SIEPR Discussion Paper No. 10-010

Heterogeneity in Financial Incentives for High and Low Income School Districts

by
George Bulman

Stanford Institute for Economic Policy Research
Stanford University
Stanford, CA 94305
(650) 725-1874

Heterogeneity in Financial Incentives for High and Low Income School Districts☆

George Bulman[a]

[a]*Stanford University, Department of Economics, 579 Serra Mall, Stanford, CA 94305. Phone: (650) 521-1603. Fax: (650) 723-2745. Email: gbulman@stanford.edu.*

Abstract

Students who attend schools in low income communities have significantly lower achievement and make smaller gains than their counterparts in high income communities. Explanations for this achievement gap are generally based on constraints in financial resources. However, when equalizing financial resources fails to close the gap in practice, the residual is attributed to student aptitude and family spillovers. This paper explicitly tests for the presence of heterogeneity in school district incentives as an alternative to these explanations. A rich data set of Ohio house sales, characteristics and construction quality is used to estimate the responsiveness of the property tax base to school quality across high and low income districts. Estimates indicate that house prices are more responsive to quality in high income suburban districts than in low income urban districts. This provides direct evidence that residents of high income districts are more willing and able to pay for education quality, which translates into stronger incentives for the districts that serve them.

Keywords: district incentives, public schools, capitalization, achievement gap
JEL codes: I22, H73, H71, H42

1. Introduction

Large gaps in achievement exist between students from poor and affluent households, and these gaps are attributable either to differences in constraints or differences in incentives. Differences in constraints tend to be the more obvious suspects, and they have been the focus of attention for most parents, educators, researchers, policy makers, and courts. Prominent examples of constraint-type explanations are as follows: poor children may be constrained to attend schools with lower per-pupil budgets; they may be constrained to enjoy fewer human

☆I am grateful to Mr. Wallace Hawley for generously supporting the Hawley-Shoven Fellowship with the Stanford Institute for Economic and Policy Research. I would also like to thank Ran Abramitzky, B. Douglas Bernheim, Seema Jayachandran, Caroline Hoxby, Arthur van Benthem, and Stanford seminar participants for providing valuable comments.

capital spillovers from their parents; they may be constrained by lower aptitude. Although evidence about the relationship between school resources and student achievement is contested, it is fair to say that the evidence does not indicate that poor children achieve as well as affluent children if they attend schools with equal financial resources (Card and Payne, 2002). Although spending gaps between schools that serve poor and affluent students have been greatly or fully closed in most states over the last four decades, the achievement gaps between such students have persisted (Vanneman et al., 2009).[1] Thus, because people have focused on constraint-type explanations for the achievement gap, the evidence has tended to make them think that family-related explanations must be important–for instance, poor students are constrained to enjoy fewer human capital spillovers from their parents or poor students are constrained by lower aptitude. These explanations are relatively unappealing because they are less susceptible to amelioration through policy. (Because human capital spillovers and aptitude are very hard to measure, they serve as residual explanations that always remain when evidence for more testable explanations fails to materialize.)

The focus on constraints to the exclusion of incentives is, however, unwarranted. There are several very plausible explanations for the achievement gap that rely on differences in incentives. For instance, compared to schools that serve affluent students, schools that serve poor students may receive less financial reward when they raise achievement. This would result in weaker incentives for administrators and make it less feasible to give staff incentives that are as high-powered as those in schools serving affluent students. Weaker financial rewards in poor students' schools are very plausible because of the combination of local property taxes and state transfers that account for most school revenue. This is explained below. As another example of an incentive-type explanation, consider that school staff who serve poorer students may receive weaker social rewards when they raise achievement. Since teachers who serve middle-income and affluent students are more likely to live in the community where they work than teachers who serve poor students, it is easy to imagine how social and other indirect rewards could be greater for teachers who serve more affluent students (Boyd et al., 2005). Parents in more affluent communities may be more capable advocates for their children, generating stronger incentives for teachers and administrators to avoid negative interactions. Differences in school district incentives implies that equal revenue does not necessarily generate equal school quality. Explanations that involve social rewards are difficult to test rigorously, but explanations that involve financial incentives are testable. This paper pro-

[1] Ohio Department of Education expenditure data indicate that total per-pupil revenue from all source between 2003 and 2008 averaged 8,987 dollars in districts classified by the state as "Urban/Suburban-high income", 9,950 dollars for "Urban/Suburban-very high income", 9,201 dollars for "Urban-high poverty", and 11,431 dollars for "Major Urban-very high poverty". Average scaled scores on the Ohio Achievement Tests and Ohio Graduation Test for these district classifications were 424, 435, 412, and 404 respectively. Likewise, the recently implemented value added measures for a year of schooling in 2008-2009 were 1.44, 2.35, 0.70, and 0.59 standard deviations respectively.

vides the first rigorous, direct test of whether schools that serve poorer students enjoy weaker financial rewards when they raise achievement.

Although constraint-based and incentive-based explanations are compatible (not mutually exclusive), the extent to which each is responsible for the achievement of poor students matters. If constraints are the primary obstacles to achievement, then the most effective policies would alleviate these constraints. For instance, school finance equalization programs that take revenue from (unconstrained) districts that serve affluent students and give the revenue to (constrained) districts that serve poor students. If incentives are important, then effective policies would equalize incentives surrounding poor and affluent students' achievement. Ironically, some policies designed to relieve constraints may unintentionally weaken incentives, and perhaps more so for districts that serve poor students. For instance, in a later section of this paper, we will see that certain types of school finance equalization essentially eliminate tax base financial incentives for schools that serve poor students.

For at least several years, researchers have theorized that schools that serve richer students face stronger incentives than schools that serve poorer students. The primary channel they have identified works as follows. Property taxes generate the majority of revenue for school districts in the U.S. Also, schooling is the only local public good supplied by independent school districts (most school districts are independent) and by far the most important local public good provided by municipalities with dependent school districts. Therefore, when a household considers residing in one jurisdiction versus another, the effect of local public schools on achievement ("school quality") is the main local public good that counterbalances a higher house price. That is, if–all else equal–the local public schools in a jurisdiction raise achievement more, households should be willing to pay more for property in that jurisdiction. Thus, a school district that raises achievement will see its property prices rise to capitalize the benefits of high quality schools and this translates mechanically into greater property tax revenue if property tax rates are stable (the process by which rates are set favors inertia in nearly all of the U.S.). Greater property tax revenue translates into more generous school district budgets. School leaders thus see their budgets expand when their schools raise achievement, generating an incentive for them, and allowing them to use their expanded budgets to reward staff via salary increases, more generous benefits, smaller class sizes, or nicer working conditions on any other dimension.

If the process described above works more elastically in districts that serve high income households, then these districts will have greater financial incentives to raise student achievement than the schools that serve poor households. That is, if a given increase in achievement, all else equal, generates a greater price increase in properties in high income districts, then districts that serve affluent households will see greater property tax revenue increases, greater budget increases, and more positive financial rewards when they raise achievement. Further, the presence of differences in financial incentives may be indirect evidence of differences in social incentives: households that are willing to pay for school quality may be more willing to apply pressure on schools and districts.

If house prices are more elastic in affluent districts, home owners may be more invested in school quality.

To test whether the financial incentive explanation for the achievement gap exists in practice, I use extremely accurate administrative data on every household and school district in 42 Ohio counties. Since Ohio finances schools in a manner that is typical of U.S. states, the evidence in this paper can reasonably be extrapolated to most parts of the country. My test employs a regression discontinuity design that very credibly identifies the effect of student achievement on property prices, as opposed to the effects of many other variables that probably affect both achievement and property prices. Specifically, I look at the prices of very similar houses just on one side and the other of school district boundaries. These houses share the same neighborhood, the same neighbors, and most local amenities (parks, roads, views, access to public transportation, and so on) except that they do not share local schools. Controls are used not only for house size, age, and location, but also for the quality of construction. I analyze whether the property price increment associated with a given increase in average student achievement is larger in absolute terms and as a percent of house value when the households are on the boundary between high income districts. If, for instance, the same standard deviation difference in student achievement generates disproportionately greater property price increases on a boundary between affluent districts than on a boundary between poor districts, I have found evidence supporting financial incentives as an explanation for the achievement gap. The district is the correct level of aggregation for examining financial incentives because it is the level at which revenue is raised, received from the state, and payrolls are administered.

It is worth noting that, for the purposes of this paper, it is not necessary to separate the effect of achievement on property prices into direct effects and endogenous effects. A "direct effect" is an effect of achievement on property prices that flows through the channel of personal, direct benefits–that is, a family values a property more because it is associated with a school that will raise its child's test scores and (thus) lifetime earnings more. An "endogenous effect" is an effect of achievement on property prices that is proximately caused by some third variable that is itself attracted by higher achievement. For instance, a family values a property more because it is associated with a school that has more desirable peers who have been attracted into the area by the local school's positive effects on achievement. Since both direct and endogenous effects are a function of school quality, they need not be separated to observe differences in incentives. It is irrelevant whether house prices are responding to the gain in student achievement directly or to a change in the composition of peers stemming from student achievement. That is, when a district superintendent sees his budget increase because local property prices have risen to capitalize his school's improved effect on achievement, he does not care how much the rise in prices is due to direct versus indirect effects. It is fortunate that dividing the effects into direct and endogenous types does not matter for this paper: previous researchers have found it very difficult to make the division without imposing many assumptions.

