DISCUSSION PAPER SERIES

DP13924

ESTIMATING AN EQUILIBRIUM MODEL OF HORIZONTAL COMPETITION IN EDUCATION

Natalie Bau

DEVELOPMENT ECONOMICS, INDUSTRIAL ORGANIZATION AND LABOUR ECONOMICS

ISSN 0265-8003

ESTIMATING AN EQUILIBRIUM MODEL OF HORIZONTAL COMPETITION IN EDUCATION

Natalie Bau

Discussion Paper DP13924
Published 11 August 2019
Submitted 08 August 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

ESTIMATING AN EQUILIBRIUM MODEL OF HORIZONTAL COMPETITION IN EDUCATION

Abstract

The quality of the match between students and schools affects learning but little is known about the magnitude of these effects or how they respond to changes in market structure. I develop a quantitative equilibrium model of school competition with horizontal competition in match quality. I estimate the model using data from Pakistan, a country with high private enrollment, and (1) quantify the importance of good matches, (2) show that profit-maximizing private schools' choices of quality advantage wealthier students, increasing inequality and reducing welfare and learning, and (3) provide intuition for when interventions in the market are valuable. I find that match matters: moving a student from her worst to best match school doubles yearly average test score gains. Setting match-specific quality socially optimally in private schools would greatly reduce inequality in learning between rich and poor students while increasing learning and welfare. These positive effects are amplified when students' enrollment decisions are more responsive to quality.

JEL Classification: I2, L1, O1

Keywords: school competition, horizontal quality in education, structural models of education markets

Natalie Bau - nbau@ucla.edu
UCLA and CEPR

Estimating an Equilibrium Model of Horizontal Competition in Education*

Natalie Bau†

August 8, 2019

Abstract

The quality of the match between students and schools affects learning but little is known about the magnitude of these effects or how they respond to changes in market structure. I develop a quantitative equilibrium model of school competition with horizontal competition in match quality. I estimate the model using data from Pakistan, a country with high private enrollment, and (1) quantify the importance of good matches, (2) show that profit-maximizing private schools' choices of quality advantage wealthier students, increasing inequality and reducing welfare and learning, and (3) provide intuition for when interventions in the market are valuable. I find that match matters: moving a student from her worst to best match school doubles yearly average test score gains. Setting match-specific quality socially optimally in private schools would greatly reduce inequality in learning between rich and poor students while increasing learning and welfare. These positive effects are amplified when students' enrollment decisions are more responsive to quality.

*I am grateful to Victor Aguirregabiria, Chris Avery, Nikhil Agarwal, Oriana Bandiera, David Baqaee, Emmanuel Farhi, Roland Fryer, Asim Khwaja, Kala Krishna, Rob McMillan, Nathan Nunn, Mar Reguant, Jesse Shapiro, Chad Syverson, and Ali Yurukoglu, as well as seminar participants at Chicago Booth, Oxford, CMU, Rochester, UPenn, Columbia, LSE, Penn State, UC San Diego, UToronto, Indiana University, UIUC, the FTC, and conference participants at BREAD, ThReD, SITE, and the Banff Empirical Micro Workshop for helpful comments. I gratefully acknowledge the support of the CIFAR Azrieli Global Scholarship.

†UCLA, CEPR, and CEGA, *contact: nbau@ucla.edu.*

1 Introduction

A student who knows calculus will not benefit from remedial math instruction, no matter how well taught. Similarly, a student who has never learned fractions is unlikely to benefit from calculus. A growing literature shows that the match between a student's instructional needs and a school's instructional level is an important determinant of learning (Arcidiacono et al., 2011; Aucejo, 2011; Arcidiacono et al., 2016; Muralidharan et al., 2019). Given the importance of this match, instructional level may be misallocated, with students attending schools whose instruction is not well-matched to their needs. Yet little is known about the extent of this misallocation, how schools choose their instructional level, and how their choices of instructional level respond to competition.

Indeed, competition may *increase* the misallocation of instructional level. A classic result in industrial organization by Spence (1975) shows that, in the presence of imperfect competition, firms will respond to the marginal consumer rather than inframarginal consumers when choosing their quality. Thus, private schools do not typically provide the socially optimal quality, and if increased competition causes the marginal consumer's preferences to be further from those of the average consumer, competition may reduce welfare. Given extremely high rates of private schooling, low levels of learning,[1] and the frequency of a single teacher teaching all the students in a grade in low-income countries, understanding how instructional match interacts with private schools' competitive incentives may be crucial for improving students' outcomes in these contexts. The increasing popularity of policies that incentivize school competition in low-income countries[2] further underscores the importance of understanding the interaction between match-specific quality and competition.

To evaluate the importance of instructional match and how it interacts with competition, I develop a novel, structural model. In this model, private schools choose their instructional levels strategically to maximize profits in response to competitive incentives, rich and poor students have different optimal instructional levels, and students choose schools based on their characteristics to maximize their perceived utilities. This approach draws on insights from a literature in industrial organization focusing on how market structure affects what products firms offer (e.g. Mazzeo (2002); Crawford and Shum (2006); Draganska et al. (2009); Crawford and Yurukoglu (2012); Fan (2013); Wollmann (2016)). I focus on the heterogeneity between richer and poorer students because wealth is among the strongest determinants of educational outcomes in low-income countries, alongside which school a student attends

[1] In Pakistan, the context of this study, Das et al. (2006) find that only 31% of third grade students can correctly form a sentence in Urdu (the vernacular) containing the word "school."

[2] For example, in 2009, India passed the Right to Education Act in India, which requires that private schools set aside 25% of their seats for poor students with vouchers. Similarly, Colombia, Chile, and Pakistan have all experimented with large-scale voucher programs.

(Das et al., 2006), and the school choice behavior of wealthier families differs from that of poorer families throughout the world.[3] I estimate both the demand and supply side of this model using data from a particularly relevant context, Pakistan's rural private schooling market. As in many low-income countries, private school enrollment in Pakistan is large and fast-growing.[4]

In line with Spence (1975), estimates show that private schools do not choose instructional levels that maximize either learning or utility. On the demand side, rich students are more responsive to match-specific school quality, as measured by their predicted test score gains in a school, when they make enrollment decisions. Thus, richer students are "more marginal." The average private school chooses an instructional level that strongly favors richer students at the expense of poorer students, even within the population that attends private schools. Furthermore, the structural estimates show that, on average, the entry of an additional private school makes wealthier students relatively more marginal, leading schools to move even closer to their optimal instructional level and exacerbating inequality in learning.

Exploiting the exit and entry of private schools over time provides a reduced-form motivation for and test of the structural model. Consistent with the structural model, I find that (1) private school entry increases inequality in test scores in the private sector, (2) this is driven by an increase in inequality *within* private schools, and (3) the increase in the inequality of test score gains is mainly due to a reduction in poor students' learning.

Estimating the supply-side of the model also allows me to identify the extent to which a student's test score gains in a school are due to match-specific quality and to what extent they are due to vertical quality (the element of quality that is the same for all students). Match is important. A school can increase test scores for poor students by as much as 0.46 s.d. (equivalent to the test score gains from 1 year of education) by moving from wealthy students' optimal instructional level to poor students' optimal instructional level.[5] Importantly, examining the correlation between the outcomes of poor and rich students within schools would not provide enough information to estimate this parameter. This is because the observed correlation between students' outcomes is affected by both the importance of instructional match for learning and schools' equilibrium choices of instructional level. If all

[3]Bayer et al. (2007) show that higher income households have a higher willingness to pay to live in neighborhoods where schools have higher average test scores, and Ajayi (2011) and Hoxby and Avery (2012) show that poor students make less sophisticated decisions when they choose schools. Relatedly, Dizon-Ross (forthcoming) shows that poor parents have less information about their children's academic performance than rich parents, and Kapor et al. (2017) show that lower income parents' have less accurate beliefs about admissions' probabilities in the New Haven school district.

[4]In rural Pakistan, 35% of students are enrolled in private schools (Andrabi et al., 2006). In rural India, private school enrollment is 28% (Pratham, 2012) and in urban India, it is 65% (Desai et al., 2008). In Sub-Saharan Africa, between 10-54% of primary school children were enrolled in private schools in 2012 (World Bank Development Indicators, 2014).

[5]This estimate is in line with the effects of a randomized interventions that improve the targeting of instructional levels in Kenya (Duflo et al., 2011) and India (Muralidharan et al., 2019).

schools choose an instructional level equidistant between the optimal instructional levels of rich and poor students, school quality for rich and poor students will be highly correlated even if instructional match is an important input in learning.[6]

Finally, to characterize the welfare and learning losses due to the Spence distortion and provide guidance for policies that improve the allocation of match-specific quality, I estimate a series of partial equilibrium counterfactuals. In these counterfactuals, the social planner chooses schools' instructional levels to maximize welfare but their other characteristics remain the same. Removing the Spence distortion alone moderately increases learning and welfare and greatly decreases inequality in learning due to school quality.

However, the benefits of reallocating match-specific quality are limited by the fact that students are not very responsive to quality when they choose schools. According to the demand estimates, even rich students are only willing to pay 1% of GDP per capita to move to school that would lead to 1 s.d. higher test score gains (equivalent to 2.5 additional years of schooling). Thus, a private school's population typically includes children with disparate instructional needs, limiting the social planner's ability to match students' with their optimal instructional level. However, low estimated responsiveness to quality doesn't necessarily mean that families do not value quality. Rather, households could have little information about quality, consistent with Andrabi et al. (2017), who show that households in this context update their beliefs about school quality when they are provided with information on test scores.

Counterfactuals where *both* rich and poor households are more responsive to quality lead to larger gains. For example, when both types are willing to pay 2% of GDP per capita for a school with a 1 test score s.d. higher value-added, the welfare of students with at least a 25% chance of attending private school increases by 7%, and their yearly test score gains increase by 0.09 s.d. (a 23% increase). Importantly, since schools' vertical quality is fixed in the counterfactuals, these gains come entirely from changing schools' match with students. Thus, to the extent that low responsiveness to quality is due to a lack of information, providing information that improves matching can magnify the effects of reducing the misallocation of match-specific quality.

Overall, this paper makes several contributions. First, it quantifies the effect of match-specific school quality in the learning production function and show that it is large. Second, it characterizes the distribution of private schools' choices of match-specific quality, which are a result of market structure, and shows that private schools choose instructional levels that strongly favor the rich half of the private school population at the expense of the poor

[6]Because equilibrium choices of instructional level may differ across markets with different market structures, these correlations could be very different across markets, even if match-specific quality enters the learning production function in the same way in every market.

half. Coupled with the importance of match-specific quality for learning, this leads to welfare losses, reduced learning, and greater inequality. In light of recent research documenting the low level of learning in low-income countries, interventions that leverage the importance of match-specific quality and reduce the misallocation of instructional level may provide a tool to improve students' outcomes. The counterfactual exercises suggest that this is especially true when these changes in instructional level are paired with interventions that improve students' sorting into better-match schools. Third, beyond showing that private schools choose instructional levels that favor wealthy students, I show that schools' incentives to do so are intensified by increased competition. These results highlight the importance of measuring within-school differences in school quality, as well as cross-school differences, when measuring the effects of school competition on learning.

This paper relates to two literatures. First, it contributes to a growing literature that estimates structural models of demand (Carneiro et al. (2016) in Pakistan, Neilson (2017) in Chile, Abdulkadiroğlu et al. (forthcoming) in NYC, Agarwal and Somaini (2018) in Cambridge, Walters (2018) among charter schools in Boston, Bayer et al. (2007) in San Francisco, Hastings et al. (2005) in North Carolina, and Dinerstein and Smith (2014) in New York) and supply (Dinerstein and Smith (2014) in New York and Singleton (2019) in Florida) in education. On the demand-side, I contribute by allowing a school's quality to be different for different students. On the supply-side, I contribute by modeling schools' endogenous choices of match-specific quality in response to student demand. Furthermore, the new framework I develop can be applied to other educational markets to estimate the importance of match-specific quality and characterize schools' choices of quality.

Second, this paper contributes to a literature on the effects of private schooling on students' outcomes in a variety of settings – Neal (1997) in the United States, Andrabi et al. (2010) in Pakistan, Muralidharan and Sundararaman (2015) and Singh (2015) in India, and Neilson (2017) and Hsieh and Urquiola (2006) in Chile. Researchers working in the same (Andrabi et al., 2010) or similar (Muralidharan and Sundararaman, 2015) South Asian contexts have demonstrated that private schools are more productive than public schools, delivering as great or greater levels of learning at lower costs. I build on this important result by showing how market structure can affect the distribution of learning *within* private schools.

The rest of the paper is organized as follows. Section 2 discusses the context and the data. Section 3 provides motivating evidence on the relationship between private school entry and test score inequality. Section 4 develops the equilibrium model of student demand and school supply, and Section 5 structurally estimates the model. Section 6 uses reduced-form techniques to test predictions from the structural estimation. Section 7 simulates the counterfactuals, and Section 8 discusses the results and concludes.

