  $D(\ell)$  is orthogonal to all student-specific investments and locational choices undertaken prior to knowledge of school construction. It follows that if families do not make location choices in response to school construction, OLS estimates of school construction effects will return unbiased estimates of  $\Delta$ . The extent to which OLS estimates are in fact biased depends on both the level of readjustment (which in turn depends on within-district relocation costs  $C$  and district arrival rate  $\gamma$ ) and on the relationship between neighborhood choice determinants  $(\theta_i, \omega_i)$  and test score determinants  $\bar{\eta}_i^g(\theta_i, \omega_i)$ . If families readjust, but do so in a way that is independent of test score inputs, OLS estimates will not be biased.

We consider three reduced-form estimators of  $\Delta$ , each of which is unbiased under different assumptions about the readjustment process and the error structure. The estimators and the assumptions required to yield unbiasedness are as follows:

- i. OLS Estimator:

$$T_{igz} = \tau_{zg} + \tau_t + \Delta D(\ell)S_z + X_i\beta + e_{iz} \quad (9)$$

$\tau_{zg}$  and  $\tau_t$  are grade-school and academic year fixed effects, respectively. We must assume  $D(\ell)$  is orthogonal to  $\bar{\eta}_i^g$  for all  $\ell$ . This assumption holds if school construction projects are assigned randomly and families either do not readjust their school choices or readjust in a way that is uncorrelated with score determinants  $\bar{\eta}_i^g$ .

- ii. Student Fixed Effect Estimator:

$$T_{igz} = \tau_{zg} + \tau_t + \tau_i + \Delta D(\ell)S_z + e_{iz} \quad (10)$$

$\tau_i$  are individual-specific fixed effects. Here,  $D(\ell)$  may be correlated with  $\bar{\eta}_i^g$  for any  $\ell$ , but we restrict  $\bar{\eta}_i^g = \bar{\eta}_i$  for all  $g$ . In this case the fixed effects estimator will recover estimates of  $\Delta$ , as the individual level heterogeneity in unobservable characteristics  $\eta_i$  will be captured in  $\tau_i$ . This assumption will hold if  $\rho_{gs} = 0$  for all  $s$  greater than the lowest  $g$  for which we have valid test score observations. Even if these assumptions fail, value added estimation will improve on OLS estimation to the extent that conditioning on past scores weakens the relationship between school construction treatment and unobservable score determinants.

iii. Value Added Estimator:

$$T_{igz} = \tau_{zg} + \tau_t + \pi T_{i,g-1,z} + \Delta D(\ell) S_z + X_i \beta + e_{iz} \quad (11)$$

This estimator is identical to OLS but controls for lagged test scores. Even if  $D(\ell)$  is correlated with  $\bar{\eta}_i^g$ , value added estimation yields unbiased estimates the effects of current-year school construction under two assumptions. The first assumption is that the effects of all test score inputs—including lagged school construction treatments—decay geometrically year to year. The second is that contemporaneous investments are orthogonal to  $D(\ell)$ . The geometric decay assumption in particular is strict and likely to be invalid (Todd and Wolpin (2007), Rothstein (2010)). Even so, value added estimation offers advantages over OLS estimation in that it controls directly for individual-specific heterogeneity in scores. Value added estimation also addresses a weakness of FE estimation by allowing for (heavily parameterized) individual effects that vary over time.

None of these estimators is perfect. OLS estimates will likely be biased if locational choices respond to school construction plans once they are revealed. Fixed effect and value added estimators allow for a flexible readjustment process, but require strong assumptions on the form of  $\bar{\eta}_i^g$ . Below, we present evidence that families do change locational choices in response to construction. We therefore focus our analysis on the fixed effect and value-added estimators rather than the OLS specification. If these estimators yield similar results, and the timing of observed effects closely follows the timing of the school construction, we will interpret our findings as strong evidence of an effect of school construction on student scores.

## 4 Data

### 4.1 Home sales data

Our home sales data is a complete record of all residential property sales that took place in the school district between January 1st, 1995 and January 31st, 2010. We obtained this data from administrative records maintained by the Office of the City Clerk. The data include sale prices as well as a variety of property and home characteristics. These characteristics include property address and acreage, home square footage, the number of bedrooms, bathrooms, and total rooms, and the ‘style’ of the property (e.g., ‘Georgian,’ or ‘multi-family’). The data also include a subjective evaluation of each home made by the town tax assessor. These evaluations are categorical and range from ‘poor’ to ‘excellent.’ The assessor’s evaluations have substantial explanatory power even after conditioning on observable home characteristics, and therefore can be interpreted as a measure of what would in most cases be deemed ‘unobservable’ home quality.

We match homes to their zoned schools using the district’s school assignment guidelines. We focus our analysis on construction of elementary and middle schools because District high schools are generally not bound by neighborhood zones and enroll students from across the district.

**Table 3** summarizes the home sales data. Between the beginning of 1995 and January 2010, there were 14,266 residential properties sold in the district. The pace of sales was relatively slow between 1995 and 1999, during which time 2,817 homes were sold, and picked up thereafter to a rate of over 5,000 homes per five year period. We were able to match nearly all of the sales records to elementary-middle-high school triplets. Non-matches were due to incomplete address records in the sales data or omissions from the school assignment list. The average price of a home sold in the district rose from \$120,301 between 1995 and 1999 to \$164,345 between 2000 and 2004 to \$245,909 between 2005 and 2010. This occurred even though characteristics of the transacted homes did not change very much: square footage, acreage, and number of rooms all remained relatively constant between 1995 and 2010. About 40 percent of homes sold in each period were deemed by the assessor to be high quality, a constructed binary designation that includes good to excellent ratings.

Table 3: Fifteen Years of District Home Sales

	1995-2010	1995-1999	2000-2004	2005-2010
Number of Homes Sold	14266	2817	5784	5665
Matched to schools	14081	2772	5718	5591
Price (mean)	188,000	120,000	164,000	246,000
Price (median)	156,000	101,000	140,000	213,000
Square feet	1956	2026	1948	1929
Acreage	0.12	0.14	0.12	0.11
Bedrooms	3.64	3.60	3.62	3.67
Bathrooms	1.88	1.87	1.89	1.89
Rooms	7.98	7.92	7.98	8.01
High quality	0.39	0.45	0.38	0.37

Note: Data describe the population of home sales in New Haven over the 1995-2010 period. Sales are counted as matched to schools if we can locate the address on the map of school zones and assign it elementary, middle, and high schools. Prices are in 2005 dollars and rounded to nearest 1000. High quality is equal to one if the it is described as ‘good,’ ‘above average,’ or ‘excellent’ in assessor’s records.

## 4.2 Student residency data

To examine the impact of the SCP on residential choices, we use data on the addresses of enrolled students for the academic years 2002-2003 through 2009-2010.<sup>10</sup> As with the home sales data, we map student addresses to neighborhoods by zoned elementary-, middle-, and high-school triplets. The idea is that public school enrollment levels reflect the outcomes of the location choice process for families with children.