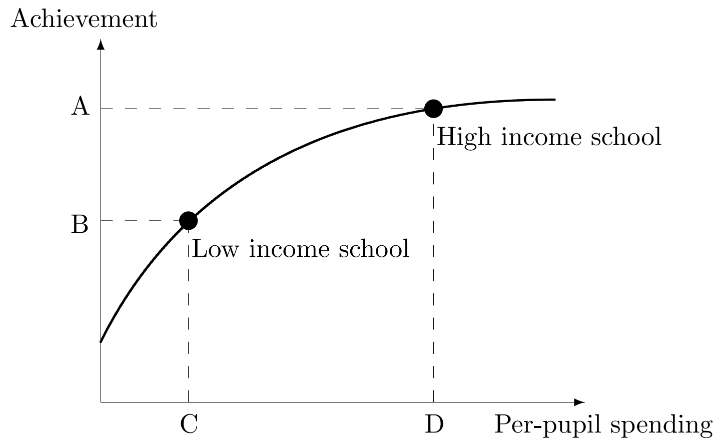
The remainder of the paper proceeds as follows. Section 2 illustrates how an incentives based explanation of the achievement gap can ease the tension between constraint based explanations based on school resources or aptitude and family spillovers. I also explain why the property tax base is especially responsive to school quality. In Section 3, I survey the extensive literature pertaining to school resource constraints and student achievement, identify papers that have suggested the potential importance of financial incentives, and review key papers that have examined the relationship between house prices and school quality. Section 4 describes the rich and extensive administrative data set of property attributes, sales, and district achievement data constructed for this paper. Section 5 discusses the primary obstacles to identification and the preferred specifications. In Section 6, I present my main results, which show that school districts that serve affluent households enjoy much greater property price increases when they have greater student achievement, all else equal. Section 7 adds several specification checks of the main results. In Section 8, I convert my property price findings into effects on school district incentives assuming no state role in district finance. I show how districts' financial incentives are affected by Ohio's statewide school finance program which redistributes revenue to districts with low property values per pupil, and I demonstrate how financial incentives for districts would be changed if Ohio were to adopt the California system of statewide school finance–a system that is of special interest because it represents an extreme form of school finance equalization.

2. Theory

Constraint based explanations generally attribute the student achievement gap to either differences in school resources or differences in student capacity to achieve (due to ability and family spillovers). Resource based explanations suggest that schools in low and high income districts would have similar student achievement if they had equal revenue (adjusting for differences in the prices of school inputs). In essence, school quality is assumed to be a function of revenue levels. Conversely, aptitude and family based explanations suggest that students in low income districts make smaller gains and have lower achievement due to factors unrelated to school quality. This paper tests for heterogeneity in incentives as an alternative to these constraint based explanations. If school districts with high income residents have systematically greater incentives to produce student achievement than districts with low income residents, a gap in school quality and thus an achievement gap can exist even if the student populations are equally capable of making gains and resources are equal. Figures 1-3 clarify the role of incentives in easing the tension between the competing constraint based explanations.

Resource based explanations of the student achievement gap assume that schools in low income districts are hindered by a lack financial resources. With greater per-pupil revenue, these schools can hire more and better teachers and administrators, buy higher quality textbooks and educational software, and thereby provide an educational product that rivals schools in wealthy districts.

Figure 1: Revenue Constraints†



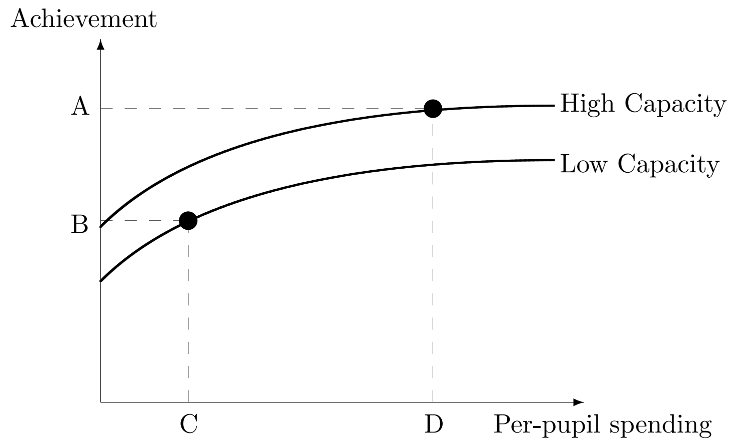
† This illustrates a resource based explanation for the achievement gap. The achievement gap A-B between high and low income districts is a function of the difference in resources D-C. In this scenario, closing the resource gap will eliminate the achievement gap.

Figure 1 represents this graphically. High and low income districts share the same achievement frontier, with per-pupil funding on the horizontal axis and achievement on the vertical axis. The difference A-B represents the achievement gap and stems solely from the difference in funding levels D-C. Reallocating revenue from high income districts to low income district will close the achievement gap. Under the assumption that the returns to spending are concave, redistribution will also increase average achievement.

Aptitude and family based explanations of the achievement gap assume that factors unrelated to school quality significantly determine student capacity to learn. Differences in innate ability and spillovers from parents determine potential achievement. Figure 2 represents this as separate achievement frontiers for high and low income districts. The achievement gap A-B and funding gap D-C are the same as those in Figure 1. The gap between the achievement frontiers implies that an achievement gap will remain even if resources are equal. Further, because high and low income districts do not share the same achievement frontier, redistributing revenue to low income districts does not necessarily raise average achievement.

Resource and aptitude based explanations are not mutually exclusive, and both are likely to play a role in explaining the achievement gap. However, when eliminating funding differences between districts with high and low income residents fails close the achievement gap, we are left to conclude that aptitude

Figure 2: Learning Capacity Constraints†



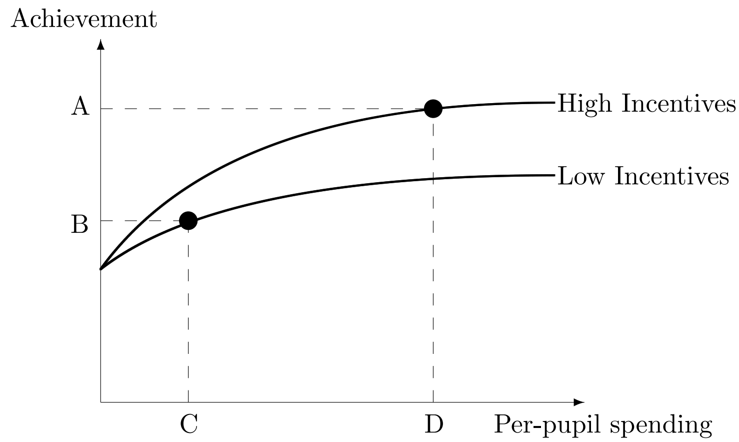
† This illustrates a learning capacity based explanation for the achievement gap. Students in low income districts perform less well due to constraints in family spillovers or aptitude. The achievement gap A-B between high and low income districts is a function of the difference in resources D-C and a difference in learning capacity. An achievement gap remains even if the resource gap is close.

and family factors are important. This implies that education policy and school finance reform might be insufficient to address the achievement gap.

The alternative explanation examined in this paper is that districts with high and low income residents face systematically different incentives. Thus districts that have equal resources are not necessarily of equal quality. A district whose funding is sensitive to the quality of the public good it provides or that serves a population with high expectations is likely to use resources more efficiently than a district that does not. Such districts might make better hiring, tenure, and curriculum decisions, and provide stronger incentives for teachers. Figure 3 exhibits how the achievement frontiers of schools that face high and low levels of incentives could differ. In this representation, schools with high incentives are more efficient at converting financial resources to student achievement. Thus schools with equal resources and children of similar aptitude can differ in average achievement. Eliminating the gap in per-pupil spending is inadequate to close the achievement gap. This paper tests explicitly for such a systematic difference in incentives across high and low income districts. The presence of this heterogeneity would imply that school quality, rather than student aptitude or family factors, could be partially responsible for the achievement gap after financial resources are equalized.

The property tax base is the primary source of school revenue and is a potentially powerful financial incentive. Two factors in particular make the property

Figure 3: Incentives†

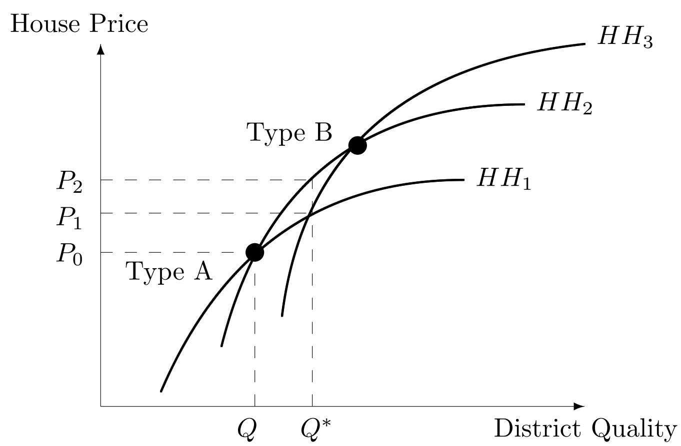


† This illustrates a school incentives based explanation for the achievement gap. The achievement gap A-B between high and low income districts is a function of the difference in resources D-C and the efficiency with which schools convert resources to student achievement. This is represented by different achievement frontiers for school districts that face high and low incentives. An achievement gap can exist even if the resource gap is eliminated and capacity to learn is not different.

tax base especially responsive to school quality. First, the assessed values of all houses in a community are a function of the prices of the most recently sold houses. When a district improves in quality and drives up the prices of houses currently on the market, assessors formulas systematically incorporate this increase into the assessed values of neighboring houses. This results in a rapid response of the entire tax base to current district quality. In contrast, the socioeconomic composition of a district changes more slowly. New residents and pupils are not weighted more heavily than existing residents when computing alternative measures of district wealth such as median household income, average household income, or the percent of students who qualify for free and reduced lunches. This difference in responsiveness between the tax base and other measures of wealth is important when considering state finance systems in Section 8.