2 Context and Data

2.1 Context

Pakistan is a natural setting to study the private schooling market in low-income countries. Like many low-income countries, Pakistan has experienced a rapid increase in low-cost, secular private schooling over the past two decades. Pakistani private schools are virtually unregulated and 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.

Private schools do not merely cater to the wealthy in Pakistan. Andrabi et al. (2008) show that grassroots, rural private schools are even affordable for day laborers, noting that the average school charges the equivalent of "a dime a day." However, private schools are more expensive than public primary schools, which are free. Private schools can afford to charge so little because they typically spend less per student than public schools, largely because public school teachers earn about five times as much as private school teachers. In rural Pakistan, there is almost no overlap in the public and private school wage distributions (Bau and Das, forthcoming). Thus, public and private schools do not compete to hire teachers from the same labor market, and unlike private schools, public schools are typically constrained to hire teachers who have completed post-secondary degrees.[7] Because the public and private sector hire from different labor markets and because public teacher hiring is centralized at the district level, changes in the local private school market are unlikely to affect public schools through the market for teachers.

Unlike private schools, public 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. Public and private schools also differ in that public schools are single-sex during the time period of the study, while most private schools are co-educational.

An advantage of studying private schooling in rural 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; Carneiro et al., 2016). Therefore, I can examine how com-

[7]On average, public school teachers have more education and more teaching training than private school teachers (Andrabi et al., 2008). However, Bau and Das (forthcoming) show that these characteristics do not correlate with teacher value-added estimates. These results suggest that though public school teachers are more qualified, they are not necessarily more effective.

petition affects equilibrium school and student outcomes at the market level. This is key for understanding how market structure affects schools' choices of match-specific quality, since each village-year observation provides a measure of market structure with no spillovers onto other villages.

The match between schools and students is also likely to be a particularly important determinant of students' learning in rural Pakistan because schools' ability to cater to specific students' instructional needs is limited. Both public and private schools have less than 1 teacher per grade on average (one teacher teaches 1.11 grades in the median private school). So, within-grade tracking is rare and cross-grade mixing within the same class is common. In this context, students requiring very different instructional levels are likely to be present in the same classroom.

Finally, neither public nor private schools in rural Pakistan typically have binding capacity constraints. The majority of schools in 2007 had some form of admissions procedure (97% of private schools and 95% of public schools), which frequently consisted of an oral exam (81% of private schools, 54% of public schools) or the perusal of previous school reports (14% of private schools and 33% of public schools). However, even if a student was deemed "weak" based on this assessment, only 11% of private schools and 3% of public schools said they would refuse admission. Based on this institutional feature, I ignore capacity constraints in my model and estimation.

2.2 Data

For this paper, I use 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 the 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 province. The LEAPS data allow for the construction of two partially overlapping samples containing data on household wealth: a sample of tested students surveyed in schools, which I refer to as the *tested sample*, and a sample of children whose parents were surveyed in their geocoded homes, which I refer to as the *household sample*. In addition to these samples, a third survey of head teachers provides data on schools' characteristics and geo-locations.

Both the household and tested student samples are necessary for the estimation of the structural model. The sample of tested students allows for the measurement of school quality, a key determinant of students' enrollment decisions, in terms of test score gains. The geocoded household sample allows for the estimation of the determinants of school choice behavior in a sample for which the distance between every school and every student is known. Since distance is a major determinant of school choice and the characteristics of students' closer schools may differ systematically by their wealth, allowing distance to affect students' enrollment decisions is key for taking into account systematic differences in poor and wealthy families' choice sets.

Tested Sample. Surveyors visited the universe of public and private schools within a 15 minute walk of the village and collected geocoded data. The initial sample included 823 schools in the first round (2003-2004), and additional schools were added to the sample as they entered the market. In the first round of data collection, all third graders in a school were tested using low-stakes tests administered by the enumerators in math, Urdu, and English. These students were followed and tested in subsequent years, and an additional sample of third graders was added in 2005-2006 and tested subsequently in 2006-2007. 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. In all, the panel includes 71,167 student-year test score observations (31,382 unique students), which comprise the tested sample.

A random sub-sample of students were also given an additional survey on their household assets, leading to a sample of 28,449 student-year observations for which both test score and survey data is available. A smaller sub-sample of the tested sample (1,269 households) can also be matched to a household survey administered to a panel of 1,740 households in each year, which is described below.

Household Sample. 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 3 in the first survey round, and 4 were chosen among households where a child eligible to be enrolled in grade 3 was not enrolled.[8] In round 1, the location of households was geocoded. As a result, I can compute the geographic distance between a household and the schools in the village for the 7,209 children aged 5-15 who appear in the household survey and comprise the household sample. Households were asked about their enrollment decisions for each of the children in each year and about their asset ownership, although the asset questionnaire differed slightly from the asset questionnaire administered to the children in the tested sample.

[8] This matches the fact that 25% of third grade-aged children are not enrolled in school.

Summary Statistics. Appendix Table A1 presents summary statistics for the four rounds of LEAPS data at the school level. Because of the panel structure of the data, each observation is a school-year. As Appendix Table A1 shows, there is substantial variation across schools in terms of facilities, and private schools generally have better facilities. Appendix Table A2 summarizes the characteristics of the children in the tested sample; 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. Finally, Appendix Table A3 reports summary statistics for the household sample, which, unlike the tested sample, includes students who are not enrolled.

3 Motivating Evidence on Competition and Inequality

In this section, to motivate the equilibrium model of educational markets, I estimate the effect of changes in market structure due to school exit and entry on students' test scores, focusing on students in the private sector. I first discuss why entry events occur. Then, I show that private school entry has no effect on mean test scores in the public or private sectors but increases the variance of test scores in the private sector. This in turn is driven by increases in within-school inequality, indicating that match-specific quality matters and responds to competition. Finally, I verify that these results do not appear to be driven by pre-trends or omitted variable bias.

Why Does Entry Occur? I argue that private school exit and entry events, while likely correlated with village characteristics, are plausibly uncorrelated with village or school-level time trends. Importantly, the LEAPs data were collected during a period of rapid private schooling expansion, during which entry was occurring in many markets.[9] During this expansion, private schools may have entered villages that were already large but not necessarily faster-growing. To examine if this is the case, I use data on village-level populations from the 1998 and 1981 population censuses of Punjab. In columns 1 and 2 of Appendix Table A4, I regress the village populations in 1998 and 1981 on the number of private schools and find that population is positively correlated with the number of private schools. However, column 3 shows that the percent change in population between 1981 and 1998 is not significantly correlated with the number of private schools. Thus, entry events did occur in larger villages, but including village or school fixed effects in my estimating equations accounts for these level differences.

[9]From 1991-2001, private school enrollment in Punjab increased from 15% to 30%. While data from later years is not readily available, the entry of private schools into the market showed no signs of plateauing by 2000 (Andrabi et al., 2008).

Effects on Test Scores Gains by Sector. Using the tested sample, I first estimate the effect of private school entry on test score gains in both the private and public sectors using the following difference-in-differences style regression

$$y_{igvt} = \rho_0 + \rho_1 num\_pri_{vt} + \alpha_t + \psi_v + \lambda_g y_{i,t-1} + \phi_g y_{i,t-1}^2 + \mathbf{\Gamma X_{vt}} + \epsilon_{igvt},$$

where i indexes students, g indexes grades, v indexes villages, and t indexes years. y_{igvt} is then a student-year level test score in math, Urdu, or English, num_pri_{vt} is the number of private schools in village v at time t, α_t are survey year fixed effects, ω_g are grade fixed effects, ψ_v are village fixed effects, and $\mathbf{X_{vt}}$ is the set of village-year controls.[10] λ_g and ϕ_g are grade specific coefficients on test scores and lagged test scores. The choice of specification is motivated by the specification for estimating teacher value-added in Chetty et al. (2014a). Like Chetty et al. (2014a), I control for grade-specific effects of lagged test scores to account for selection into particular schools. In this case, these controls account for the correlation between student selection into a sector and private school entry and exit. Standard errors are clustered at the village-level.

Table 1 reports the results of these regressions for the public and private school populations. The results suggest that the degree of competition in a market does not affect the average student's test scores in either sector. However, if competition has heterogeneous effects, individual students' outcomes may change even if on average students' test scores remain the same.

Effects on Inequality by Sector. To see if the null average effect masks heterogeneous effects, I estimate the effect of exit and entry events on inequality in the private sector using the specification:

$$inequality_{vt} = \rho_0 + \rho_1 num\_pri_{vt} + \alpha_t + \psi_v + \mathbf{\Gamma X_{vt}} + \epsilon_{vt}, \tag{1}$$

where $inequality_{vt}$ is a village-year level measure of inequality in test scores in the private or public sector, given by the variance of test scores in the sector. Now, an observation is at the village-year level rather than the student-year level.

Table 2 reports the effect of private school entry on the variance of test scores in the public and private sectors. Outcomes within the public schools again appear to be unaffected by competition. In contrast, the entry of a private school increases the variance in private sector test scores by a statistically significant 0.08 test score standard deviations. To show that this

[10] These controls consist of a control for whether the village was treated in a village-level randomized report card intervention, which supplied information on schools' average test scores and student's test scores, studied by Andrabi et al. (2017). I control for this intervention throughout and allow its effects to vary by year.

pattern is robust to the measure of inequality, Appendix Table A5 investigates the effect of entry on the gap between students' test scores at the 90th and 10th percentile. Inequality again widens with private school entry.

Parallel Trends Assumption. One potential concern is that the results in Table 2 are driven by pre-existing time trends. Figure 1 provides evidence that this is not the case. The figure plots the γ coefficients and their 95% confidence intervals from the regression

$$\begin{aligned} inequality_{vt} =& \rho_0 + \gamma_1 event_{v,s-3} + \gamma_2 event_{v,s-2} + \gamma_3 event_{v,0} + \gamma_4 event_{v,s+1} + \gamma_5 event_{v,s+2} \\ &+ \gamma_6 event_{v,s+3} + \alpha_t + \psi_v + \mathbf{\Gamma X_{vt}} + \epsilon_{vt}, \end{aligned} \tag{2}$$

for the private sector, where s denotes the years since an exit or entry event occurred, the *event* variables takes the value 1 if the event was an entry event, -1 if it was an exit event, and 0 if no event occurred, and the outcome variable is the variance in test scores. In villages with multiple events, the first event is used for the coding.[11] The event study graph shows that, prior to an entry event, there is no increase in inequality in test scores in villages where entry events occurred. Following the event, there is a significant increase in inequality.

The event study graph suggests that the growth in inequality is not driven by pre-trends in inequality and test scores. As a further check, in Appendix Table A4, I also estimate the association between private school exit and entry and other village characteristics that may affect inequality in test scores. I first regress a village-year measure of village wealth on num_pri_{vt}, controlling for village and year fixed effects.[12] The coefficient for number of private schools (0.017) is small and statistically insignificant (see column 4). Using the Gini coefficient for wealth as the outcome variable or the percent of households that own land in a village-year yields similar results (columns 5 and 6). The coefficients are again small (0.001 and -0.001 respectively) and insignificant. Thus, there is no strong relationship between the number of private schools in a village and village-level trends in wealth or inequality.

What Drives the Increase in Inequality? Finally, in Table 3, I investigate the drivers of the increase in test score inequality in the private sector. Specifically, I examine whether the increase in inequality is due to changes in school composition or due to changes in match-specific school quality. To test whether the increase is driven by a change in the composition

[11] As entry and exit events occur in different years in different villages, the coefficients for +2 and -3 years are identified from the relatively smaller set of events. It is impossible for an event to occur in round 1 of the data collection, and villages with events in rounds 3 and 4 are not observed 2 years after the event.

[12] To create this measure, I follow Filmer and Pritchett (2001) and predict the first principal component from a principal components analysis over indicator variables for asset ownership for the different assets in the household survey (see Appendix Table A2 for a list of assets).

of the private sector or a change in the composition of the private schools due to entry, I re-estimate equation (1) using an alternative sample to calculate the variance in test scores. The sample is now restricted to private school students who never change schools *and* were observed in a private school prior to the entry event. This ensures that students switching into the private sector or between schools are not driving the effects on inequality. As the odd columns of Table 3 show, measuring inequality in this way only strengthens the association between entry and inequality.

The even columns of Table 3 investigate two other possible drivers of the growth in inequality. In these columns, I recalculate the variance in test scores at the *school-year* level instead of the sector-year level and use this as the outcome variable in equation (1), controlling for school fixed effects instead of village fixed effects. If the growth in learning inequality is driven by a growth in inequality *across* private schools, the number of private schools should not affect this outcome. Additionally, I control for two conventional measures of peer effects – the average of the lagged test scores of the students in a school and the variance in the lagged test scores – as well as student-teacher ratios, which vary across years. As the even columns of Table 3 show, I find that the growth in the inequality in test scores is mainly driven by growth in *within-school* inequality, and the peer effects controls do not account for the growth in inequality. A new private school increases within-school test score variance by 0.061 s.d. (column 8). This specification also suggests that the growth in inequality is driven by the pre-existing schools in the market responding to competition rather than the quality of the new school. Because the specification controls for school fixed effects and entering schools are not observed prior to the entry event, variation in the entering schools' students' test scores associated with market structure is absorbed by the school fixed effects.

