School Competition and Product Differentiation

Natalie Bau*

Abstract

In this paper, I establish that schools respond to increased competition by catering to wealthier students, at the expense of poorer students. I develop a Hotelling-style model of school competition in the presence of differential attentiveness or search costs. If (1) the match between a school and a student affects learning and (2) wealthier students are better at identifying their closest-match schools, then schools cater more to the needs of wealthy students. Therefore, the entry of an additional school reduces the quality of the match between existing schools and poorer students. Using panel data from rural Pakistan, I show that the match between schools and students is important for learning and school competition. Exploiting the exit and entry of private schools and an instrumental variable that affects the cost of opening a private school, I find that all the academic benefits of competition accrue to wealthier students. A structural model of school choice demonstrates that wealthier students are relatively better at sorting into schools with higher predicted test score gains for them. Given the rise of private schools in the developing world and the increasing promotion of school competition in the United States, these results point to potential adverse welfare effects of school competition.

1 Introduction

Papers on education and school competition have typically focused on school quality as single-dimensional characteristic (hereafter, vertical quality) and assumed that all students benefit from attending a "better" school. In a classic paper, Hoxby (2000) argues that increased school competition leads to better outcomes for students on average. Similarly, Neilson (2013) models school quality as a single characteristic, which he approximates with mean test scores. However, educational quality also has a match-specific component (hereafter, horizontal quality). For example, in a classroom tracking experiment, Duflo et al. (2011) show that different students benefit from

*Harvard University, Cambridge, MA, 02138, *Contact: Natalie_Bau@hks.harvard.edu.* I gratefully acknowledge the support of the NSF Graduate Research Fellowship and the Harvard Inequality and Social Policy Fellowship. I am deeply indebted to my advisors, Chris Avery, Roland Fryer, Asim Khwaja, and Nathan Nunn for their guidance and encouragement. Alberto Abadie, Nava Ashraf, David Baqaee, Raj Chetty, Sarah Cohodes, Dave Deming, Glenn Ellison, James Feigenbaum, Rema Hanna, Nathan Hendren, Christopher Jencks, Larry Katz, Greg Lewis, Ariel Pakes, Rohini Pande, Martin Rotemberg, Doug Staiger, and Bryce Millett Steinberg also provided very helpful comments. I am also grateful to participants in the Harvard Development and Industrial Organization Workshops and the Labor and Fiscal Policy Lunch for their helpful comments.

different instructional levels. Duflo et al. (2011) found that forming classes of students based on ability led to test score gains of between 0.1 and 0.2 standard deviations relative to the control group, and the effects were not concentrated among the high-ability students. These results suggest that the match between a school's level of instruction and a student's entering ability or knowledge of the material is an important determinant of student outcomes. High ability students will benefit from a high level of instruction at the expense of low ability students, and low ability students will benefit from a lower level of instruction at the expense of high ability students. Moreover, since schools in developing countries typically have 1 teacher (or less) per grade, it is difficult for a school to simultaneously teach at multiple instructional levels by sorting students into different classrooms.

More broadly, there are many important ways that schools can horizontally differentiate. Different students are likely to benefit from different teacher characteristics, course offerings, and extracurricular activities, and offering any given bundle of teacher characteristics, course offerings and extracurricular activities has an opportunity cost for a school. Even in the unlikely event that all of a school's inputs or offerings benefit *all* students, as long as school characteristics have heterogeneous effects, for a given cost of providing education, different bundles will maximize gains for different subsets of students. For example, Glewwe et al. (2009) found that providing textbooks to students in rural Kenya did not help the median student but did benefit more advanced students. Providing textbooks to students does not harm poor students directly, but it comes at a cost since the same money could be used to provide remedial tutoring or some other school characteristic that benefits disadvantaged students.

In a standard Hotelling model of competition with horizontal differentiation, such as the one developed by Eaton and Lipsey (1975), students are arrayed on a unit line and attend the school that places closest to them on that line. Then, as the number of schools increases, competition between schools seeking to maximize their shares of students will cause schools to product differentiate, spreading out on the line, and students will sort perfectly into their preferred schools, increasing net student welfare (Eaton and Lipsey, 1975). Milton Friedman cites this very effect as one of the benefits of school competition, writing, "The injection of competition would do much to promote a healthy variety of schools" (1962). However, Friedman's assertion implicitly relies on rational students with perfect information sorting into the best schools for them. In fact, students, particularly more disadvantaged students, may have limited information about their match with all the available schools and little opportunity to experiment with different schools. For example, as research by Mullainathan and Shafir (2013) shows, disadvantaged students and their families may have fewer cognitive resources to devote to enrollment decisions.

As I show in this paper, when students' ability to identify their match school is related to their location on the Hotelling line, competition can have adverse consequences since schools will be incentivized to cater to students who are better at identifying their best-match schools. I develop the first model of horizontal product competition where an individual's ability to identify the best product for her is correlated with her location on the unit line or, equivalently from the perspective of schools, how much an individual values school quality is correlated with her location on the line. A key implication of this model is that, when advantaged students are better able to identify their best-match schools or advantaged students (or their parents) value their match more on average, this imposes a negative externality on less advantaged students by shifting the school offerings toward the righthand side of the Hotelling line. This model has two key testable predictions. First, consistent with a more standard Hotelling model with greater than 2 firms, where students do not differ in their responsiveness to their match to schools, school competition increases product differentiation (Eaton and Lipsey, 1975). Second, inconsistent with the Hotelling model, school competition will increase average inequality between advantaged and disadvantaged students within the same school. Increased competition causes welfare losses for students who are very bad at sorting, and theoretically, these losses may be so great that competition decreases net utility.

I verify the predictions of this model using the Learning and Educational Attainment in Punjab Schools data (LEAPS), a panel data set from 2003-2007 including geocoded data on Pakistani students and schools. The private schooling market in Pakistan is a natural place to study competition and horizontal differentiation. First, there is a large, low cost private schooling market, with 35 percent of students enrolled in private schools as of 2005 (Andrabi et al., 2006). Second, horizontal differentiation is likely to be particularly important to learning outcomes in developing countries. Along these lines, in a review of the literature, Kremer et al. (2013) find that it is generally difficult to shift student test scores by increasing the quality of classroom inputs but that interventions that more closely tailor the instructional level to a student's own level are highly effective in developing countries. While I test the implications of this model for education in the developing world, this model may also apply to more general settings. It may be relevant for understanding competition in other environments where consumers have little opportunity to experiment with products and information is limited, such as health care and voting.

I use the LEAPS data to perform two related empirical tests. First, I estimate a structural model of school choice to test the fundamental modeling assumption that advantaged students are much better than disadvantaged students at identifying their best-match schools. Consistent with the key assumption of the theoretical model, I find that advantaged students respond much more to their predicted type-specific test score gains than do disadvantaged students when they make enrollment decisions. Second, I use two separate identification strategies to test the theoretical model's prediction that an increase in competition benefits advantaged students but harms disadvantaged students. Consistent with this prediction, I find that an (exogeneous) increase in the number of private schools in a village yields test score gains for advantaged students, but reduces test scores gains for less advantaged students.

Overall, this paper has two key contributions. First, it establishes the importance of taking into account horizontal quality when studying school competition. Under this view, economic stratification across schools may actually be beneficial if poorer students have different needs than richer students. Second, it demonstrates that the greater ability of advantaged students to sort into their best match schools imposes a negative externality on less advantaged students by incentivizing schools to skew their quality toward more advantaged students. I show that this mechanism has real consequences for the learning of disadvantaged students.

These results shed light on the grassroots private schooling market in the developing world. Understanding the private schooling market in the developing world is increasingly important. Recent estimates put private school enrollment at 28 percent in rural India (Pratham, 2012) and at 65 percent in urban India (Desai et al., 2008). The statistics for sub-saharan Africa are similar. For example, in the Gambia, 28 percent of enrolled children attended private sector schools in 2012 (World Bank Development Indicators, 2014). Understanding the mechanisms underlying school competition in the developing world opens the door to policies that will harness the benefits of competition for disadvantaged students as well. For example, supplying students with information on their type-specific predicted gains within a school may allow disadvantaged students to make better decisions and incentivize schools to cater to them more. Because making disadvantaged students more informed decision-makers will impact schools' choices of type-specific quality, even disadvantaged students who do not receive information or do not respond to the information will benefit from these policies.

These results may also be relevant to the public school market in the United States, where school funding relies on the number of students a school can attract. Even in public school systems, this financial pressure may incentivize schools to compete for students. Thus, the increasing popularity of school choice systems within districts – for example, in Boston and New York City – may cause schools to disproportionately cater to the needs of wealthier students. In this paper, I test whether competition disproportionately benefits advantaged students in an unregulated, highly privatized educational market. A natural next question for future research is whether the same mechanisms are at work in the public sector in the United States, where choice policies replicate some elements of private sector competition.

This paper also highlights a potential methodological issue in past work on school competition and accountability. My results suggest that small net effects of school competition (in either direction) or null results may mask important heterogeneity in the effects of competition. Therefore, this paper points to the importance of using theory to suggest what heterogeneous effects researchers should test for when evaluating the impacts of changing schools' competitive environments.

The paper is organized as follows. Section 2 presents a theoretical model of school competition. Section 3 discusses the context and the data. Section 4 presents a brief roadmap of the empirical strategy. Section 5 describes a structural model which tests the theoretical model's assumptions and presents results. Section 6 describes reduced form empirical tests of the theoretical model's implications and presents results. Section 7 provides robustness tests, and section 8 concludes.

2 Model

In this section, I develop a Hotelling-style model of school product differentiation with competition, imperfect information, and heterogeneous sorting. A student's type is her optimal bundle of investments by the school (represented by a location on the unit line), and schools maximize their shares of students. A school's location on the line can naturally be thought of as its instructional level, and a student's location can be thought of as her optimal instructional level. The key assumption is that students from similar socioeconomic backgrounds have similar optimal instructional levels. In this model, a student's ability to sort into her best-match school (or how much she values her match to a school) is correlated with her location on the unit line. Intuitively, it is less costly for more advantaged students or parents to attain information about school quality. For example, if more advantaged parents are more educated, they will be better able to assess their children's learning in subjects like English and math. Alternatively, disadvantaged parents may value learning less than more advantaged parents; from the perspective of schools, these models of school choice are equivalent. However, this model has different implications for welfare, since poor parents may legitimately value learning less. Despite this, from a policy perspective, it is still important to understand how school competition can promote inequality. Poor parents may not maximize their children's long-term welfare, and the state may care about promoting the education of poor students even if parents themselves do not.[1]

There are $N \in \{1, 2\}$ schools and $T = 3$ student types on a line. Schools first sequentially choose their locations on the line s_k, and then students sort into the schools. The utility of a student of type i in a school k is

$$u_{ik} = -(x_i - s_k)^2$$

However, her perceived utility v_{ik} includes an error drawn from the type 1 extreme value distribution. Her perceived utility in school k is

$$v_{ik} = -(x_i - s_k)^2 + \frac{\epsilon_{ik}}{\delta_i}$$

where ϵ_{ik} is the error and δ_i is a type-specific scalar capturing the size of the error's effect. This formulation is similar to the Hotelling model developed by De Palma et al. (1987) and De Palma et al. (1985), where consumers idiosyncratically value another product dimension besides match and price. However, this model differs substantially from the models in those papers because there is a correlation between an individual's type and the effect of ϵ_{ik} on her perceived utility.

[1] Additionally, the fact that these parents send their children to private schools at all, instead of enrolling them in free government schools or not enrolling them in school suggests that they *do* value learning.

Students have an outside option with utility $-\frac{1}{4} < u_o < 0$. I assume that $u_o < 0$ so that the outside option is not more attractive than attending even a perfect-match school. I also assume it is greater than $-\frac{1}{4}$ so that the outside option is binding. In the Pakistan context, the outside option can be thought of as attending a government school or not enrolling at all. A student chooses a school or the outside option to maximize her perceived utility. We can preserve the ranking of a type's utility in different schools and the outside option by multiplying through by δ_i for $\delta_i < \infty$. This in turn yields the familiar logit shares, so a school k's share of type i is:

$$p_{ik} = \frac{e^{-\delta_i(x_i - s_k)^2}}{e^{\delta_i u_o} + \sum_j e^{-a\delta_i(x_i - s_j)^2}}.$$

In the mathematical appendix, I micro-found this expression for the shares with rational inattention (Sims, 2003), where the δ_i's represent different constraints on attention. Rational inattention also captures the idea that more advantaged students have more attention to devote to the school choice problem. This idea is consistent with the work of Mullainathan and Shafir (2013), which shows that attention constraints are less binding for the more advantaged.

Schools sequentially choose their locations on the line to maximize their total students so, in the subgame perfect equilibrium, a school k chooses s_k to maximize

$$p_k = \sum_i p_{ik}(s_k, s_{-k}),$$

and then individuals sort into schools. I characterize the equilibrium for an illustrative special case. I assume there are 3 types located at $(0, \frac{1}{2}, 1)$, and let $\delta_1 = 0$ and δ_3 be ∞, so that $\delta_1 < \delta_2 < \delta_3$. Letting $\delta_1 = 0$ captures the idea of types with no information sorting randomly. Letting δ_3 be ∞ captures the idea of conventional Hotelling-style types with perfect information. In the following two propositions, I characterize the equilibrium for $N = 1$ and $N = 2$ for different values of δ_2. Understanding the equilibrium in the $N = 1$ and $N = 2$ cases will help illustrate the mechanisms that lead increased competition to increase inequality.

Proposition 2.1. *For N=1, there is a unique equilibrium at* $s^* = 1 - (-u_o)^{1/2}$.

Proof. When there is one school, it maximizes its share when it minimizes the share lost to the outside option. It has no effect on the share of the lowest types lost to the outside option, so we can ignore these. The school will discontinuously receive all the highest types as long as $-(1-l)^2 \geq u_o$. It can never receive all the middle types, so it will always choose to place to receive all the high types at $l \geq 1 - (-u_o)^{1/2}$. The school minimizes the loss of the middle types subject to this constraint at $l = 1 - (-u_o)^{1/2} \geq \frac{1}{2}$ by the assumption that $u_o > -\frac{1}{4}$. □

Proposition 2.2. *For N=2, for a given u_o, if δ_2 is sufficiently large, the unique pure strategy equilibrium is $(\frac{1}{2}, 1)$. Otherwise, it is $(1, 1)$.*

Proof. In a pure strategy equilibrium, the first school will always place at 1 since otherwise the other school could deviate by ϵ up to take all the high types. Given that the first school places at 1, the only possible maxima in the other school's payoff function are at $\frac{1}{2}$ and 1, since it will lose all of the high types if it places below 1, and its share of the middle types increases as it approaches $\frac{1}{2}$. Then the equilibrium is $(1,1)$ if the second school is better off pooling with the first school, which is true if

$$\frac{1}{1+e^{\delta_2 u_o}+e^{\frac{-\delta_2}{4}}} > \frac{1}{2} + \frac{e^{\frac{-\delta_2}{4}}}{e^{\delta_2 u_o}+2e^{\frac{-\delta_2}{4}}}.$$

When $\delta_2 \to \infty$, there will always be an equilibrium where the second school places at $\frac{1}{2}$ and as $\delta_2 \to 0$, she will only place at 1 in equilibrium. Then, to show there is a cutoff δ_2 above which the school always places at $\frac{1}{2}$ and below which it always places at 1, it is sufficient to show that the left hand side (LHS) and the right hand side (RHS) of the inequality only cross once. To show single crossing, it is sufficient to show that $\frac{\partial LHS}{\partial \delta_2} > 0 \ \forall \delta_2$ and $\frac{\partial RHS}{\partial \delta_2} < 0 \ \forall \delta_2$. Taking the derivative of the LHS and simplifying produces:

$$\frac{\partial LHS}{\partial \delta_2} = \frac{\frac{\delta_2}{4}e^{\frac{-\delta_2}{4}} - \delta_2 u_o e^{\delta_2 u_o}}{(e^{\delta_2 u_o}+1+e^{\frac{-\delta_2}{4}})^2} > 0$$

since $u_o < 0$.

Taking the derivative of the RHS, we can focus on simplifying its numerator since its denominator is always positive by the quotient rule. After simplifying, the numerator of $\frac{\partial RHS}{\partial \delta_2}$ is

$$-\delta_2 e^{\delta_2 u_o - \frac{\delta_2}{4}}(\frac{1}{4}+u_o)$$

and this expression is less than or equal to 0 if $u_o \geq -\frac{1}{4}$. □

Comparing the equilibrium for $N=1$ and $N=2$ shows that the implications of this model are very different from a more standard Hotelling model, such as the one developed by Eaton and Lipsey (1975). As in this more standard Hotelling model, schools differentiate when δ_2 is sufficiently large and the equilibrium position changes from $\frac{1}{2} \leq l \leq 1$ to $(\frac{1}{2},1)$. However, the implications of the two models for the average match quality between schools and low types differ.