Second, the prices of houses in a district are not determined by the average resident of that district, but by the marginal resident who is close to indifferent about living in that district or another. When a school district improves in quality, it will attract those residents who are were previously on the margin about living there. Theory indicates that these residents will, on average, be willing to pay more for the change in school district quality than would the average resident. This is illustrated in Figure 4.

Figure 4: Marginal Mover Between Districts†



† The indifference curves for three types of households are displayed. When a Type A district improves its quality from Q to Q*, it attracts Type 2 households who were previously indifferent between District A and District B. These households place a higher value on the change in education quality than Type 1 households: $P_2 > P_1$.

School district quality is on the horizontal axis and the average market price of houses in that district is on the vertical axis. The indifference curves for three households are shown, with utility increasing in school district quality and decreasing in house prices. For simplicity we assume that there are two types of districts, Type A and Type B, with the latter being higher quality. Household 1 has a low valuation of school quality and resides in a Type A district; Household 3 has a high valuation of school quality and resides in a Type B district; and Household 2 has an intermediate valuation of school quality and is indifferent between living in a Type A or Type B district. If a Type A district improves in quality from Q to Q*, Household 2 can pay as much as P_2 to reside in the improved district without a loss in utility. By comparison, Household 1 could pay only P_1 without a loss of utility. The amount that Household 2, the marginal resident, will pay for an improvement in school quality exceeds the amount that Household 1, the infra-marginal resident, values the change. In brief, an improvement in school quality can generate a relatively large increase in house price because it attracts households that place a relatively high value on education and this large increase is mechanically factored into the assessed value of neighboring houses.

3. Literature

An extensive literature has examined the role of school resources (most notably per-pupil spending and the student-teacher ratio) on student test scores, educational attainment, and future earnings. The findings of this important and influential literature vary with the decade of the cohort examined, the aggregation level of the data, and the empirical approach. It is generally agreed that the literature as a whole has produced inconclusive or limited evidence of the positive effects of additional resources alone. Several handbook chapters summarize the findings in excellent detail (Betts, 1996; Card and Krueger, 1996; Hanushek, 1997, 2006).

A related literature examines the effectiveness of state finance equalization formulas for generating equal revenue and equal achievement in high and low income districts (Card and Payne, 2002; Downes et al., 1997; Evans et al., 1997; Hoxby, 2001; Wenglinsky, 1998). Centralized school finance formulas have unambiguously produced more equitable revenue across districts. As in the broader literature, there is some evidence that this revenue equalization has produced achievement gains for students who attend low income schools, but the magnitude of the effect leaves much of the achievement gap unexplained. As a result, it is suggested that differences in ability and family spillovers are an important factor (Duncan and Magnuson, 2005). In reference to the effects of various education reforms on racial and economic achievement gaps, Fryer observes that: "The lack of progress has fed into a long-standing and rancorous debate among scholars, policymakers, and practitioners as to whether schools alone can close the achievement gap, or whether the issues children bring to school as a result of being reared in poverty are too much for even the best educators to overcome." (Fryer, 2010). In this debate, constraints in family spillovers represent the alternative when school reforms fail to close the achievement gap.

Others attribute the failure of resources to radically improve schools to the economic environment in which schools operate. As Hanushek posits: "The presumption of efficient provision is suspect when government produces the services. And, if the resource use is inefficient, the relationship between added resources and outcomes is unclear."(Hanushek, 2006) This thinking has lead policy makers to implement programs intended to mimic more competitive markets using tools such as vouchers and charter schools. As with the resource constraints literature, the effects of such programs are widely contested.

The reasoning that incentives and efficiency may be an important factor in determining quality has typically referred to public school inefficiency in general, but some have noted that the environment in which schools operate varies systematically between high and low income communities (Bayer and McMillan, 2005). Systematically weaker incentives (financial or social) for districts that serve low income students could lead to less efficient use of resources and thus promote an achievement gap even when resource constraints are eased. It is this hypothesis, that high and low income school districts face fundamentally different incentives, that is tested in this paper.

The majority of school revenue comes from property taxes retained at the district level or redistributed through a centralized state fund, so it is natural to examine the elasticity of property values (the tax base) to school quality to determine if high and low income districts face heterogeneous financial incentives. Sandra Black's seminal paper pioneered the use of school attendance borders as a means of controlling for unobserved community quality, the primary challenge to identification in these hedonic regressions, to estimate the average responsiveness of house prices to elementary school test scores (Black, 1999). One concern with using these estimates to infer district incentives is that they rely on a static difference in scores rather than a response to a dynamic change. The ideal experiment would exploit an exogenous change in school quality, but such an opportunity is unlikely to emerge in practice. However, Figlio and Lucas were able to examine the capitalization over time of a change in *perceived* school quality, controlling for true school quality, after the release of state assigned school performance grades in Florida (Figlio and Lucas, 2000). They find large and rapid average changes in house prices in response to the release of school grades, indicating that inertia might not be an important concern.

Bayer et al. estimate the elasticity of house prices with respect to school quality, as well as other community characteristics, using several variations of a structural approach (Bayer and McMillan, 2005; Bayer et al., 2007). Their model reveals heterogeneous rates of capitalization across various socioeconomic groups in the Bay Area of California, but the results do not translate naturally or easily into a measure of how financial incentives vary for districts and schools that serve more affluent populations. This is especially true due to the public school finance system employed in California, which is unique in creating a disconnect between local property values and local district revenue. Bayer et al. wish to separate the direct and endogenous effects of school quality on property prices, which requires a structural model and a number of assumptions. As discussed in the introduction, this separation is unnecessary for examining the incentives faced by schools. Controlling for socio-demographic variables on either side of the border may indicate the extent to which the endogenous effects are an important consideration (though controlling for all such variables is impossible). I add demographic variables to gauge the signficance of the endogenous effect in my estimates, but find it to be a modest factor, which is likely due to the rich set of controls for house and neighborhood quality.

4. Data

The ideal experiment for testing household willingness to pay for student achievement and the resulting incentives for districts would exploit a rapid and exogenous shock to school quality. Enough districts would need to be affected by exogenous shocks to reliably measure how willingness to pay and thus incentives vary across districts of differing wealth. This would allow us to use variation in the prices of all houses in a district since all houses would be either treated or untreated. Unfortunately, such a randomized experiment would likely require the implementation of a statewide program specifically designed for research purposes.

The border discontinuity approach provides a feasible alternative, but is quite data intensive. Only houses in close proximity to borders are useful, and the nature of regression discontinuity requires dense data, so having the population of house sales is important. Likewise, to avoid attributing systematic differences in house quality to differences in district quality, it is important to have a rich set of house trait and quality controls. For the purpose of external validity, the state considered should be representative of other states in terms of its socio-demographic composition, housing market, and the structure of its public school finance system. For example, California would be relatively undesirable: school revenue does not depend on local property values and house prices are high relative to national averages (so capitalization may represent a lesser percent of house value). Each border represents an experiment, so the identification of capitalization across heterogeneous districts relies on having a large number of borders: affluent suburban districts bordering other affluent suburban districts, low income urban districts bordering other urban districts, and so on. Finally, to have an objective measure of district performance, and preferably one that is salient to parents, the state should have a long and consistent history of standardized exams.

Ohio was selected for this study because it satisfies each of these criteria. Ohio is home to a number of large and moderate sized metropolitan areas: Cincinnati, Cleveland, and Columbus are among the forty most populous metropolitan areas in the country, and Dayton, Akron, Toledo, and Youngstown are among the one hundred most populous. Thus the results in this paper are based on an average across several metropolitan areas. Further, Ohio is characterized by relatively small districts (e.g. Cleveland and its immediate suburbs are home to more than 30 districts), so each metropolitan area provides a network of borders between suburban and urban school districts–providing more border observations and the heterogeneity necessary for identification.

Historical house prices place Ohio slightly below the national median, so estimates of capitalization as a fraction of house price should be representative of other states. And, while convoluted ideosyncrasies make each state finance system unique, Ohio's underlying formula is based on the concept of foundation aid–the most common form of revenue equalization. The share of school revenue coming from local, state, and federal sources is close to national means.

County auditors in Ohio maintain precise, uniform databases of housing characteristics, sale prices, and sale validity codes, which they use to estimate property values and tax liabilities. These databases contain the population of properties and sales and have the accuracy and detail needed for precise estimation. The data set of Ohio properties, collected by the author from each county, spans the population of sales for 42 counties, 250 school districts, 1.3 million properties, and 2.2 million arms length sales. After restricting attention to only properties within 400 meters of a district border, the data includes more than 45,000 arm length transactions between 2003 and 2008. From the set of variables available from all counties, eight property characteristics were identified as having independent explanatory power of house prices (in approximate order of importance): finished living area, full bathrooms, construction quality, age at the time of sale, acres of land, bedrooms, rooms, and stories. The construction quality grade in particular is irreplaceable because no set of physical characteristics acts as suitable proxy. Estimates that do not include an equivalent measure are likely to be biased.