The fact that private school entry increases test score inequality in the private sector and that this increase is driven by increasing within-school inequality suggests that match-specific quality matters and that it reacts to market structure. Additionally, the fact that mean test scores in the private sector do not improve implies that schools are not just improving with some faster-learning students benefiting more from these improvements than others. Rather, some students' learning is increasing while others' is declining.[13] Motivated by these results, in the next section, I develop an equilibrium structural model to better understand the importance of match-specific quality and test whether strategic incentives can help explain these reduced-form findings.

[13]In Section 6, I directly verify that this is the case.

4 Equilibrium Model

In this section, I develop a structural model of students sorting into schools and schools choosing their quality to test whether increased competition indeed changes private schools' incentives. Based on the evidence in the previous section, I allow schools' effects on students' test scores to have a match-specific, as well as a vertical component. I also allow students to belong to one of two types, depending on their wealth. I choose wealth as my key source of heterogeneity since it is the major, non-school driver of students' learning (Das et al., 2006) and is a well-known determinant of school choice behavior in the literature. Additionally, Appendix A verifies that wealthy parents are more knowledgeable about the educational marketplace and self-report placing more weight on school quality when they make enrollment decisions.

To allow school quality to have both a vertical and match-specific component, I model poor students as having an optimal instructional level of 0 on a unit line, while rich students have an optimal instructional level of 1. School j chooses both its location on the line ($h_j \in [0, 1]$) and its vertical quality, v_j. Then, if a school chooses h_j, a poor student's learning (or test score gains) in the school will be

$$VA_{j,poor} = v_j - \beta(o_{poor} - h_j)^2 \tag{3}$$

and a rich student's learning will be

$$VA_{j,rich} = v_j - \beta(o_{rich} - h_j)^2, \tag{4}$$

where o_z is the optimal instructional level of a student of type $z \in \{poor, rich\}$. Here, β captures the importance of horizontal quality relative to vertical quality in the learning production function, and is a parameter of interest that I will be able to estimate in the next section. Importantly, the model does not assume that match-specific quality matters, and only vertical quality matters in the special case where $\beta = 0$.[14]

In the first subsection of this section, I specify student demand for schools as a function of students' types and schools' characteristics, and in the next subsection, I specify the school's problem. In the final subsection, I use a simplified version of the model to develop intuition for why increased competition can lead to greater inequality in learning within private schools.

[14]One limitation of this model is that it does not allow for direct peer effects in the learning production functions in equations (3) and (4). Allowing for direct peer effects substantially complicates the estimation of the model. Since the results in Table 3 show that conventional measures of peer effects do not account for the association between private school entry and inequality that motivates the model, I abstract away from them.

4.1 Student Decision Problem

A student i of type z chooses the school j that maximizes her perceived utility

$$u_{ijz} = \delta_z VA_{jz} + \mathbf{\Gamma_z} X_{ij} + \xi_j + \epsilon_{ij}, \tag{5}$$

where X_{ij} is the set of other student-school characteristics, such as fees and geographic distance, ξ_j is unobserved, type-invariant school-specific quality, and ϵ_{ij} is an idiosyncratic shock drawn from the type 1 extreme value distribution that captures both parents' misperception of school quality and idiosyncratic taste. $\mathbf{\Gamma_z}$ and δ_z are parameters that govern the effect of school quality and other school characteristics on a student's perceived utility. A student's choice set includes both private and public schools, and a student may also choose the outside option (not enrolling in school), which is normalized to have a utility of zero. If poor students are less responsive to school quality than rich students, as the descriptive statistics in Appendix A suggest, $\delta_{low} < \delta_{high}$.

Since ϵ_{ij} is drawn from the type 1 extreme value distribution, the probability that a student i of type z attends a school j can be written as

$$p_{ijz} = \frac{e^{\delta_z VA_{jz} + \mathbf{\Gamma_z} X_{ij} + \xi_j}}{\sum_k e^{\delta_z VA_{kz} + \mathbf{\Gamma_z} X_{ik} + \xi_k}},$$

where k indexes the other schools in a student's choice set. In practice, a student's choice set consists of the schools in her village and the outside option, but including characteristics like distance in the student's utility allows the attractiveness of different schools to depend on where the student lives within a village.

4.2 School Decision Problem

A private school j chooses its characteristics (including its fee, fee_j, and its vertical quality, v_j) simultaneously with other private schools in a static equilibrium to maximize its profits. The profits of a private school j are given by

$$\pi_j = fee_j \times s_j(v_j, h_j, fee_j, \mathbf{v}, \mathbf{h}, \mathbf{fee}) - c(v_j, h_j, s_j(v_j, h_j, fee_j, \mathbf{v}, \mathbf{h}, \mathbf{fee})),$$

where c is the school's cost function, $s_j = \sum_i p_{ijz}$ is the number of students attending j, and the bolded variables are vectors of these variables for the schools in a student's choice set.

Differentiating π_j with respect to h_j^*, the profit-maximizing value of h_j, yields

$$\frac{\partial \pi_j}{\partial h_j^*} = \left(fee_j - \frac{\partial c}{\partial s_j}\right)\frac{\partial s_j}{\partial h_j^*} - \frac{\partial c}{\partial h_j^*} = 0.$$

If we assume that changing a school's instructional level is costless and that c is weakly concave as a function of s_j,[15] for h_j^* to be profit-maximizing, it must be the case that $\frac{\partial s_j}{\partial h_j^*} = 0$. As we will see in Section 5, imposing this assumption makes the problem of identifying and estimating β and h_{jt}^* separable while making only limited assumptions about the shape of c.

In contrast to private schools, I do not make any assumptions about public schools' objective functions. Assuming that private schools are profit-maximizing is sufficient to estimate β and the equilibrium value of h_j for every private school in the market.

4.3 Equilibrium Behavior in a Simplified Model

In equilibrium, private schools choose the vector of characteristics that will maximize their profits and students choose the schools that will maximize their perceived utility. A simplified version of the model is useful to develop intuition for how new entry can incentivize the average private school to choose a h_j closer to 1, increasing learning for wealthy students and decreasing it for poor students. For simplicity, let $\mathbf{\Gamma_z}$ and ξ_j be 0, so that students only respond to horizontal quality, and fix fee_j. Also let h_j be the only characteristic that a school can choose and assume there is no outside option besides a passive public school. Then, a private school j maximizes s_j, and a student i of type z's utility in a school is

$$u_{ijz} = -\delta_z \beta (o_z - h_j)^2 + \epsilon_{ij}.$$

Additionally, assume that δ_{rich} is infinite so that high types always select their best-match option (as in the traditional Hotelling model), while $0 < \delta_{poor} < \infty$.[16] Finally, assume that there are equal numbers of rich and poor types in the market.

To see how competition can affect a private school's incentives, consider the cases with one ($N = 1$) or two ($N = 2$) private schools in the market, and a public school that gives utility $u_0 \leq 0$. The following two propositions characterize schools' equilibrium instructional levels in these two cases.

[15] The cost, c, will be weakly concave in the number of students if there are constant or increasing economies of scale in providing education. This is almost certainly the case since there are large fixed costs to providing education, such as hiring teachers and building a school.

[16] This model, in which students choose their best-match school subject to an idiosyncratic shock, relates to work on the Hotelling model by De Palma et al. (1987) and De Palma et al. (1985), who develop Hotelling models where consumers also have an idiosyncratic shock in their utility function. However, in this case, the relative importance of the shock is correlated with a student's location on the Hotelling line.

Proposition 4.1. *For N=1, there is a unique equilibrium where the single private school chooses* $h^* = \max(1 - (-u_o)^{1/2}, 0)$.

Proof. See Appendix B.

Proposition 4.2. *For N=2, if* δ_{low} *is sufficiently small, the unique equilibrium is* $(h_1, h_2) = (1, 1)$.

Proof. See Appendix B.

Comparing the equilibrium for $N = 1$ and $N = 2$ shows that the implications of this model are very different from a standard Hotelling model (e.g. Eaton and Lipsey, 1975). In standard Hotelling models, increasing the number of schools in the market typically leads to symmetric product differentiation. However, this is no longer the case when there is heterogeneous sorting by rich and poor students.[17]

When there is only one school in the market, it caters more to the poor students because it is more likely to lose them. As long as the rich students prefer the single school to the outside option, the school gains nothing by becoming more attractive to them. On the other hand, it gains poor students continuously as it moves h_1 closer to zero. However, when a second school enters, it changes the first school's incentives, in line with the intuition from Spence (1975). Now, it must compete aggressively to retain the rich students, and since they are more responsive to quality, competing for them has a higher payoff. As a result, in this simplified version of the model, rich students benefit from increasing competition while the poor students' learning falls, increasing inequality. As in Spence (1975), this change in schools' incentives due to competition does not necessarily make the average consumer better off (as measured by learning) because schools choose h_j in response to the marginal consumer rather than the average consumer. While the rich types' learning unambiguously increases, the poor types' decrease in learning can more than cancel out these increases in the average private school. Estimating the more general structural model will allow us to examine whether the same dynamics are at play in the real world.

5 Estimation of the Equilibrium Model

In this section, I estimate the parameters of the structural model. This section is divided into four parts. In the first part, I use students' asset data to determine which students are

[17]For simplicity, in this section, I have focused on cases where the pure strategy Nash equilibrium exists. When δ_{low} is not sufficiently low, in the special case of the simplified model where $\delta_{high} = \infty$, the pure strategy Nash equilibrium does not exist for $N = 2$. For a proof of this, see Appendix B. However, in the next section, where I estimate δ_{high} and δ_{low} and then compute the equilibrium for the full model, I do not impose any restrictions on δ_{high} and δ_{low}. In this setting, Schauder's fixed point theorem guarantees that there is an equilibrium in mixed strategies, and when I estimate the model, I estimate each school's equilibrium h_j directly, demonstrating that an equilibrium in h_j exists.

rich and poor types. In the second part, combining these estimates of students' types with panel data on students' test scores, I estimate the predicted test score gains of a student of type z in a school j, estimating VA_{jz} from the previous section. Using these estimates, in the third subsection, I estimate the parameters of students' demand for schools, given by equation (5) in the previous section. Finally, in the fourth subsection, using the estimates of the parameters of the demand system, I estimate β, the importance of match-specific quality, and schools' equilibrium choices of h_j. With these estimates in hand, I can (1) quantify the importance of the match between instructional level and a students' instructional needs in the learning production function, (2) characterize the distribution of instructional level in the private sector, and (3) test whether h_j moves closer to 1 when there are more private schools in a market, indicating that competition incentivizes schools to choose instructional levels that favor wealthy students on average.

5.1 Assignment of Rich and Poor Types

It is necessary to assign types to both the students in the data set of children who were surveyed in schools (tested sample) and to the children of parents who were surveyed in their homes (household sample). Recall that while these data sets overlap, most children in one data set are not in the other. The panel of tested students is needed for estimating schools' quality since it contains data on students test scores over time. The geocoded panel of households is needed for estimating the determinants of students' demand for schools since it contains data on the distance between every household and every school.

Assigning Types in the Tested Sample. Neither the LEAPS questionnaire administered to the tested sample nor the one administered to the household sample directly measures income. Even if the questionnaires had, measures of income are unlikely to measure socioeconomic status well in rural Pakistani villages where the majority of the population is engaged in agriculture. However, the LEAPS survey administered to tested children does ask "yes" or "no" questions about asset ownership for beds, radios, TVs, refrigerators, bicycles, ploughs, small agricultural tools, tables, chairs, fans, tractors, cattle, goats, chickens, watches, motor rickshaws, motorcycles, cars, telephones, and tubewells. To synthesize this data into a single latent wealth factor, I conduct a factor analysis of these assets. I first regress an indicator variable for ownership of each asset on year, child age, and village fixed effects and predict the residual from this regression. This helps ensure that my measure picks up differences in wealth within villages and years and is not simply picking up parents later in their life cycle. From the perspective of the schools in my model, which compete for students within a village, across-village or across-year differences in wealth are not relevant.

I then predict the first factor from this factor analysis, which explains 60% of the residual variation in asset ownership.

I use a simple rule to determine which students are rich or poor based on their value for the first factor. I code students as rich if their first factor value is above the median for the private school population and as poor otherwise. For children whose status changes across years, I assign them to be rich types if they are "rich" the majority of the time and poor types otherwise. Therefore, by construction, 50% of private school students in the tested data are poor types. Since private school students are wealthier than the general population, in the full data set, 69% of all students are poor types.[18]

Assigning Types in the Household Sample. Now, I also need to assign types in the household data, where test scores are not available for most children. The household survey asks about a slightly different set of assets than the tested child survey. For example, the household survey asks about VCR, gun, and thresher ownership, while the tested child survey does not. To generate types for this data set, I again regress the indicator variables for each of these assets on village, child age, and year fixed effects and create residualized asset ownership variables. Then, for the 1,269 children who appear in both data sets, I use a lasso logistic regression to select the covariates that are most predictive of their assigned types.[19] I use the estimated coefficients to predict the probability of being a rich type for all the children in the household sample. For the children who are assigned both a type from the first procedure and a probability of being a rich type from the second procedure, the two measures are strongly related with a correlation of 0.51. So, to summarize, in the tested sample, children are assigned to their type with certainty, while in the household sample, children are assigned a probability of being a rich type.