Table 4 presents the means and standard deviations of neighborhood-year observations of school enrollment levels and flows. Over the period in question, average enrollment in a district neighborhood was 740, and trended downward from an average of 794 students per neighborhood in 2004 to an average 693 students per neighborhood in 2010. The lower rows of Table 4 decompose this change into cross-district and within-district inflows and outflows. Inflows count the number of enrolled students living in a neighborhood in a given year who were not enrolled in the district in the previous year, or who were enrolled but lived in a different neighborhood. Outflows count the number of students who

<sup>10</sup>We refer to academic years using the spring year from this point forward.

lived in a given neighborhood and were enrolled in a district school last year, but this year either were not enrolled or remained enrolled but moved to a different neighborhood. On average, 215 students are new to a neighborhood each year, while 233 students leave. This is consistent with the observed pattern of decreasing enrollment over time. Note that both inflow and outflow counts capture a wide variety of student movements. Inflows include district residents entering public school for the first time, students arriving from out of town in time for fifth grade, and students whose families move across town at any point during a students' time in the district. Outflows include students who drop out, graduate, switch neighborhoods within the district, or leave the district entirely.

We further decompose inflows (outflows) by students' starting point (destination). Specifically, we define intra-district inflows and outflows as those in which a student switches neighborhoods within the district but remains enrolled in school. Interdistrict inflows then consist of students who are new to district enrollment, while interdistrict outflows consist of students who leave the district. Intradistrict inflows and outflows must be equal in any given year,<sup>11</sup> so changes in total district enrollment are determined by interdistrict flows. On average, about 89 students— 12 percent of the average enrollment level— arrive and leave neighborhoods each year through intradistrict flows. We see in [Table 4](#) that on average a neighborhood gains 126 enrollees through interdistrict inflows in a given year, but loses 140 enrollees through interdistrict outflows. The overall picture painted by data on neighborhood-level enrollment and flows is one of a school district that is shrinking in size and in substantial residential flux, as students enter and leave the district and move within it.

---

<sup>11</sup>Differences in standard deviations are expected and reflect the fact that intra-neighborhood inflows and outflows can have different distributions even if they have the same mean.

Table 4: Mean and Standard Deviation of School enrollment by neighborhood

	Total	2004	2007	2010
Enrollment	740	794	712	693
	386	413	374	379
Inflows	215	233	209	200
	113	118	113	107
Inflows: inter.	126	138	114	122
	66	69	60	68
Inflows: intra.	89	95	96	78
	52	54	58	44
Outflows	229	230	262	199
	121	128	130	99
Outflows: inter.	140	135	166	122
	75	72	84	63
Outflows: intra.	89	95	96	78
	53	62	55	41
N	200	25	25	25

Note: Student enrollment in district public schools by neighborhood-year. Within each row, the upper number is the variable mean and the lower number is the standard deviation. Inflows represent students new to a neighborhood between current year and previous year. See text for a description of the distinction between inter- and intra-district in- and out-flows.

### 4.3 Test score data

To assess the effect of the SCP on academic outcomes, we use student microdata for the academic years 2004 through 2010. Key variables include student race, English Language Learner (ELL) status, special education status, free or reduced-price lunch status, and student scores on state-mandated assessment tests (the Connecticut Mastery Test, or CMT), which we standardize using state-level means and standard deviations within grade-year cells. [Table 5](#) shows summary statistics for the students in our data. We have data on 152,151 student-years over the seven-year window, reflecting a district size of about 22,000 students. Black students make up roughly half of all students, and Hispanic students account for another 35 percent. Because the proportion of free lunch students is so high, all district students receive free lunch at school. Each year, the district sends home a survey requesting income data so that they can renew district-level free lunch eligibility, and our data reflects the results of this survey. Generally about 80 percent of students report being

free or reduced price eligible.<sup>12</sup> Mean reading and math scores in the district were approximately two thirds of a student-level standard deviation below state means throughout the period.

*Table 5: School district demographic profile*

	Total	Matched	In-district matched	FE sample	VA sample
N	152151	136883	123275	38191	20584
Black	0.52	0.52	0.53	0.51	0.51
Hisp.	0.35	0.35	0.36	0.38	0.39
ELL	0.11	0.12	0.13	0.12	0.11
Spec. Ed.	0.11	0.1	0.1	0.07	0.06
F/R lunch	0.74	0.76	0.77	0.85	0.92
Reading	-0.66	-0.65	-0.69	-0.65	-0.63
Math	-0.63	-0.62	-0.66	-0.59	-0.57
PK-2	0.31	0.26	0.27	0	0
Gr. 3-8	0.41	0.45	0.46	1	1
Gr. 9-12	0.27	0.29	0.27	0	0

Note: Characteristics of student population observed in microdata. Unit of observation is the student-year. ‘Total’ column includes all students in district. ‘Matched’ column includes student-years with matched addresses. ‘In-district matched’ includes student years with matched addresses for in-district students only (i.e., not students from neighboring towns). ‘FE sample’ column describes student-year obs. with current-year scores for students who are never enrolled in transitional schools, and have test scores less than three standard deviations above or below district mean. ‘VA sample’ column introduces lag-score requirement. Reading scores and math scores standardized using state-level means and standard deviations.

When conducting our analysis of test scores, we restrict our student sample in a number of ways. Since treatments take place at the neighborhood level, we eliminate enrollment records that cannot be matched to addresses. As shown in the second column of [Table 5](#), matched students tend to be older than the student body as a whole but are otherwise demographically indistinguishable. We also eliminate out-of-district students who enroll in district schools, because these students cannot be matched to neighborhood-level treat-

<sup>12</sup>In 2005, the district school lunch survey estimated that only 35 percent of students were eligible, a sharp break with both earlier and later datapoints. District officials attribute this to poor survey design in that year, combined with the limited incentives parents have to fill out the forms. We experimented with dropping the free lunch dummy from our analysis. This generally increased both the magnitude and sign of our results, so to be conservative we chose to include the variable despite its inconsistency across time.

ments.<sup>13</sup> The third column of **Table 5** describes these students, who again resemble the broader student population.

To construct our analysis sample from the sample of in-district students with matched addresses, we make several further sample trims. We eliminate students who attend ‘transitional’ schools— schools specifically for struggling students— in any of our data years. We eliminate these students because transitional schools are not tied to specific school zones, and because we are interested in the effects of school construction on students in standard academic programs.<sup>14</sup> We also drop student-year observations with test scores more than three standard deviations above or below the mean. The goal of this cut is to limit the impact of score outliers on our analysis, but in practice our results are not sensitive to changing or eliminating the threshold.

In our main analysis sample, used for fixed effect estimation, we include all remaining student-year observations with valid scores. This requirement eliminates students in non-tested grades: the CMT was administered in grades three through eight between 2006 and 2010, and in grades four, six, and eight prior to 2006. Students in other grades are dropped. This sample is described in the fourth column of **Table 5**. In our value added analysis, we include only students with nonmissing current- and prior-year scores. This sample is described in the fifth column of **Table 5**. The prior-year score requirement eliminates all students in academic years 2004 and 2005, third and fourth graders in 2006, and third graders between 2007 through 2010. Though requiring the presence of baseline scores reduces our sample size from 38,191 to 20,584, students in the value added sample do not differ substantially from students in the fixed effects sample in terms of their observable characteristics.

## 5 Effects on home prices

**Table 6** reports our estimates of three versions of equation 2. We report results for elementary school and middle school construction only, since high school assignment is generally not neighborhood based. The first column includes year effects, seasonal effects, observable home covariates, neighborhood intercepts and slopes, and high school construction treatment variables as controls. We find that sale prices rise by 0.29 percent per \$10,000

---

<sup>13</sup>These students enroll in District schools through regional school choice programs.