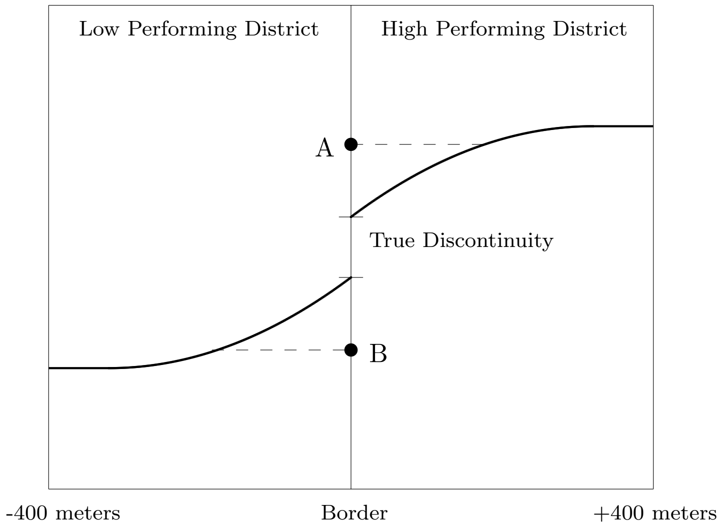
Only arms length sales of single family homes, which constitute the majority of residential property in Ohio, are included in the analysis. The paper uses sales between 2003 and 2008 because a single version of the state standardized tests was in place during these years. In 2003-2004, the state transitioned from the Ohio Proficiency Tests to the Ohio Achievement Tests (OATs) for grades 3 through 8, and the Ohio Graduation Test (OGT) for grade 10. District performance is computed as the average score in math and reading across all tested grade levels, weighted by the number of students taking the exam. These core exams were administered to all eligible grades for all years of interest, and their average should provide an accurate measure of student achievement across districts.

When using a regression discontinuity polynomial, it is important to have an accurate measure of the distance between the property and the neighboring district to estimate a distance trend. The precise location of each property is found using ESRI's ArcGIS geocoding software. Any properties for which the software was unable to find an exact match for the street number, street name, street direction, or town name, were dropped from the data set. The latitude and longitude of two properties are used to find the distance between them. District borders aligned with county borders are excluded to reduce the probability that a border is associated with a change in public services other than schooling, and because such borders are more likely to be delineated by a physical border or road that creates a discontinuity in community desirability. Attention is restricted to properties within 400 meters of district borders when a polynomial in distance is not included. The average distance between these properties is 220 meters.

5. Empirical Methods

Precisely measuring the capitalization of school quality is difficult due to the likelihood that many determinants of house price are correlated with school quality. For example, communities with less crime, better views, shorter commute times, more desirable businesses, and nicer parks may also have higher quality schools. A hedonic estimate that fails to control adequately for these factors may overestimate the effect of school quality on house prices. The border discontinuity design pioneered by Sandra Black, whereby attention is restricted to houses that are very close to one another along enrollment zone borders, addresses many of these concerns. Houses in the same neighborhood or on the same street share essentially the same determinants of community desirability, but the quality of the districts students are permitted to attend changes discontinuously at the border. The strategy is implemented by restricting attention to properties that abutt a district border and including a fixed effect for each border. Border fixed effects control for the desirability of the community due to

Figure 5: Unexplained House Value with Continuous Trends†



† This displays how continuous, unobserved factors could bias estimates of school quality capitalization. Unobserved community desirability is assumed to increase from the side of the border with the low performing district to the side with the high performing district. Failing to account for this trend results in an estimate of capitalization of A-B rather than the true discontinuity.

physical factors (climate, elevation, access to public transportation, proximity to work places and parks), economic factors that affect house prices (regional housing markets, employment rates), and factors that depend on the quality of neighboring residents (crime rates).

Challenges to this method of identification must stem from either a continuous or discontinuous change in desirability that is correlated with school quality. Examples of factors that are likely to change continuously are elevation, distance from businesses or public transportation, and crime rates. Concerns about the factors are essentially concerns that the bands around the border are too wide. Figure 5 illustrates a problematic scenario, with a rapid increase in unobserved community quality from the side of the border with the low performing district to the side with the high performing district generating upward bias in the estimates. Thus, in addition to border fixed effects, I include a polynomial in distance from the border, with negative distances on the low achieving side and positive distance on the high achieving side. This should reduce or eliminate bias due to continuous changes. The details of how this is implemented across multiple heterogeneous borders is discussed later in this section.

Challenges to identification due to a discontinuous changes at the border are almost certainly endogenous, such as the quality of other public goods, since exogenous factors tend to be continuous in nature. However, schools are often the only public good provided at the district level, and are certainly the most important. Many public goods also benefit adjacent houses relatively equally (e.g. roads, police presence, and fire department responsiveness). Discontinuous changes in desirability that are a response to school quality, such as a difference in the desirability of peers in schools, are not problematic in terms of measuring district incentives. In fact, to the extent that the polynomial in distance controls for some endogenous factors that change continuously at the border, the estimates could understate the incentives districts face. The quality of housing on either side of the border represents a possible exception to the clear differentiation between exogenous and endogenous discontinuities. Both are a concern: an exogenous difference would bias the estimates, while an endogenous difference might take effect too slowly to be an effective incentive. The full set of house characteristics is included in every specification to account for differences in the size and age of houses. Further, a measure of construction quality is included to capture variation that housing size measure can not.

Two outcome variables are used in this analysis: house sale price and annual housing payment including taxes. House sale price indicates the size of the change in sale price attributable to higher district performance. This estimate is useful for computing district incentives, as school revenue is a function of local property values. Annual payment for housing, which is the sum of the estimated annual mortgage payment and local property taxes, identifies the willingness of a household to pay for school district achievement.

Of specific interest to this paper are differences in the capitalization rates faced by districts with high and low income residents. Because the data set includes hundreds of districts of varying wealth and the borders between them, changes in the rate of capitalization of school performance can be measured in this dimension. Each district border is assigned an index, computed using the median incomes of the abutting districts. Thus borders between two high income districts will have a high index and the border between two low income school districts will have a low index. This index is used to split the borders in to two broad categories: low income borders (all borders that fall below the median index), and high income border (all borders that lie above the median index). The estimates for low income borders should be representative of the capitalization faced by districts with low income residents surrounded by other low income districts (such as those found in poor urban communities), while the estimates for high income borders should be representative of the capitalization faced by districts with high income residents surrounded by other high income districts (such as those found in high income suburban communities).

To identify which borders are driving the observed heterogeneity in capitalization, the borders are further divided into ten deciles. Decile 1 is comprised of borders between two very low income districts while decile 10 is comprised of the border between two very high income districts.

$$saleprice_{i,b,d,y,m} = \alpha_b + \alpha_y + \alpha_m + X_i'\delta_D \cdot Decile + \beta_1 \cdot \bar{T}_d \cdot LowIncome + \beta_2 \cdot \bar{T}_d \cdot HighIncome_b + \epsilon_{i,b,d,y,m} \quad (1)$$

Specification 1 illustrates how the traditional border discontinuity approach is extended to observe heterogeneity in capitalization between high and low income districts. The dependent variable $saleprice_{i,b,d,y,m}$ is the sale price of property i in district d on border b that was sold in year y and month m. The sale price is regressed on a border fixed effect α_b, year and month fixed effects α_y and α_m, the interaction of housing characteristics with dummies for the ten income deciles, and district test score performance $\bar{T}_d$ interacted with a dummy indicating if the border is categorized as low or high income. $\bar{T}_d$ is the average performance of district d on standardized exams in terms of standard deviations from the mean. High income borders are those in deciles 6-10 as determined by the median incomes of the districts that form it, and low income borders are those in deciles 1-5. Year and month fixed effects allow average house sales prices to vary both seasonally and non-monotonically across years. Border fixed effect acts as a community quality control and produce unbiased results under the assumption that variation in the quality of the border community does not vary meaningfully between the higher and lower performing district within the narrow band.

The preferred specifications interacts the housing controls and the construction quality grade with the ten income deciles, allowing the estimated effect of each to vary freely across the income distribution. If the estimated coefficients on house traits were restricted to be the same for the entire sample, they would likely under control for the contribution of traits to house prices in more expensive communities. As a result, if houses on the high performing side of the border are slightly larger on average, their predicted value will be systematically underestimated in high income communities, generating an unexplained gap that could be spuriously attributed to district quality. Thus the coefficients are estimated separately for each decile. If the coefficients on housing characteristics differ between borders, then the standard errors at each border are likely to be positively correlated. Thus standard errors are clustered at the border level.