5.2 Estimation of Match-Specific School Quality

To measure match-specific quality for each type at a given school (VA_{jz}), I draw on the value-added literature in education economics (for example, see Rockoff (2004), Rivkin et al. (2005), Kane and Staiger (2008), Chetty et al. (2014a), and Chetty et al. (2014b)), to estimate the predicted test score gains to a rich or a poor type from attending a given school

[18]While, in principle, the model could be generalized to have more than two types, in the next section, I will use the type assignments to estimate each school's type-specific quality. To obtain these (non-parametric) estimates, I must observe sufficiently many students of each type in each school. Therefore, with more and more types, the estimates of type-specific school quality become less credible and these parameters may not be identified.

[19]Lasso regression is a technique for selecting the most predictive covariates when there are a large number of possible covariates. It selects the covariates that minimize the sum of squared residuals subject to an L1 regularization penalty. For more details, see Tibshirani (1996).

conditional on the number of competing schools. Hereafter, this measure is referred to as a school's type-specific value-added. I allow a school's value-added to vary over time with how many private schools are in the village since schools may be incentivized to change their instructional level when the number of private schools changes. To calculate a school's type-specific value-added, using the tested data, I estimate the following regression for math, English, and Urdu:

$$y_{ijgztn} = \tau_{gz,1} y_{i,t-1} + \tau_{gz,2} y_{i,t-1}^2 + \omega_{gz} + \alpha_{zt} + \eta_{zjn} + \mathbf{\Gamma X_{vzt}} + \epsilon_{ijgztn} \quad (6)$$

where y_{ijgztn} is the outcome variable consisting of normalized test scores in math, English, or Urdu, i indexes an individual, g indexes a grade, z indexes a type, j indexes a school, n indexes the number of private primary schools in the market place, and t indexes a year. $\tau_{gz,1}$ is a grade-type specific coefficient on the lagged test score, $\tau_{gz,2}$ is a grade-type specific coefficient on the lagged test score squared, α_{zt} is a year-type fixed effect, ω_{gz} is a grade-type fixed effect, η_{zjn} is a fixed effect at the school-type-number of schools in the village level, and $\mathbf{X_{vzt}}$ is a set of village-type-year level controls.[20] Then, the value-added for a type z in a school j under competitive regime n is given by η_{zjn}. To construct a single measure of quality for mean scores, I average across a school's estimated type-specific value-addeds in math, Urdu, and English.

The goal of this method is to estimate the causal effect of attending a school on test scores separately for rich and poor students. The controls 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 is then attributed to the school that students attended. Intuitively, the fixed effect η_{zjn} 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 accounts for the selection of students into schools.

In Appendix Table A6, I test whether these measures are in fact strong predictors of student test scores when students change schools.[21] I show that the type-specific value-added measures are indeed highly predictive of a student's out-of-sample gains from attending a given school. In fact, in most cases, the coefficient is approximately 1, as we would expect if equation (6) is the correctly specified equation for student test score outcomes. Appendix Table A6 also indicates that the type-specific value-addeds are similarly predictive of test

[20]These again consist of controls for the report card intervention, whose effects are allowed to be different for each type in each year.

[21]To avoid any spurious correlation between the student's own test scores and the school's estimated value-added, I re-calculate the value-addeds leaving the sample of school-switchers in these regressions out of the sample used for estimation.

score gains for both rich and poor types and that, while own type-specific value-added is a strong predictor of a child's test scores gains, the value-added for the other type has no additional predictive power.

With these value-added estimates, I can also directly compare value-addeds for rich and poor types in the same schools to see if match-specific quality appears to affect students' outcomes. Figure 2 plots each private school's value-added estimate for rich types against its value-added estimate for poor types. There is a strong correlation between the two (0.72), but the fact that this value is not equal to 1 is not merely due to measurement error. An F-test of the estimated interactions between school fixed effects and an indicator variable for being a rich type in equation (6) rejects the possibility that these interaction terms are jointly equal to 0 with a F-statistic of 20,312 when the outcome variable is mean test scores. While this suggests that match-specific quality plays a role in students' outcomes, it is important to keep in mind that the correlation between value-addeds for rich and poor students is *not* a sufficient statistic for the importance of match-specific quality since it is also affected by schools' equilibrium choices of h_{jt}. In an extreme example, if every school selected a h_{jt} of 1/2 in the model, the correlation between rich and poor students' value-added would be 1 regardless of the size of β.

5.3 Determinants of School Demand

In this section, I estimate the parameters of equation (5) using a discrete choice model with unobserved school quality in the spirit of Berry et al. (2004). Since my data has repeated observations of schools' characteristics and students' enrollment decisions over time, I now add the t subscript to equation (5).

Then, the new equation for the utility of a student i of type z in school j and year t is

$$u_{ijzt} = \delta_z VA_{jzn} + \mathbf{\Gamma_z^{indiv}} X_{ijt}^{indiv} + \zeta_{jn} + \epsilon_{ijt}, \tag{7}$$

where X_{ij}^{indiv} is the set of characteristics affecting school choice that vary at the individual-level, consisting of the interaction of school fees and an indicator variable for being a rich type, the effect of distance on rich and poor types, a control for a child having been in school j in the previous period interacted with type, and controls for a boy attending a school marked as an all boys school or a girl attending a school marked as an all girls school.[22] ζ_{jn} is the school fixed effect, which is allowed to vary non-parametrically with the number of

[22] $\mathbf{\Gamma_z^{indiv}}$ does not include the interaction between school fees and being a poor type since including both this control and the interaction of school fees with being a high type would be collinear with the school fixed effect. Instead, I estimate the baseline effect of school fees on school choice, $\mathbf{\Gamma_z^{school}}$, separately.

private schools in the market, n, and is equal to

$$\zeta_{jn} = \xi_{jn} + \mathbf{\Gamma_z^{school}} X_{jn}^{school}. \tag{8}$$

Here, ξ_{jn} is the school's unobserved quality and X_{jn}^{school} is a school j's average, inflation-adjusted fees under competitive regime n. Therefore, my parameters of interest are the coefficients in a student's utility function: $\{\mathbf{\Gamma_z^{school}}, \mathbf{\Gamma_z^{indiv}}, \delta_{low}, \delta_{high}\}$.

Understanding the distinction between equation (7) and equation (8) is important. The variables $VA_{j,poor,n}$, $VA_{j,rich,n}$ and X_{ijt}^{indiv} vary at the individual-level, while X_{jn}^{school} varies at the school-level. Therefore, the coefficients $\{\mathbf{\Gamma_z^{indiv}}, \delta_{poor}, \delta_{rich}\}$ can be estimated jointly with the school fixed effects, ζ_{jn}, using a maximum likelihood procedure with individual-year level choice data. Intuitively, including the school fixed effects accounts for any unobserved characteristics of the school that may affect school choice and be correlated with a school's type-specific value-added and which would otherwise bias the estimates of δ_z. For example, if schools with higher value-addeds also have other attractive features like toilets, the estimates of δ_{poor} and δ_{rich} would be positively biased in the absence of school fixed effects.

$\mathbf{\Gamma_z^{school}}$ is not identified in this procedure since X_{jn}^{school} is collinear with the school fixed effects. For this reason, I separate my estimation into two stages. In the first stage, I estimate $\{\mathbf{\Gamma_z^{indiv}}, \delta_{poor}, \delta_{rich}\}$ using maximum likelihood, and in the second stage, I estimate $\mathbf{\Gamma_z^{school}}$ using the general method of the moments. In the second stage, geographic variation in the cost of teachers allows me to instrument for price, allowing me to identify the effect of fees. As I will show in Section 5, it is not necessary to estimate $\mathbf{\Gamma_z^{school}}$ to estimate private schools' equilibrium choices of horizontal quality or β since the baseline effect of school fees is subsumed by the estimates of ζ_{jn}. Nonetheless, estimates of $\mathbf{\Gamma_z^{school}}$ are useful since they allow us quantify welfare effects.

To allow students' choices of schools to depend on distance, I use the household data rather than the tested data to estimate the demand for schooling. Additionally, to ensure that there is no correlation between the estimated value-addeds and the errors in the discrete choice model, I drop children in the household survey data who also appear in the tested sample. In other words, I estimate the school characteristics $VA_{j,poor,n}$, $VA_{j,rich,n}$ using the tested sample but use the child characteristics and observed choices in the non-overlapping portion of the household sample to estimate the determinants of school demand.

Identification of $\{\mathbf{\Gamma_z^{school}}, \mathbf{\Gamma_z^{indiv}}, \delta_{poor}, \delta_{rich}\}$

Since I previously assumed that ϵ_{ijt} is a type 1 extreme value error, the probability that a student i of type z attends school j in year t can be written as

$$p_{ijzt} = \frac{e^{\delta_z VA_{j,z,n} + \mathbf{\Gamma_z^{indiv}} X_{ijt}^{indiv} + \zeta_{jn}}}{\sum_k e^{\delta_z VA_{k,z,n} + \mathbf{\Gamma_z^{indiv}} X_{ikt}^{indiv} + \zeta_{kn}}}.$$

However, recall that for each child in the household survey, I estimated the probability of being a rich type rather than assigning a binary type. Therefore, I write the expression for the probability that a child i attends a school j in year t as:

$$p_{ijt} = P(type_i = rich)p_{ij,rich,t} + (1 - P(type_i = rich))p_{ij,poor,t}, \tag{9}$$

where $P(type_i = rich)$ is the probability that i is a rich type that was previously estimated with the logistic lasso regression.

Using equation (9), I choose the parameters $\{\mathbf{\Gamma_z^{indiv}}, \delta_{poor}, \delta_{rich}, \zeta_{zn}\}$ that maximize the log likelihood function

$$\sum_{ijt} \mathbb{1}_{ijt} log(p_{ijt}),$$

where $\mathbb{1}_{ijt}$ is an indicator variable equal to 1 if i attends school j in year t. Intuitively, this estimation procedure chooses the parameters that make students' observed enrollment decisions most likely. More details of the estimation procedure are described in Appendix C.

Identification of $\mathbf{\Gamma_z^{school}}$

Equation (8) appears to be a linear regression. If school fees were unrelated to a school's unobserved quality, ξ_{jn}, then I could estimate $\mathbf{\Gamma_z^{school}}$ by regressing the estimated school fixed effects $\widehat{\zeta_{jn}}$ on X_{jn}^{school}. However, this assumption is unlikely to be satisfied. Profit-maximizing schools with higher ξ_{jn} should charge higher prices. Therefore, to identify $\mathbf{\Gamma_z^{school}}$, I need an instrument for school fees.

Following Hausman (1996) and Nevo (2001), I use a variable that shifts the cost of providing education as my instrument for price. In particular, I use residual geographic variation in teacher salaries (similar to Neilson (2017)), after controlling for teacher characteristics, as an instrument for private schools' prices. I create a sub-district measure of the mean of residual teacher salaries, leaving the school's own village out. The leave-one out estimator ensures that differences in teacher salaries are not driven by competition over teachers in village v, which may also be related to ξ_{jn}. The key assumption is that any one village is too small to change prices in other villages in the sub-district, but villages in the same sub-district market are likely to have the same systematic differences in teacher labor supply.

The final instrument is then the interaction of this leave-one out estimate $cost_v$ with an indicator variable for whether a school j is private $I_j^{private}$, controlling for $cost_v$ and $I_j^{private}$ separately. Therefore, variation in the instrument comes from being a private school in a sub-district where private school teachers are more expensive. With this instrument, the parameters of equation (8) can be estimated with the general method of the moments. The estimation procedure is described in more detail in Appendix C.

Estimates

Table 4 reports the estimates for the key parameters of the utility function. Reassuringly, the directions and relative magnitudes of the coefficients match standard economic intuitions. Both types respond negatively to distance, but poor types are much more sensitive to distance. This is consistent with descriptive statistics from Appendix A, which indicate that poor households are more likely to list distance as the most important factor in choosing a school. Both types also respond negatively to fees (measured in 1,000s of Rupees), but the wealthy types are somewhat less sensitive.

The estimates confirm that poor types are much less responsive to match-specific quality relative to rich types ($\delta_{rich} > \delta_{poor}$), suggesting that schools have an incentive to choose instructional levels that benefit rich types. While poor types do respond positively to their predicted test score gains in a school (although the effect is not statistically significant), rich types are much more responsive. An increase in predicted test score gains of 1 test score s.d. increases the utility of attending a school for rich types by two times as much as it increases it for poor types.