<sup>14</sup>Including these students does not affect our findings.

of per-capita construction expenditure upon filing, 0.13 percent upon the beginning of construction, and a further 0.85 percent upon occupancy. The price changes at filing and construction start are not significantly different from zero, but we reject the hypothesis that the change at occupancy is zero at the five percent level. The estimated total effect of construction— the sum of the score gains at each project phase— is 1.27 percent per \$10,000 in per capita expenditure, and is significantly different from zero at the five percent level. Multiplying the total effect by the average level of per capita expenditure across all houses in the regression sample (roughly 8.1) shows that the average effect of elementary and middle school construction was to raise home prices by a total of 10.2 percent.

In column II, we add controls for assessor-measured ‘unobservables’ to the regression. This causes our estimated effects to rise slightly: prices increase by 0.95 percent per \$10,000 in per-capita expenditure upon occupancy, and by 1.38 percent in total. The estimated average effect of construction on home prices rises to 11.1 percent. In column III, we add an interaction term between dollars of per-student expenditure, a post-occupancy dummy, and years post occupancy. This allows the post-occupancy effect to deteriorate or increase over time. The estimated effect, labeled  $\delta_o^s$ , is small and statistically insignificant. It appears that school construction has a significant and large effect on home prices, and that our results are not driven by neighborhood-specific trends or by changes in the unobservable characteristics of homes sold, and that post-occupancy effects do not decay as time passes.

To put the size of these effects into perspective, it is useful to construct a back-of-the-envelope estimate of the total effect of elementary and middle school construction on tax revenues and compare it to the total cost of the project. Assume for the purposes of this exercise that the average effect of school construction on the value of all district real estate (residential and non-residential) is the same as its average effect on the price of transacted homes. Assume further that there are no spillover effects, so that school construction in one neighborhood does not affect the prices of homes in other neighborhoods. Then the change in annual tax revenues resulting from construction of elementary and middle schools can be written

$$\delta_{tax} = \text{Real estate value} \times \text{Tax rate} \times \text{Average effect}$$

Table 6: Elementary and middle school construction and home prices

	I	II	III
$\delta_f$	0.0029 (0.0027)	0.0033 (0.0027)	0.00322 ( 0.0024)
$\delta_c$	0.0013 (0.0021)	0.0011 (0.0021)	0.0010 ( 0.0025 )
$\delta_o$	0.0085** (0.0038)	0.0095** (0.0037)	0.0095** (0.0035)
$\delta_o^s$			-0.0010 (0.0033)
$\delta_{pre}$	0.0042 (0.0040)	0.0044 (0.0040)	0.0042 (0.0040)
$\delta_{tot}$	0.0127** (0.0058)	0.0138** (0.0057)	0.0137** (0.0059)
Average Effect	<b>0.102</b>	<b>0.111</b>	<b>0.110</b>
Year effects	Yes	Yes	Yes
Season effects	Yes	Yes	Yes
House Covariates	Yes	Yes	Yes
High School const.	Yes	Yes	Yes
Neighborhood Fixed Effects	Yes	Yes	Yes
Neighborhood Slope	Yes	Yes	Yes
Quality	No	Yes	Yes
Neighborhood Clustering	Yes	Yes	Yes
N	13559	13551	13551

\*: Significant at the 10% level. \*\*: Significant at the 5% level. \*\*\*: Significant at the 1% level. Results from a regression of log home sale price in 2005 dollars on time relative to filing, occupancy, construction of neighborhood elementary and middle schools (per \$10,000 in per-capita expenditure), and the indicated covariates.  $\delta_f$ ,  $\delta_c$ , and  $\delta_o$  refer to effects observed upon project announcement, construction start, and occupancy, respectively.  $\delta_o^s$  is a slope parameter that allows  $\delta_o$  to change with years post-occupancy.  $\delta_{pre} = \delta_f + \delta_c$ , and  $\delta_{tot} = \delta_f + \delta_c + \delta_o$ . Average effects are equal to  $\delta_{tot}$  multiplied by average per-capita expenditure weighted by home sales. HS construction treatment variables are included but not reported. Standard errors are clustered at the neighborhood level.

The most recent assessed value of the stock of real estate in New Haven is \$5.2 billion (in 2005 dollars), and the property tax rate is 0.0439. Combining these values with the

estimated average effect from column II (11.1 percent) indicates that  $\delta_{tax}$  is equal to roughly \$25 million. If, as reported in column III, effect sizes do not diminish significantly over time, and if city government can borrow and lend at a five percent rate of interest, then the costs of delaying investment in school construction are quite high: a ten year postponement of the stream of \$25 million payments has a present value cost of \$218 million. This represents more than a fifth of the \$1.07 billion of total expenditures on elementary and middle school construction. The key point here is not that this value represents a precise estimate of the cost of postponing investment, but rather that the value District residents place on school construction investment is large enough that the District can expect to mechanically recoup at least some portion of its costs through increased tax revenue.

Before moving on, it is worth noting that these results do not rule out forward-looking behavior on the part of home-buyers. Roughly a third of the total effect of school construction accrues prior to building occupancy, though these pre-effects are not statistically significant by themselves. And even if pre-occupancy price effects are in fact zero, this is consistent with a model in which home-buyers are forward-looking but skeptical about the prospects for project completion.

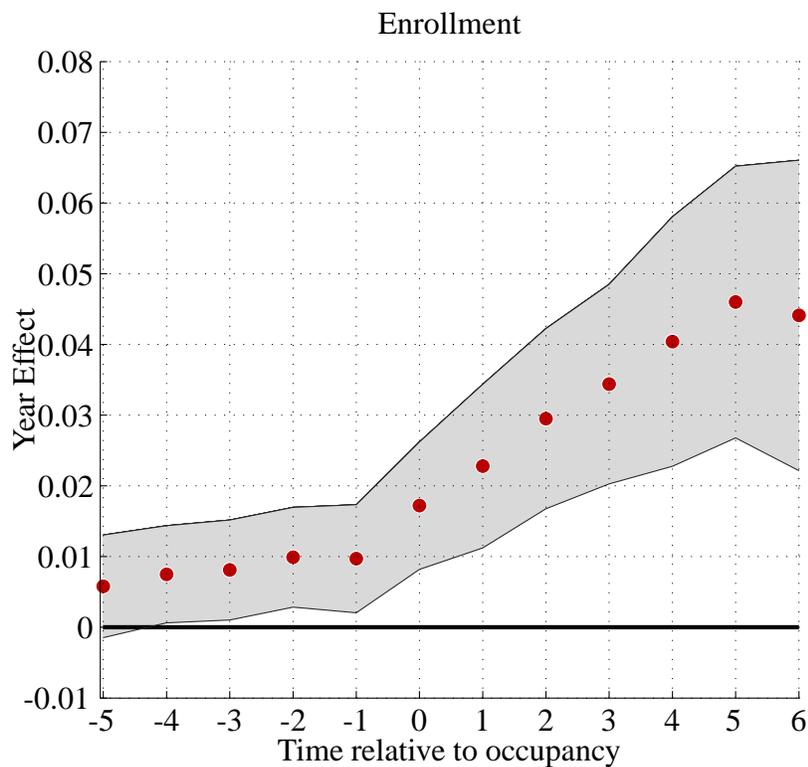
## 6 Effects on school enrollment

At least two distinct stories are consistent with our finding that school construction increases home prices. The first is that school infrastructure is a selling point for homeowners regardless of whether they have children they intend to enroll in a neighborhood school. For example, homeowners may value local amenities like swimming pools or playing fields. The second is that price increases are driven by the desire of homeowners to enroll their children in the rebuilt schools. These stories are not mutually exclusive, but have different implications for the effects of infrastructure investment on schooling demand and community demographics. In this section, we examine the relationship between schooling demand and the residency patterns for district students, and find support for the second story, though we cannot rule out the first using the data at hand.