The coefficient β_1 is the capitalization of a one standard deviation increase, 8 scaled score points on an exam with an average score of 420, in district performance for a border in decile 1-5 (i.e. a border between districts with relatively low median income). The coefficient reflects both the capitalization of the direct effect and the endogenous effect of school quality. β_2 is the capitalization per standard deviation experienced along borders between high income districts. The average sale price for a house on a border in deciles 1-5 was 113,178 dollars, and 193,742 for a house on a border in decile 6-10. Thus, if capitalization increases proportionally with the sale price of a home, we would expect the capitalization level β_2 to be approximately 1.7 times larger than β_1.

$$
\begin{aligned}
& saleprice_{i,b,d,y,m} = \alpha_b + \alpha_y + \alpha_m \\
& + X_i' \delta_D + \beta_D \cdot \bar{T}_d \cdot Decile + \epsilon_{i,b,d,y,m}
\end{aligned} \tag{2}
$$

Specification (2) isolates the capitalization of school performance at each income decile (*Decile* is an indicator for each decile), revealing which parts of the income distribution are responsible for the estimates $\beta 1$ and β_2 in Specification 1. Systematically higher estimates of β at higher wealth deciles in absolute dollars and as a proportion of house value imply that capitalization increases disproportionately with wealth. If the controls in the above specifications are inadequate to ensure that bias from neighborhood desirability are not driving the results, then we would expect the strongest bias between the most unlike districts–where a high income district borders a low income district. Examining the borders that comprise each income decile reveals that borders between two high wealth districts comprise the top two deciles and borders between two low wealth districts comprise the bottom two deciles, so these deciles should be relatively free of this form of bias.

The close proximity of the houses in the sample should allay most concerns that unobserved community quality unassociated with district performance is biasing the estimates. Nonetheless, the primary specification check adds a regression discontinuity polynomial in distance from the border. The distances associated with properties on the less desirable side of the border are made negative and those on the more desirable side are positive, creating a continuum from least to most desirable. Ideally, separate regression discontinuities would be run on each border, allowing the effect of distance and the discontinuity at the cutoff to vary across each. However, most borders do not have the density of data necessary to make this a powerful estimation technique (especially while also controlling for housing characteristics, construction quality, and tax rates). Instead, to avoid bias across the income distribution, a separate polynomial is computed for each decile. The principle behind this is the same as for computing separate coefficients on each house trait for each decile: coefficients on the polynomial computed using the entire sample might over or underestimate the trends in very high or very low income districts. By allowing the polynomial to vary by decile, this form of bias is avoided.

Since there are many borders within each decile, they are merged according to the expected gap in desirability between the districts on either side (no unobserved trend would be expected if the communities on either side of the border were equally desirable). The cross-validation criterion is used to determine the optimal bandwidth for Specifications 1-3. (Imbens and Lemieux, 2008; Lee and Lemieuxa, 2010) I extend this procedure to choose between alternative proxies for the expected gap in desirability at the border (e.g. the gap in median and average income). This method indicates that, given the rich set of controls, the optimal bandwidth in the traditional specifications (without the distance control) is 400 meters. With the polynomial control in distance, the suggested bandwidth is larger. This is consistent with the expectation that adding data results in more precise estimation of the coefficients on the distance trend, which in turn reduce bias at the discontinuity. The regression discontinuity polynomial included in Specification 3 is cubic in distance from the border and interacted with the gap in median income between the abutting districts.

$$\begin{aligned} &saleprice_{i,b,d,y,m} = \alpha_b + \alpha_y + \alpha_m \\ &+ X_i'\delta + poly(distance) \cdot gap + \beta_1 \cdot \bar{T}_d + \beta_2 \cdot \bar{W}_b \cdot \bar{T}_d + \epsilon_{i,b,d,y,m} \end{aligned} \tag{3}$$

Each of the above specifications is intended to identify how house price capitalization of school district quality varies across the income distribution. To estimate the annual payment households are willing to make for school quality, house sale price is replaced with an estimate of the annual mortgage payment and local property taxes. The estimated annual payment is the sum of the estimated mortgage and tax payments. Tax rates in Ohio are assessed on 35 percent of the market value of a property. All property is eligible for a 10 percent reduction in taxes and all owner occupied houses are eligible for an additional 2.5 percent reduction, so most single family homes receive a 12.5 percent reduction in local property taxes. Residential property taxes for school districts range from 20 to 71 mills, or 0.7 to 2.49 percent of the full market value of the property. The mortgage is assumed to be a 30 year fixed rate mortgage with an annual interest rate of 6 percent compounded monthly. This is the average mortgage rate during the period from 2003 to 2008. Each specification is rerun with annual payment as the dependent variable. The coefficient on district performance reflects both house price capitalization of district quality and local tax rates.

6. Results

The estimates reveal strong heterogeneity in the capitalization of school district performance across the income distribution. High income suburban districts are characterized by disproportionately higher rates of capitalization of school performance than are low income urban districts. In both dollar terms, and as a percent of house value, residents in high income communities pay more for each standard deviation increase in school performance on state tests. A one standard deviation difference in test score performance (the average gap between neighboring districts is 1.2 standard deviations) is associated with an increase in house sale price of approximately 2,000 dollars along borders between low income districts, and more than 10,000 dollars along borders between high income districts. This fivefold difference cannot be explained by the difference in house prices, which average 113,490 and 194,908 dollars respectively.

When capitalization and tax rates are merged to form an annual payment for each standard deviation, the average household in one of the five poorest deciles pays about 140 dollars per standard deviation increase in school test score average and the average household in the top five deciles pays more than 800 dollars. Again, this difference is disproportionately greater than the difference

Table 1: Capitalization at Low and High Income Borders[a]

	House Price		Annual Payment	
	Traditional	RD	Traditional	RD
Scaled score (Low Income)	1,808	-716	142	-108
	(896)	(1,322)	(64)	(118)
Scale score (High Income)	10,251	10,129	849	779
	(1,270)	(1,634)	(102)	(144)
Polynomial in Distance		x		x
House Characteristic * Decile	x	x	x	x
Construction Quality * Decile	x	x	x	x
Tax Rate * Decile	x	x		
Year fixed effects	x	x	x	x
Month fixed effects	x	x	x	x
Border fixed effects	x	x	x	x
R2	0.80	0.79	0.80	0.79
Observation	47,600	107,025	47,600	107,025

[a] The dependent variable in columns 1 and 2 is the sale price of the house in 2008 dollars. The dependent variable in columns 3 and 4 is the estimated annual mortgage payment and the estimated annual property taxes paid for local schools in 2008 dollars. Columns 1 and 3 use the traditional border discontinuity design with no trend in distance, while columns 2 and 4 control for a polynomial trend in distance interacted with the size of the gap in median income between the bordering districts. A bandwidth of 0.25 miles is used for the traditional border discontinuity design and a bandwidth of 0.40 miles is used when controlling for the polynomial. These bandwidths were selected using the cross-validation criterion.

in house prices. Though the magnitudes of these estimates vary some across specifications, the finding that high income communities are characterized by disproportionately greater capitalization is robust.

The specifications in Table 1, which include only two border classifications, reveal that capitalization along the borders between districts that serve low income students is quite small in magnitude, capitalization is much larger for districts that serve high income students, and the gap in capitalization between these districts exceeds the expected difference due to house prices. The first row shows the capitalization level of a one standard deviation difference in test scores along a borders of below average income (i.e. borders between districts that have low median incomes). The coefficient in the second row is the capitalization level along borders of above average income(i.e. borders between district that have high median incomes).

The estimated capitalization of 1,808 dollars per standard deviation for the low decile borders, as seen in column of Table 1, represents only 1.6 percent of the typical house sale price of 113,178 dollars. Conversely, the high decile borders have estimated capitalization of 10,251 dollars per standard deviation, or 5.3 percent of the average sale price of 193,742 dollars. Capitalization rates

Table 2: Capitalization by Decile[a]

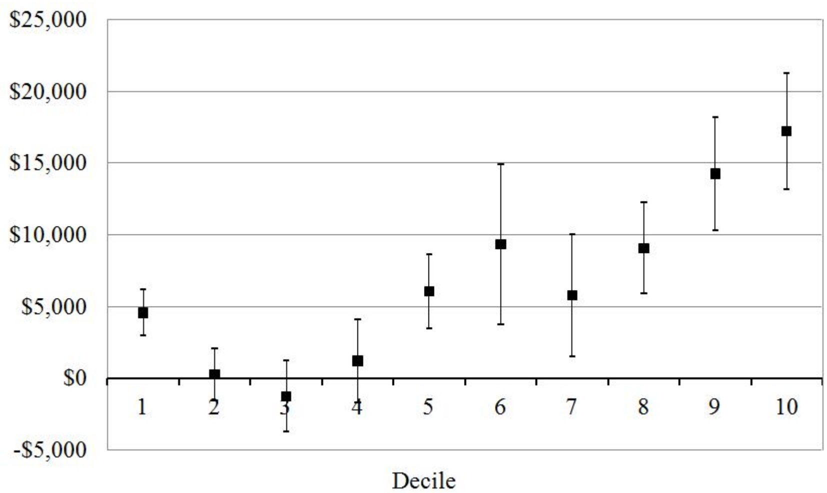
	House Price		Annual Payment	
	Traditional	RD	Traditional	RD
Scale score * Wealth Decile 1	4,567	1,319	385	83
	(814)	(1,882)	(70)	(142)
Scale score * Wealth Decile 2	260	-3,865	31	-361
	(935)	(1,302)	(74)	(419)
Scale score * Wealth Decile 3	-1,261	-3,373	-107	-392
	(1,262)	(1,302)	(123)	(128)
Scale score * Wealth Decile 4	1,201	-448	148	45
	(1,489)	(2,201)	(131)	(213)
Scale score * Wealth Decile 5	6,047	4,939	225	166
	(1,313)	(2,133)	(101)	(82)
Scale score * Wealth Decile 6	9,339	13,055	637	1,063
	(2,850)	(3,322)	(201)	(320)
Scale score * Wealth Decile 7	5,777	1,069	539	94
	(2,163)	(2,932)	(207)	(260)
Scale score * Wealth Decile 8	9,075	7,571	902	493
	(1,614)	(1,329)	(182)	(216)
Scale score * Wealth Decile 9	14,260	11,681	1,021	907
	(2,021)	(4,078)	(174)	(270)
Scale score * Wealth Decile 10	17,225	14,683	1,380	1,169
	(2,066)	(4,016)	(162)	(362)
Polynomial in Distance		x		x
House Characteristic * Decile	x	x	x	x
Construction Quality * Decile	x	x	x	x
Tax Rate * Decile	x	x		
Year fixed effects	x	x	x	x
Month fixed effects	x	x	x	x
Border fixed effects	x	x	x	x
R2	0.80	0.79	0.80	0.79
Observation	47,600	107,025	47,600	107,025