Since coefficients in the utility function are difficult to interpret, I also compare the effects of type-specific value-added for rich and poor types in two other ways. First, I can compare what change in fees is equivalent to a 1 test score s.d. increase in value-added. For rich types, an increase in value-added of 1 s.d. is equivalent to a reduction in fees of 303 Rupees. For poor types, it is only 144 Ruppees. Second, I can calculate the average derivative of each school's enrollment with respect to type-specific value-added and compare these derivatives for the value-addeds of rich and poor types. On average, a school will increase its student population by nine times as much by increasing value-added for rich types relative to poor types.

5.4 Equilibrium Choice of Horizontal Quality

To estimate h_{jt} and β, I assume that schools are observed choosing their equilibrium characteristics. The estimation strategy then consists of two stages. In the first stage, the condition $\frac{\partial s_{jt}}{\partial h^*_{jt}} = 0$ is used to identify h_{jt} for every private school-year observation jt. In the second stage, the estimates of h_{jt} are used to identify β.

In the first stage, I begin by noting that, in equilibrium, for each jt,

$$\frac{\partial s_{jt}}{\partial h_{jt}} = \sum_{it} P(type_i = rich) \frac{\partial p_{ij,rich,t}}{\partial h_{jt}} + (1 - P(type_i = rich)) \frac{\partial p_{ij,poor,t}}{\partial h_{jt}} = 0,$$

where

$$\begin{aligned} \frac{\partial p_{ij,poor,t}}{\partial h_{jt}} &= 2\delta_{poor} \beta h_j (p^2_{ij,poor,t} - p_{ij,low,t}) \\ \frac{\partial p_{ij,rich,t}}{\partial h_{jt}} &= \delta_{rich} p_{ij,rich,t} (2\beta - 2\beta h_j)(1 - p_{ij,rich,t}). \end{aligned} \tag{10}$$

Dividing through by 2β produces the expression used to identify h_{jt}:

$$\begin{aligned} &\sum_{it} P(type_i = rich) \delta_{rich} (1 - h_{jt}) p_{ij,rich,t} (1 - p_{ij,rich,t}) \\ &+ (1 - P(type_i = rich)) \delta_{poor} h_{jt} (p^2_{ij,poor,t} - p_{ij,poor,t}) = 0 \quad \text{for each } jt. \end{aligned} \tag{11}$$

Equation (11) is now independent of β, and all the terms besides h_{jt} are observed. The probability that a student is a rich type, $P(type_i = rich)$, is given by the logistic lasso regression, and $p_{ij,rich,t}$ and $p_{ij,poor,t}$ can be estimated using the demand system estimates in the previous subsection. Then, I estimate h_{jt} for all jt by solving for the h_{jt} that satisfies (11).

Using these estimates of h_{jt}, it is straightforward to estimate β. Manipulating equations (3) and (4) results in the expression

$$VA_{j,rich,n} - VA_{j,poor,n} = \beta(2h_{jt} - 1), \tag{12}$$

and using this expression, β is identified by a regression of the difference between the estimated value-addeds for high and low types on $(2\widehat{h}_{jt} - 1)$. Equation (12) is intuitive: it uses the correlation between schools' equilibrium, profit-maximizing choices of horizontal quality and the difference in test score gains in a school for rich and poor students to identify β. Thus, if match-specific quality has no effect on test scores, the model can estimate that $\beta = 0$. More details of the estimation procedure are discussed in Appendix D.

Estimates

I estimate that β is 0.461.[23] This implies that horizontal quality can play a large role in students' outcomes; choosing an instructional level that is optimal for rich types will reduce poor types' test scores by 0.461 s.d. relative to choosing the instructional level that is optimal for poor types. This loss is equivalent to the average test scores gains from 1.15 years of schooling.

Figure 3 plots the distribution of the estimates of schools' equilibrium horizontal qualities. The average value of h^*_{jt} is 0.702, implying that private schools typically choose instructional levels that strongly advantage rich types over poor types. Figure 3 also reveals that schools' choices of horizontal quality are strongly skewed toward rich types with almost no schools selecting a horizontal quality below 0.5. Given the estimate of β, moving a poor student from a private school with the average horizontal quality to one at her optimal instructional level would improve her test scores by 0.227 s.d.

The estimates of h^*_{jt} are also consistent with the event study results from Section 3, which show that increasing competition increases inequality in test scores in the private sector. A regression of the estimated h^*_{jt} on the number of private schools in the market shows that – on average – an additional private school in the marketplace is associated with a increase in h^*_{jt} of 0.005 (with a standard error of 0.002). An increase in h_{jt} would in turn increase test score inequality between rich and poor students, increasing inequality in the private sector. Appendix Figure A1 plots the local linear regression of the relationship between h^*_{jt} and the number of schools. The figure shows that the positive linear relationship in the regression masks a non-linear relationship, where additional schools are most associated with increases in h_{jt} in markets that already have several private schools.

Finally, to understand whether differential responsiveness to quality drives schools to choose instructional levels that favor rich students, I re-estimate schools' equilibrium values of h^*_{jt} after setting $\delta_{low} = \delta_{high}$. I find that if poor students were equally responsive to school quality, the average value of h_{jt} would be 0.55 instead of 0.70. Thus, the fact that schools choose horizontal qualities that so strongly advantage rich students appears to be largely driven by the fact that rich students are more responsive to quality.

[23] The standard error from the OLS regression is 0.02, but this does not account for error coming from estimation error in the h_{jt}.

6 Reduced-Form Test of the Structural Model's Predictions

Before conducting counterfactual exercises in the next section, in this section, I report the results of a model-independent test of the structural model's prediction that school entry on average increases inequality between rich and poor types. In the first subsection, I again exploit the exit and entry of private schools into the education market over time to test how increased school competition affects the different types' test scores in the private sector. I also provide evidence on how different types of students perform on hard and easy questions that is consistent with schools responding to competition by moving to more advanced instructional levels. In the second subsection, I provide evidence that the estimates from the exit-entry regressions are not driven by selection or omitted variable bias.

6.1 Evidence From School Entry and Exit

Empirical Strategy. To test whether competition increases inequality between rich and poor types, I estimate the heterogeneous effects of the number of private schools on the learning of rich and poor types using

$$\begin{aligned} y_{ijgzt} =& \rho_0 + \rho_1 num\_pri_{vt} + \rho_2 num\_pri_{vt} \times \mathbb{1}_{rich} + \eta_{zj} + \alpha_{zt} + \omega_{gz} \\ &+ \lambda_{gz} y_{i,t-1} + \phi_{gz} y^2_{i,t-1} + \mathbf{\Gamma X_{vzt}} + \epsilon_{ijgzt}, \end{aligned} \tag{13}$$

where i indexes students, g indexes grades, z indexes types, v indexes villages, j indexes schools, and t indexes years. y_{ijgzt} is then a student-year level test score in math, Urdu, or English, α_{zt} are year-type fixed effects, ω_{gz} are grade-type fixed effects, and η_{zj} are school-type fixed effects. λ_{gz} and ϕ_{gz} are grade-type specific coefficients on test scores and lagged test scores, and $\mathbf{X_{vzt}}$ includes additional controls at the village-type-year level.[24] Standard errors are clustered at the village-level.

In some specifications, I also include controls for student-teacher ratios and peer controls, all of which are allowed to have different effects for rich and poor types. I include these controls since exit and entry may affect both the number of students in a school and the composition of the schools. Peer effects controls consist of a control for the mean and variance of the lagged test scores of the students in a school and a control for the percent of rich students in the school.

This regression allows me to estimate the change in the test scores of rich and poor

[24] Just as in the value-added estimation, these controls include controls for the report card intervention, whose effects are allowed to vary at the year-type level.

individuals in the private sector induced by the entry of a new private school. The coefficients of interest are ρ_1 and ρ_2. Controlling for the school-type fixed effects and the rich function of lagged test scores means that ρ_1 and ρ_2 are identified by changes to the value-added for a type within a school when another school enters or exits the market.

Results. Table 5 reports the results from this specification for math, Urdu, English, and mean test scores. The odd columns report the results for the basic specification, while the even columns add the additional peer effects and student-teacher ratio controls. Adding an additional private school has a negative effect on mean test scores for poor types and a positive effect for rich types (columns 7 and 8). Including the additional controls has little effect on this pattern. An additional private school in the marketplace increases inequality in yearly test score gains by 0.08 standard deviations, and this is mainly driven by a 0.07 s.d. fall in poor types' scores.

The increase in inequality between rich and poor types documented here is consistent with the predictions from the structural model. The fact that this inequality is mainly driven by a reduction in poor students' test scores is also consistent with the model. In the model, students suffer convex losses the farther away a school is from their optimal instructional level. Since schools are usually much closer to rich types' optimal instructional level, moving their instructional level even closer to rich students will reduce poor students' test scores more than it increases rich students'.

Additional Evidence on Mechanisms. To tease out the proposed mechanism in the structural model – that schools change their instructional level to advantage rich students who desire more advanced instruction at the expense of poor students, I now analyze data on how students perform on specific questions. For the set of questions asked in all four years, I rank questions in difficulty based on the percent that were answered correctly in the first year.[25] The one-third of questions that the fewest students answered correctly are coded as hard, the middle one-third is coded as medium, and the one-third that the most answered correctly is coded as easy. Now, I re-estimate equation (13) with the share of questions of each type answered correctly as the outcome variables. Since I am combining questions across subjects, I control for lagged mean test scores.

If schools respond to competition by moving their instructional levels closer to the optimal level for more advanced, wealthy students, we expect rich types to perform better on hard questions when there are more private schools in the market and poor types to perform worse on relatively easier questions. Table 6 provides suggestive evidence that this is the

[25] Data from the first year does not appear in the regressions, since lagged test score controls are not available in the first round of data collection.

case. While competition reduces the share of easy questions answered correctly by poor types, it has the strongest positive effects on the share of hard questions answered correctly by rich types.

6.2 Robustness of School Exit and Entry Results

In this subsection, I discuss two possible sources of bias in the estimation of ρ_1 and ρ_2 and test whether my results are robust. I show that (1) the results are not driven by pre-trends, and (2) they are not driven by new students entering the private sector following an entry event.

To test whether the results are biased by pre-trends, I include the forward lag of the number of private schools and its interaction with being a rich type in equation (13). These forward lags are placebo tests; they test whether the number of private schools in year $t+1$ had an effect on outcomes in year t before the entry or exit event took place. If the main effects are indeed driven by pre-trends, one would expect the estimates of the forward lags to be similar to the estimates of ρ_1 and ρ_2. Appendix Table A7 shows that there is no evidence that this is the case. Across all subjects and for both high and low types, the forward lags are small and statistically insignificant.[26]

Additionally, the estimates of ρ_1 and ρ_2 may also be biased if the creation of new private schools leads new students to enroll in private schools who would otherwise attend public schools. These new students may be unobservably worse than the poor students already attending private schools, leading to a negative correlation between the number of private schools in the market and the performance of poor types in private schools. To ensure that my results are robust to this possibility, I re-estimate equation (13), restricting my sample to students who always attend private schools and are observed prior to the exit and entry events. Appendix Table A8 reports the results of this regression. The results are similar: the addition of a private school increases inequality in test scores between rich and poor types by 0.1 s.d.

7 Counterfactuals

To characterize the welfare and learning losses due to the misallocation of match-specific quality, I now consider the effects of three types of partial equilibrium counterfactuals. In these counterfactuals, the social planner chooses private schools' instructional levels and students re-sort into schools. These counterfactuals are partial equilibrium in the sense

[26] The smaller sample size in Appendix Table A7 relative to Table 5 is due to the fact that the forward lag for the number of private schools is missing in 2007, causing these observations to be dropped.

that they fix schools' characteristics besides h_{jt} to be the same as in the data. In the first counterfactual, I solve for the social planner's choice of instructional level when the objective is to maximize social welfare, and utility is given by the demand-side estimates in equation (7). This gives a lower bound estimate of the effects of the Spence distortion, since it assumes that poorer types' relatively low value for δ_{poor} captures the true value poor types' place on quality, and δ_{poor} is not negatively biased by lack of information. The second counterfactual imposes that the social planner weights school quality the same way for poor types as rich types (setting $\delta_{poor} = \widehat{\delta_{rich}}$) but still assumes poor types have the same information/preferences as before, so they choose schools to maximize the preceived utility given by equation (7). Finally, in the third set of counterfactuals, I estimate the effects of making both poor and rich students more responsive to quality when they make enrollment decisions. This final set of counterfactuals shows how increased information can amplify the positive effects of allowing the social planner to choose h_{jt}.