In a standard Hotelling model, only the match between a type and the closest school is important for that type's welfare. However, this is no longer the case with heterogeneous sorting. Since students make mistakes, the distance between their position and that of the remaining schools also impacts their utility. When this distance is greater, mistakes are more costly. As the next proposition shows, since the $N=2$ equilibrium $(\frac{1}{2},1)$ may increase the average distance between the low types and a school, low types may be made worse off, even though the minimum distance to a school is reduced.

Proposition 2.3. *For a sufficiently small u_o, $\delta_2 < \infty$, the average match between a low type and* *a school will decrease and the average match between a high type and a school will increase when the number of schools increases from $N = 1$ to $N = 2$.*

Proof. This follows from propositions 1 and 2. From proposition 1, when $N = 1$, $s^* < \frac{3}{4}$ if

$$1 - (-u_o)^{\frac{1}{2}} \leq \frac{3}{4}.$$

Rearranging, this is true if

$$u_o < -\frac{1}{16}.$$

From proposition 2, the equilibrium at $N = 2$ is $(1, 1)$ or $(\frac{1}{2}, 1)$. If it is $(1, 1)$, the average match to the low type will have decreased with the addition of another school no matter what the outside option is. Even if the new equilibrium is $(\frac{1}{2}, 1)$, if $u_o < -\frac{1}{16}$, the average match between the lowest type and schools decreases. Symmetrically, this means that the average match between the highest type and a school increases. □

Corollary. *The welfare of the lowest type will be lower when the number of schools increases from 1 to 2 if $u_o \leq -\frac{1}{16}$. Since the lowest type chooses randomly regardless of the location of the schools, they will be worse off if the schools are farther away on average.*

In fact, the negative effect of the low types' mistakes on social welfare can be so large that, as the next proposition shows, even in the asymmetric equilibrium at $(\frac{1}{2}, 1)$, consumer welfare is reduced.

Proposition 2.4. *For a given u_o, $\delta_2 < \infty$, increasing the number of schools from $N = 1$ to $N = 2$ can lower consumer welfare (the sum of the types' utilities), even in the asymmetric equilibria.*

Proof. Consider the case where $u_o = -\frac{1}{16}$ so that the equilibrium at $N = 1$ is at $\frac{3}{4}$. Then the gains to the high types will be $\frac{1}{16}$ and the maximal gains to middle types will be less than $\frac{1}{16}$ (the gains if none of the middle types chose the outside option when the equilibrium is $(\frac{1}{2}, 1)$ in the $N = 2$ asymmetric case. It is straightforward to calculate that the utility of the low types changes from $-\frac{5}{16}$ to $-\frac{7}{16}$ when N changes from 1 to 2, so their net losses are $\frac{1}{8}$, which is equal to the bound on the sum of the gains to the high and the middle type. □

Overall, the intuition of this model is quite straightforward. When there is only one school in the market, it caters more aggressively to the middle types (whose ideal instructional level is closer to that of the low types), because it is more likely to lose them to mistake-making. As long as the high types prefer the single school to the outside option, the school gains nothing by becoming more attractive to them. On the other hand, it gains middle types continuously as it moves closer to the center of the Hotelling line. However, when a second school enters, it changes the first school's incentives. Now, it must compete aggressively to retain the high types, and since they do not make mistakes, competing for them has a higher payoff. As a result, high types benefit disproportionately from increasing school competition.

The implications of the heterogeneous Hotelling model are quite different from those of a more conventional Hotelling model. In the model of Eaton and Lipsey (1975), for cases where pure strategy equilibria exist, increasing the number of schools increases both welfare and product differentiation, and the effect is symmetric for types on the left and the right of the line. One implication of the heterogeneous Hotelling model mirrors the standard Hotelling model: increasing the number of schools can lead to product differentiation. However, in the heterogeneous Hotelling model, competition increases average inequality within a school between advantaged and disadvantaged students and can even decrease the sum of the students' utilities. Table 1 breaks down these different testable predictions of the Hotelling and heterogeneous Hotelling models.

One concern is that, in this heterogeneous Hotelling model, adding another private school to the market does not always benefit high types relative to low types. For example, for some δ_2 and u_o, the equilibrium at $N = 2$ for this model is $(1, 1)$, but the equilibrium for $N = 3$ is $(\frac{1}{2}, 1, 1)$. In this case, additional competition benefits low types. While this shows that more general forms of the heterogeneous Hotelling model have ambiguous effects on different types of students, it is reasonable to think that the relevant margin of competition is the change from $N = 1$ to $N = 2$ in my data. First, the median village in the data has two private schools. Second, even in villages where there are more than 2 private schools, since students are very responsive to distance, typically only 1 to 2 of these private schools will be relevant for a given student. 88 percent of students have 2 or less private schools within 1 kilometer of their households. More broadly, the model I present here is quite stylized. Because there are only two types who respond at all to their match to schools, for $N > 1$, there are only two locations where schools will place on the Hotelling line, and the average match between a low type and schools can never be worse than in the case where 2 schools place at $(1, 1)$. However, this is not true in a more general model with more types.

In the following sections, I test the heterogeneous Hotelling model's predictions. However, before doing so, I verify that more advantaged students are more responsive to schools' type-specific quality than less advantaged students. When I classify students by their advantage, using a factor analysis of indicator variables for asset ownership (a proxy for their x_i), I show that less advantaged students are less responsive to the predicted gains of attending a school than more advantaged students are, consistent with the assumption that δ_i is increasing in x_i. Then, I test how an additional private school impacts the test scores of disadvantaged (low x_i students) and advantaged (high x_i students) in the same school. Since the model predicts that the average match between advantaged students and their schools increases when N increases and the average match between disadvantaged students and their schools decreases, I expect a greater N to lead to more inequality within schools in the data. Finally, I test whether a greater N leads to greater product differentiation, as both the heterogeneous Hotelling and the Hotelling model suggest.

The implications of the heterogeneous Hotelling model also differ from those of a model where schools can influence a single quality characteristic (vertical characteristic) that impacts all students equally and maximize their profits subject to this characteristic and price. In an alternative model with a single quality characteristic, schools may vertically differentiate in response to competition such that some schools compete as "high price/high quality" schools, while others compete as "low price/low quality" schools. First, while vertical differentiation could still result in greater variation in school characteristics, this variation would occur because some schools have "low cost" characteristics while others have "high cost" characteristics. Therefore, in the vertical differentiation case, variation in the cost or amount of characteristics should account for variation in characteristics across schools. Second, if increased competition results in greater vertical differentiation, it may increase inequality *across* schools, but it should not increase inequality *within* schools. Disadvantaged and advantaged students alike should benefit from being in a "good school" and suffer from being in a "bad school."

3 Context and Data

3.1 Context

Pakistan is an ideal setting to study the private schooling market in the developing world. Like many developing countries, Pakistan has experienced a rapid increase in low-cost, secular private schooling over the past two decades. According to Andrabi et al. (2008), roughly 35 percent of rural students are enrolled in private schools. Pakistani private schools are virtually unregulated and are, for the most part, unsupported by the government. As a result, they offer us a glimpse into what mechanisms may be important in a purely private market for education. In my rural sample, the private school system is quite competitive. In 2007, the median village in my sample had 2 private schools (the average had 2.8).

Private schools do not merely cater to the wealthy. Andrabi et al. (2008) show that grassroots, rural private schools are even affordable for day laborers. The median private school charges only $20 per year for students in grades 1-3. Figure 1 is a kernel density plot of the annual fees charged by private schools in Pakistani rupees. Indeed, private schools typically spend less per student than public schools, largely because public school teachers earn about five times as much as private school teachers (Andrabi et al., 2010). Public and private schools do not compete to hire teachers from the same labor market since public schools are typically constrained to hire teachers who have completed post-secondary degrees.

Unlike private schools, government schools face relatively little competitive pressure. Historically, most teachers were hired on permanent contracts and are difficult to fire. School budgets are not determined by the number of students enrolled, and schools face little threat of closure if enrollments drop, in contrast to the public sector in the United States.

An additional advantage of studying private schooling in Pakistan is that, at the primary level, villages act as closed educational markets. Villages are typically far apart or separated by natural barriers, and students are very sensitive to distance when they make enrollment decisions (Andrabi et al. (2010), Carniero et al. (2010)). Therefore, I can consider how competition effects equilibrium school and student outcomes at the market level.

Instructional level is likely to be a particularly important match-specific determinant of student outcomes in this context. Both public schools and private schools have less than 1 teacher per grade on average, suggesting that within-grade tracking is extremely rare and cross-grade mixing within the same class is quite common. Figure 2 shows the distribution of the ratio of the number of grades offered by a school to the number of teachers working in the school in public and private schools. In both cases, the mass of the data is below 1. In this context, students requiring many different instructional levels are likely to be present in the same classroom.

Some descriptive statistics from my data suggest that more advantaged parents may have more information about school quality and be more responsive to school quality than disadvantaged parents are.[2] Advantaged parents are more likely to know the name of a child's teacher than disadvantaged parents (55 versus 51 percent), and advantaged parents are also more likely to have an opinion on a child's teacher's quality (91 percent are willing to rate the teacher's quality versus 89 percent of disadvantaged parents). Moreover, advantaged parents beliefs about teacher quality are more strongly associated with students' predicted test score gains. Parents rated teachers overall on a scale of 1-5 (very poor, poor, average, above average, highly above average). In a regression of a students' predicted test score gains in a school on indicator variables for each teacher rating and their interactions with whether a student's parents are advantaged, each interaction variable enters positively, indicating that at every point in the teacher performance distribution, advantaged parents beliefs about teacher quality are more predictive of school quality. An F-test of these interaction terms indicates that they are jointly marginally significant ($p < 0.10$).

The schools in the sample typically do not exclude students. The majority of schools in 2007 had some form of admissions procedure (97 percent of private schools and 95 percent of government schools), which typically consisted of an oral exam (81 percent of private schools, 54 percent of government schools) or the perusal of previous school reports (14 percent of private schools and 33 percent of government schools). However, even if a student was deemed "weak" based on this assessment, only 11 percent of private schools and 3 percent of government schools said they would refuse admittance. Therefore, a student typically attends their household's first choice school, conditional on other school characteristics such as distance and price. Furthermore, a student's primary school is unlikely to impact whether she is admitted to a given secondary school apart from its impact on her learning.

[2] Advantage is calculated using a factor analysis of household assets. High advantage households are those whose first factor estimate is above the median for parents of children attending private school.

3.2 Data

For this paper, I mainly draw upon the Learning and Educational Achievement in Pakistan Schools study (LEAPS). The LEAPS data consists of four rounds of data collected between 2004 and 2007 in a stratified random sample of 112 rural villages in Attock, Rahim Yar Khan, and Faisalabad districts of Punjab. To be included in the sample, villages were required to have at least one private school in 2003. Therefore, the sampled villages are somewhat more populous and wealthier on average than the average village in the state.

The data are particularly rich, since surveys were administered to schools, children attending the schools, and to a sub-sample of households. Surveyors collected geocoded data from the universe of schools within a 15 minute walk of the village. The initial sample included 823 schools in the first round (2004), and additional schools were added to the sample as they entered the market. These data included questionnaires administered to head teachers and school owners on school facilities, how time at school was spent, and what characteristics they value when they hire teachers. In addition, in the first round of data collection, all third graders in a school were tested using low-stakes tests administered by the surveyors in math, Urdu, and English. These students were followed and tested in subsequent years. In all, the panel includes 70,493 student-year test score observations (31,338 unique students). Test scores on the exams were calculated using item response theory (see Das and Zajonc (2010) for details), so that the mean test score in the population is 0 and the standard deviation is 1. A random sub-sample of students were also administered a survey including details on household assets and anthropometrics.

A smaller sample of the tested students can be matched to a household survey administered to a panel of 1740 households in each year. Drawing from a baseline census of the villages, 16 households were sampled in each village. 12 households were randomly chosen among those who had a child attending grade three in the first survey round and 4 were chosen among households where a child eligible to be enrolled in grade 3 was not enrolled. Importantly, in round 1, the location of households was geocoded. As a result, I can impute the geographic distance between a household and the schools in the village for children who appear in the household survey.

Table 2 presents summary statistics of the four rounds of LEAPs data at the school level. Because of panel structure of the data, each observation is at the school-year level. As table 2 shows, there is substantial variation across schools in terms of facilities, time use, and desired characteristics of teachers. Table 3 summarizes the characteristics of the tested children; each observation is at the child-year level. While students in private schools tend to have higher test scores and more assets, there is again substantial variation in student performance and wealth.

I supplement the LEAPS data with data from a complete census of private schools carried out by the Pakistani Federal Bureau of Statistics in 2000, administrative data on the location and construction of public schools in Punjab province, data from the National Educational Census (2005), and data from the 1998 and 1981 population censuses of Punjab, which include data on village population and characteristics.[3]

4 Empirical Roadmap

In the next two sections of this paper, I outline the two key parts of the empirical strategy and present results before turning to robustness checks. Section 5 focuses on testing the key assumption of the theoretical model - that disadvantaged students are less responsive to their match to a school than advantaged students are when they make enrollment decisions. In contrast, section 6 tests the implications of the theoretical model about the effects of competition on students' test scores using quasi-random variation in the number of private schools in a village. Therefore, the tests in section 6 do not rely on the results or methods in section 5 or vice versa.

In section 5, to test whether disadvantaged students are indeed less responsive to their learning in a school than more advantaged students, I estimate a structural model of school choice where students' school choices depend in part on their match to the school. To do this, I first must determine which students are rich and poor (akin to the higher and lower types in the model), and then, I need to develop a proxy for the match between a student and a school. To determine students' latent socioeconomic advantage, I estimate the first factor of a factor analysis of household assets and child characteristics. I then separate children into low and high types based on their score on this measure. Next, to estimate the match between a child and a school, I predict the test score gains to a child of a low or high type from attending a given school, following the value-added literature in education (for example, see Chetty et al. (2014), Rivkin et al. (2005), Kane and Staiger (2008), and Rockoff (2004)). Finally, I estimate a structural model of school choice, allowing a student's utility in a school to depend on the interaction between her type and school fees, distance, and type-specific value-added. Comparing the coefficients on type-specific value-added for low and high types allows me to directly test the theoretical model's assumption that low types are less responsive to their match to a school than high types.

In section 6, I test predictions from the theoretical model about school competition. Following the theoretical model, I expect an additional private school to have a negative impact on the average match between private schools and low types and a positive effect on the average match between private schools and high types. Using two identification strategies, I exploit time-varying and cross-sectional variation in the number of private schools in a village to test these predictions. In my first identification strategy, I take advantage of variation in the number of schools in a village over time due to exit and entry of private schools over the course of the panel. Exit and entry allows me to test for heterogeneous effects of the number of private schools on more and less advantaged students in the same school by controlling for school-by-type fixed effects. Conceptually, the exit or entrance of a private school leads schools to reposition themselves in the new equilibrium. In

[3] I am very grateful to Andrabi et al. (2013a), who assembled this data and kindly shared it with me.

response, students may also change their school selections. Thus, overall, the entrance or exit of a private school impacts student test scores by changing both which schools students attend and how much schools cater to different types. However, the inclusion of the school-by-type fixed effects means that my identifying variation comes from changes in the match between a type and a school over time (presuming there is no unobserved variation in student composition with types that is not captured by the type-by-school fixed effects).[4]). Therefore, this strategy eliminates the second channel through which school competition impacts student learning – the re-sorting of students into schools – and directly tests the theoretical model's predictions about the *average match* between students and schools.

In the second identification strategy, I take advantage of the fact that the presence of private schools is constrained by the availability of secondary-school educated young women. These women are not very mobile and have few outside options, so they form a cheap labor force for private schools. Moreover, private schools do not compete with government schools for these teachers, since civil service teaching jobs typically require an advanced degree. Therefore, the presence of private schools in a village today is predicted by the past presence of a girls' government secondary school in a village (Andrabi et al., 2013a). Historically, the allocation of government girls' secondary schools to villages was based on village population ranking within a patwar circle, a sub-district administrative region typically consisting of 3 to 5 villages. Therefore, whether a village had the highest historical population in its patwar circle predicts its number of private schools today. After conditioning on flexible controls for the village population in the 1981 and 1998 censuses, the variation in likelihood of having a girls' government secondary school comes from a discontinuous change in whether a village had a greater population than its neighbors. Thus, I instrument for number of private schools using an indicator variable equal to one if a village had the highest population in its patwar circle in 1981 conditional on a control for village area and semi-parametric controls for population and village level socioeconomic status. Using this instrumental variables strategy, I regress a school's type-specific value-added on the number of private schools and its interaction with type. This provides a robustness test of the estimates that exploit changes in the number of private schools over time. Even if both strategies are biased, they are unlikely to be biased by the same phenomena since they exploit different sources of variation.