[a] The dependent variable in columns 1 and 2 is the sale price of the house in 2008 dollars. The dependent variable in columns 3 and 4 is the estimated annual mortgage payment and the estimated annual property taxes paid for local schools in 2008 dollars. Columns 1 and 3 use the traditional border discontinuity design with no trend in distance, while columns 2 and 4 control for a polynomial trend in distance interacted with the size of the gap in median income between the bordering districts. A bandwidth of 0.25 miles is used for the traditional border discontinuity design and a bandwidth of 0.40 miles is used when controlling for the polynomial. These bandwidths were selected using the cross-validation criterion (see Appendix B for details).

appear to be highly heterogeneous across districts.

Column 2 shows the results of an unparsimonious specification that, in addition to the controls in column 1, adds a polynomial in distance from the border to control for trends in community desirability. Interestingly, while this estimation technique results in statistically insignificant capitalization for low income deciles, it has little effect on capitalization estimates for high income deciles. It

Figure 6: Capitalization of Performance by Decile†



† This figure shows the average difference in house sale price attributable to a one standard deviation increase in district test scores. Results are shown by the income decile of the border, where decile 1 is comprised of borders between districts with the lowest median incomes and decile 10 is comprised of borders between districts with the highest median incomes.

is important to remember that the specification includes border fixed effects, and the contribution of house characteristics and distance from the border are estimated separately for each decile. Thus these results do not stem from over or under controlling for these factors in some deciles more so than in others.

Columns 3 and 4 of Table 1 estimate the annual payment for district performance by summing the annual mortgage payment with annual public school taxes. As with capitalization, there is a disproportionately larger annual payments in the high income districts: households on the borders between low income districts pay 142 dollars annually for a one standard deviation increase in performance, while households between two high income districts pay 849 dollars.

The source of the variation between high and low income districts becomes clearer when the effect is estimated separately for each income decile in Table 2. This specification allows capitalization to differ more freely across the wealth distribution. Average capitalization per standard deviation is 2,163 dollars among the poorest five deciles, and is statistically insignificant for three of the four poorest deciles. In contrast, capitalization averages 11,135 among the richest five deciles of community income and is statistically significant for each. Thus capitalization represents about 1.9 percent of house price for the

Table 3: Decile Characteristics[a]

	Avg Price	Median Inc	Urban Low Inc	Suburban High Inc
Decile 1	79,365	23,774	1.00	0.00
Decile 2	111,855	25,775	0.92	0.06
Decile 3	113,678	27,574	0.71	0.25
Decile 4	127,075	29,141	0.63	0.33
Decile 5	135,479	30,396	0.54	0.44
Decile 6	149,062	31,845	0.31	0.62
Decile 7	173,258	34,301	0.19	0.76
Decile 8	180,278	37,485	0.24	0.76
Decile 9	230,254	39,302	0.14	0.86
Decile 10	241,686	45,318	0.09	0.91

[a] House price is in terms of 2008 dollars. Urban Low Income is the proportion of houses that are in a district labeled by the state as either "Major Urban-very high poverty" and "Urban-high poverty". Suburban High Income is the proportion of houses that are in a district labeled by the state as either "Urban/Suburban-high income" or "Urban/Suburban-very high income".

bottom five deciles and 5.7 percent for the top five deciles. The capitalization rate for the two highest income deciles is 14,260 and 17,225 dollars per standard deviation, or 6.2 and 7.1 percent of house price respectively. Controlling for a polynomial in distance trends reduces the estimates more for the lowest deciles than the highest, indicating that the heterogeneity in capitalization of district quality is not the result of larger unobserved difference in community desirability between the wealthiest districts.

Examining annual payment reveals a similar phenomenon. The third and fourth columns of Table 1 reveal that annual payment is about 136 dollars per standard deviation among the bottom five deciles and 896 dollars among the top five deciles. This peaks at 1,021 and 1,380 dollars per standard deviation for the two highest income deciles. An interesting aspect of these results is the composition of the districts that form the borders in each decile. The Ohio Department of Education has placed each district into one of seven categories based on whether they are urban, rural, or a small town, and whether they have high or low average average income and poverty. The top two deciles are primarily comprised of borders between two high income suburban districts while the bottom two deciles are primarily comprised of borders between two low income urban districts. Borders between one rich and one poor district make up the majority of borders in the third and fourth poorest deciles. The third and fourth poorest deciles have essentially no evidence of payment for, or capitalization of, school district quality, while the highest deciles have the most in both absolute and relative terms.

These results indicate that estimates of *average* capitalization, as computed by existing studies, reveal little of the underlying phenomenon. Capitalization of quality is essentially undetectable in low income, largely urban districts, but exceeds previous estimates in high income suburban districts. Estimating capitalization as an average across heterogeneous borders is a misspecification and is likely to result in estimates that are sensitive to specification error. Conversely, the results found in this section are quite robust to a diverse set of specifications, as shown in the next section. Systematically greater capitalization and willingness to pay for school quality has implications for financial incentives for schools. I examine the financial implications under several state finance formulas in Section 8.

7. Alternative Specifications

This section details three specification checks that are run to ensure the robustness of the results in Section 6: polynomial controls for house traits are used as an alternative to linear controls by decile; the natural log of the sale price is used as the dependent variable; and allowing for separate trends in distance on each side of the border. The approximate magnitudes and observed phenomena across the income distribution are found to be robust to these checks, implying that the preferred controls are sufficient to produce reliable results. Finally, I add block group controls for race and educational attainment to assess the importance of endogenous sorting into neighborhoods on the estimates of capitalization.

Polynomial Controls for House Characteristics

The preferred specification allows the value of housing characteristics to vary by income decile as discussed in Section 5. If the contribution of housing characteristics is based on the average across all deciles it may underestimate the contribution in high income districts. Avoiding this form of bias is important when trying to differentiate capitalization rates across the income distribution. However, the literature generally includes a polynomial in house characteristics instead. Based on the reasoning above, specifications using only polynomial controls should result in estimates that have an upward bias for capitalization on high income borders and downward bias for low income borders. This bias is less of a concern to previous studies that measure average capitalization rather than heterogeneity in capitalization across the income distribution.

Column 1 of Table 4 reveals similar capitalization to that in the preferred specification: 1,585 dollars per standard deviation among the poorer deciles and 11,148 among the richer deciles. Table 5, which computes each decile separately, indicates capitalization of 1,544 dollars for the bottom five deciles (smaller than with the preferred specification) and 11,636 dollars for the top five deciles (larger than with the preferred specification). Thus polynomial controls for house characteristics produce similar results to those using the preferred specifications, but are biased (as predicted) in favor of even greater heterogeneity across the income distribution.

Natural Log Specification

A specification in which the natural log of the sale price is used as the dependent variable is useful for several reasons. First, it allows house characteristics (and all other controls) to affect the house price in a fundamentally different way

Table 4: Specification Checks[a]

	House Price		ln[House Price]
	1	2	3
Scaled score (Low Income)	1,585	-259	.0293
	(803)	(1,161)	(.0091)
Scale score (High Income)	11,148	9,023	.0575
	(1,417)	(1,295)	(.0084)
Polynomial of Characteristics	x		
Separate Distance Trends		x	
House Characteristic * Decile		x	x
Construction Quality * Decile		x	x
Tax Rate * Decile		x	x
Year fixed effects	x	x	x
Month fixed effects	x	x	x
Border fixed effects	x	x	x
R2	0.76	0.79	0.72
Observation	47,600	107,025	47,600

[a] The dependent variable in columns 1-4 is the sale price of the house in 2008 dollars. The dependent variable in column 5 is the natural log of the sale price of the house in 2008 dollars. Column 1 includes a linear trend in distance interacted with the gap in median income between the bordering districts. Column 2 includes a polynomial in each house characteristic as an alternative to allowing the coefficients to vary by wealth decile. Column 3 omits those borders for which the average house on the side of the border with higher performing schools is statistically significantly larger than those on the low performing side (regardless of whether or not this differences is the result of a continuous trend). Column 4 omits any district border that is also a town name or zip code border. A bandwidth of 0.25 miles is used for all specifications. Ths bandwidth was selected using the cross-validation criterion (see Appendix B for details).

than the specification in price levels. Each house characteristic adds a percent to the value of a house rather than a dollar value. This addresses the concern that house features are worth more in areas with higher house prices. A second advantage of this specification is that it is immediately clear if capitalization is disproportionately larger along borders between high income districts because the estimates represent a fraction of the house price. A larger coefficient for high income borders implies disproportionately greater capitalization. Column 3 of Table 4 indicates a capitalization rate of 2.9 percent for poorer borders and 5.8 percent for wealthier borders. Thus a higher percentage of house price is attributable to local district quality in high income communities. Table 5 shows average capitalization of 2.9 percent for the bottom five deciles and 6.1 percent of house value for the top five deciles, which is similar to the 1.8 and 5.7 percent estimated in Section 6.