7.1 Eliminating the Spence Distortion

Following Small and Rosen (1981), under the assumption that ϵ_{ijt} is drawn from a type 1 extreme value distribution and the demand estimates from equation (7) give the true parameters of the utility function, expected total welfare in Ruppee terms is given by

$$\begin{aligned} W_1 = & \sum_i \frac{P(type_i = rich)}{\alpha_{rich}} \log\left(\sum_j e^{\delta_{rich} VA_{j,rich,n} + \mathbf{\Gamma}^{\mathbf{indiv}}_{\mathbf{rich}} X^{indiv}_{ijt} + \zeta_{jn}}\right) \\ & + \frac{1 - P(type_i = rich)}{\alpha_{poor}} \log\left(\sum_j e^{\delta_{poor} VA_{j,poor,n} + \mathbf{\Gamma}^{\mathbf{indiv}}_{\mathbf{poor}} X^{indiv}_{ijt} + \zeta_{jn}}\right), \end{aligned} \tag{14}$$

where α_{rich} and α_{poor} are the coefficients on school fees in the utility function for rich and poor types respectively. To capture the lower bound effect of eliminating the Spence distortion, I solve for the choices of private schools' h_{jt} that would maximize W_1. To allow students to fully sort in the counterfactual equilibrium, I also eliminate the effect of past schools on welfare by setting the variable for previously attending a given school to 0. To calculate changes in welfare, I calculate W_1 under the new allocation of h_{jt} and the existing allocation in the data. More details of the estimation procedure for this and the other two types of counterfactuals are described in Appendix E.

Eliminating the Spence distortion decreases the average value of h_{jt} from 0.70 to 0.57. Panel A of Figure 4 shows the original distribution of h_{jt} and the counterfactual distribution. Reflecting the relatively low weight poor types place on quality in the demand estimates, the effect on total welfare is small, and welfare only increases by 0.1%. Table 7 reports the distribution of the welfare changes and shows that the changes are heterogeneous and highly right-skewed, reflecting the fact that much of the population does not attend private schools. While a student at the 50th percentile experiences 0 gains, a student at the 90th percentile's welfare increases by 0.5% and one at the 95th percentile's welfare increases by 0.8%. Welfare gains to poor students come at a relatively low cost to richer students, with a student at the 10th percentile only experiencing a welfare loss of 0.001%. Turning to learning, across all children, average yearly learning increases by 0.004 s.d. every year, and inequality in the school value-added experienced by rich and poor types declines by 15%.

These effects are estimated across the whole market, but recalling that only a minority of children attend private schools in the data, I next estimate the effects for the "most affected sample." This group has at least a 25% chance of a attending a private school in either the original or counterfactual setting (23% of total students). In this group, welfare increases by 0.3%, and learning increases by 0.01 s.d. per year. Inequality in test score gains due to school quality falls by 37%. Thus, at a lower bound, removing the Spence distortion leads to moderate average test score gains but meaningfully reduces inequality in learning, both overall and particularly in the private sector.

The fact that overall welfare gains are relatively small in this counterfactual is consistent with the fact that, in the demand estimates, poor types place a low weight on school quality. However, this may reflect lack of information by poor types. Additionally, even if poor parents place a low weight on school quality, the social planner may place a higher weight. This is especially true if poor parents do not internalize the benefits to education, either because there are externalities or because there are incomplete contracting problems between parents and children (Ashraf et al., forthcoming; Bau, 2019). In the next counterfactual, I consider a case where the social planner weights school quality equally for rich and poor children.

7.2 Equally Weighting School Quality for Rich and Poor Types

In this counterfactual, I assume that poor types still use the demand estimates from equation (7) to make enrollment decisions, but I replace δ_{poor} with δ_{rich} in the utility they experience once they attend a school. Thus, the social planner cares about rich and poor students' learning the same amount, but poor students are still worse at sorting into higher quality schools. The new expected welfare formula follows Train (2015) and is

$$W_2 = \sum_i \frac{P(type_i = rich)}{\alpha_{rich}} \log\Big(\sum_j e^{\delta_{rich} VA_{j,rich,n} + \mathbf{\Gamma^{indiv}_{rich}} X^{indiv}_{ijt} + \zeta_{jn}}\Big)$$
$$+ \frac{1 - P(type_i = rich)}{\alpha_{poor}}\Big(\log\Big(\sum_j e^{\delta_{poor} VA_{j,poor,n} + \mathbf{\Gamma^{indiv}_{poor}} X^{indiv}_{ijt} + \zeta_{jn}}\Big)$$
$$+ \sum_j p_{ijt}(\delta_{rich} - \delta_{poor}) VA_{j,poor,n}\Big). \quad (15)$$

The social planner chooses private schools' h_{jt} to maximize W_2. To calculate changes in welfare, I calculate W_2 under the new allocation of h_{jt} and the existing allocation in the data.

Under this counterfactual, the average private schools' h_{jt} is now 0.384, indicating that h_{jt} moves substantially closer to poor types' optimal instructional levels, as shown in Panel B of Figure 4. Total welfare increases by 0.4%, and per-student yearly test score gains rise by 0.01 s.d. Inequality in yearly test score gains between rich and poor types due to school quality falls by 34%. Table 7 reports the distribution of welfare gains. Losers experience larger losses now, but losses are still relatively small, with students at the 10th percentile only experiencing losses of 0.4%. In contrast, there is a long tail in gains, with students at the 75th percentile experiencing gains of 1%, students at the 90th percentile experiencing 2% gains, and students at the 95th percentile experiencing 3.3% gains.

Focusing on students with at least a 25% chance of attending a private school in the original or counterfactual case (23% of the sample), welfare increases by 1%. Yearly test score gains increase by 0.02 s.d. and inequality in test score gains due to school quality falls by 82%. Since average yearly test score gains in Pakistan are 0.40 s.d., a gain of 0.02 s.d. is equivalent to one-twentieth of an extra year of schooling every year. At the end of 5 years of primary school, this is equivalent to a fourth of an extra year of education.

7.3 Improving Sorting

Finally, I consider a set of counterfactuals where I vary both types' responsiveness to school quality when they make enrollment decisions. This set of counterfactuals is motivated by the idea that improved sorting could increase the social planners' scope for improving welfare, since better sorting allows for more product differentiation and better matches between schools and students. Additionally, the low estimated responsiveness to quality by both rich and poor students may be driven by lack of information about quality rather than preferences.[27] If this is the case, policymakers can undertake interventions – like providing information – that improve sorting. Thus, these counterfactuals can be thought of as capturing the effects of the social planner pairing changes in private schools' match-specific quality with informational interventions that inform parents about match-specific school quality.

For these counterfactuals, I replace both δ_{high} and δ_{low} in the demand estimates with $\widehat{\delta_{high}} \times m$, where the multiplier m is allowed to vary from 1 to 3 at increments of 0.1. Even the maximum value for m, $m = 3$, does not imply a huge valuation of quality. It indicates that parents are willing to give up 3% of GDP per capita for a school with 1 test score s.d. higher value-added (equivalent to the test score gains from 2.5 extra years of education). The social planner then chooses the set of h_{jt} for private schools that maximizes W_1. To calculate changes in welfare, I calculate baseline welfare under the original allocation using a similar formulation to W_2. In this formulation, W_2 is modified so that poor and rich students sort in the same way as in the original demand estimates, but preferences for school quality are the same as in the counterfactual so that changes in welfare are not driven by changes in preferences.

Figure 5 plots the counterfactual average values of horizontal quality, welfare gains, and learning gains for each value of m. Even for $m = 1$, sorting leads to meaningful gains. Net welfare now increases by 0.5%, and across all students, yearly test score gains increase by 0.02 s.d. Inequality in test score gains across students falls by 55%. As Table 7 shows, welfare changes are yet again right-skewed, with gains of 1% at the 75th percentile and 2% at the 90th percentile. Welfare losses are also smaller than in the previous counterfactual, with a student at the 10th percentile only experiencing a 0.2% loss in welfare. This reflects the fact that improved sorting means that some schools need to cater less aggressively to poorer types than they did before. For the most affected population (24%), yearly welfare increases by 1%, test score gains increase by 0.04 s.d., and poor types' test score gains due to school quality completely catch up with rich types. Thus, by the end of primary school, this group receives the equivalent of 0.5 more years of schooling.

Turning to larger values of m, the gains are more dramatic. For example, when $m = 2$, indicating that students are willing to pay 2% of GDP per capita for a 1 test score s.d. better school, welfare gains to choosing the optimal instructional level for the full population are 3% and yearly test score gains increase by 0.05 s.d. For those with a 25% chance of attending private school, welfare increases by 7%, and test score gains increase by 0.09 s.d. As Table 7 shows, almost all students benefit in the counterfactual. Even students at the 10th percentile experience gains. Thus, improved sorting also almost entirely mitigates any

[27] Appendix A provides descriptive evidence that parents may lack information about school quality. For example, while parents' ranking of school is correlated with schools' match-specific quality, their ranking only explains 3% of the variation in school quality.

losses from moving away from rich types' optimal instructional levels. These gains occur in part because higher values of m allow schools to product differentiate more, better matching poor students' optimal instructional levels. Appendix Figure A2, which reports kernel density plots of private schools' horizontal quality under $m = 1$ and $m = 2$, illustrates this effect. Thus, altogether, the results from these counterfactuals suggest that reducing misallocation in instructional levels in conjunction with improved sorting can have meaningful effects on welfare and learning.

8 Discussion & Conclusion

In this paper, I estimate a structural model where schools compete for students by choosing both match-specific and vertical quality. Importantly, the demand estimates show that wealthy students are more responsive to their predicted test score gains when they choose schools, while poor students are much less responsive. Following the intuition of Spence (1975), this causes schools to respond to competition by choosing instructional levels closer to wealthier students' optimal instructional level to the detriment of poor students.

When a new private school enters the market, within-school inequality in test scores increases in the private sector. This is mainly driven by a decline in poor students' test scores, although wealthy students do benefit. My model-independent estimates imply that an additional private school in the market leads the gap in test score gains between rich and poor students to grow by 0.08 s.d. This result highlights the fact that inequality in test scores between rich and poor students can be driven by inequality *within* schools, as well as inequality across schools.

The estimation of the structural model delivers an important additional set of results. First, private schools choose instructional levels that are highly skewed toward the optimal instructional levels of wealthy students. The average private school in the data chooses a horizontal quality of 0.7, even though half of the private school population desires an optimal instructional level of 0. Second, the structural estimation delivers an estimate of β, the importance of instructional match in the learning production function. According to this estimate, a student in a school whose instructional level is the opposite of her optimal instructional level will have 0.46 s.d. lower yearly test score gains than a student in a school that perfectly matches her optimal level.

I conduct several partial equilibrium counterfactuals to examine the effects of the misallocation of instructional level on students' outcomes. The first two counterfactuals show that removing misallocation increases total welfare by 0.1-0.4%, increases learning by as much as 0.01 s.d. for every child in every year, and reduces inequality in learning by 15-34%. The human capital effect is moderate but not trivial: by the end of primary school, a yearly 0.01 s.d. gain is equivalent to one-eighth of a year of additional education. Removing the Spence distortion would thus greatly reduce inequality in learning while improving overall learning and welfare.

Yet, in light of the large estimate of β, one might ask why removing misallocation in instructional levels doesn't have more dramatic effects on welfare and test scores. One reason is that private schools still make up the minority of enrollment in Pakistan. Thus, addressing misallocation in instructional levels in public schools as well would likely lead to meaningful gains. Interventions that improve the match between instructional level and students, such as tracking by ability (Duflo et al., 2011), tutoring, and educational technology (Muralidharan et al., 2019), could therefore yield high returns when applied in the public sector.

However, the third set of counterfactuals indicates that there is another reason why the effects in the first two counterfactuals are not dramatic, which may point to areas for future research. In markets with many schools, if students were very responsive to quality, schools should strongly product differentiate, and rich and poor types would attend schools that closely match their optimal instructional levels. However, in reality, students only weakly respond to quality. Poor students are only willing to walk an extra fifth of a kilometer to go to a school with 1 test score s.d. higher value-added. Even wealthy students are only willing to give up 1% of GDP per capita to attend a school that has a 1 s.d. higher value-added. As a result, school populations are quite mixed,[28] and the scope for matching students to their optimal instructional level is limited.

More dramatic gains in the third set of counterfactuals come from improved sorting. In these counterfactuals, poor types are as responsive to quality as rich types, and both types' responsiveness to quality is allowed to increase. When students are more responsive to quality, the social planner can differentiate instructional levels more, and students can sort into better match schools. As a result, welfare and learning gains are much larger. For example, when rich and poor types are both willing to trade-off 2% of GDP per capita to attend a school that has a 1 s.d. higher value-added ($m = 2$ in counterfactual #3), total welfare increases by 3% and yearly test score gains increase by 0.05 s.d. Among the most affected population, welfare increases by 7% and test score gains increase by 0.09 s.d. (a 23% increase in average yearly test score gains).

Thus, an important question for policy is why students are so unresponsive to quality. One possibility is that low responsiveness simply reflects a low valuation of quality by parents

[28]A private school at the 25th percentile in the data has a population that is 30% rich types, while one at the 75th percentile has 56% rich types.

in Pakistan. However, an alternative possibility is that the low estimated response to school quality reflects limited information about quality. This is consistent with Andrabi et al. (2017), who show that when parents in Pakistan are informed about average test scores, they update their beliefs about school quality *and* schools respond by changing their quality. Thus, pairing changes in match-specific quality with information that improves household sorting may be a promising path for improving learning. Indeed, counterfactuals where households are more responsive to quality can deliver more dramatic gains.