Finally, I test whether, as the theoretical model predicts, increased school competition results in more product differentiation. I develop indices of school product differentiation that proxy for the difference between a school's characteristics and the average characteristics of a school in a given village-year. Using these indices as an outcome variable, I use the instrumental variables strategy to test whether product differentiation increases with the number of schools.

[4]In practice, I control extensively for selection in my main results.

5 Empirical Analysis of School Choice

In this section, I describe my estimation strategy and results for two models of school choice, a multinomial logit model and a discrete choice model with school quality that is unobserved by the econometrician. The multinomial logit model showcases interesting correlations in the data, but its coefficients may be biased if I do not account for all the drivers of school choice in the model. In contrast, the discrete choice model explicitly accounts for the possibility that schools have an unobserved characteristic that influences enrollment decisions. Including this characteristic in the model requires an additional exogeneous instrument for price to identify the coefficients in the utility function. Therefore, I present results from both models to show that (1) there appears to be a correlation between a student's socioeconomic status and her responsiveness to her match to a school, and (2) that this correlation is robust to explicitly accounting for correlations between school price and unobserved school quality.

5.1 Framework of the Multinomial Logit Model

The theoretical model is predicated on the idea that more advantaged students are better at identifying their best-match schools than less advantaged students are. I first test this assumption by estimating a multinomial logit discrete choice model, which includes several of the major determinants of school choice. To estimate this model, I must first identify more and less advantaged students and develop a proxy for the match between a type of student and a school.

To identify more advantaged students, I combine my student characteristic and test score data into a single index of advantage. To do so, I regress indicator variables for all the available asset variables on village-year and child age fixed effects, and use the residual variation that is unrelated to child age or idiosyncratic variation in types of assets held at the year or village level to construct a within village-year advantage measure. I similarly residualize a top-censored measure of the child's body mass index, which is set to be 0 if a child's body mass index is greater than the U.S. average. Finally, I residualize a child's lagged test scores in Urdu, English, and math by village-year fixed effects and lagged grade fixed effects.

Residualizing these indicator variables is important for two reasons. First, it reduces the correlation between these measures and other child characteristics that could drive children's learning outcomes. Child age is likely correlated with parent age, which in turn, will be correlated with greater asset ownership. Similarly, test scores will be related to years of school, which will also be related to learning outcomes. Second, in the latter part of this paper, I will study how within village-year competition impacts schools decisions about horizontal differentiation. From a school's point of view, the relevant distribution of types for its competitive decision-making is the within village-year distribution. Since each of the asset, bmi, and test score variables is correlated with underlying advantage, I conduct a factor analysis and predict the latent advantage index. "High types" are defined as those above the fiftieth percentile for this index among private school students, and "low types" are below the fiftieth percentile. This approach echoes the methodology of Filmer and Pritchett (2001), who show that the first factor in a principal components analysis of household assets proxies well for wealth, predicting child enrollment in India as well as or better than expenditure data does.

To proxy for the match between a low type or a high type and a particular school, I draw on the value-added literature in education economics (for example, see Chetty et al. (2014), Rivkin et al. (2005), Kane and Staiger (2008), and Rockoff (2004)) to estimate the predicted test score gains of a low or a high type attending a given school (from hereafter, referred to as a school's type-specific value-added). To calculate a school's type-specific value-added, I estimate the following regression for math, English, and Urdu:

$$y_{igzt} = \eta_{zs} + \sum_{g'} \sum_{z'} \left(\lambda_{g'z'} \mathbb{1}_{z'} \times \mathbb{1}_{g'} \times y_{igz,t-1} + \phi_{g'z'} \mathbb{1}_{z'} \times \mathbb{1}_{g'} \times y^2_{igz,t-1} \right) + \omega_{gz} + \alpha_{zt} + \epsilon_{igzt}$$

where y_{igzt} is the outcome variable consisting of normalized test scores in math, English, or Urdu, i indexes an individual, t indexes a year, z indexes a type, and g indexes a grade. $\sum_{z'}$ sums over types and $\sum_{g'}$ sums over grades. α_{zt} is a year-by-type fixed effect, ω_{gz} is a grade-by-type fixed effect, and η_{zs} is a type-by-school fixed effect. Then, the type-specific value-added for a type z in a school s is given by η_{zs}.

The goal of this method is to estimate the causal effect of attending a school on test scores. The initial regressions account for variation in test scores that is explained by year of test-taking, grade of test-taking, and a student's past performance. The remaining unaccounted for variation in test scores is then attributed to the school students attended. Then, the fixed effect η_{zs} is the average of the unaccounted for variation in test score gains for different types of students in different schools. Therefore, for these measures to be unbiased, the underlying assumption is that controlling for the lagged test scores and fixed effects account for most of the selection of students into schools. In a robustness test, I show that these type-specific value-added measures are indeed highly predictive of a student's out-of-sample gains from attending a given school. To proxy for the type-specific value-added of a school with a single variable, I average the math, Urdu, and English type-specific value-addeds.

For the discrete choice model, it is desirable to estimate type-specific value-addeds for the universe of schools. To identify value-addeds in this way, I must observe both high and low types in the same school. In cases where it is impossible to estimate type-specific value-addeds from test scores, but where school characteristics are available, I extrapolate type-specific value-addeds using school characteristics, regressing my previous estimates on a complete set of school facilities, rankings of desired teacher characteristics according to the school headmaster, and time use variables. This method is used to estimate the type-specific value-added for schools attended by less than 5 percent of the sample, the bulk of which are government schools with no high types.

With these estimates of a student's type and a school's type-specific value-addeds, I can now estimate a multinomial logit discrete choice model, where a student i of type z's utility in a school s in year t is given by

$$u_{ist} = \sum_z \Big(\sum_k \beta_{1zk} \mathbb{1}_z \times \overline{VA_{ks}} + \beta_{2z} \mathbb{1}_z \times fee_{st} + \beta_{3z} \mathbb{1}_z \times distance_{ist} \Big) + \gamma X_{ist} + \epsilon_{ist},$$

and ϵ_{ist} is a draw from the type 1 extreme value distribution (consistent with my underlying model of school choice), $\mathbb{1}_z$ is an indicator variable equal to 1 if student i is of type z, fee_{st} is the fee in 1000s of rupees charged by school s in year t, residualized by round to account for inflation, and $distance_{ist}$ is the distance between student i's household and school s. $\overline{VA_{ks}}$ is the average value-added of a school s for type k. X_{ist} consists of school-student-year level controls, including indicator variables for a student choosing the same school she chose in the previous year, for a girl attending a boy's school, for a boy attending a girls' school, and for a student attending a school whose grade-level exceeds the school's self-reported maximum grade.

Assuming that errors are drawn from a type 1 extreme value distribution produces a closed form expression for the probability that a student i attends a school s in year t. Let

$$v_{ist} = \sum_z \Big(\sum_k \beta_{1zk} \mathbb{1}_z \times \overline{VA_{ks}} + \beta_{2z} \mathbb{1}_z \times fee_{st} + \beta_{3z} \mathbb{1}_z \times distance_{ist} \Big) + \gamma X_{ist}.$$

Then, the probability a student i attends a school s in year t is

$$p_{ist} = \frac{\mathbb{1}_{village_{st}=village_{it}} e^{v_{ist}}}{\sum_j \mathbb{1}_{village_{jt}=village_{it}} e^{v_{ijt}}},$$

where $\mathbb{1}_{village_{st}=village_{it}}$ is an indicator variable equal to 1 if a school s is located in the same village as a student i in year t and 0 otherwise. This excludes schools outside of a student's village from her choice set. Therefore, the multinomial logit model can be estimated with maximum likelihood by maximizing the log likelihood objective function

$$\sum_i \sum_t \sum_s \mathbb{1}_{its} \times log(p_{ist}),$$

where $\mathbb{1}_{its}$ is an indicator variable equal to 1 if a student i enrolls in a school s in year t and 0 otherwise.

5.2 Results From Multinomial Logit Model

Before turning to the results of the multinomial logit model, it is helpful to understand the factor analysis that produces high and low types, the composition of schools, the distribution of the estimated type-specific value-addeds, and the choice sets of high and low types within a village. Table 4 presents the factor loadings of each variable included in the factor analysis that determines high and low types. Ownership of each asset, BMI, and lagged test scores all load positively on the first factor, consistent with the idea that the factor is proxying for underlying advantage. The first factor accounts for 45 percent of the correlation between the component variables. Figure 3 presents the distribution of the percentage of high types in private and public schools according to this calculation. Since the socioeconomic factor is positively related to attending a private school, there are relatively fewer high types in government schools. While there is some segregation by types in private schools, overall the composition is quite mixed. The median private school is 47 percent high types.

Figure 4 plots the distribution of value-addeds by type in private and government schools. For both types, private schools appear to be better than public schools, consistent with other work on private schooling in Pakistan and India by Andrabi et al. (2010) and Muralidharan and Sundararaman (2013). However, in general, value-addeds are much lower for low types in both types of schools, with very little overlap of the distributions of value-added for high and low types. Finally, figures 5 and 6 plot the distributions of the choice sets for low and high types within a village. In figure 5, the kernel density plots represent the distribution of the worst match choice in a village, the median match choice, and the best match choice for a low type. Figure 6 plots the same set of choices for high types. These figures suggest that there is substantial heterogeneity in match quality *within* villages. In a world where there is no within-village heterogeneity in match quality, the three curves would be the same.

The type-specific value-added estimates also allow us to plot value-added for high types against value-added for low types and directly verify that these value-added estimates differ by type. Figure 7 displays this relationship. I find that, while the value-addeds are correlated, the correlation coefficient is only 0.4. While this correlation may seem high, if school quality is purely vertical, the true correlation would be 1. This low correlation is not merely driven by attenuation bias. In a regression of student test scores on a rich function of lagged student test scores, school fixed effects, and school fixed effects interacted with indicator variables for high socioeconomic status, an F-test strongly rejects the possibility that the school-by-type fixed effects are jointly equal to 0. While the positive relationship between the two suggests that elements of school quality that impact all students are important, the large degree of unexplained variation suggests that match-specific quality is also an important component of the learning production function.

Table 5 presents parameter estimates from the multinomial logit model of school choice. The directions of the parameter estimates match our economic intuitions. The effects of higher fees and greater distances are negative, and children from poorer households are particularly price sensitive. The key take-away is that low types do not respond positively to their expected gains within a school. The coefficient on $\overline{VA_{low,s}} \times \mathbb{1}_{low}$ is actually negative. While low types do not respond positively to their own predicted gains, high types respond to both their and low types' predicted gains, but by far, they respond the most to their own predicted gains. These results suggest that the heterogeneous Hotelling model is a useful framework for studying school competition. More advantaged students are more responsive to their match to a school when they sort into schools. In fact, disadvantaged students are *negatively* responsive to their type-specific value-added, which, if we take the theoretical model seriously, suggests that they will be very likely to make costly mistakes. However, this relationship may be driven by bias, since the multinomial logit model does not account for unobserved school quality, which, in turn, may be correlated with price. In the next subsection, I also take unobserved school quality into account.

5.3 Framework for Discrete Choice Model

One limitation of the multinomial logit discrete choice model is that parents are likely to make enrollment decisions based on school characteristics that are unobservable to the econometrician. Since schools will set price as a function of these characteristics, the coefficient estimates from the original multinomial logit model may be biased. To ensure that this bias is not driving my results, I estimate a structural model of school choice following Berry et al. (2004), which incorporates school-specific unobserved quality. In this model,

$$u_{ist} = \beta_1 \overline{fee_s} \times \mathbb{1}_{high} + \sum_k \beta_{2k} \mathbb{1}_{high} \times \overline{VA_{ks}} + \sum_z \beta_{3z} \mathbb{1}_z \times distance_{ist} + \gamma X_{ist} + \delta_s + \epsilon_{ist}$$

and δ_s is:

$$\delta_s = \beta_4 \overline{fee_s} + \beta_5 \overline{VA_{low,s}} + \beta_6 \overline{VA_{high,s}} + \xi_s.$$

Here, ξ_s is a school-specific unobservable characteristic, and $\overline{fee_s}$ is the mean across years of fee_{st}. $\sum_z$ sums over types of students and $\sum_k$ sums over types of value-added. X_{ist} is a set of student-school level controls, including the interaction of the student's gender and the sex of the school (if the school is not co-educational) and interactions between student age fixed effects and indicator variables for whether the highest grade offered by the school is at the primary, junior secondary, higher secondary, or greater level. In this model, students still choose schools based on the same set of observed characteristics, but now their decisions also depend on a school-specific unobserved quality ξ_s.

Estimating this model requires estimating δ_s, a school-specific indicator variable, for each school. There are only 913 children in the sub-sample of children who appear in the household survey (distance data are available) and who are also tested (lagged test score data are available). As a result, of over 800 total schools, students are only observed attending $\frac{1}{3}$ of the schools in the sub-sample data, making it impossible to identify δ_s for two thirds of this sample. To avoid this problem, I instead use the full sample of children 5-15 in the survey administered to households to estimate the model. Since test scores and anthropometric measurements are not available for most of this sample, and the survey questions administered to households were different from those administered to children within schools, I proxy for advantage using a factor analysis of 25 indicator variables for household assets normalized at the village-year and child-age level and define high types as those with first factors above the fiftieth percentile for private school enrollees. Among the sub-sample of children who appeared in both surveys, the two measures are quite correlated (correlation = 0.37, $p < 0.01$). Table 6 presents the factor loadings of the 25 assets.

I estimate the model using a two-stage procedure. First, I estimate $\theta_1 = (\delta_s, \beta_1, \beta_{2k}, \beta_{3z}, \gamma)$ using the same maximum likelihood strategy as before. Then, it remains to estimate $\theta_2 = (\beta_4, \beta_5, \beta_6)$. Since schools set $\overline{fee_s}$ as a function of ξ_s, estimating θ_2 requires additional instruments. To instrument for price, I create a proxy for variation in teacher salaries across villages that is independent of village-level demand or teacher quality. I first estimate residual variation in private teacher salaries at the teacher level with the following regression:

$$salary_{ist} = qualification\_FE_{ist} + training\_FE_{ist} + exper\_FE_{ist} + age\_FE_{ist} + female_{ist} + \alpha_s + \mu_{ist},$$

where $salary_{ist}$ is the salary of a teacher i in a year t and school s, $qualification\_FE_{ist}$ are fixed effects for the teacher's education level, $training\_FE_{ist}$ are fixed effects for type of teacher training (if any), $exper\_FE_{ist}$ are fixed effects for groupings of experience levels (0-1 years, 2-3 years, 3+ years), age_FE_{ist} are teacher age fixed effects, and α_s is a school fixed effect. For each teacher i in year t and school s, I predict the residual variation $\alpha_s + \mu_{ist}$. To create a proxy for differences in the cost of teachers in each village, I calculate the village-level mean of this measure, weighting each year equally, and then take the mean over the villages at the sub-district administrative level (called a tehsil), leaving out the village of interest.

The validity of the instrument rests on two assumptions. First, it assumes that the teacher characteristic controls effectively control for variation in teacher quality, which might drive salaries and be geographically correlated. Second, it assumes that factors that impact demand for school quality in a village (and therefore, potentially, teacher salaries) do not impact teacher salaries in other villages in the same tehsil. This variable is strongly related to private school teacher salaries (coefficient=0.25, $p < 0.01$). I interact an indicator variable for whether a school is private with this measure since it is only relevant for private teacher salaries (and private school fees). Then, I residualize this new interaction by an indicator variable for whether the school is private and by the original measure to create the final instrument $salary\_inst_s$. The goal of this approach to allow the effect of the instrument on school fees to be greater for private schools while removing any correlation between the size of the instrument and the presence of private schools in the village. The final instrument is also relevant to prices. In a regression of mean prices on the final instrument, the coefficient is 0.225 ($p < 0.1$), indicating that of a 1 R. increase in teacher salaries, 0.2 R.s are passed on to consumers in prices.