**Figure 2** and the first column of **Table B -3** display estimated enrollment elasticities computed using a regression of log public school enrollment by neighborhood on per capita infrastructure investment, year fixed effects, and neighborhood fixed effects. The coefficients on treatment dummies for years more than five years prior to building occupancy are

restricted to be zero, and coefficients on treatment dummies corresponding to six or more years after treatment are restricted to be the same. In the four years prior to occupancy, the effects of school construction on neighborhood enrollment differ significantly from enrollment levels six or more years pre-occupancy, but effect sizes are small. Beginning in the year of occupancy, the enrollment effects begin to rise, and continue to do so for the next five years. The trend flattens out six or more years post-occupancy. By this point, \$10,000 of per capita school construction expenditure raises neighborhood enrollment by 4.4 percent. Note that because these findings reflect changes in where students live, they cannot be a mechanical result of changes in school capacity.

Figure 2: The effects of school construction on neighborhood residency



Note: The figure shows the effect of \$10,000 per capita school construction treatment on log neighborhood enrollment by year relative to occupancy. The estimates are obtained from Equation 2. See Table B -3. Shaded area represent 95% confidence intervals.

To raise neighborhood enrollment, school construction must either increase neighborhood inflows or reduce neighborhood outflows. Table B -3 reports estimated effects obtained using an identical specification with the logs of these two quantities as the dependent variables. Effects on inflows follow a similar pattern to those on overall enrollment: no dis-

cernible trend pre-occupancy, and then a steady post-occupancy increase. By the sixth year post occupancy, \$10,000 of per capita expenditure raises neighborhood inflows by 6.7 percent. This estimate differs significantly from zero at the five percent level. For outflows, the effect of school construction is delayed. Effects are approximately zero until three years after occupancy, when they begin to rise. Six or more years post occupancy, \$10,000 of per capita expenditure raises outflows by 5.0 percent. As reported in columns four through seven of [Table B -3](#), inflows and outflows reflect a balance of inter- and intra-district movers. Changes in enrollment are thus the result of post-occupancy increases in student ‘churn’; inflow effects are larger and begin earlier than outflow effects.

Who are the new arrivals to neighborhoods with rebuilt schools? To answer this question we construct indices of observable test score determinants by regressing reading and math scores on race dummies, a sex dummy, and free lunch status and computing predicted test score values for each student. We do not include ELL and special education status in the indices because these outcomes may be endogenous to education quality. [Table 7](#) reports results from a regression of math and reading score indices on a post-occupancy dummy multiplied by per-student building expenditures (in \$10,000 per-student units), controlling for year and neighborhood/grade fixed effects. The coefficient on the expenditure-scaled post-occupancy dummy is labeled  $\delta_{post}$ , and differs significantly from zero in both the math and reading specifications. In the Supplemental Appendix we show that these increases are due to growth in index values beginning at least one year post-occupancy. Covariate-specific regressions (available upon request) show that these effects are driven largely by changes in the free lunch status of neighborhood students. School construction brings many new students into neighborhoods, and the choice to move is correlated with demographics that predict high test scores.

Table 7: Selection on indices of observable test score determinants

	Reading	Math
$\delta_{post}$	0.00184**	.00138*
SE	0.00084	0.00071
N	50,145	50,145
Year FE	Yes	Yes
School/grade FE	Yes	Yes

\*\* : significant at 5% level. \* : Significant at the 10% level. Results from a regression of a linear index of test score observables on year FEs, neighborhood/grade FEs, and dollars of per-student investment interacted with a post-occupancy dummy. The linear index includes race dummies, sex dummies, and free lunch status. Weights are determined by a regression of test scores on these variables. Standard errors clustered at the neighborhood-year level.

The finding that neighborhood-specific school enrollment begins to rise at the time of school occupancy is consistent with the finding of rising home prices at that juncture. In the context of our conceptual model, it suggests that migration rates are high enough or fixed costs low enough that readjustment in response to school construction is feasible; families willing to pay for school infrastructure move in, while families not willing to pay move out. Demographically advantaged families with school-age children with appear to value the change in neighborhood amenities associated with school construction more than other types of households. One possible reason for this is that families with children benefit directly from any test score gains associated with construction, and families from high-scoring demographic groups either place an especially high value on these gains or can more easily pay the costs associated with moving. The next section assesses the size of the test score gains caused by school construction.

## 7 Effects on test scores

Table 8 presents results from estimates of equations 9 (OLS), 10 (fixed effects), and 11 (value added) for reading and math scores. For each subject area, the first column presents the OLS specification, the second the value added specification, and the third column the fixed effect specification. We restrict effects to be zero more than five years prior to building occupancy, and to be equal six or more years post-occupancy. OLS estimates of reading score effects are generally insignificant but trend upwards; point estimates are less than 0.0043 for all years pre-occupancy, but are equal to at least 0.0095 four or more years post-occupancy. The theoretically preferable value added and fixed effects measures show dis-

tinct trend breaks beginning in the year of building occupancy. Value added and fixed effect estimates rise in parallel for at least the first six years after building occupancy.<sup>15</sup> By six or more years post-occupancy, \$10,000 of per-student infrastructure investment raises reading scores by 0.0265 standard deviations in the fixed effect specification and 0.0236 standard deviations in the value added specification. In both the fixed effect and value added specifications, coefficients become significant at the ten percent level in the year of occupancy and are significant at the five percent level four or more years post-occupancy.<sup>16</sup>

We interpret the trend break in estimated effects at the time of building occupancy as strong evidence that school construction caused reading scores to rise in affected neighborhoods. That we observe this pattern even when controlling for individual-specific heterogeneity in multiple ways indicates that the estimated effects are appropriately viewed as the causal impact of school construction on the individual education production function (whether through motivational or direct pedagogical channels), not as a consequence of selection into treated neighborhoods. Arguing against this interpretation requires a story in which selection into treated neighborhoods is correlated with time-varying individual-specific effects conditional on prior year scores.