Altogether, these results reveal both the promise and limitations of improving the match between students and schools. When there are multiple options and students are informed about and responsive to school quality, product differentiated schools can provide a path to improved learning. When students cannot or do not respond to quality, policy-makers may want to focus on interventions that either improve match within schools or provide students with more tailored instruction in addition to regular class time.

References

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A Pathak, "The welfare effects of coordinated assignment: Evidence from the New York City high school match," *American Economic Review*, 2017, *107* (12), 3635–89.

Agarwal, Nikhil and Paulo Somaini, "Demand Analysis using Strategic Reports: An application to a school choice mechanism," *Econometrica*, 2018, *86* (2), 391–444.

Ajayi, Kehinde F, "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana," *Working Paper*, 2011.

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja, "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets," *American Economic Review*, 2017, *107* (6), 1535–1563.

_ , _ , and Asim Khwaja, "A dime a day: The possibilites and Limits of private schooling in Pakistan," *Comparative Education Review*, 2008, *52* (3), 329–355.

_ , _ , _ , and Tristan Zajonc, "Religious school enrollment in Pakistan: A Look at the Data," *Comparative Education Review*, 2006, *50* (3), 446–477.

_ , Natalie Bau, Jishnu Das, and Asim Khwaja, "Are Bad Public Schools public bads? Test Scores and civic values in public and private schools," *Working Paper*, 2010.

Arcidiacono, Peter, Esteban M Aucejo, and V Joseph Hotz, "University differences in the graduation of minorities in STEM fields: Evidence from California," *The American Economic Review*, 2016, *106* (3), 525–562.

_ , _ , Hanming Fang, and Kenneth I Spenner, "Does affirmative action lead to mismatch? A new test and evidence," *Quantitative Economics*, 2011, *2* (3), 303–333.

Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena, "Bride Price and the Returns to Education for Women," *Journal of Political Economy*, forthcoming.

Aucejo, Esteban M, "Assessing the role of teacher-student interactions," *Working Paper*, 2011.

Bau, Natalie, "Can Policy Change Culture? Government Pension Plans and Traditional Kinship Practices," 2019.

_ and Jishnu Das, "Teacher Value-Added in a Low-Income Country," *AEJ: Economic Policy*, forthcoming.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 2007, *115* (4), 588–638.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 2004, *112* (1), 68–105.

Carneiro, Pedro Manuel, Jishnu Das, and Hugo Reis, "The Value of Private Schools: Evidence from Pakistan," *IZA Discussion Paper*, 2016.

Chetty, Raj, John Friedman, and Jonah Rockoff, "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *American Economic Review*, 2014, *104* (9), 2593–2632.

_ , _ , and _ , "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood," *The American Economic Review*, 2014, *104* (9), 2633–2679.

Cragg, John G, "More efficient estimation in the presence of heteroscedasticity of unknown form," *Econometrica*, 1983, pp. 751–763.

Crawford, Gregory S and Ali Yurukoglu, "The welfare effects of bundling in multi-channel television markets," *The American Economic Review*, 2012, *102* (2), 643–685.

_ and Matthew Shum, "The Welfare Effects of Endogenous Quality Choice: The Case of Cable Television," *Working Paper*, 2006.

Das, Jishnu and Tristan Zajonc, "India Shining and Bharat Drowning: Comparing two Indian States to the Worldwide Distribution in Mathematics Achievement," *Journal of Development Economics*, 2010, *92* (2), 175—-187.

_ , Priyanka Pandey, and Tristan Zajonc, "Learning levels and gaps in Pakistan," 2006.

Desai, Sonalde, Amaresh Dubey, Reeve Vanneman, and Rukmini Banerji, "Private Schooling in India: A New Educational Landscape," *India Human Development Survey Working Paper No. 11*, 2008.

Dinerstein, Michael and Troy Smith, "Quantifying the Supply Response of Private Schools to Public Policies," *Working paper*, 2014.

Dizon-Ross, Rebecca, "Parents' Beliefs About Their Children's Academic Ability: Implications for Educational Investments," *American Economic Review*, forthcoming.

Draganska, Michaela, Michael Mazzeo, and Katja Seim, "Beyond plain vanilla: Modeling joint product assortment and pricing decisions," *Quantitative Marketing and Economics*, 2009, *7* (2), 105–146.

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," *American Economic Review*, 2011, *101* (5), 1739–1774.

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

Fan, Ying, "Ownership consolidation and product characteristics: A study of the US daily newspaper market," *The American Economic Review*, 2013, *103* (5), 1598–1628.

Filmer, Deon and Lant H Pritchett, "Estimating wealth effects without expenditure data-Or tears: An application to educational enrollments in states of India," *Demography*, 2001, *38* (1), 115–132.

Hansen, Lars Peter, "Large sample properties of generalized method of moments estimators," *Econometrica*, 1982, pp. 1029–1054.

Hastings, Justine S, Thomas J Kane, and Douglas O Staiger, "Parental preferences and school competition: Evidence from a public school choice program," *NBER Working Paper #11805*, 2005.

Hausman, Jerry A, "Valuation of new goods under perfect and imperfect competition," in "The economics of new goods," University of Chicago Press, 1996, pp. 207–248.

Hoxby, Caroline M and Christopher Avery, "The missing "one-offs:" The hidden supply of high-achieving, low income students," *NBER Working Paper #18586*, 2012.

Hsieh, Chang-Tai and Miguel Urquiola, "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program," *Journal of public Economics*, 2006, *90* (8), 1477–1503.

Kane, Thomas J and Douglas O Staiger, "Estimating teacher impacts on student achievement: An experimental evaluation," *NBER Working Paper #14607*, 2008.

Kapor, Adam, Christopher A Neilson, and Seth D Zimmerman, "Heterogeneous Beliefs and School Choice Mechanisms," 2017.

Mazzeo, Michael J, "Product choice and oligopoly market structure," *RAND Journal of Economics*, 2002, pp. 221–242.

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J Ganimian, "Disrupting education? Experimental evidence on technology-aided instruction in India," *American Economic Review*, 2019, *109* (4), 1426–60.

_ **and Venkatesh Sundararaman**, "The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India," *Quarterly Journal of Economics*, 2015, *130* (3), 1011–1066.

Neal, Derek, "The effects of Catholic secondary schooling on educational achievement," *Journal of Labor Economics*, 1997, pp. 98–123.

Neilson, Christopher, "Targeted vouchers, competition among schools, and the academic achievement of poor students," *Working Paper*, 2017.

Nevo, Aviv, "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 2001, *69* (2), 307–342.

Palma, André De, Victor Ginsburgh, and Jacques-François Thisse, "On existence of location equilibria in the 3-firm Hotelling problem," *Journal of Industrial Economics*, 1987, *36* (2), 245–252.

_ **,** _ **, Yorgo Y Papageorgiou, and J-F Thisse**, "The principle of minimum differentiation holds under sufficient heterogeneity," *Econometrica*, 1985, *53* (4), 767–781.

Pratham, "Annual Status of Education Report," 2012.

Rivkin, Steven G, Eric A Hanushek, and John F Kain, "Teachers, schools, and academic achievement," *Econometrica*, 2005, *73* (2), 417–458.

Rockoff, Jonah E, "The impact of individual teachers on student achievement: Evidence from panel data," *American Economic Review*, 2004, *94* (2), 247–252.

Singh, Abhijeet, "Private school effects in urban and rural India: Panel estimates at primary and secondary school ages," *Journal of Development Economics*, 2015, *113*, 16–32.

Singleton, John D, "Incentives and the supply of effective charter schools," *American Economic Review*, 2019, *109* (7), 2568–2612.

Small, Kenneth A and Harvey S Rosen, "Applied welfare economics with discrete choice models," *Econometrica*, 1981, *49* (1), 105.

Spence, A Michael, "Monopoly, quality, and regulation," *The Bell Journal of Economics*, 1975, pp. 417–429.

Tibshirani, Robert, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1996, pp. 267–288.

Train, Kenneth, "Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples," *Journal of choice modelling*, 2015, *16*, 15–22.

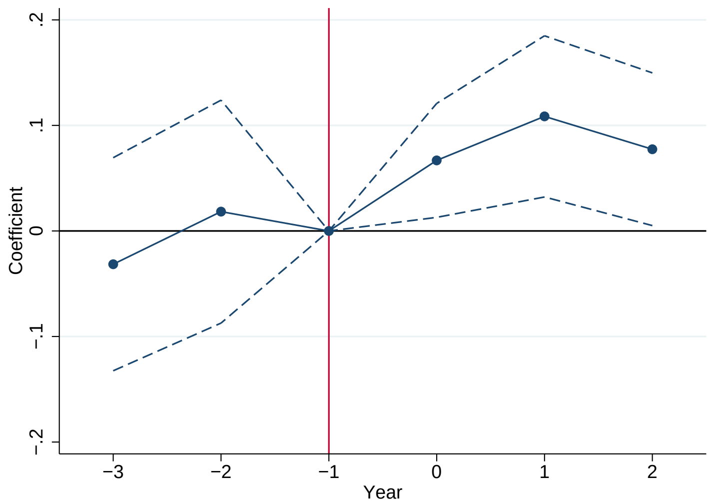
Walters, Christopher R, "The demand for effective charter schools," *Journal of Political Economy*, 2018, *126* (6), 2179–2223.

Wollmann, Thomas, "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles," *Working Paper*, 2016.

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

Figures

Figure 1: Event Study Graph of the Effect of Exit and Entry on the Variance of Private Sector Test Scores

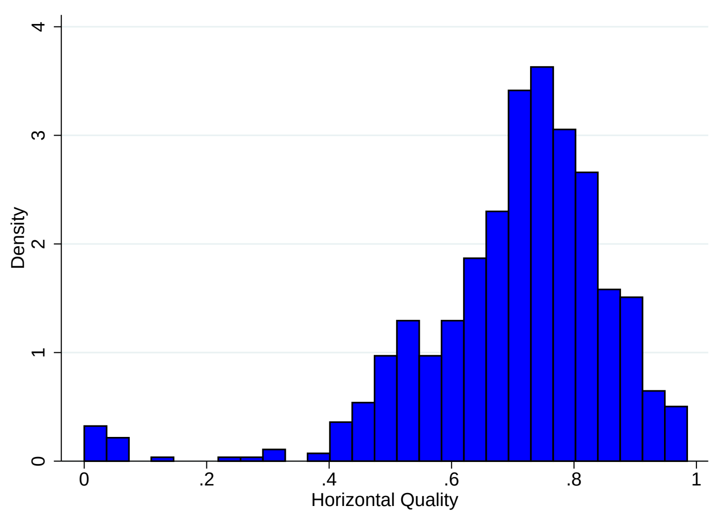


This figure plots the γ coefficients from equation (2), which estimate the effect of exit/entry events by year 3 years and 2 years before the event occurred, the year of the event, and 1 and 2 years afterwards. The x-axis is the year, where the event is normalized to take place in year 0. Coefficient estimates are relative to the effect the year before the event. The dotted lines plot the 95% confidence interval.

Figure 2: Comparison of School Value-Addeds for Rich and Poor Types

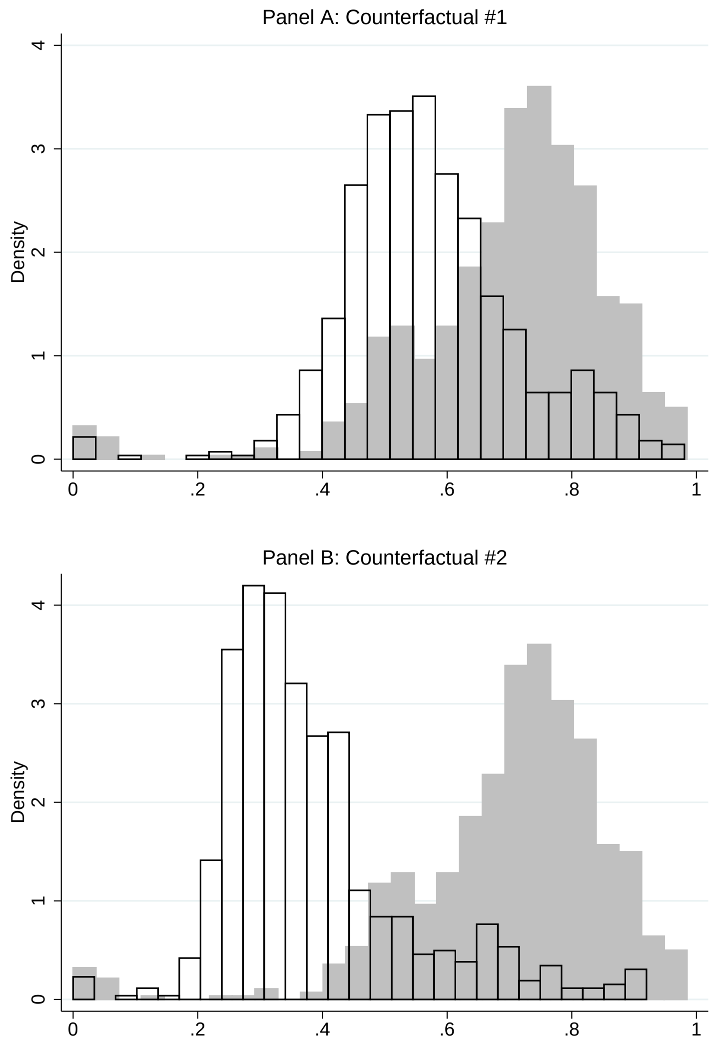
This figure plots the estimates of schools' value-added for rich types against the estimates of their value-added for poor types. The estimates come from equation (6).