Now, I am able to define six moment conditions and estimate θ_2 using generalized method of the moments, where my moment conditions are given by

$$\Phi(\xi'\mathbf{z}) = \begin{pmatrix} \xi'\overline{VA_{low,s}} \\ \xi'\overline{VA_{high,s}} \\ \xi' salary\_inst_s \\ \xi'\overline{VA_{low,s}}^2 \\ \xi'\overline{VA_{high,s}}^2 \\ \xi' salary\_inst_s^2 \end{pmatrix}$$

and $\mathbf{z}$ is the set of identifying instruments.

5.4 Results From Discrete Choice Model

Table 7 presents parameter estimates from a discrete choice model of school choice that allows schools to have an unobserved characteristic. These estimates are qualitatively similar to those in the multinomial logit model. High types are quite sensitive to a school's type-specific value-added; low types are insensitive to their own type specific value-added. In a departure from the previous estimates, low types do not respond strongly *negatively* to their own value-added or positively to high types' value-added, suggesting that these estimates may have been biased. As we might expect once we control for unobserved variation in school quality, the coefficient on $\overline{fees_s}$ is also greater in magnitude in this model. Overall, these parameter estimates are again consistent with the theoretical model's assumption that less advantaged students will be less responsive to their own type-specific value-added than more advantaged students when they make enrollment decisions. Figure 8, which reports quantiles of the relationship between the actual shares of students within a village attending a school and the shares predicted by the model, shows that the model also fits the data closely.

In table A1, I re-estimate the structural model using only children who do not also appear in the tested sample and, therefore, were not used to calculate the value-addeds. The goal of this exercise is to ensure the point estimates in table 7 are not biased by correlations between school choice and biases in the type-specific value-added estimates. In fact, the point estimates in A1 are virtually the same as those in table 7.

6 Empirical Analysis of Student Achievement and Product Differentiation

6.1 Student Achievement Empirical Strategy

If more advantaged students (high types) are more able to sort into their best match schools, the heterogeneous Hotelling model suggests that on average the match between more advantaged students and schools could increase at the expense of the less advantaged (low types) in response to competition. To test whether competition makes high types better off in the same school relative to low types, I first exploit variation in the number of private schools due to private school entry and exit. I examine whether this variation in the number of private schools has heterogeneous effects on low and high types.

The identifying variation is driven by changes in the number of schools over the course of the panel in 40 of the 112 villages. One concern about this strategy is that the coefficient estimates may be biased by different time trends in villages that gain or lose private schools. To test this, I create a mean asset index by conducting a factor analysis of indicator variables for household-level asset ownership in the LEAPS household surveys. I calculate the mean of the first component of this index for each village. Figure 9 plots the mean of this index across villages over the 4 years for the sets of villages in which the number of private schools declines (21 villages), increases (19 villages), and stays the same (72 villages). As the figure shows, there is little difference at baseline between the mean asset indices for each type of village and in all three cases, the asset index is quite flat over time.

Formally, to estimate the heterogeneous effects of the number of private schools on the learning of high and low types, I run the regression

$$y_{igzmst} = \beta_0 + \beta_1 num\_pri_{mt} + \beta_2 num\_pri_{mt} \times \mathbb{1}_{high} + \eta_{zs} + \alpha_{zt} + \omega_{gz}$$
$$+ \sum_{g'} \sum_{z'} \left(\lambda_{g'z'} \mathbb{1}_{z'} \times \mathbb{1}_{g'} \times y_{igzms,t-1} + \phi_{z'g'} \mathbb{1}_{z'} \times \mathbb{1}_{g'} \times y^2_{igzms,t-1} \right) + \epsilon_{igzmst},$$

where i indexes students, t indexes years, m indexes villages, z indexes types, and g indexes grades. y_{it} is then a test score of a student i in year t, α_{zt} are year-by-type fixed effects, ω_{gz} are grade-by-type fixed effects, η_{zs} are school-by-type fixed effects, $\mathbb{1}_{g'}$ is an indicator variable equal to 1 when a student attends grade g' and 0 otherwise, and $\mathbb{1}_{z'}$ is an indicator variable equal to 1 if a student is type z' and 0 otherwise. This regression allows me to estimate the change in the test scores of low and high type individuals in the same school induced by the entry of a new private school. The coefficients of interest are β_1 and β_2. Intuitively, controlling for the school-by-type fixed effects and the rich function of lagged test scores suggests that β_1 and β_2 are identified by changes to the value-added for a type within a school when another school exits or enters the market.

As a robustness test on the regressions that rely on school exit and entry, I also instrument for the number of private schools, using the fact that historical receipt of a girls' government school increases private school entry today (Andrabi et al., 2013a). The instrument follows from rules regarding the allocation of girls' government secondary schools in Pakistan. The schools were built in villages with the highest populations in their patwar circle; to my knowledge, no other policies follow this rule. Therefore, variation in a village's own population relative to that of its neighbors, induces quasi-random variation in the construction of a girls' secondary school. In turn, a girl's government secondary school increases the supply of educated young women. Teaching in a private school is one of the few options open to young women without a post-secondary degree, so increasing the supply of these women reduces the cost of offering private schooling. Therefore, my instrument $rank_m$ is an indicator variable for whether a village had the highest population in its patwar circle in 1981. To ensure the instrument's validity, I condition on village area, cubics in 1998 and 1981 population, and quadratics in percent permanent houses, percent houses with electricity, percent houses with water, and female literacy in 1981. Thus, the first stage regression is

$$num\_pri_{mt} = \delta_0 + \delta_g rank_m + \gamma_3 Z_m + \mu_{mt},$$

where Z_m includes the village-level controls. Unfortunately, the area variable is missing for the entire district of Attock, so this specification only uses two-thirds of the sample villages. Since this specification only uses two-thirds of the sample and cannot take advantage of time variation, the resulting point estimates are less precise. $rank_m$ does not vary within villages or schools, so I can no longer run regressions exploiting variation in the individual outcomes of students within the same school. The school-by-type fixed effects in the previous regression would be collinear with the village-level instrument. Therefore, to compare the effect of competition on different types *in the same school*, I again construct a type-specific value-added measure. However, this time I construct a separate measure for each school-year.[5] This instead allows me to estimate the effect of number of private schools on predicted gains for a given type within a school. To do so, for each year, I run a separate regression

$$y_{igzst} = \eta_{zst} + \sum_{g'} \sum_{z'} \left(\lambda_{g'z'} \mathbb{1}_{z'} \times \mathbb{1}_{g'} \times y_{igzs,t-1} + \phi_{z'g'} \mathbb{1}_{z'} \times \mathbb{1}_{g'} \times y^2_{igzs,t-1} \right) + \phi_{gzt} + \epsilon_{igzst}$$

where η_{zst} is a fixed effect for school-by-type-by-year and ϕ_{gzt} is a fixed effect at the grade-by-type-by-year level. In this case, the type-specific value-added for a type z in a school s in year t is η_{zst}.

To estimate the effect of number of private schools on a school's type-specific value-added, I estimate

$$VA_{zst} = \gamma_0 + \gamma_1 num\_pri_{mt} + \gamma_2 num\_pri_{mt} \times \mathbb{1}_{high} + \zeta Z_{zm} + \omega_{zt} + \epsilon_{zst}$$

where an observation is at the school-by-type-by-year level, and I instrument for num_pri_{mt} using $rank_m$ and for $num\_pri_{mt} \times \mathbb{1}_{high}$ using $rank_m \times \mathbb{1}_{high}$. Z_{zm} consists of the same set of village-level controls from the census and administrative data as before interacted with indicator variables for type, and ω_{zt} is a year by type fixed effect. γ_1 and γ_2 are the coefficients of interest and have the same interpretation as β_1 and β_2. While the outcome variables are now value-addeds instead of

[5] Estimating a different type-specific value-added for each year was not feasible for estimating the discrete choice model, since the model requires knowledge of type-specific value-added for a student's entire choice set and missing school-years are very problematic.

test scores, the units of the coefficients are the same: they are changes to test score gains in terms of standard deviations of the population test score distribution.

6.2 Student Achievement Results

The first column of 8 shows the effect of the instrument $rank_m$ on the number of private schools. Conditional on polynomials in village population, a village gains an additional 1.421 ($p < 0.01$) private schools when it moves from being the second ranked to first ranked in population in 1981, suggesting that the instrument is relevant. The remaining columns of table 8 show the differential effects of number of private schools on high and low type students in the same schools. The same pattern emerges across the two identification strategies and across outcomes. The point estimates for the effect of number of private schools on low types are consistently negative, with the exception of Urdu, for which they are positive but small in magnitude. In contrast, the point estimates for the interaction between number of private schools and being a high type are consistently positive and relatively large in magnitude. Combining these outcomes to form an average effect size suggests that high types make gains due to competition, while low types in the same school are negatively impacted on average.

The strongest evidence of this differential is in math, the subject which is typically most affected by educational interventions in both the developing world and the United States. According to the results exploiting the entry and exit of private schools, an additional private school leads low types to have test scores that are -0.083 standard deviations lower ($p < 0.01$), while an additional private school leads to gains of 0.012 for high types in the same school. The instrumental variables results, while less precisely estimated, suggest an even larger differential with low types losing 0.084 standard deviations in test score gains per year and high types gaining 0.077 standard deviations. In all, the exit-entry point estimates suggest that the gap between low types and high types attending the same school in math will grow about 0.1 standard deviations per year over the course of 5 years of primary school in response to another competitor entering the market. To put this magnitude into perspective, the resulting gap in test score levels between advantaged and disadvantaged students will be half the size of the black-white achievement gap in the United States (Fryer and Levitt, 2004). These results are consistent with the Hotelling model with heterogeneous sorting but inconsistent with a standard Hotelling model. Consistent with the Hotelling model with heterogeneous sorting, the match between less advantaged types and schools decreases when competition increases. These results are also inconsistent with vertical differentiation and price competition, since vertical differentiation does not increase within-school inequality.

6.3 Product Differentiation Empirical Strategy

Both the Hotelling model of Eaton and Lipsey (1975) and my heterogeneous Hotelling model predict that a greater number of schools will lead to more product differentiation. Even though this prediction does not separate the different Hotelling models of school competition, it is of interest, since, if it is true, it provides additional evidence that horizontal competition is an important aspect of school competition in rural Pakistan. In this section, I test this prediction. To measure variation in school characteristics, I develop three indices of school characteristics: facilities, time use, and desired teacher characteristics. The facilities index consists of data on a school's number of permanent rooms, semi-permanent rooms, and blackboards, as well as data on whether the school has a store, toilets (or a latrine), wall, hall, fan, library, sports, computer and electricity. The time-use index consists of minutes spent in class, on break, playing sports, in play time, on extra teaching, on arts, on extracurriculars, on prayer, and on homework. Finally, the teacher characteristic index consists of the head teacher or school owner's rankings of how much he or she values a teacher's ability to enroll new students, whether a teacher is highly educated, a teacher's training, whether a teacher is young, whether a teacher is experienced, whether a teacher comes from a wealthy family, and whether a teacher is local.

For each index, I normalize each of the component variables and then take their mean at the village-year level. To capture the extent of variation between schools in a village-year, I sum the squared distance between each component of the school-year characteristic vector and each component of the mean school-year characteristic vector in the village. Thus,

$$Index_{st} = \sum_j (x_{stj} - \frac{1}{S_{mt}} \sum_i x_{itj})^2$$

where s indexes a school, t is a year, j is the component of the index, S_{mt} is the number of schools in the village m in the year t, and the i indexes the schools which are summed over at the village-year level. In cases where there is only one school in a village in a given year, $Index_{st}$ will always be zero since the mean village-year characteristic vector will be equal to the single school-year characteristic vector. Since including these observations will introduce a mechanical correlation between $Index_{st}$ and the number of private schools, I drop them out of the analysis.

To test whether the number of private schools impacts school differentiation in the private market, I first run the OLS regression

$$Index_{smt} = \beta_0 + \beta_1 num\_pri_{mt} + \gamma_1 X_{smt} + \epsilon_{smt}$$

on the sample of private schools, where num_pri_{mt} is the number of private schools in a village m in period t, and X_{smt} is a vector of village and school characteristics, which may vary at the year level, including fixed effects for the maxmimum grade offered by the school, survey round fixed effects, and district fixed effects. To account for correlation in villages over time, I cluster my standard errors at the village level. Since number of private schools may be endogeneous, as in the previous section, I instrument for number of private schools with $rank_m$. Since the index measure is undefined in villages where there is only 1 private school, I can no longer exploit the exit-entry identification strategy. Unfortunately, much of the variation in number of private schools is between 1 and 2 schools, and therefore, dropping these villages greatly reduces the identifying variation.

School differentiation in terms of characteristics may also be caused by competition and vertical differentiation. For example, Andrabi et al. (2013b) provides evidence that private schools compete on price and vertical quality in the presence of asymmetric information. To account for this possibility, in a subset of the regressions, I include controls that account for the vertical quality of the school. In my time-use differentiation regressions, I control for total number of minutes in the school day, while in my facilities differentiation regressions, I control for the number of facilities. If the variation in school characteristics is predominantly between "good" high cost schools with longer school days and more facilities and "bad" low cost schools with few facilities and shorter school days, controlling for these variables should account for much of the variation.

6.4 Product Differentiation Results

Table 9 presents the results of the ordinary least squares and instrumental variables regressions of the differentiation indices on number of private schools with standard errors clustered at the village level. The table shows that there is a strong relationship between the number of private schools and the three differentiation indices. In the ordinary least squares regressions, an increase in the number of private schools leads to a substantial increase in the variation in the three characteristic indices ($p < 0.01$). Unsurprisingly, the instrumental variables estimates are less precise. However, the point estimates from the instrumental variables regressions are even larger than the ordinary least squares estimates and typically significant at the 5 percent level. Including controls for the total minutes of school time or the total count of facilities has little effect on the point estimates of interest in either the ordinary least squares or the instrumental variables regressions, suggesting that variation in school characteristics is not predominantly driven by vertical competition. Overall, these results are consistent with both the standard Hotelling model and the heterogeneous Hotelling model, both of which suggest that greater competition will incentivize schools to differentiate. Therefore, these results provide additional evidence in favor of the idea that competition in horizontal quality is an important component of school competition.

7 Robustness

In this section, I present several robustness checks verifying the validity of the instrument and the exit-entry identification strategy. I also test the sensitivity of the main results to the definition of types.

7.1 Sensitivity to Definition of Types

First, instead of using factor analysis, I use a simple heuristic to assign types. Once I residualize lagged test scores by village-year and lagged grade fixed effects, I assign private school students above the median for the mean lagged test score to be high types and those below the median to be low types. I then repeat the student outcomes analysis in table 8 using these new types. Table 10 presents the results. While the results are less precise, the coefficients largely follow the same pattern as in table 8. According to the instrumental variables identification strategy, the average effect size for the effect of an additional private school on high types is positive and significant ($p < 0.05$), while the effect for low types is negative.

7.2 Validity of Exit-Entry Strategy

The exit-entry strategy relies on the assumption that time trends in type-specific value-added are the same in villages where the number of private schools changed and in villages where it stayed the same. To test if this is the case, I re-run the exit-entry regressions including a forward lag in the number of private schools and the interaction between whether an individual is a high type and the forward lag for the number of private schools. If the results in table 8 are biased by differential time trends, the coefficients on the forward lag should be negative and significant, while the coefficient on the forward lag interacted with being a high type should be positive and significant. Instead, as table 11 shows, both coefficients are insignificant, and the forward lag coefficient is positive on average while the interaction is negative on average. Including the forward lags in the regressions has little effect on the key results. Overall, this table suggests that the exit-entry coefficient estimates are not driven by differential time trends in villages that gained or lost private schools.

An additional concern about identification is that the point estimates in table 8 are driven by compositional effects. If the specifications in table 8 do not adequately control for unobserved student characteristics, and changing the number of private schools in a village changes the underlying composition of the students I have classified as high and low types, this will bias my point estimates. To show that this bias does not drive my results, I restrict the sample of my exit-entry specifications to students who were enrolled in private schools before the number of private schools changed. Table 12 reports the results of these regressions. The results are largely unchanged.

Alternatively, the effects of private schools on outcomes for high and low types may be "causal," but the mechanism may be the change in the composition of the peer group rather than the change in schools' incentives. To rule out this alternative explanation, I re-run the regressions in table 8, including a control for the percentage of high types in school s in year t whose effect is allowed to vary by type. Table 13 reports the results of these regressions in both identification strategies. The instrumental variables regression results are less precise, but overall, the point estimates for both identification strategies are virtually the same.