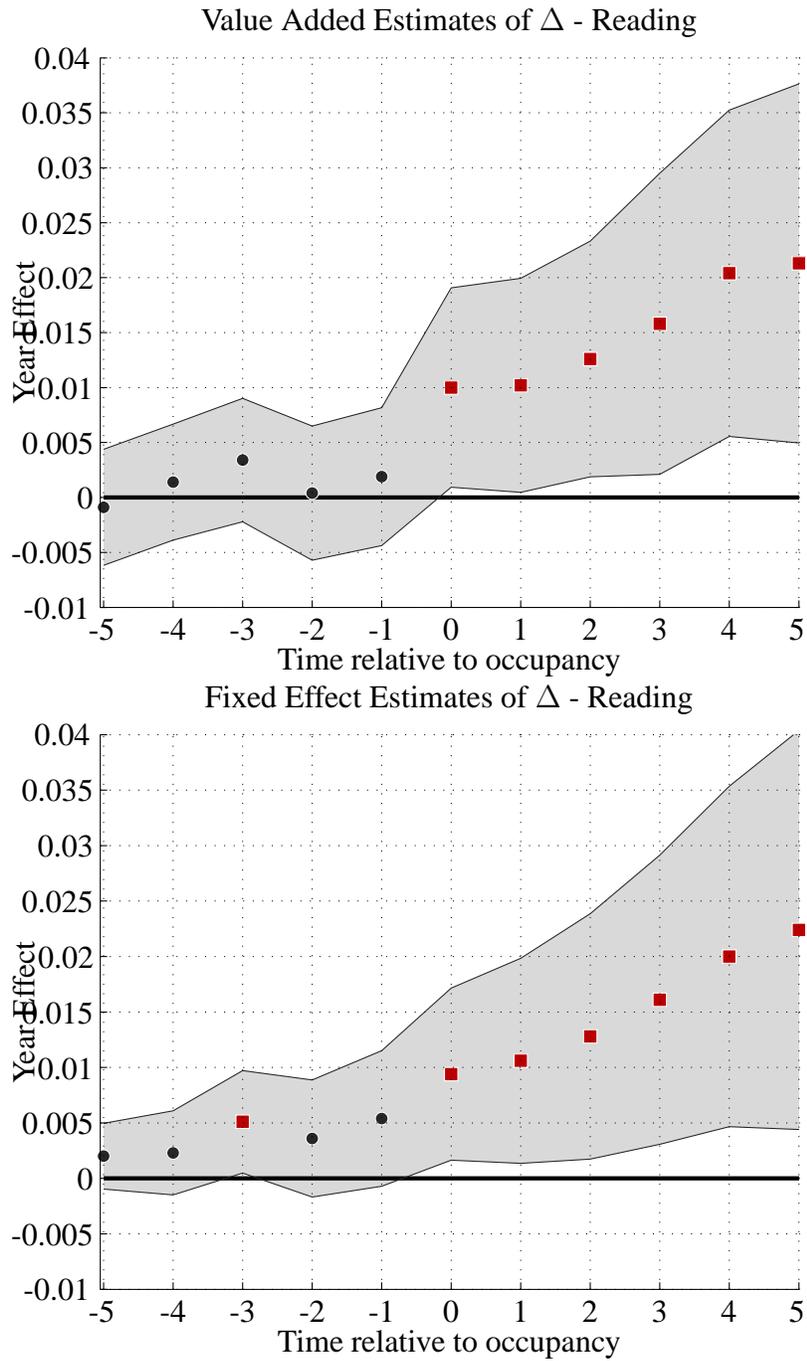
Estimates of math score effects do not paint as clear a picture. There is some evidence of a positive effect on math scores in the year of occupancy, as both the value added and fixed effect estimates differ significantly from zero at the ten percent level. In addition, estimates from the value added specification begin to slope upward three years after construction, reaching a level of 0.0172 six or more years post-occupancy. This effect is relatively close to the estimate in the reading specification, but does not differ significantly from zero at the ten percent level. OLS and fixed effect estimates show no discernible trend. Our estimates are noisy, so we cannot reject large effect sizes. However, compared to the strong evidence we find for reading score effects, evidence of math score effects is very weak.

---

<sup>15</sup>Note, however, that value added and fixed effect estimates are not strictly comparable. The value added estimates imply dynamic effects of lagged treatments on current scores, while fixed effect estimates do not.

<sup>16</sup>Out of the thirty pre-occupancy effects we estimate (five coefficients in each of six specifications), four are significant at the ten percent level. This is consistent with what we would expect given a true effect of zero in these years.

Figure 3: Effect of school construction on score levels



Note: Estimates of effects of per-capita school construction spending on student score levels by year relative to treatment, as described in equation 5. Shaded areas show a 90 percent confidence interval, allowing for clustered errors at the neighborhood-year level. Squares indicate significance at the 10% level. Controls include student characteristics, year effects, and school-grade fixed effects. Estimates reported in Tables 8.

Table 8: Effect of school construction on test scores

$t - t_0$	Reading			Math		
	OLS	VA	FE	OLS	VA	FE
-5	0.002	-0.0009	0.002	0.0026	0.0007	0.0038**
	0.0016	0.0032	0.0018	0.0019	0.0031	0.0019
-4	0.0038**	0.0014	0.0023	0.0006	-0.0009	0.0015
	0.0017	0.0032	0.0023	0.002	0.0036	0.0022
-3	0.0031	0.0034	0.0051*	0.0017	0.0021	0.0037
	0.002	0.0034	0.0028	0.0024	0.004	0.0026
-2	0.0013	0.0004	0.0036	0.0004	0.0003	0.0029
	0.0019	0.0037	0.0032	0.0025	0.0042	0.0029
-1	0.0043*	0.0019	0.0054	0.0031	0.0012	0.0033
	0.0024	0.0038	0.0037	0.0027	0.0043	0.0034
0	0.0062*	0.0100*	0.0094**	0.0056	0.0103*	0.0068*
	0.0033	0.0055	0.0047	0.0039	0.0055	0.0041
1	0.0046	0.0102*	0.0106*	-0.0007	0.0024	-0.0006
	0.0037	0.0059	0.0056	0.0046	0.0065	0.0049
2	0.0064	0.0126*	0.0128*	-0.0018	0.0055	-0.0032
	0.0045	0.0065	0.0067	0.005	0.0072	0.0056
3	0.0044	0.0158*	0.0161**	0.0013	0.0114	-0.0001
	0.0054	0.0083	0.0079	0.0064	0.0085	0.0064
4	0.0095	0.0204**	0.0200**	0.0038	0.0147	0.0021
	0.006	0.009	0.0093	0.0072	0.0099	0.0078
5	0.0136**	0.0213**	0.0224**	0.0075	0.012	0.001
	0.0065	0.0099	0.0109	0.0077	0.011	0.009
6+	0.0123	0.0236**	0.0265**	0.0061	0.0172	0.0029
	0.0076	0.0113	0.012	0.009	0.0128	0.0104
Student Cov.	Yes	Yes	No	Yes	Yes	No
Student FEs	No	No	Yes	No	No	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Nbd./Grade FEs	Yes	Yes	Yes	Yes	Yes	Yes
Sch./Grade FEs	Yes	Yes	Yes	Yes	Yes	Yes
Lag scores	No	Yes	No	No	Yes	No
N	38617	20584	38191	38694	21022	38991

\*: significant at the 10 percent level. \*\*: significant at the five percent level. Column 'FE' displays estimates of effects of per-capita school construction expenditure by year relative to treatment obtained using equation 10 with reading/math z-scores as the dependent variable in the FE analysis sample. Column 'VA' displays estimates of score gains obtained using equation 11 in the VA analysis sample. Controls include observable student covariates, year FEs, neighborhood/grade/school FEs, and lag scores (as indicated). Standard errors allow for clustering at the neighborhood-year level.

How large are the results presented here relative to the effects of other test score determinants? The average student lived in a neighborhood that received \$76,800 of per-capita school construction expenditure.<sup>17</sup> Multiplying this value by the estimated reading score effect six or more years post-occupancy in the fixed effect added specification, we estimate that on average school construction raised reading scores by 0.21 standard deviations. These effects are large, but not implausibly so. Lottery-based evaluations of the effects of attending high-performing KIPP charter schools find score gains of 0.12 standard deviations for reading and 0.35 standard deviations for math per year of enrollment (Angrist et al. (2010)). Between 6th and 8th grade, students enrolled in Harlem Children’s Zone (HCZ) schools experience reading score gains of one quarter to one third of a standard deviation and math score gains of 0.8 standard deviations relative to HCZ non-participants (Dobbie and Roland G. Fryer (2009)).