Alternative Regression Discontinuity Polynomial

Table 5: Specification Checks by Decile[a]

	House Price		ln[House Price]
	1	2	3
Scale score * Wealth Decile 1	4,008	1,798	.0640
	(780)	(1,480)	(.0113)
Scale score * Wealth Decile 2	-259	-3,692	.0115
	(982)	(3,295)	(.0076)
Scale score * Wealth Decile 3	-1,252	-2,878	.0080
	(1,448)	(1,067)	(.0112)
Scale score * Wealth Decile 4	1,723	560	.0170
	(1,613)	(2,167)	(.0126)
Scale score * Wealth Decile 5	3,500	4,738	.0457
	(865)	(1,819)	(.0153)
Scale score * Wealth Decile 6	7,680	9,968	.0556
	(2,798)	(2,266)	(.0229)
Scale score * Wealth Decile 7	7,098	2,537	.0371
	(2,539)	(2,039)	(.0143)
Scale score * Wealth Decile 8	11,878	6,574	.0540
	(3,593)	(1,279)	(.0115)
Scale score * Wealth Decile 9	15,538	9,672	.0771
	(2,363)	(2,981)	(.0087)
Scale score * Wealth Decile 10	15,988	16,589	.0797
	(4,518)	(3,123)	(.0109)
Polynomial of Characteristics	x		
Separate Distance Trends		x	
House Characteristic * Decile		x	x
Construction Quality * Decile	x	x	x
Tax Rate * Decile	x	x	x
Year fixed effects	x	x	x
Month fixed effects	x	x	x
Border fixed effects	x	x	x
R2	0.76	0.79	0.72
Observation	47,600	107,025	47,600

[a] The dependent variable in columns 1-4 is the sale price of the house in 2008 dollars. The dependent variable in column 5 is the natural log of the sale price of the house in 2008 dollars. Column 1 includes a linear trend in distance interacted with the gap in median income between the bordering districts. Column 2 includes a polynomial in each house characteristic as an alternative to allowing the coefficients to vary by wealth decile. Column 3 omits those borders for which the average house on the side of the border with higher performing schools is statistically significantly larger than those on the low performing side (regardless of whether or not this differences is the result of a continuous trend). Column 4 omits any district border that is also a town name or zip code border. A bandwidth of 0.25 miles is used for all specifications. The bandwidth was selected using the cross-validation criterion (see Appendix B for details).

The preferred specification has a third order polynomial in distance. This is undesirable if the trends on each side of the border are fundamentally different.

Table 6: **Direct Versus Endogenous Effects**[a]

	Traditional	Race	Race and Education
Scale score * Wealth Decile 1	4,567	2,953	2,206
	(814)	(861)	(725)
Scale score * Wealth Decile 2	260	-1,920	-521
	(935)	(2,018)	(1,473)
Scale score * Wealth Decile 3	-1,261	-1,918	-1,251
	(1,262)	(1,432)	(1,046)
Scale score * Wealth Decile 4	1,201	-2	-688
	(1,489)	(1,512)	(1,417)
Scale score * Wealth Decile 5	6,047	2,785	2,557
	(1,313)	(1,182)	(1,545)
Scale score * Wealth Decile 6	9,339	7,423	7,017
	(2,850)	(2,511)	(2,403)
Scale score * Wealth Decile 7	5,777	5,204	4,104
	(2,163)	(2,218)	(1,999)
Scale score * Wealth Decile 8	9,075	7,366	6,530
	(1,614)	(1,193)	(1,128)
Scale score * Wealth Decile 9	14,260	13,764	10,538
	(2,021)	(1,996)	(2,271)
Scale score * Wealth Decile 10	17,225	16,332	15,611
	(2,066)	(2,296)	(2,057)
Race		x	x
Education			x
House Characteristic * Decile	x	x	x
Construction Quality * Decile	x	x	x
Tax Rate * Decile	x	x	x
Year fixed effects	x	x	x
Month fixed effects	x	x	x
Border fixed effects	x	x	x
R2	0.80	0.80	0.81
Observation	47,600	47,310	47,310

[a] The dependent variable is the sale price of the house in 2008 dollars. Column 1 uses the traditional border discontinuity design. Column 2 adds the proportion of residents in the block group who are white, black, and hispanic. Column 3 includes both the race controls and the proportion of adult residents in the block group who did not graduate from high school, a high school diploma, or college degree.

Thus, as an alternative, I include a separate distance trend on each side of the border for each decile. This process produces very similar results to those using a single polynomial for each decile: the borders between low income districts do not exhibit capitalization that is statistically significant from zero, while borders between high income districts exhibit significant capitalization.

Block Group Demographics
When a district improves in quality, it attracts households that are willing to pay more for high quality schools. These households may, on average, have characteristics that make them more attractive to other prospective residents.

This endogenous sorting will result in an increase in house value above the direct effect of school quality for two reasons: these households may make more desirable neighbors, and the children in these households may make more desirable school peers. The former effect should have been largely eliminated as a side effect of controlling for location desirability with border fixed effects and distance controls. The sorting of more desirable peers into schools that perform well is pertinent to understanding the extent to which house price capitalization of school performance stems from a direct effect (valuing a higher quality school) versus an endogenous effect (valuing higher quality peers). As discussed in 1, this distinction is unimportant for the purpose of understanding district incentives: district administrators are not concerned with whether an improvement in the quality of their schools increases local house prices due to the direct or endogenous effect. Nonetheless, estimating the relative importance of the two effects is an interesting exercise, so controls for block group race and education are added to the specification. If the endogenous effect is in fact the dominant force behind differences in house prices, then controlling for the composition of the block group might significantly reduce estimated capitalization of district scores.

Column 1 of Table 6 shows the estimates from the baseline specification in Table 2. Capitalization averages 2,163 per standard deviation for the lowest fives deciles and 11,135 dollars for the highest five deciles. The estimates in column two are from a specification that controls for the proportion of the block group that is black, white, and hispanic. Interestingly, adding controls for race reduces estimated capitalization for borders between low income districts but has little effect for borders between high income districts. Capitalization for the lowest deciles is reduced to 380 dollars, which is statistically insignificant from zero. In contrast, the estimates for higher deciles is 10,018 dollars, indicating that race only reduces the estimates by 10 percent. Including controls for the proportion of residents in the block group who did not complete high school, have a high school degree, or have a college degree has little effect on the estimates for the lowest five deciles. Estimated capitalization among the richest five deciles with both race and education controls is 21 percent lower than without either. This indicates that while the endogenous composition of neighborhoods does have an effect on the estimates of capitalization in this paper, there is not evidence that this is the dominant effect.

8. Finance

Students in low income communities performed less well in school than their wealthier counterparts under decentralized public education. The responsibility for raising revenue and implementing public education was delegated to local governments and large differences in local property wealth and income translated into large differences in revenue between districts that served high and low income households. In response to court rulings that such delegation was unconstitutional, states implemented centralized school finance systems designed to close the gap in revenue levels between high and low income districts. However, this paper suggests that high and low income district differ not only in terms of their capacity to raise revenue, but also in the incentives they face to provide a high quality public good. Districts that serve high income households have a greater capacity to raise revenue and face a more elastic tax base. Thus it is natural to consider how state finance systems affect not only revenue levels, but also financial incentives.

In this section I examine how the estimates in Section 6 might translate into financial incentives under the decentralized local finance systems employed by most states forty years ago. I then examine how these incentives would change under a system like that used in Ohio–a foundation aid system typical of many other states, whereby districts are guaranteed a "foundation" level of funding per student. I evaluate the effect of this finance formula on incentives for low income districts, high income districts, and very high income districts. Districts in very high income communities are evaluated separately because they are often "off-formula" in the sense that they are financially independent and do not receive any state aid. I also consider how these incentives would change if Ohio adopted the finance system employed by California–perhaps the most controversial system employed by any state.

Table 7 presents the capitalization rate per standard deviation increase in district test score based on the estimates in Table 2. An urban district that serves households of below average income and is neighbored by similar districts would be expected to have capitalization rates as shown in the first column. Likewise, a wealthy suburban district neighboring other wealthy suburban districts could have capitalization as high as that shown in the third column. The second row shows the corresponding annual housing payments.

First consider the case of pure local finance. The amount of capitalization per standard deviation in test score is converted to a per-household annual increase in revenue. With a property tax rate of 1.26 percent (the average in Ohio), the low capitalization rate translates into an additional 27 dollars per student per year for each standard deviation improvement in test scores, while the highest capitalization rate translates into 217 dollars. This is in the absence of a centralized finance system.