7.3 Validity of the Instrument

My fourth robustness test concerns the validity of the instrument. First, I perform a placebo test of the instrument's identifying assumption. For the instrument to be valid, conditional on a flexible function of population, there is no discontinuity in a village's number of private schools as it passes from being the second to the first most populous village in its patwar circle apart from that induced by the allocation of girls' government secondary schools. Therefore, there should be no discontinuity in the number of private schools as a village passes from being the third to the second most populous village. The first column of table 14 tests this by regressing the number of private schools on an indicator variable for whether a village is the second or first most populous village in its patwar circle ($rank\_placebo_m$), controlling for the standard controls from the instrumental variables specification and an indicator variable for whether the village is the most populous village in its patwar circle. The coefficient estimate for $rank\_placebo_m$ is relatively small and insignificant, indicating that there is no discontinuity as villages move between the second and third population rankings.

A second concern about the instrumental variables strategy is that villages that had girls' government secondary schools in the past may have larger populations of educated adult women today, which may impact educational outcomes through other channels. To test whether this is the case, I use data from rounds 3 and 4 of the LEAPS survey (the only rounds in which questions about literacy were asked) to calculate the female literacy rate in each village. Then, I repeat the first stage of the instrumental variables regression, replacing number of private schools with adult female literacy as the outcome variable. The second column of table 14 presents the results of this regression. The effect of the rank indicator variable on female literacy is small, negative and insignificant, indicating that the instrument is not associated with large increases in female literacy. This is not wholly surprising since out-marriage, where females marry outside of their birth villages, is common in rural Pakistan, and few females reach the secondary schooling level. According to Andrabi et al. (2011), 50 percent of mothers in this sample live in a different village from their birth village (86 percent of these are due to out-marriage), and only 10 percent of mothers report obtaining greater than primary schooling (75 percent report no education at all).

A third concern is that having a girls' government secondary school in the past may predict having a girls' government secondary school today. If this is the case, the presence of a girls' secondary government school may impact the outcomes of female students through their aspirations even though the sample consists entirely of private school students. The remaining columns of table 14 use the instrumental variables strategy to estimate the heterogeneous effect of number of private schools on type-specific value-added using a male-only sample. As table 14 shows, while the estimates are again less precise, the point estimates are qualitatively and quantitatively the same as before.

7.4 Predictive Power of Type-Specific Value-Addeds

The final robustness test verifies that school type-specific value-added is good proxy for predicted test score gains for both types. I adapt the out-of-sample teacher value-added validation test developed by Chetty et al. (2014). For the sample of school-changers, I regress their test scores in year t on the value-added of their new school s (calculated not including own test score) for their own type and the opposite type. Table 15 presents the results. Across all tested subjects, I find that own type-specific value-added is highly correlated with a student's predicted test score gains in a school and is much more predictive of a student's gains in the school than the value-added for the other type. Moreover, type-specific value-added is similarly predictive for both types, though slightly *more* predictive for low types, who are *less* responsive to it. Therefore, we can reject the possibility that low types' lower responsiveness to their own type-specific value-added in the discrete choice model is driven by a lower correlation between their type-specific value-added and their test score gains.

8 Conclusion

In this paper, I develop a heterogeneous Hotelling model of spatial competition where consumers' information about their match with a product is correlated with their type. The work of Mullainathan and Shafir (2013), which provides evidence that more advantaged consumers also have greater reserves of attention, suggests that rational inattention with differential attention constraints is a natural micro-foundation for this model. While I test the implications of this model in educational markets in Pakistan, the theoretical contribution is more general. When individuals are differentially informed and their informedness is related to their optimal instructional level, increased competition can have adverse impacts on less informed types. I show that, depending on the distribution of information in the population, increased competition can *decrease* net welfare. These results may apply to both political markets and other product markets, such as health, where information about horizontal quality is limited, and opportunities to experiment with products are rare.

Education is a natural place to test the implications of the heterogeneous Hotelling model. Work by Duflo et al. (2011) and Kremer et al. (2013) suggests that horizontal differentiation is very important to students' learning outcomes in the developing world. Duflo et al. (2011) show that tracking, which allows a teacher to better tailor her instructional level to her students, increases learning for all students. In a review of the literature on education and developing countries, Kremer et al. (2013) conclude that interventions that better match schools' instructional levels to students' needs are among the most effective.

I show that a model where schools compete on horizontal quality is consistent with school choice behavior and educational outcomes in rural Pakistan, where there is a thriving low cost private schooling market. Consistent with the assumptions of the model, high types are quite responsive to their predicted test score gains when they choose schools, while low types are not responsive. I find that while wealthier "high" types benefit from increased competition, "low" types in the same schools suffer. My more conservative estimates imply that an additional private school in the market will lead to a nearly 0.1 standard deviation gap in the gains of low and high types. Over the course of five years of primary schooling, the magnitude of this gap will be $\frac{1}{2}$ the size of the black-white gap in the United States. I also find that schools tend to differentiate more in terms of their characteristics in the presence of competition, providing additional evidence that schools compete horizontally.

Improving the ability of disadvantaged students to sort into their best match schools will have long-run supply side effects on schools' horizontal quality. Interventions which help disadvantaged types sort will even aid those who don't respond to the intervention by changing the distribution of the horizontal quality of the available schools. In general, these results suggest that as some policymakers move toward offering vouchers on a large scale, as they have already done in Chile, they may also want to adopt informational interventions.

The recent growth of low cost private schools in Pakistan mirrors that of much of the rest of the developing world. Educational markets and enrollment rates in private schooling are similar across South Asia and Sub-Saharan Africa. Given the growth of private schooling markets in the developing world, it is important to understand how the interaction of market mechanisms and information can impact the educational outcomes of different groups of students. This model also may have implications for schooling in the United States, which has increasingly adopted a more pro-choice, high accountability approach to education. As schools compete for students in districts with school choice like Boston and New York, some of the same market mechanisms may play out. This is an area that warrants further research.

Finally, this paper raises an important methodological point. As my results show, estimates of the effect of school competition that do not take into account the heterogeneity of the effects of competition on poorer and wealthier students may find moderate or null effects. However, these moderate effect sizes may mask large positive effects on wealthy students and negative effects on poorer students.

9 Tables

Table 1: Predictions of a Hotelling Model With Perfect Sorting and the Hotelling Model With Heterogeneous Sorting

	Hotelling	Heterogeneous Hotelling
Average Match, Low	$\uparrow$	$\downarrow$
Average Match, High	$\uparrow$	$\uparrow$
Product Differentiation	Yes	Yes
Consumer Utility	$+$	Ambiguous

This table outlines the different predictions of the Hotelling model of Eaton and Lipsey (1975) and a Hotelling model which incorporates heterogeneous sorting.

Table 2: Summary Statistics of Schools in LEAPS Rounds 1-4

	Government			Private		
	Mean	SD	N	Mean	SD	N
Number of Private Schools	3.000	2.596	1,977	4.630	3.356	1,197
Number of Government Schools	6.364	3.675	1,977	4.916	3.181	1,197
Maximum Grade Offered	5.888	2.000	1,972	7.524	1.997	1195
Permanent Rooms	3.336	3.033	1,975	4.204	4.110	1195
Semi Permanent Rooms	0.658	1.514	1,976	1.821	2.966	1,197
Staff Rooms	0.261	0.475	1,975	0.530	0.532	1,197
Store	0.266	0.620	1,976	0.428	0.569	1,197
Toilets	0.579	0.494	1,976	0.857	0.350	1,197
Number of Blackboards	5.214	4.156	1,977	6.958	4.451	1,197
Library	0.222	0.415	1,977	0.390	0.488	1,197
Computer	0.010	0.100	1,977	0.266	0.442	1,197
Sports	0.108	0.311	1,977	0.351	0.477	1,197
Hall	0.068	0.252	1,977	0.195	0.396	1,197
Wall	0.655	0.475	1,977	0.962	0.192	1,197
Fans	0.471	0.499	1973	0.940	0.237	1193
Electricity	0.535	0.499	1,977	0.958	0.200	1196
Ranking of Teachers Enrolling New Students	5.722	2.017	1976	5.995	1.822	1196
Ranking of Teacher Highly Educated	2.842	1.730	1,977	2.617	1.538	1,197
Ranking of Teacher Training	3.460	2.233	1,977	3.820	2.829	1,197
Ranking of Teacher Young	4.751	1.537	1,977	5.368	1.549	1,197
Ranking of Teacher Experienced	3.009	1.657	1,977	2.760	1.650	1196
Ranking of Teacher Female	5.539	2.015	1,977	4.629	1.722	1,197
Ranking of Teacher Wealthy	7.233	1.281	1,974	7.371	1.217	1196
Ranking of Local	4.320	1.828	1,977	4.680	1.671	1,197
Time Spent in Class	256.409	49.973	1,977	248.073	48.006	1,197
Time Spent on Break	37.157	26.659	1,977	30.515	20.723	1,197
Time Spent on Sports	0.789	8.102	1,977	0.982	8.167	1,197
Time Spent on Play Time	0.627	6.194	1,977	0.564	6.154	1,197
Time Spent on Extra Teaching	1.022	9.796	1,977	2.477	18.401	1,197
Time Spent on Arts	0.258	5.741	1,977	0.439	5.915	1,197
Time Spent on Extracurriculars	0.549	9.401	1,977	0.401	5.108	1,197
Time Spent on Prayer	2.484	10.806	1,977	2.247	13.672	1,197
Time Spent on Homework	3.945	17.864	1,977	5.050	18.564	1,197

This table presents summary statistics for the facilities, time use, and the desired qualities of teachers in private and public schools in rounds 1-4 of the LEAPS study. Time use is measured in minutes, and the desirability of teacher characteristics is ranked by head teachers or school owners from 1-8 (with 1 being the most desirable).

Table 3: Summary Statistics for Tested Children in LEAPS Rounds 1-4

	Government			Private		
	Mean	SD	N	Mean	SD	N
Female	0.449	0.497	47,438	0.443	0.497	18,993
Age	10.485	1.785	48,836	10.273	1.761	19,478
Grade	3.981	0.936	48,837	4.027	1.010	19,479
Mom Finished Primary	0.225	0.417	33,599	0.400	0.490	16,154
Dad Finished Primary	0.502	0.500	33,595	0.683	0.465	16,156
Beds	0.996	0.059	27,607	0.998	0.044	14,088
Radio	0.539	0.498	27,607	0.633	0.482	14,088
Television	0.585	0.493	27,607	0.761	0.427	14,088
Refrigerator	0.319	0.466	27,607	0.601	0.490	14,088
Bicycle	0.710	0.454	27,607	0.746	0.435	14,088
Tables	0.851	0.356	27,607	0.952	0.214	14,088
Chairs	0.842	0.365	27,607	0.952	0.214	14,088
Fans	0.922	0.268	27,607	0.973	0.161	14,088
Watches	0.957	0.203	27,607	0.972	0.166	14,087
Scooter	0.168	0.374	27,607	0.295	0.456	14,086
Car	0.048	0.215	27,607	0.125	0.330	14,087
Telephone	0.357	0.479	27,607	0.577	0.494	14,087
Plough	0.219	0.413	27,607	0.250	0.433	14,088
Agricultural Tools	0.717	0.450	27,607	0.697	0.460	14,088
Tractor	0.115	0.319	27,607	0.160	0.367	14,087
Goats	0.663	0.473	27,607	0.528	0.499	14,087
Cattle	0.599	0.490	27,607	0.530	0.499	14,087
Chickens	0.649	0.477	27,607	0.573	0.495	14,087
Richshaw	0.038	0.192	27,607	0.040	0.196	14,087
Tubewell	0.158	0.365	27,607	0.234	0.423	14,087
BMI	-0.776	1.225	26,820	-0.768	1.237	13,685
Math	-0.154	0.971	48,837	0.322	0.838	19,479
Urdu	-0.182	0.964	48,837	0.390	0.853	19,479
English	-0.245	0.939	48,837	0.563	0.762	19,479
Change in Math	0.378	0.722	26,050	0.400	0.654	9,995
Change in Urdu	0.435	0.671	26,050	0.437	0.587	9,995
Change in English	0.387	0.691	26,050	0.359	0.595	9,995

This table presents summary statistics for the assets, parental education, and test scores for students enrolled in public and private schools in rounds 1-4 of the LEAPS study. Test scores have been normalized using item response theory and have a mean of zero and a standard deviation of 1 across the whole sample. BMI is in US population standard deviations.

Table 4: Factor Analysis of Student Characteristics

	Factor Loading
$english_{i,t-1}$	0.524
$math_{i,t-1}$	0.453
$urdu_{i,t-1}$	0.504
$bmi_{it} \times \mathbb{1}_{bmi_{it} \leq 0}$	0.033
$beds_{it}$	0.026
$radio_{it}$	0.280
$television_{it}$	0.340
$refrigerator_{it}$	0.491
$bicycle_{it}$	0.199
$tables_{it}$	0.474
$chairs_{it}$	0.483
$fans_{it}$	0.199
$watches_{it}$	0.094
$scooter_{it}$	0.405
car_{it}	0.257
$telephone_{it}$	0.436
$plough_{it}$	0.364
$agricultural\_tools_{it}$	0.188
$tractor_{it}$	0.388
$goats_{it}$	0.084
$cattle_{it}$	0.203
$chicken_{it}$	0.100
$rickshaw_{it}$	0.057
$tubewell_{it}$	0.344
Eigenvalue	2.637
Proportion	0.453

This table presents the variable loadings for the first factor from a factor analysis of student assets, lagged test scores, and body mass index. Each component of the factor analysis has been normalized at the village-year and student age levels, except for lagged test scores, which are normalized at the lagged grade level.

Table 5: Coefficients in the Multinomial Logit Model of School Choice

	(1) Parameter	(2) Standard Error
$fee_{st} \times \mathbb{1}_{low}$	-0.4903***	0.0573
$fee_{st} \times \mathbb{1}_{high}$	-0.0009	0.0641
$\overline{VA_{s,low}} \times \mathbb{1}_{low}$	-0.8878***	0.1405
$\overline{VA_{s,low}} \times \mathbb{1}_{high}$	-0.1244	0.2344
$\overline{VA_{s,high}} \times \mathbb{1}_{low}$	0.7076***	0.0938
$\overline{VA_{s,high}} \times \mathbb{1}_{high}$	1.2602***	0.1806
$distance_{is} \times \mathbb{1}_{low}$	-0.5461***	0.0473
$distance_{is} \times \mathbb{1}_{high}$	-0.5998***	0.0657
$\mathbb{1}_{boy_i} \times \mathbb{1}_{girls\_school}$	-0.7186***	0.2557
$\mathbb{1}_{girl_i} \times \mathbb{1}_{boys\_school}$	-0.5335**	0.2621
$\mathbb{1}_{grade_{it} > max\_grade_s}$	-2.1142***	0.2559
$\mathbb{1}_{school_{it} = school_{i,t-1}}$	4.2428***	0.0913
$\mathbb{1}_{school_{it} = school_{i,t-1}} \times \mathbb{1}_{grade_{it} > max\_grade_s}$	-3.1678***	0.3463

This table presents parameter estimates from a multinomial logit model of school choice. The parameters are estimated using maximum likelihood. Fees are normalized by year and are in 1000s of ruppees units.

Table 6: Factor Analysis of Household Characteristics

	Factor Loading
$beds_{it}$	0.066
$tables_{it}$	0.545
$chairs_{it}$	0.572
$fans_{it}$	0.267
$sewing\_machine_{it}$	0.442
air_cooler_{it}	0.418
$air\_conditioner_{it}$	0.218
$refrigerator_{it}$	0.494
$radio_{it}$	0.327
$television_{it}$	0.371
vcr_{it}	0.332
$watches_{it}$	0.199
$guns_{it}$	0.299
$plough_{it}$	0.477
$harvester_{it}$	0.403
$tractor_{it}$	0.478
$tubewell_{it}$	0.397
$agricultural\_machinery_{it}$	0.373
$hand\_tools_{it}$	0.284
$scooter_{it}$	0.376
car_{it}	0.282
$bicycle_{it}$	0.187
$cattle_{it}$	0.224
$goats_{it}$	0.112
$chicken_{it}$	0.154
Eigenvalue	3.184
Proportion	0.641

This table presents the variable loadings for the first factor from a factor analysis of household assets. Each component of the factor analysis has been normalized at the village-year and student age levels.