It is also useful to relate these results to changes in home prices and school enrollment by computing the implied elasticities of home prices and school enrollment with respect to changes in test scores. If school construction only altered home prices through its effects on student test scores, we could reasonably compute the elasticity of home prices with respect to school quality by dividing the percentage change in home prices post-construction by the per-capita change in test scores post-construction. Of course, school construction may affect home prices through other channels, like neighborhood aesthetics or access to public facilities. We conduct the exercise in spite of this limitation and interpret our results as upper bounds on the true elasticities. We further assume that long-term test score effects are immediately capitalized into home prices.

From [Table 6](#), we know that the estimated effect of \$10,000 of per student expenditure on home prices was 1.3 percent. From [Table B -3](#), we know that the estimated effect on neighborhood enrollment counts six or more years post-occupancy was 4.4 percent. The estimated effect of school construction on reading scores six or more years post-occupancy in the fixed effects specification was 0.027 standard deviations. These values imply that a 0.1 standard deviation increase in a school’s effect on reading scores would raise home prices by 4.7 percent and public school enrollment amongst neighborhood children by 16.2 percent. These elasticities should not be compared directly to those presented in Black (1999) or Bayer et al. (2007), because both the numerator and denominator differ in criti-

---

<sup>17</sup>This number is smaller than the home-sale weighted average expenditure of \$81,000. Relative to home sales, students were disproportionately concentrated in low-investment neighborhoods.

cal ways. In the denominator, we use student-level standard test score standard deviations while Bayer et al. and Black use percent changes in school average scores. In the numerator, we use changes in the causal effect of schools on test score production, while they use school average scores which incorporate both school causal effects and student selection into schools.

## 8 Possible Mechanisms

### 8.1 How do school construction projects help students?

Having documented the test score gains that accompany the construction of new school buildings, it is natural to ask why this might occur. Thus far, we have remained agnostic about whether school construction affects test scores through the direct pedagogical effect of improved facilities or through improved motivation for students and teachers. In practice, it may be difficult to distinguish between the two pathways. If a student's access to a computer within the classroom encourages him to read news online when at home and this improves his reading score, it is unclear whether to attribute the gain to the specific feature of the facilities or to the change in investment. Still, some pathways can be clearly categorized, and if one or the other plays a dominant role it would be valuable to know this.

A related question with important implications for policymakers is *which* building features are associated with score gains. Even if the pedagogical impacts of a given feature could not be distinguished from the motivational effects, future construction programs might like to design buildings with features that have large total effects. Unfortunately, we do not have consistent data on the characteristics of the newly-constructed buildings, and therefore cannot examine the heterogeneity of construction effects across building features in a quantitative way.

In lieu of a quantitative analysis, we address the motivation versus pedagogy issue and the specific building features question using a survey of district principals. We surveyed principals at 22 district schools about their experiences before, during, and after school construction. We chose to interview school principals rather than teachers or students because we believe principals' experiences are likely to be the most representative of school climate as a whole. Of the 22 principals we surveyed, ten were in office at the time of school

construction; we restrict our discussion to the responses of these ten individuals.

Our questionnaire asked principals to rate the contribution of the SCP to student, parent, and teacher motivation, and the timing of any observed changes. We also asked about the improvements they observed in different facility attributes, like libraries, classrooms, and ventilation, and about how much they believed each improvement type contributed to academic performance. We then asked principals to weigh the relative importance of indirect motivational effects and direct 'new facilities' effects in improving students' scores. Copies of the survey are available upon request.

Principals agreed that the school construction project raised motivation at home and at school. All of the surveyed principals reported moderate to large effects of school construction on parent involvement, and nine of the ten reported large effects on student motivation. All principals reported moderate or large effects on teacher motivation. When asked to compare the importance of motivational effects to the importance of direct infrastructure effects for raising test scores, nine out of ten principals believed that the motivational effects of the SCP were at least as important as the direct effects of improved infrastructure on pedagogy. Though principals likely faced the same difficulties we do when trying to separate motivational from pedagogical effects, the surveys indicate that, at minimum, observed school construction effects are not entirely due to direct pedagogical changes. This is consistent with the emphasis placed on community, student, and teacher involvement in the construction process, and with a growing economic literature on the importance of intrinsic motivation in determining student outcomes (see, e.g., Heckman et al. (2006)).

When asked to identify specific building features that were important to student success, principals pointed to library improvements and heating, air conditioning and ventilation. Particularly interesting in this part of the survey were the responses principals gave to an open-ended question in which they were asked to identify important pathways through which school construction affected student outcomes that had not been identified elsewhere on the survey. Several responses focused on 'student and teacher pride,' while others identified important but subtle building features, such as a glass wall which allowed teachers to observe student activities in hallways while standing in a central courtyard location. An implication is that some of the infrastructure features that determine student achievement may be a) relatively inexpensive, but b) difficult to measure or categorize. This presents both an opportunity and a challenge to designers of future infrastructure improvements.

## 9 Discussion

This paper describes the effects of a comprehensive school construction program in a poor urban district on student and community outcomes. We find that school construction had substantial positive effects on home prices in affected neighborhoods, and led to increases in the population of families with children attending public schools. These effects coincided with increases in student reading scores on the order of those typically observed in students who win lotteries to attend high-performing charter schools. Given the pressing need for large-scale investment in school infrastructure at the national level, and in poor, urban areas in particular, our findings are important for assessing the costs and benefits of potential infrastructure policies.

The evidence presented here also links prior work on the home price effects of school construction to a broader literature on the way that housing markets capture school quality. We document for the first time the way that dynamic changes in school quality (and other amenities associated with school construction) impact home prices and patterns of public school enrollment, and in doing so help explain how the social stratification along school boundaries described in Bayer et al. (2007) could arise over time due to local changes in education policy. We innovate further with respect to the housing market literature by focusing on the price effects of changes in the causal effect of schools on student scores, not on school average scores that mix differences in education production with student selection. This distinction is important if we wish to separate the value families place on school quality and the value they place on attending schools that students with high levels of observable and unobservable test score determinants also attend.

Our work has number of important limitations. We cannot determine whether school construction affects test scores through specific changes to the built environment that enhance pedagogy, or through more generalized changes in student, parent, and teacher motivation that accompany the project. A survey of school principals suggests that both physical and motivational changes play an important role. A corollary is that we cannot identify specific building features that are particularly important for improved educational outcomes, though more expensive projects are associated with larger effects. Our attempts to compute the elasticities of home prices with respect to changes in school value added are hampered by the fact that we cannot determine the extent to which home buyers value other amenities associated with school construction. We therefore interpret the estimated elasticities as

upper bounds on the true effects of test scores on prices.

We conclude with a broader discussion of the relationship between school construction and other policy interventions aimed at helping students in poor urban districts. The basic challenge in these districts is to help students from low-SES backgrounds succeed in school despite limits on local resources. Many current policies aim to help students who have the wherewithal to seek out educational opportunity leave troubled schools or districts for privately-run alternatives. In at least some instances— notably a subset of high-achieving charter schools— students who win admissions lotteries realize large score gains. What is unclear is the extent to which these policies are scalable: straightforward models of economic behavior suggest that students who do not opt in to high-achieving charters would benefit less from attendance than those who do. Further, there may be negative spillovers from choice-based policies if the students who exercise choice no longer positively influence those who do not, though empirical evidence suggests that these effects are not large (Altonji et al. (2010)).

School construction differs from choice-based policies because students do not have to opt in. With this in mind, the observed score gains may be even more impressive, because they are not limited to students who express an interest in improving their academic outcomes. The sticker price of school construction projects is much higher than the price of choice-based reforms in almost every case. But, given the poor state of infrastructure in poor urban districts, some school construction costs are fixed in the sense that they must surely be undertaken at some point in the relatively near future. At minimum, the results we present here indicate that when this construction occurs, it should be viewed not as an unfortunate necessity but as key part of the broader school reform toolkit.