Ohio's foundation aid system guarantees a "foundation" level of revenue per pupil on the condition that a district levies a minimum property tax of 0.7 percent. State aid per pupil equals the foundation level less 0.7 times the per-pupil property value in that district. For example, if the state deems 6,000 dollars to be the foundation revenue level, but a district can only raise 3,200 dollars of revenue per pupil with a tax rate of 0.7 percent, then the state will provide 2,800 dollars in aid. Thus, at a tax rate of 0.7 percent, any increase in local revenue due to an increase in the local property tax base is offset exactly by a reduction in state aid. Only revenue generated from taxes above the 0.7 percent minimum can raise additional revenue above the foundation level. The average district in Ohio levies a tax of rate of 1.26 percent and therefore exceeds the required tax rate by 0.56 percent. Increases in the local property tax base will produce additional revenue through this additional tax. In Table 7 this is reflected as a decrease in the per-pupil revenue return to an improvement in

Table 7: **Finance Systems and Incentives**[a]

	Low Income (lowest 50 %)	**High Income** (highest 50%)	**Highest Income** (highest 10%)
Capitalization	2,162	11,135	17,225
Annual Payment	136	896	1,380
Annual Revenue Per-Pupil:			
Pure Local Finance	27	140	217
Ohio Finance Formula	12	62	217
California Finance Formula	0	0	< 217

[a] Each column shows the estimated house price capitalization, household annual payment, and additional revenue per standard deviation increase in the test performance of a school district.

school performance relative to pure local finance.

Property rich districts that raise more than the foundation level with the minimum property tax rate are typically considered "off-formula". They raise school revenue independent of state aid, so a change in the size of the tax base is reflected proportionally in district revenue. For simplicity, assume that districts with borders in the highest decile are off-formula, and thus their per-pupil revenue is unchanged under Ohio's public school finance system. Under these assumptions, low income districts have small revenue incentives initially and these are reduced significantly when the state guarantees per-pupil funding levels. High income districts that are not off-formula have relatively large incentives initially and see the most significant reduction (in dollar terms) when they receive state aid. Conversely, the wealthiest districts remain autonomous from the system and their incentives remain the largest of all districts.

California also employs what is termed a foundation aid system. However, the tax rate is legislatively set at 1 percent and can rise no higher, and per pupil revenue is dictated by a function that is dependent on revenue in the previous year. Thus an increase in the size of the tax base increases a district's contribution to the state but does not result in greater local revenue. Only districts that are off-formula have revenue that is responsive to the local property tax rate. This phenomenon is further complicated because houses are not reassessed for tax purposes on a regular basis. Ohio, like most states, regularly updates the assessed value of all houses using trends in sales prices of neighboring houses. Thus an increase in sale prices is quickly reflected in the tax base. In California, houses are only assessed at their market value when they are sold, while those that are not sold are assumed to increase in value by only 2 percent per year. If only a small fraction of houses have recently changed hands, then the capitalization of district quality will have a limited effect on the property tax base. Thus even off-formula districts see little revenue response to improved performance. The reduced role of tax base incentives is accompanied by greater importance for enrollment levels. State aid is typically allocated on a per-pupil basis, and thus attracting more students will increase revenue (though it will not increase revenue per-pupil).

As discussed in Section 2, the inherent strength of house prices as a financial incentive for districts is that the entire tax base (the assessed value of all property) depends on the most recent sale prices. An increase in school quality that has a large effect on current sale prices will in turn have a large effect on the tax base as a whole (because all houses are assessed at a higher value even if they have not been sold). However, this process is muted when changes in the tax base are offset by a reduction in state aid. A system of finance that bases state aid on another measure of district wealth, such as the median income, will not have this effect. An increase in school quality may attract wealthier households on the margin, thus driving up the market price for houses and the tax base, but these new residents will have little effect on the median or average income since they constitute a small fraction of the population. In short, the marginal mover has a large effect on the property tax base, but a small effect on the average characteristics of residents. Thus if redistribution is based on median income and local revenue is raised through property taxes, a district that improves its quality will reap additional revenue through the property tax while receiving only slightly less state aid due to attracting wealthier residents.

9. Conclusion

Students in high income school districts have higher achievement and make greater gains than their counterparts in low income districts. This achievement gap has generally been attributed to variation in school quality stemming from large and salient differences in school resources. Districts with greater revenue can hire better teachers, staff, and administrators, build better facilities, and buy better educational materials (or any other school resource). The logic of centralized school finance plans is that the quality of districts in low income communities is constrained primarily by a lack of financial resources. However, the achievement gap has narrowed little even as revenue differences have been eliminated. Thus speculation naturally turns to the capacity of students to learn: low income students may be constrained by their home environment or natural aptitude.

This paper explicitly tests for the existence of heterogeneous incentives as an alternative explanation for the difference in school quality between high and low income districts. If districts in high income communities serve more discriminating parents who have greater willingness to pay for school quality, these districts will face stronger financial (and possibly social) incentives. Two districts with equal revenue will not necessarily be of equal quality. This explanation eases the tension between constraint based explanations that rely on revenue or student capacity to learn.

I find that the capitalization of district test performance into local house prices is much greater in high income communities than in low income communities. As a percent of house price, capitalization in low income communities is approximately 2 percent per standard deviation increase in test score. In contrast, capitalization is approximately 6 percent per standard deviation in high income communities. The population of data from 42 counties and 250 school districts in Ohio has the richness to reliably estimate capitalization across the income distribution with an extensive set of controls and alternative specifications. The finding of heterogeneous capitalization is robust to the inclusion of a regression discontinuity polynomial in distance, alternative controls for house quality, and a natural log specification. Because Ohio has a reasonably standard system of district revenue redistribution, and the results are based on data that span seven metropolitan areas, I believe that the results in this paper have external validity.

These findings suggest that school districts in high income communities face inherently stronger financial incentives than districts in low income communities, and this could be responsible for some of the unexplained achievement gap. Though only financial incentives were tested explicitly, it is likely that the social incentives are stronger in high income communities as well. Households that are more willing to pay for high quality schools may hold those schools to a higher standard. Further, since all homeowners in high income communities are affected by the greater rate of capitalization, they may be more invested in seeing that local schools are of high quality.

The empirical findings in this paper are intended to suggest an alternative explanation for gaps in school quality and achievement. Thus the results are interpreted primarily in the context of their implications for district incentives. A related question, that is beyond the scope of this paper but is interesting in its own right, is why high income households are willing to pay disproportionately more for higher quality schools and better peers. The typical investment based explanation of educational attainment (where marginal cost equals expected marginal return) would suggest that parents in high income communities perceive greater returns to high quality schooling than do low income parents. Alternatively, schooling is at least in part a consumption good, and high income parents may have a taste for school quality or more disposable income to purchase it.

P. J Bayer and R. McMillan. Choice and competition in local education markets. *NBER working paper*, 2005.

P. J Bayer, F. V Ferreira, and R. McMillan. A unified framework for measuring preferences for schools and neighborhoods. *NBER Working Paper*, 2007.

J. R Betts. Is there a link between school inputs and earnings? fresh scrutiny of an old literature. *Does money matter?: the effect of school resources on student achievement and adult success*, page 1141, 1996.

S. E Black. Do better schools matter? parental valuation of elementary education*. *Quarterly Journal of Economics*, 114(2):577599, 1999. ISSN 0033-5533.

D. Boyd, H. Lankford, S. Loeb, and J. Wyckoff. The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management*, 24(1):113132, 2005. ISSN 1520-6688.

D. Card and A. B Krueger. School resources and student outcomes: an overview of the literature and new evidence from north and south carolina. *The Journal of Economic Perspectives*, 10(4):3150, 1996. ISSN 0895-3309.

D. Card and A. A Payne. School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics*, 83(1):4982, 2002. ISSN 0047-2727.

T. A Downes, D. N Figlio, and University of WisconsinMadison. Institute for Research on Poverty. *School Finance Reforms, Tax Limits, and Student Performance: Do Reforms Level Up or Dumb Down?* Institute for Research on Poverty, University of WisconsinMadison, 1997.

G. J Duncan and K. A Magnuson. Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children*, 15(1):3554, 2005. ISSN 1054-8289.

W. N Evans, S. E Murray, and R. M Schwab. Schoolhouses, courthouses, and statehouses after serrano. *Journal of Policy Analysis and Management*, 16 (1):1031, 1997. ISSN 1520-6688.

D. N Figlio and M. E Lucas. What's in a grade? school report cards and house prices. *NBER working paper*, 2000.

R. G Fryer. Financial incentives and student achievement: Evidence from randomized trials. *NBER Working Paper*, 2010.

E. A Hanushek. Assessing the effects of school resources on student performance: An update. *Educational evaluation and policy analysis*, 19(2):141, 1997. ISSN 0162-3737.

E. A Hanushek. School resources. *Handbook of the Economics of Education*, 2: 865908, 2006. ISSN 1574-0692.

C. M Hoxby. All school finance equalizations are not created equal*. *Quarterly Journal of Economics*, 116(4):11891231, 2001. ISSN 0033-5533.

G. W Imbens and T. Lemieux. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615635, 2008. ISSN 0304-4076.

D. S Lee and T. Lemieuxa. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281355, 2010. ISSN 0022-0515.

A. Vanneman, L. Hamilton, J. B Anderson, and T. Rahman. Achievement gaps: How black and white students in public schools perform in mathematics and reading on the national assessment of educational progress. statistical analysis report. NCES 2009-455. *National Center for Education Statistics*, page 80, 2009.

H. Wenglinsky. Finance equalization and within-school equity: The relationship between education spending and the social distribution of achievement. *Educational Evaluation and Policy Analysis*, 20(4):269, 1998. ISSN 0162-3737.