Table 7: Coefficients From the Discrete Choice Model of School Choice With Unobservable Characteristics

	(1) Parameter	(2) Standard Error
$\overline{VA_{low,s}} \times \mathbb{1}_{high}$	-0.1184	0.0806
$\overline{VA_{high,s}} \times \mathbb{1}_{high}$	0.5321***	0.0438
$\overline{VA_{low,s}}$	-0.0205	0.0889
$\overline{VA_{high,s}}$	-0.0532	0.1361
$\overline{fee_{s,t}} \times \mathbb{1}_{high}$	0.4200***	0.0374
$\overline{fee_{s,t}}$	-0.8856***	0.0694
$distance_{is} \times \mathbb{1}_{high}$	-0.9612***	0.0349
$distance_{is} \times \mathbb{1}_{low}$	-0.9850***	0.0290
$\mathbb{1}_{school_{it}=school_{i,t-1}}$	4.9244***	0.1618
$\mathbb{1}_{boy_i} \times \mathbb{1}_{girls\_school_s}$	-2.8107***	0.0798
$\mathbb{1}_{girl_i} \times \mathbb{1}_{boys\_school_s}$	-2.8126***	0.0710

This table presents parameter estimates from a discrete choice model of school choice with unobserved school quality. The individual by school level parameters are estimated using maximum likelihood. The school level parameters are estimated using generalized method of the moments. Fees are normalized by year and are in 1000s of ruppees units. The model also includes a set of interactions between a student's age and whether a school's highest grade is at the primary, secondary, upper secondary, or greater level.

Table 8: Effect of Number of Private Schools on Test Scores for High and Low Types

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Number Pri. Schools	Math		Urdu		English		AES	
	First Stage	Exit-Entry	IV	Exit-Entry	IV	Exit-Entry	IV	Exit-Entry	IV
$rank_m$	1.421***								
	(0.530)								
num_pri_{mt}		-0.083***	-0.084	0.024	0.025	-0.057*	-0.009	-0.039	-0.022
		(0.032)	(0.086)	(0.030)	(0.047)	(0.030)	(0.053)	(0.025)	(0.052)
$num\_pri_{mt} \times \mathbb{1}_{high}$		0.095**	0.161*	0.016	0.052	0.107***	0.103*	0.073**	0.104*
		(0.044)	(0.093)	(0.041)	(0.064)	(0.040)	(0.60)	(0.034)	(0.060)
Grade FE × Type FE × Lagged Score	N	Y	N	Y	N	Y	N	Y	N
Grade FE × Type FE × Lagged Score Sq.	N	Y	N	Y	N	Y	N	Y	N
School FE × Type FE	N	Y	N	Y	N	Y	N	Y	N
Year FE × Type FE	N	Y	Y	Y	Y	Y	Y	Y	Y
Type FE × IV Controls	N	N	Y	N	Y	N	Y	N	Y
IV Controls	Y	N	N	N	N	N	N	N	N
N	786	8,463	988	8,463	988	8,463	988		
Adjusted R^2	0.710	0.583	-0.012	0.635	0.030	0.597	0.032		
Clusters	73	5,346	205	5,346	205	5,346	205		

This table exploits exit and entry of schools and an instrumental variables strategy to examine the effect of school competition on test scores for "high" and "low" types in private schools. Types were determined by a factor analysis of the students' residualized assets, BMI, and lagged test scores. The first column demonstrates that the instrument predicts number of private schools, and standard errors in this column are clustered at the village level. Standard errors are clustered at the student level in the exit-entry regressions and the school level in the instrumental variables regressions. In the exit-entry regressions, the outcome variable is the test score of the student i in the year t. In the instrumental variables regressions, it is the value-added of a school s in the year t for a type z. Instrumental variables controls consist of variables from the Pakistani census and administrative data from the state of Punjab including village area, cubics in total population in 1998 and 1981, and quadratics in percent of houses permanent, percent of houses with electricity, percent of houses with water, and the female literacy rate in 1981. Average effect sizes are the averages of the coefficients in the regressions for math, English, and Urdu with appropriately adjusted standard errors.

Table 9: Effect of Number of Private Schools on School Differentiation Indices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		Class Time Index			Teacher Index			Facilities Index		
	OLS	OLS	IV	IV	OLS	IV	OLS	OLS	IV	IV
		Vertical Control		Vertical Control				Vertical Control		Vertical Control
num_pri_{mt}	0.858***	0.945***	1.798**	1.843**	0.432***	0.771**	0.664***	0.574***	0.902	0.750
	(0.275)	(0.276)	(0.846)	(0.876)	(0.069)	(0.360)	(0.162)	(0.166)	(0.560)	(0.598)
$total_{st}$		0.060***		0.062***				0.149***		0.143***
		(0.023)		(0.022)				(0.050)		(0.054)
District FE	Y	Y	N	N	Y	N	Y	Y	N	N
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Max Grade Offered FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
IV Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean	7.362				4.590		6.770			
N	700	700	700	700	699	699	696	696	696	698
Adjusted R^2	0.080	0.109	0.056	0.086	0.075	0.052	0.182	0.201	0.133	0.154
Clusters	60	60	60	60	60	60	60	60	60	60

This table shows the relationship between number of private schools and the differentiation indices. Column 1 presents the first stage of the instrumental variables regression. Columns 1, 2, 5, 7, and 8 present the results from an ordinary least squares regression of each of the different indices on the number of private schools in the village in year t. Columns 3, 4, 6, 9, and 10 present the instrumental variables estimates. To account for vertical differentiation, columns 2, 4, 8, and 10 include $total_{st}$, a control for either a count of the school's total facilities or the total minutes of school time. In all the school differentiation regressions, the sample excludes village-years where there was only one private school in the village. Standard errors are clustered at the village level. Instrumental variables controls consist of variables from the Pakistani census and administrative data from the state of Punjab, including village area, cubics in total population in 1998 and 1981, and quadratics in percent of houses permanent, percent of houses with electricity, percent of houses with water, and female literacy rate in 1981.

Table 10: Robustness of Effect of Number of Private Schools on Test Scores to Varying the Definition of Types

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Math		Urdu		English		AES	
	Exit-Entry	IV	Exit-Entry	IV	Exit-Entry	IV	Exit-Entry	IV
num_pri_{mt}	-0.055*	-0.073	0.033	0.004	-0.034	-0.019	-0.019	-0.028
	(0.033)	(0.079)	(0.030)	(0.048)	(0.032)	(0.049)	(0.026)	(0.049)
$num\_pri_{mt} \times \mathbb{1}_{high}$	0.016	0.154*	-0.031	0.058	0.033	0.129**	0.006	0.113**
	(0.042)	(0.083)	(0.040)	(0.053)	(0.040)	(0.052)	(0.033)	(0.52)
Grade FE × Type FE × Lagged Score	Y	N	Y	N	Y	N	Y	N
Grade FE × Type FE × Lagged Score Sq.	Y	N	Y	N	Y	N	Y	N
School FE × Type FE	Y	N	Y	N	Y	N	Y	N
Year FE × Type FE	Y	Y	Y	Y	Y	Y	Y	Y
Type FE × IV Controls	N	Y	N	Y	N	Y	N	Y
N	9,995	999	9,995	999	9,995	999		
Adjusted R^2	0.593	-0.014	0.653	0.053	0.608	0.059		
Clusters	6,625	209	6,625	209	6,625	209		

This table exploits the exit and entry of schools and instrumental variables strategies to examine the effect of school competition on test scores for "high" and "low" types in private schools. Types were determined by whether a student i in year z was above or below the median value for the mean of lagged test scores residualized at the village and lagged grade level. Standard errors are clustered at the student level in the entry-exit regressions and at the school level in the instrumental variables regressions. In the exit-entry regressions, the outcome variable is the test score of the student i in the year t. In the instrumental variables regressions, it is the value-added of a school s in the year t for a type z. Instrumental variables controls consist of variables from the Pakistani census and administrative data from the state of Punjab, including village area, cubics in total population in 1998 and 1981, and quadratics in percent of houses permanent, percent of houses with electricity, percent of houses with water, and the female literacy rate in 1981. Average effect sizes are the averages of the coefficients in the regressions for math, English, and Urdu with appropriately adjusted standard errors.

Table 11: Test for Differential Time Trends in Villages that Gained or Lost Private Schools

	(1) Math	(2) Urdu	(3) English	(4) AES
num_pri_{mt}	-0.093***	-0.013	-0.046	-0.051*
	(0.035)	(0.035)	(0.037)	(0.029)
$num\_pri_{mt} \times \mathbb{1}_{high}$	0.100**	0.052	0.090*	0.081**
	(0.049)	(0.046)	(0.047)	(0.038)
$num\_pri_{m,t+1}$	0.030	0.038	-0.044	0.008
	(0.038)	(0.031)	(0.057)	(0.036)
$num\_pri_{m,t+1} \times \mathbb{1}_{high}$	-0.069	-0.041	-0.012	-0.040
	(0.048)	(0.037)	(0.062)	0.041
Grade FE × Type FE × Lagged Score	Y	Y	Y	Y
Grade FE × Type FE × Lagged Score Sq.	Y	Y	Y	Y
School FE × Type FE	Y	Y	Y	Y
Year FE × Type FE	Y	Y	Y	Y
N	6,845	6,845	6,845	
Adjusted R^2	0.601	0.643	0.606	
Clusters	4,621	4,621	6,621	

This table tests whether forward lags of the number of private schools and the number of private schools interacted with being a high type predict test score gains for high types and test score losses for low types. Types were determined by a factor analysis of the students' residualized assets, bmi, and lagged test scores. Standard errors are clustered at the student level. In the exit-entry regressions, the outcome variable is the test score of the student i in the year t. Average effect sizes are the averages of the coefficients in the regressions for math, English, and Urdu with appropriately adjusted standard errors.

Table 12: Exit-Entry Identification Strategy With Non-Entrant Sample

	(1) Math	(2) Urdu	(3) English	(4) AES
num_pri_{mt}	-0.063*	0.055*	-0.023	-0.011
	(0.037)	(0.032)	(0.034)	(0.027)
$num\_pri_{mt} \times \mathbb{1}_{high}$	0.115**	0.018	0.104**	0.079**
	(0.049)	(0.043)	(0.044)	(0.036)
Grade FE × Type FE × Lagged Score	Y	Y	Y	Y
Grade FE × Type FE × Lagged Score Sq.	Y	Y	Y	Y
School FE × Type FE	Y	Y	Y	Y
Year FE × Type FE	Y	Y	Y	Y
Type FE × IV Controls	Y	Y	Y	Y
N	7,325	7,325	7,325	
Adjusted R^2	0.590	0.653	0.616	
Clusters	4,290	4,290	4,290	

This table exploits exit and entry of schools to examine the effect of school competition on test scores for "high" and "low" types in private schools. Types were determined by a factor analysis of the students' residualized assets, bmi, and lagged test scores. Standard errors are clustered at the student level. The outcome variable is the test score of the student i in the year t. Average effect sizes are the averages of the coefficients in the regressions for math, English, and Urdu with appropriately adjusted standard errors. The sample is restricted to students who were enrolled in private schools before exit or entry occurred in the village.

Table 13: Effect of Number of Private Schools on Test Scores for High and Low Types, Including Peer Effects Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Math		Urdu		English		AES	
	Exit-Entry	IV	Exit-Entry	IV	Exit-Entry	IV	Exit-Entry	IV
num_pri_{mt}	-0.082**	-0.050	0.020	0.036	-0.056*	-0.009	-0.039	-0.007
	(0.032)	(0.072)	(0.030)	(0.046)	(0.030)	(0.054)	(0.025)	(0.049)
$num\_pri_{mt} \times \mathbb{1}_{high}$	0.094**	0.087	0.020	0.002	0.105***	0.066	0.073**	0.050
	(0.044)	(0.072)	(0.041)	(0.050)	(0.040)	(0.052)	(0.034)	(0.047)
Grade FE × Type FE × Lagged Score	Y	N	Y	N	Y	N	Y	N
Grade FE × Type FE × Lagged Score Sq.	Y	N	Y	N	Y	N	Y	N
School FE × Type FE	Y	N	Y	N	Y	N	Y	N
Year FE × Type FE	Y	Y	Y	Y	Y	Y	Y	Y
Type FE × IV Controls	N	Y	N	Y	N	Y	N	Y
Peer Effects Controls	Y	Y	Y	Y	Y	Y	Y	Y
N	8,463	982	8,463	982	8,463	982	8,463	982
Adjusted R^2	0.582	0.005	0.636	0.092	0.597	0.095	0.687	0.080
Clusters	5,347	206	5,347	206	5,347	206	5,347	206

This table exploits exit and entry of schools and instrumental variables strategies to examine the effect of school competition on test scores for "high" and "low" types in private schools. Types were determined by a factor analysis of the students' residualized assets, bmi, and lagged test scores. Standard errors are clustered at the student level in the exit-entry regressions and at the school level in the instrumental variables regressions. In the exit-entry regressions, the outcome variable is the test score of the student i in the year t. In the instrumental variables regressions, it is the value-added of a school s in the year t for a type z. Instrumental variables controls consist of variables from the Pakistani census and administrative data from the state of Punjab including village area, cubics in total population in 1998 and 1981, and quadratics in percent of houses permanent, percent of houses with electricity, percent of houses with water, and the female literacy rate in 1981. Average effect sizes are the averages of the coefficients in the regressions for math, English, and Urdu with appropriately adjusted standard errors.

Table 14: Test of the Instrumental Variables Strategy's Validity

	(1) Number of Pri. Schools Full Sample	(2) Female Literacy Full Sample	(3) Math Males	(4) Urdu Males	(5) English Males	(6) AES Males
$rank\_placebo_{mt}$	0.325					
	(0.627)					
$rank_{mt}$		-0.015				
		(0.024)				
num_pri_{mt}			0.005	0.022	0.028	0.010
			(0.52)	(0.051)	(0.050)	(0.055)
$num\_pri_{mt} \times \mathbb{1}_{high}$			0.112	0.044	0.089	0.088
			(0.071)	(0.061)	(0.055)	(0.065)
IV Controls	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Number of observations	786	382	834	834	834	
Adjusted R^2	0.710	0.547	0.024	0.030	-0.035	
Clusters	73	71	197	197	197	

This table contains three tests of the validity of the instrument $rank_{mt}$ for number of private schools. The first column tests whether the placebo instrument $rank\_placebo_{mt}$, which is an indicator variable equal to 1 if a village is the first or second most populous village in its patwar and 0 otherwise, predicts the number of private schools in a village conditioning on the instrumental variables controls and a control for $rank_{mt}$. The second column tests whether the instrumental variable predicts adult female literacy at the village-year level using data from rounds 3 and 4 of the LEAPS survey. The remaining columns replicate the instrumental variables results in table 8 in the restricted sample of private school-attending males. Types were determined by a factor analysis of the students' residualized assets, bmi, and lagged test scores. Standard errors are clustered at the village level in columns 1 and 2 and at the school level in the remaining regressions. Average effect sizes are the averages of the coefficients in the regressions for math, English, and Urdu with appropriately adjusted standard errors.

Table 15: Out of Sample Validation of Type-Specific Value-Addeds

	(1)	(2)	(3)	(4)	(5)	(6)
	Math		Urdu		English	
	Low Type	High Type	Low Type	High Type	Low Type	High Type
$\overline{VA_{s,low}}$	0.892***	0.049	0.951***	0.146*	0.908***	0.147**
	(0.074)	(0.078)	(0.071)	(0.084)	(0.056)	(0.069)
$\overline{VA_{s,high}}$	0.031	0.916***	-0.021	0.906***	0.054	0.822***
	(0.055)	(0.085)	(0.055)	(0.145)	(0.054)	(0.139)
Grade FE × Lagged Score	Y	Y	Y	Y	Y	Y
Grade FE × Lagged Score Sq.	Y	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y	Y
N	1,507	749	1,507	749	1,507	749
Adjusted R^2	0.608	0.580	0.652	0.628	0.595	0.615
Clusters	356	290	356	290	356	290

This table presents correlations between student test score outcomes and school type-specific value-added for each type. In columns 1, 3, 5, the sample consists of school changers who are low types. In columns 2, 4, and 6, it consists of school changers who are high types. Standard errors are clustered at the school level.