## References

- Altonji, Joseph G., Ching-I Huang, and Christopher R. Taber**, “Estimating the Cream Skimming Effect of School Choice,” NBER Working Papers 16579, National Bureau of Economic Research, Inc December 2010.
- Angrist, Joshua D., Susan M. Dynarski, Thomas J. Kane, Parag A. Pathak, and Christopher R. Walters**, “Who Benefits from KIPP?,” NBER Working Papers 15740, National Bureau of Economic Research, Inc February 2010.
- Bayer, Patrick, Fernando Ferreira, and Robert McMillan**, “A Unified Framework for Measuring

Preferences for Schools and Neighborhoods,” *Journal of Political Economy*, 08 2007, 115 (4), 588–638.

**Black, Sandra E.**, “Do Better Schools Matter? Parental Valuation Of Elementary Education,” *The Quarterly Journal of Economics*, May 1999, 114 (2), 577–599.

**Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein**, “The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design,” *The Quarterly Journal of Economics*, February 2010, 125 (1), 215–261.

**Dobbie, Will and Jr Roland G. Fryer**, “Are High Quality Schools Enough to Close the Achievement Gap? Evidence from a Social Experiment in Harlem,” NBER Working Papers, National Bureau of Economic Research, Inc November 2009.

**Duflo, Esther**, “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment,” *American Economic Review*, September 2001, 91 (4), 795–813.

**Filardo, Mary**, “GROWTH and DISPARITY : A Decade of U.S. Public School Construction,” Technical Report, Building Education Success Together – BEST 2006.

**Hanushek, Eric A.**, “Assessing the Effects of School Resources on Student Performance: An Update,” *Educational Evaluation and Policy Analysis*, Summer 1997, 19 (2), 141–164.

**Heckman, James J., Jora Stixrud, and Sergio Urzua**, “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior,” *Journal of Labor Economics*, July 2006, 24 (3), 411–482.

**Kennan, John and James R. Walker**, “The Effect of Expected Income on Individual Migration Decisions,” *Econometrica*, 2011, 79 (1), 211–251.

**Meghir, Costas and Steven G. Rivkin**, “Econometric Methods for Research in Education,” NBER Working Papers 16003, National Bureau of Economic Research, Inc May 2010.

**Rothstein, Jesse**, “Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement,” *The Quarterly Journal of Economics*, February 2010, 125 (1), 175–214.

**Todd, Petra E. and Kenneth I. Wolpin**, “The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps,” *Journal of Human Capital*, 2007, 1 (1), 91–136.

## Appendix

### Appendix A Descriptive Statistics

*Table A -1: State of Service Systems in Connecticut Schools: Principals Survey*

Less than good	2001	2009
Hartford	0.54	0.30
New Haven	0.53	0.14
Connecticut Average	0.32	0.18

Note: Percentages of school systems deemed 'less than good' in survey of public school principals. Categories include : Internal Communications , Interior Lighting, Technology Infrastructure, Exterior Lighting, Air Conditioning, Roadways and Walks, Heating, Plumbing/Lavatories (Scale: 4 = excellent, 3 = good, 2 = fair, 1 = poor, 0 = missing.)

*Table A -2: School construction project summary*

Number of schools	Elem/MS	HS
Total schools	33	9
Planned	31	6
Constructed	27	6
Occupied	25	5
Expenditures (millions of 2005 dollars)		
Mean	34.07	50.09
Median	37.60	48.35
75th percentile	40.60	64.07
25th percentile	29.79	35.16
Duration (in years)		
Filing to occupancy	6.08	6.93
Construction to occupancy	1.74	2.64

Source: NHPS. Counts exclude transitional schools and count each school address as a separate school.

## Appendix B Results - Residential Choices

Table B -3: School enrollment by neighborhood

$t - \tau_o$	Enrolled	Inflows	Outflows	New dist.	New nbd.	Leave dist.	Leave nbd.
-5	0.0058	0.0071	0.0011	0.0088	0.0106	-0.0021	0.0095
	0.0037	0.0064	0.0059	0.0078	0.0122	0.0077	0.0103
-4	0.0075**	0.0087	0.0056	0.0025	0.0115	0.0051	0.0083
	0.0035	0.0061	0.0056	0.0074	0.0116	0.0073	0.0098
-3	0.0081**	0.0101	0.0108*	0.0081	0.0119	0.0047	0.0188*
	0.0036	0.0062	0.0057	0.0075	0.0117	0.0074	0.0099
-2	0.0099**	0.0078	0.0068	0.0104	0.0113	-0.0018	0.0198*
	0.0036	0.0065	0.0059	0.0079	0.0123	0.0077	0.0104
-1	0.0097**	0.006	0.0035	0.0047	-0.0051	-0.0041	0.0086
	0.0039	0.0066	0.0061	0.0081	0.0126	0.008	0.0107
0	0.0172**	0.0181**	0.0089	0.0179*	0.0174	0.0017	0.0226*
	0.0046	0.0078	0.0072	0.0096	0.0149	0.0094	0.0126
1	0.0228**	0.0210**	0.0043	0.0121	0.0339*	0.0058	0.003
	0.0059	0.0102	0.0094	0.0125	0.0195	0.0123	0.0165
2	0.0295**	0.0332**	0.0064	0.0285**	0.024	0.0053	0.0064
	0.0065	0.0113	0.0104	0.0138	0.0215	0.0136	0.0182
3	0.0344**	0.0357**	0.0253**	0.0384**	0.0385	0.018	0.0425**
	0.0072	0.0126	0.0116	0.0154	0.024	0.0152	0.0203
4	0.0404**	0.0494**	0.0272*	0.0514**	0.0518*	0.0263	0.0304
	0.009	0.0157	0.0144	0.0191	0.0298	0.0188	0.0251
5	0.0460**	0.0558**	0.0326**	0.0459**	0.0776**	0.0315	0.0426
	0.0098	0.0171	0.0157	0.0209	0.0325	0.0205	0.0275
6	0.0441**	0.0671**	0.0497**	0.0565**	0.0855**	0.0494**	0.0602*
	0.0112	0.0197	0.0181	0.0241	0.0375	0.0237	0.0317
N	200	175	175	175	175	175	175
N	200	175	175	175	175	175	175
Nbd. FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: \*\*: significant at 5% level. \*: significant at 10% level. Effects of per-capita construction expenditures by year relative to treatment on neighborhood-level log enrollment flows. Observations are at the neighborhood year level. See section 6 for detailed explanation and definitions

## Supplemental Appendix - For Online Publication

### Appendix A Description of school construction projects.

This supplemental appendix describes two representative school construction projects. An official description of the SCP as a whole and photographs of many of the new buildings are available online at <http://nhps.net/SchoolConstruction>. The text quoted here comes from a District-provided summary of SCP status as of May 2010.

The Fair Haven School at 164 Grand Avenue in New Haven was completed in 2004. The changes at this school illustrate the SCP's focus on improving the school environment and community access as well as adding basic amenities like heat and air conditioning.

The District describes the improvements as follows.

- 'When built in the late 1920s, this school provided an architectural anchor to the neighborhood... its disrepair was extensive by the end of the century.'
- '[C]lassrooms were enlarged, updated to current technology standards, and, heating and air conditioning were installed.'
- 'The interior of the school has been completely reordered and the library, cafeteria, and gymnasium replaced with those accessible to the students and the community. [A]n addition was required to build a regulation size gymnasium.'
- 'The light wells introduced natural light throughout the 4-story structure through the skylights and the stairwells, and terminated in the newly designed and greatly expanded library and cafeteria spaces at their bases. Each room, though located at the center of the building, receives natural light from two wells.'

The Mauro-Sheridan School at 191 Fountain Street in New Haven was completed in 2009. The changes at this school illustrate the ways in which the SCP improved students' access to technology. The District describes the changes as follows.

- 'The 1954 addition has been demolished and a new addition built... The 1922 building has been renovated and features such diverse technology offerings as robotics, high tech graphic arts studies, instrumental and electronic studies, as well as advanced computer studies.'

- ‘The auditorium has been renovated as a flat-floor, multipurpose music space... The conversion of the balcony to a video lab allows for the recording of performances.’
- ‘[A] diverse technology curriculum includes: the NASA partnership (applied technology laboratory featuring mechanical design and robots); a video production lab that supports literacy skills; a distance learning lab for interactive global learning; computer labs; science labs; electronic music laboratories; and the integration of technology throughout the building to enhance instruction.’

## Appendix B Student selection

In this appendix we examine the relationship between selection into neighborhoods affected by school construction and the observable determinants of student test scores. If students moving into new neighborhoods are better students than the students there prior to occupancy, we might attribute changes in mean school scores across time to school construction when in fact they are a product of the changing student population. We address this question by looking at how the characteristics of neighborhood populations relate to time relative to occupancy.

To do this, we create an index of observable and predetermined characteristics weighted by the role each plays in determining test scores, and document how this index varies with time relative to treatment. We construct the index using the regression

$$T_{igz} = x'_{ig}\beta + e_{igz}, \quad (\text{Appendix B.1})$$

where  $x_{ig}$  is a vector consisting of dummies for race, sex, and free lunch status. We exclude other observable characteristics like special education status and ELL status because these are learning outcomes that could be endogenous to school construction.  $x'_{ig}\hat{\beta}$  thus forms an index of observable score determinants. We assess selection on the observables by estimating the equation

$$x'_{ig}\hat{\beta} = \gamma_t + \gamma_z + \gamma_g + \delta^{sel}D(\ell) + e_{izg}. \quad (\text{Appendix B.2})$$

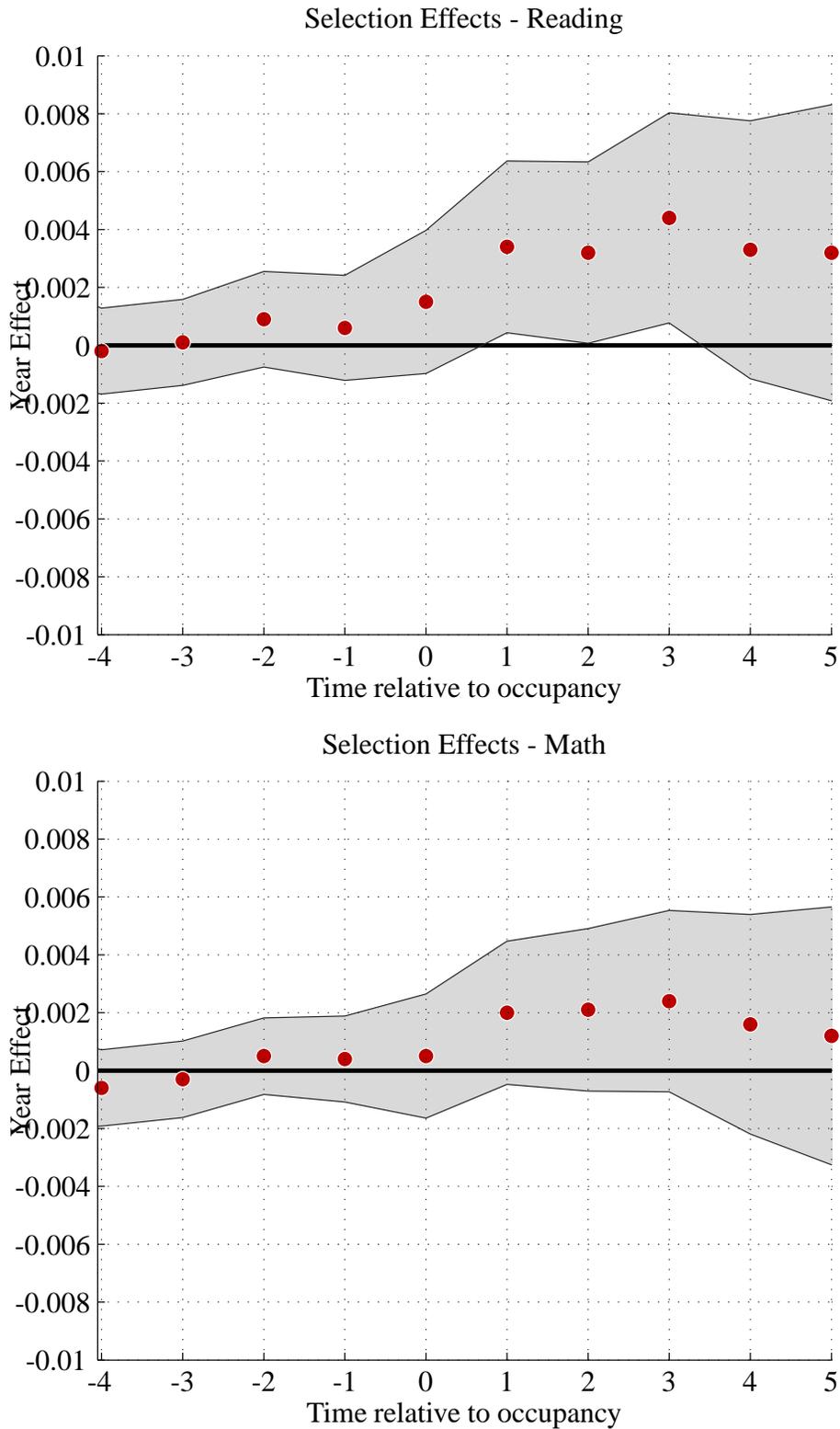
The  $\gamma_t$  are year fixed effects, the  $\gamma_g$  are grade effects, and the  $\gamma_z$  are neighborhood fixed

effects. We do not correct our standard errors for sampling error in  $\hat{\beta}$ ,<sup>18</sup> but do allow arbitrary correlation in error terms within neighborhood-year cells. We graph our results in **Figure B -1**. We find that levels of observed selection on the reading and math indices are close to zero and insignificant until one year post-occupancy. In that year, values of both of the indices jump substantially, and the estimated coefficient in the reading specification becomes statistically significant at the ten percent level. The reading and math selection indices both remain at high levels for several years, though estimates become noisy. As we discuss in the main text, aggregating the year-specific effects into pre-and post-occupancy groups yields a finding of statistically significant post-occupancy selection in both reading and math.

---

<sup>18</sup>Intuitively, correcting for sampling error could only reduce the statistical significance of evidence for selection. In practice we find that such corrections have little effect because the  $\beta$  are tightly estimated.

Figure B -1: Predicted Test Scores by Year Relative to Occupancy



Note: Estimates of effects of per-capita school construction spending on indices of observable score determinants by year relative to treatment, as described in equations C1 and C2. Shaded areas show a 90 percent confidence interval, allowing for clustered errors at the neighborhood-year level.