10 Figures

Figure 1: Distribution of Private School Fees for Grades 1-3 (U.S. Dollars)

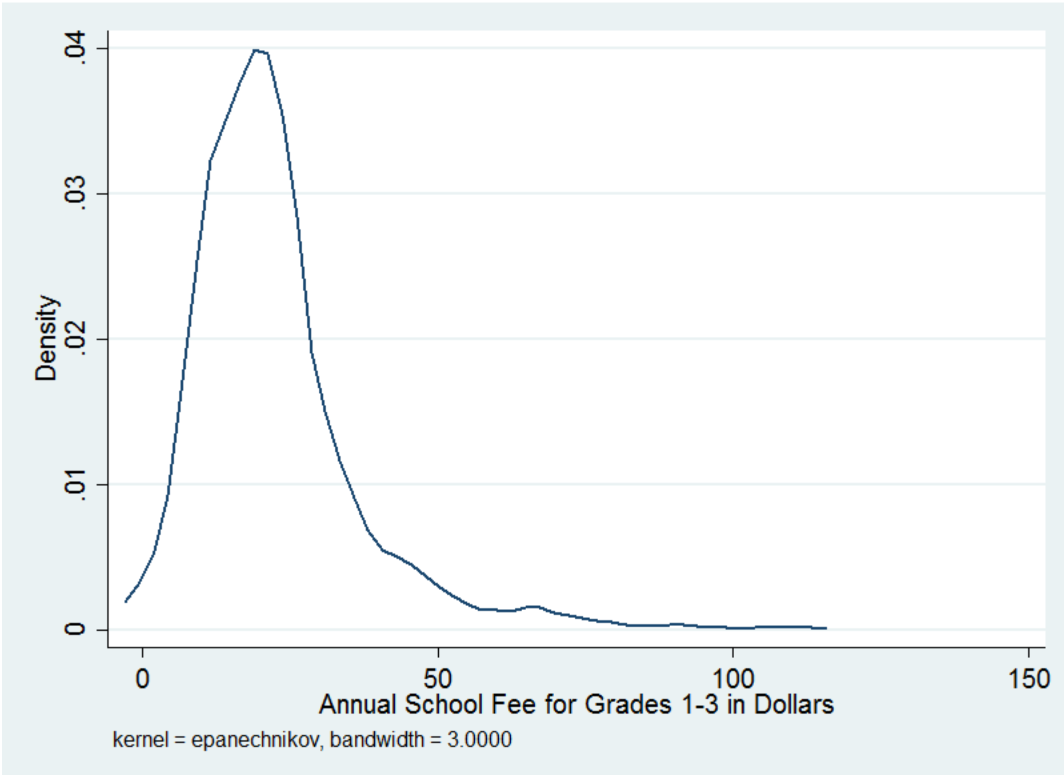


Figure 2: Distribution of Teachers per Grade in Government and Private Schools

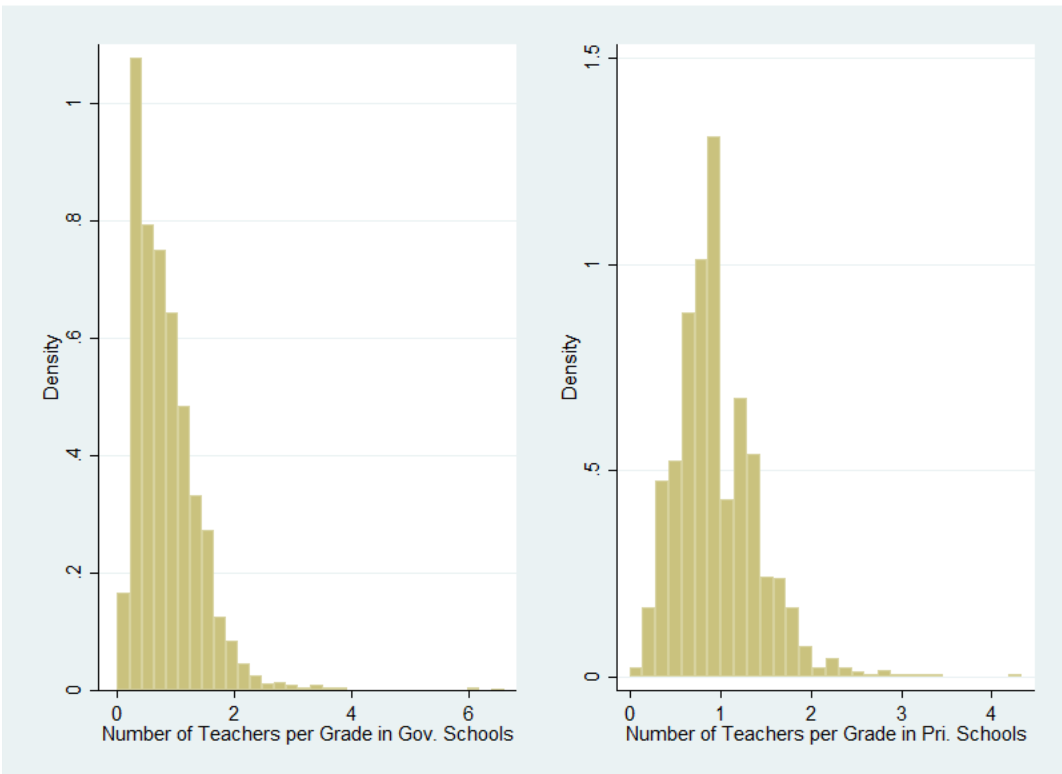


Figure 3: Composition of Government and Private Schools

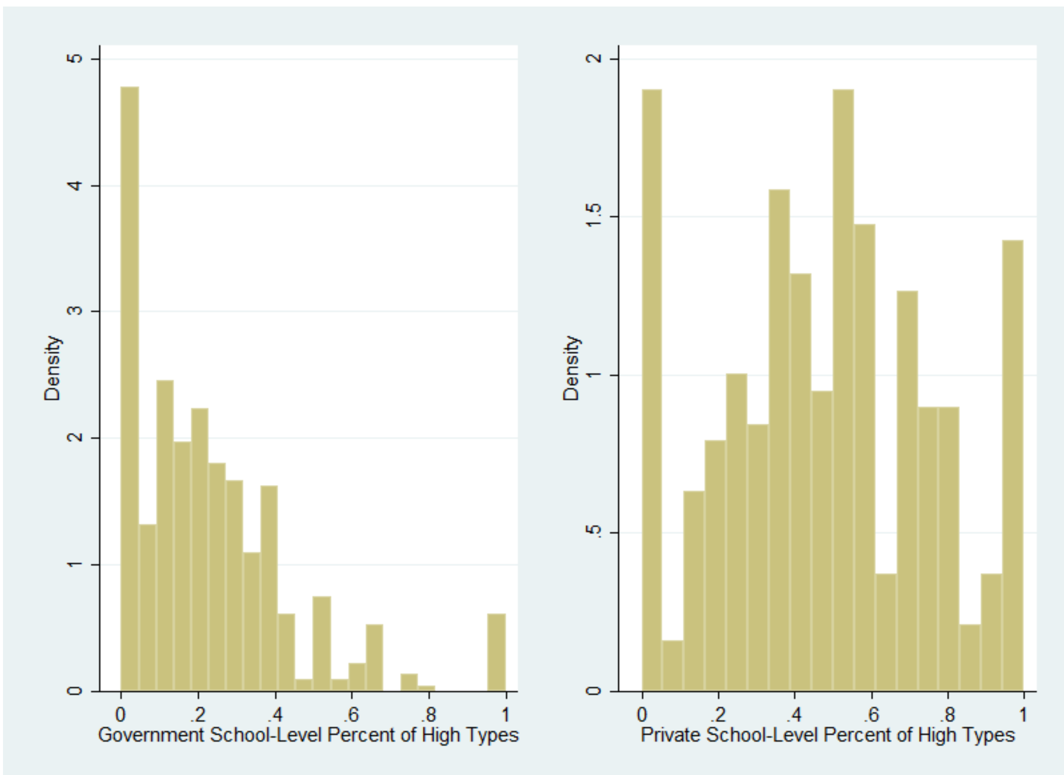


Figure 4: Kernel Density Plots of Mean Value-Added by Type in Government and Private Schools

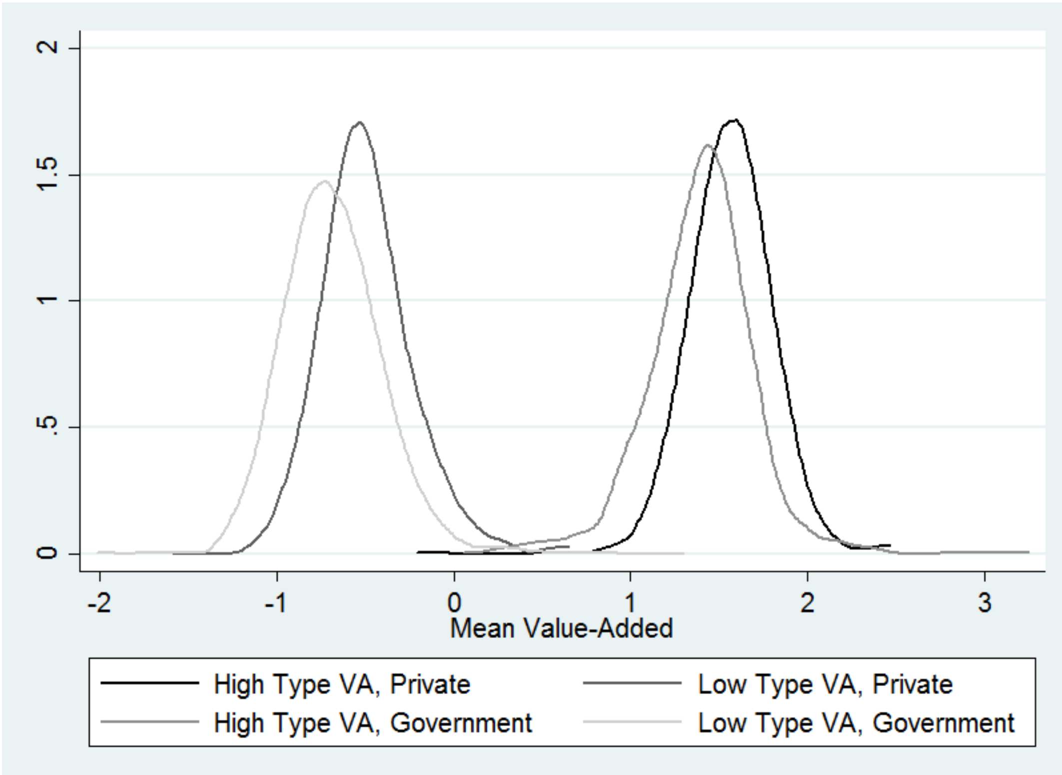


Figure 5: Kernel Density Plots of Worst, Median, and Best Match Schools Within a Village for Low-Types

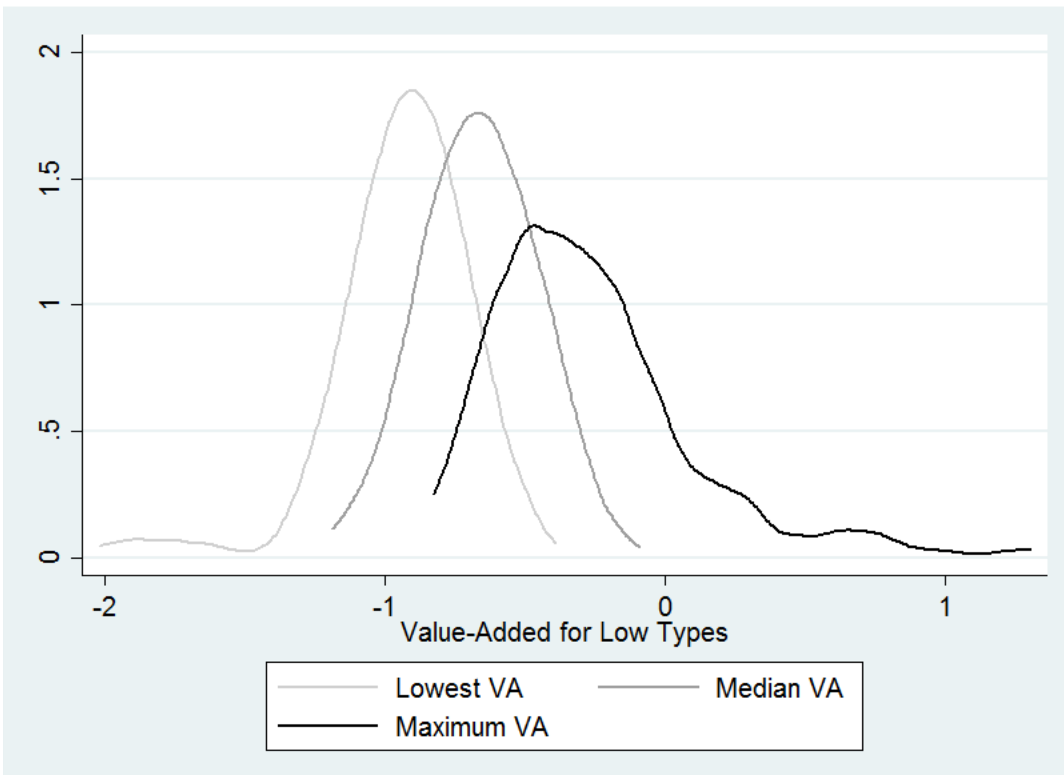


Figure 6: Kernel Density Plots of Worst, Median, and Best Match Schools Within a Village for High-Types

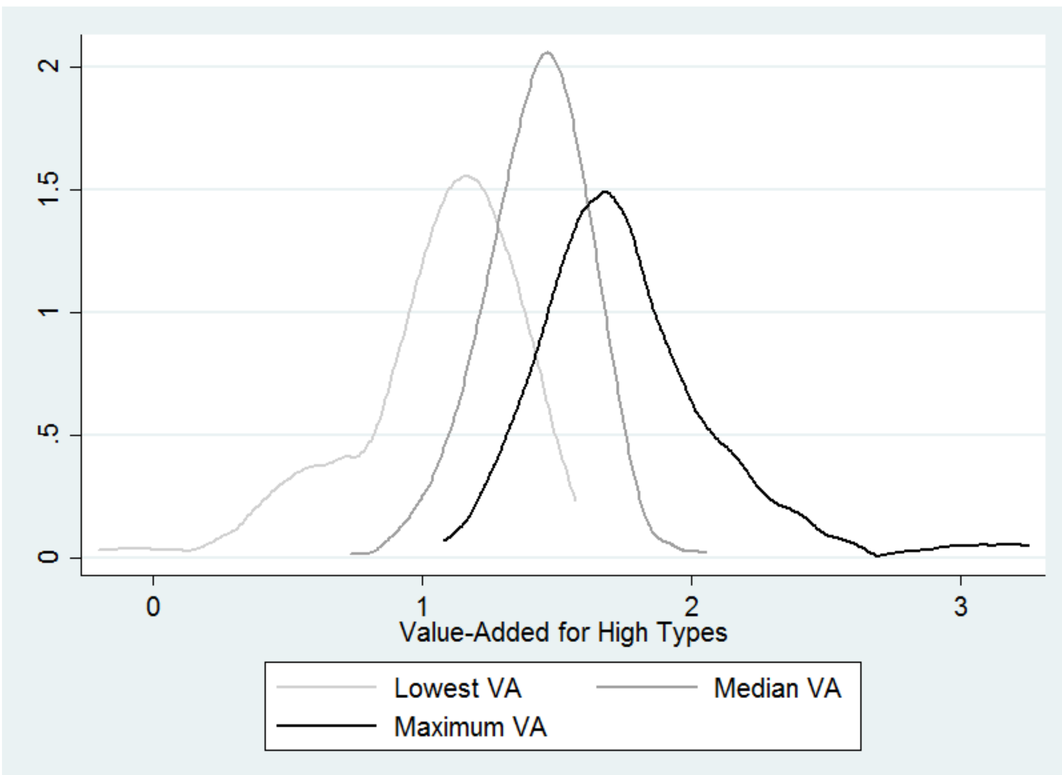


Figure 7: Relationship Between Mean Type-Specific Value-Addeds

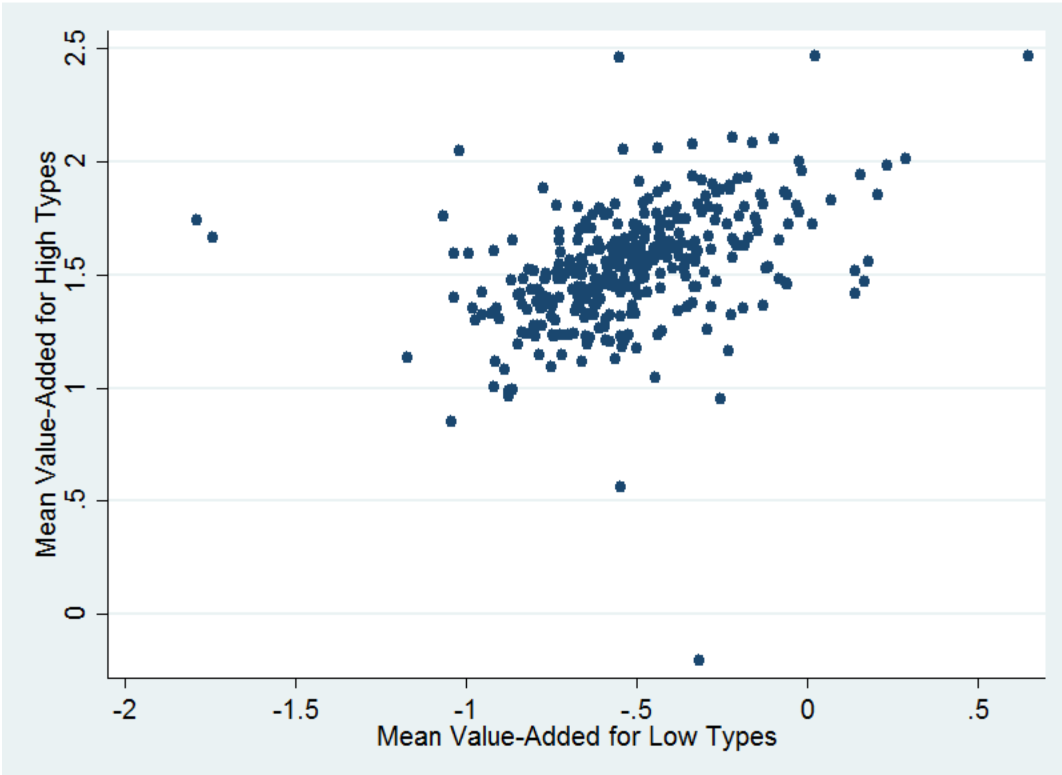


Figure 8: A Binned Scatter Plot of Actual School Shares on School Shares Predicted by Micro-BLP

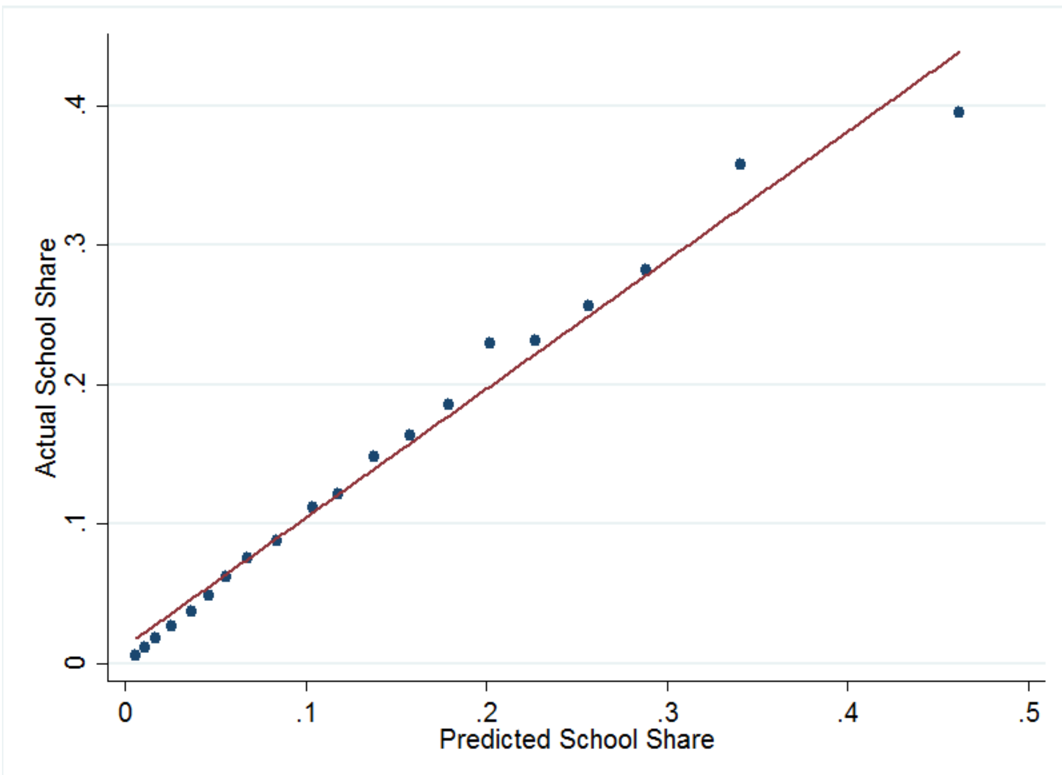
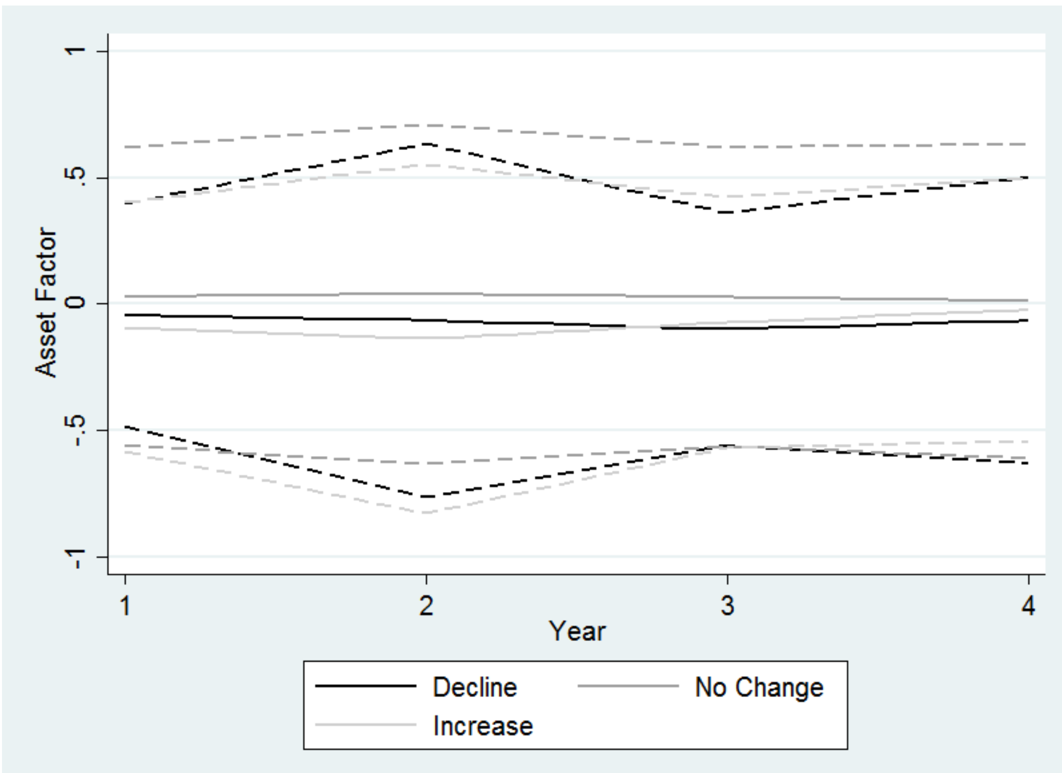


Figure 9: Time Trends in Assets by Change in the Number of Private Schools



11 Mathematical Appendix

11.1 Rational Inattention

A type x_i is unsure about the state $(x_i - s_k)$ if she attends school k. Define $\bar{s}$ to be the state and $\bar{b}$ to be x_i's action. Then, x_i solves the utility maximization problem:

$$\max_{p(\bar{b}|\bar{s})} \sum_{\bar{b},\bar{s}} -(\bar{s}^2)p(\bar{b}|\bar{s})p(\bar{s})$$

such that

$$\sum_{\bar{s},\bar{b}} p(\bar{s})p(\bar{b}|\bar{s})log(\bar{b}|\bar{s}) - \sum_{\bar{s},\bar{b}} p(\bar{b})log(p(\bar{b})) \leq \bar{I}$$

and

$$\sum_{\bar{b}} p(\bar{b}|\bar{s}) = 1 \; \forall \bar{s}.$$

The interpretation of the first constraint is that the reduction of entropy of x_i's knowledge of the state is bounded (the signal can only increase knowledge of a state at most by $\bar{I}$). The interpretation of the second constraint is that the probability distribution of states must sum to 1. Then, the lagrangian is:

$$\mathcal{L} = -\sum_{\bar{s},\bar{b}} p(\bar{s})p(\bar{b}|\bar{s})\bar{s}^2 - \theta \left\{ \sum_{\bar{s},\bar{b}} p(\bar{s}p(\bar{b}|\bar{s})log(p(\bar{b}|\bar{s}) - \sum_{\bar{b}} (\sum_{\bar{s}} p(\bar{b}|\bar{s})p(\bar{s}))log[\sum_{\bar{s}} p(\bar{a}|\bar{s})p(\bar{s})] \right\}$$

$$+ \sum_{\bar{s}} p(\bar{s})\psi(\bar{s}) \sum_{\bar{b}} p(\bar{b}|\bar{s}) - 1),$$

where θ and $\psi(\bar{s})$ are the lagrange multipliers. Taking the first order condition of the lagrangian gives:

$$\frac{\partial \mathcal{L}}{\partial p(\bar{b}|\bar{s})} = -p(\bar{s})\bar{s}^2 - \theta p(\bar{s})log(\bar{b}|\bar{s}) - \theta p(\bar{s}) + \theta[p(\bar{s})log(p(\bar{b}) + p(\bar{s})] + \psi(\bar{s})p(\bar{s}) = 0.$$

Re-arranging leads to the relationship:

$$p(\bar{b}|\bar{s}) = exp(-\frac{\bar{s}^2}{\theta})p(\bar{b})exp(-\frac{\psi(\bar{s})}{\theta}).$$

Substituting this back into the constraint $\sum_{\bar{b}} p(\bar{b}|\bar{s}) = 1$ leads to:

$$\sum_{\bar{b}} exp(\frac{-\bar{s}^2}{\theta})p(\bar{b})exp(\frac{-\psi(\bar{s})}{\theta}) = 1.$$

Re-arranging this expression leads to the expression

$$exp(\frac{-\psi(\bar{s})}{\theta}) = \frac{1}{\sum_{\bar{b}} exp(-\frac{\bar{s}^2}{\theta})p(\bar{b})}.$$

Substituting this back into the first order condition produces:

$$p(\bar{b}|\bar{s}) = \frac{exp(\frac{-\bar{s}^2}{\theta})p(\bar{b})}{\sum_{\bar{b}} exp(\frac{-\bar{s}^2}{\theta})p(\bar{b})}.$$

Note that $p(\bar{b})$ is the marginal probability of selecting each action without any information. We can choose "uniformed" priors to make the action probability uniform. In this case, the $p(\bar{b})$ terms cancel. Recall that θ is the lagrangian multiplier of the attention constraint; it represents the shadow price of attention. Letting θ for a type i be $\frac{1}{\delta_i}$ produces the same relationship for the shares as in section 2.

12 Appendix Tables

Table A1: Coefficients from Discrete Choice Model with Unobservable Characteristics in Restricted Sample

	(1) Parameter	(2) Standard Error
$\overline{VA_{low,s}} \times \mathbb{1}_{high}$	-0.1679*	0.0923
$\overline{VA_{high,s}} \times \mathbb{1}_{high}$	0.5006***	0.0498
$\overline{VA_{low,s}}$	-0.0003	0.0579
$\overline{VA_{high,s}}$	-0.0707	0.1180
$\overline{fee_{s,t}} \times \mathbb{1}_{high}$	0.4349***	0.0440
$\overline{fee_{s,t}}$	-0.8060***	0.0713
$distance_{is} \times \mathbb{1}_{high}$	-0.9254***	0.0405
$distance_{is} \times \mathbb{1}_{low}$	-0.9407***	0.0338
$\mathbb{1}_{school_{it}=school_{i,t-1}}$	4.6806***	0.1654
$\mathbb{1}_{boy_i} \times \mathbb{1}_{girls\_school_s}$	-2.7408***	0.0931
$\mathbb{1}_{girl_i} \times \mathbb{1}_{boys\_school_s}$	-2.7188***	0.0809

This table presents parameter estimates from a discrete choice model with unobservable school quality. The individual-by-school level parameters are estimated using maximum likelihood. The school-level parameters are estimated using generalized method of the moments. Fees are normalized by year and are in 1000s of ruppees units. The model also includes a set of interactions between a student's age and whether a school's highest grade is at the primary, secondary, upper secondary, or greater level. The sample is restricted to students who were not used to calculate school's type-specific value-addeds.

References

Andrabi, T., N. Bau, J. Das, and A. Khwaja (2010). Are bad public schools are public bads? Test scores and civic values in public and private schools. *Working Paper*.

Andrabi, T., J. Das, and A. Khwaja (2008). A dime a day: The possibilites and limits of private schooling in Pakistan. *Comparative Education Review 52*(3), 329–355.

Andrabi, T., J. Das, and A. Khwaja (2011). What do you do all day? maternal education and child outcomes. *Journal of Human Resources 47*(4), 873–912.

Andrabi, T., J. Das, and A. Khwaja (2013a). Students today, teachers tomorrow: Identifying constraint on the provision of education. *Journal of Public Economics 100*, 1–14.

Andrabi, T., J. Das, A. Khwaja, and T. Zajonc (2006). Religious school enrollment in Pakistan: A look at the data. *Comparative Education Review 50*(3), 446–477.

Andrabi, T., J. Das, and A. I. Khwaja (2013b). Report cards: The impact of providing school and child test scores on educational markets.

Berry, S., J. Levinsohn, and A. Pakes (2004). Estimating differentiated product demand systems from combined micro and macro data: The new car model. *Journal of Political Economy 112*(1), 68–105.

Carniero, P., J. Das, and H. Reis (2010). Estimating the demand for school attributes in Pakistan. *Working Paper*.

Chetty, R., J. Friedman, and J. Rockoff (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*.

Das, J. and T. Zajonc (2010). India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement. *Journal of Development Economics 92*(2).

De Palma, A., V. Ginsburgh, Y. Y. Papageorgiou, and J.-F. Thisse (1985). The principle of minimum differentiation holds under sufficient heterogeneity. *Econometrica: Journal of the Econometric Society*, 767–781.

De Palma, A., V. Ginsburgh, and J.-F. Thisse (1987). On existence of location equilibria in the 3-firm hotelling problem. *The Journal of Industrial Economics*, 245–252.

Desai, S., A. Dubey, R. Vanneman, and R. Banerji (2008). Private schooling in India: A new educational landscape. *India Human Development Survey Working Paper No. 11*.

Duflo, E., P. Dupas, and M. Kremer (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American Economic Review 101*, 1739–1774.

Eaton, C. and R. Lipsey (1975). The principle of minimum differentiation reconsidered: some new developments in the theory of spatial competition. *The Review of Economic Studies 42*(1), 27–49.

Filmer, D. and L. H. Pritchett (2001). Estimating wealth effects without expenditure dataor tears: An application to educational enrollments in states of india*. *Demography 38*(1), 115–132.

Friedman, M. (1962). *Capitalism and freedom.* University of Chicago press.

Fryer, R. and S. Levitt (2004). Falling behind: New evidence on the black-white achievement gap. *Education Next 4*(4), 64–71.

Glewwe, P., M. Kremer, and S. Moulin (2009). Many children left behind? Textbooks and test scores in Kenya. *American Economic Journal: Applied Economics 1*(1), 112–135.

Hoxby, C. M. (2000). Does competition among public schools benefit students and taxpayers? *American Economic Review*, 1209–1238.

Kane, T. J. and D. O. Staiger (2008). Estimating teacher impacts on student achievement: An experimental evaluation. Technical report, National Bureau of Economic Research.

Kremer, M., C. Brannen, and R. Glennerster (2013). The challenge of education and learning in the developing world. *Science 340*(6130), 297–300.

Mullainathan, S. and E. Shafir (2013). *Scarcity: Why Having Too Little Means So Much.* Times Books.

Muralidharan, K. and V. Sundararaman (2013). The aggregate effect of school choice: Evidence from a two-stage experiment in india. *NBER Working Paper No. 19441*.

Neilson, C. (2013). Targeted vouchers, competition among schools, and the academic achievement of poor students.

Pratham (2012). Annual status of education report.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain (2005). Teachers, schools, and academic achievement. *Econometrica 73*(2), 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 247–252.

Sims, C. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

World Bank Development Indicators (2014). http://data.worldbank.org/indicator/SE.PRM.PRIV.ZS. Accessed: 2014-06-19.