Figure 3: Distribution of Private Schools' Instructional Levels



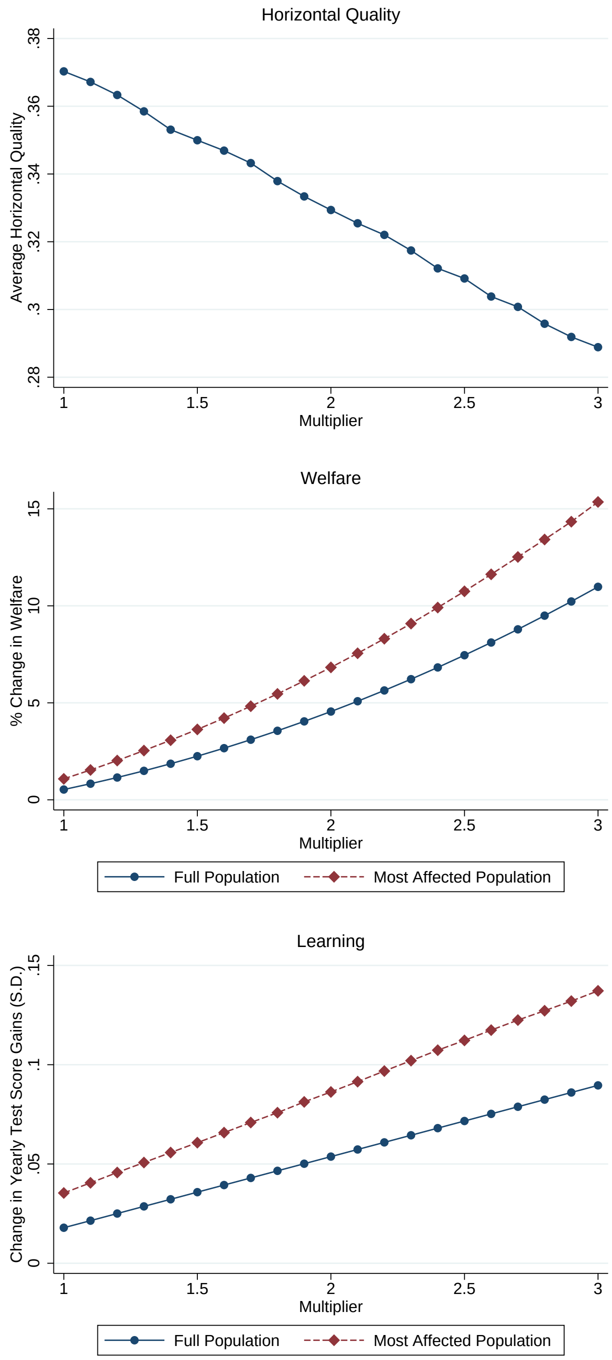
This figure plots the distribution of private schools' estimated values of h_{jt} in the data.

Figure 4: Removal of the Spence Distortion and Private Schools' Counterfactual Instructional Levels



This figure plots the distribution of private schools' values of h_{jt} under the two counterfactuals that evaluate the effects of only removing the Spence distortion (clear columns) against the distribution of h_{jt} in the data (gray columns). In counterfactual #1, the social planner maximizes social surplus based on the demand estimates. In counterfactual #2, the social planner weights school quality the same amount for rich and poor students.

Figure 5: Effects of Improved Sorting on Horizontal Quality, Welfare, and Learning



This figure plots the counterfactual horizontal quality, welfare gains, and learning gains when the both rich and poor types become better at sorting (counterfactual #3). Each point represents a counterfactual allocation chosen by the social planner to maximize welfare when both rich and poor types' coefficient on quality in the demand estimates (δ_{rich} and δ_{poor}) is replaced with $\widehat{\delta_{rich}}$ multiplied by the multiplier given in the x-axis. The most affected population is the population with at least a 25% chance of attending a private school.

Tables

Table 1: Effect of Private School Entry on Test Scores in the Public and Private Sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Math		English		Urdu		Mean	
	Private	Public	Private	Public	Private	Public	Private	Public
num_pri_{vt}	-0.088	0.014	-0.079	0.001	-0.056	-0.016	-0.074	-0.0003
	(0.067)	(0.042)	(0.054)	(0.028)	(0.064)	(0.031)	(0.061)	(0.033)
Lagged Test Scores	Y	Y	Y	Y	Y	Y	Y	Y
Grade FE	Y	Y	Y	Y	Y	Y	Y	Y
Village FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Number of observations	8,732	27,748	8,732	27,748	8,732	27,748	8,732	27,748
Clusters	109	112	109	112	109	112	109	112
Adjusted R^2	0.221	0.174	0.247	0.259	0.237	0.189	0.258	0.231

This table reports estimates of the effect of the number of private schools in the village on test scores for students in the private and public sector separately. The regressions use data from the LEAPS tested sample. Lagged test score controls consist of the relevant lagged test score and its square interacted with grade fixed effects. Observations are at the student-year level. Standard errors are clustered at the village level. The number of clusters differs by sector because not all villages have a private school in operation in the years that lagged test scores are available. *, **, and *** denote 10, 5, and 1% significance respectively.

Table 2: Effect of Private School Entry on the Variance of Test Scores in the Public and Private Sectors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Variance Math		Variance English		Variance Urdu		Variance Mean	
	Private	Public	Private	Public	Private	Public	Private	Public
num_pri_{vt}	0.084***	0.020	0.095**	0.045	0.098***	0.042	0.077***	0.034
	(0.030)	(0.029)	(0.037)	(0.036)	(0.037)	(0.034)	(0.026)	(0.029)
Village FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Number of observations	426	447	426	447	426	447	426	447
Clusters	109	112	109	112	109	112	109	112
Adjusted R^2	0.301	0.400	0.252	0.364	0.246	0.313	0.329	0.413

This table reports estimates of the effect of the number of private schools in the village on the variance of test scores in the public and private sectors separately. An observation in the regression is a village-year. The regressions use data from the LEAPS tested sample. Standard errors are clustered at the village level. The number of clusters differs by sector because not all villages have a private school in operation in the years that lagged test scores are available. *, **, and *** denote 10, 5, and 1% significance respectively.

Table 3: What Drives the Increase in Test Score Inequality Associated with Entry?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Variance Math		Variance English		Variance Urdu		Variance Mean	
	No Switchers	School-Level Variance	No Switchers	School-Level Variance	No Switchers	School-Level Variance	No Switchers	School-Level Variance
num_pri_{vt}	0.096**	0.089**	0.075	0.062*	0.123**	0.054	0.085**	0.061**
	(0.044)	(0.039)	(0.047)	(0.034)	(0.054)	(0.047)	(0.036)	(0.030)
School FE	N	Y	N	Y	N	Y	N	Y
Village FE	Y	N	Y	N	Y	N	Y	N
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Peer Controls	Y	N	Y	N	Y	N	Y	N
Student-Teacher Ratio Control	Y	N	Y	N	Y	N	Y	N
Number of observations	405	801	405	801	405	801	405	801
Clusters	108	108	108	108	108	108	108	108
Adjusted R^2	0.282	0.188	0.217	0.226	0.222	0.269	0.298	0.242

This table examines drivers of the increase in inequality associated with private school entry. Odd columns report estimates of the effect of the number of private schools in the village on the village-level variance of test scores in the private sector, restricting the sample used to calculate the variances to students who never change schools and who were observed in school prior to an entry event. In even columns, the outcome is the school-level variance of test scores in the private sector. So, in odd columns, an observation is at the village-year and in even columns, it is at the private school-year level. Peer controls consist of controls for the variance and average of the lagged test scores of students in a school. The regressions use data from the LEAPS tested sample. Standard errors are clustered at the village level. *, **, and *** denote 10, 5, and 1% significance respectively.

Table 4: Determinants of Demand in the Equilibrium Model

	(1) Coefficient	(2) Se
$VA_{j,poor,n} \times \mathbb{1}_{poor}$	0.418	0.294
$VA_{j,rich,n} \times \mathbb{1}_{rich}$	0.846***	0.240
$distance_{ij} \times \mathbb{1}_{poor}$	-1.844***	0.097
$distance_{ij} \times \mathbb{1}_{rich}$	-0.410***	0.115
fee_{jn}	-2.912***	1.005
$fee_{jn} \times \mathbb{1}_{rich}$	0.123	0.115

This table reports estimates of the determinants of school choice using a discrete choice model where schools are allowed to have time-varying unobserved quality. $VA_{j,poor,n}$ is the average of a school j's value-added for poor types in math, Urdu, and English under competitive regime n, while $VA_{j,rich,n}$ is the average value-added for rich types. Distance is measured in kilometers, and fees are measured in 1,000s of Rupees. The coefficients are estimated using the LEAPS household survey data. *, **, and *** denote 10, 5, and 1% significance respectively.

Table 5: Effect of Number of Private Schools on Test Scores by Type

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Math		English		Urdu		Mean	
$\mathbb{1}_{rich} \times num\_pri_{vt}$	0.114*	0.111*	0.065	0.063	0.062	0.048	0.083**	0.076*
	(0.064)	(0.061)	(0.043)	(0.045)	(0.046)	(0.048)	(0.041)	(0.041)
num_pri_{vt}	-0.110***	-0.118***	-0.059	-0.061	-0.024	-0.032	-0.062*	-0.069**
	(0.038)	(0.040)	(0.052)	(0.053)	(0.042)	(0.041)	(0.035)	(0.033)
Peer Controls	N	Y	N	Y	N	Y	N	Y
Student-Teacher Ratio Controls	N	Y	N	Y	N	Y	N	Y
Lagged Test Score Controls	Y	Y	Y	Y	Y	Y	Y	Y
School by Type FE	Y	Y	Y	Y	Y	Y	Y	Y
Grade by Type FE	Y	Y	Y	Y	Y	Y	Y	Y
Year by Type FE	Y	Y	Y	Y	Y	Y	Y	Y
Number of observations	6,788	6,754	6,788	6,754	6,788	6,754	6,788	6,754
Clusters	108	108	108	108	108	108	108	108
Adjusted R^2	0.568	0.581	0.597	0.603	0.630	0.640	0.685	0.703

This table reports estimates of the effect of the number of private schools in the village on test scores for rich and poor types attending private schools. The regressions use data from the LEAPS tested sample. The peer controls consist of the school-level mean and variance of lagged test scores in year t, as well as percent of rich types in a school, all of which are allowed to have different effects for rich and poor types. The student-teacher ratio controls consist of a control for the school's student-teacher ratio in year t, which is allowed to have different effects for rich and poor types as well. Lagged test score controls consist of the relevant lagged test score and its square interacted with grade-type fixed effects. Observations are at the student-year level. Standard errors are clustered at the village level. *, **, and *** denote 10, 5, and 1% significance respectively.

Table 6: Effect of Number of Private Schools on Performance on Hard, Medium, and Easy Questions

	(1) % Easy Questions Correct	(2) % Medium Questions Correct	(3) % Hard Questions
$\mathbb{1}_{rich} \times num\_pri_{vt}$	0.017	0.029*	0.031***
	(0.011)	(0.015)	(0.011)
num_pri_{vt}	-0.020**	-0.019	-0.016*
	(0.009)	(0.013)	(0.009)
Peer Controls	Y	Y	Y
Student-Teacher Ratio Controls	Y	Y	Y
Lagged Test Score Controls	Y	Y	Y
School by Type FE	Y	Y	Y
Grade by Type FE	Y	Y	Y
Year by Type FE	Y	Y	Y
Mean	0.862	0.615	0.306
Number of observations	6,754	6,754	6,754
Clusters	108	108	108
Adjusted R^2	0.503	0.624	0.658

This table reports estimates of the effect of the number of private schools in the village on performance on easy, medium, and hard questions for rich and poor types attending private schools. The regressions use data from the LEAPS tested sample. Lagged test score controls consist of the mean lagged test score and its square interacted with grade by type fixed effects. The peer controls consist of the school-level mean and variance of lagged test scores in year t, which are allowed to have different effects for rich and poor types, as well as the percent of rich types in a school. The student-teacher ratio controls consist of a control for the school's student-teacher ratio in year t, which is allowed to have different effects for rich and poor types. Observations are at the student-year level. Standard errors are clustered at the village level. *, **, and *** denote 10, 5, and 1% significance respectively.

Table 7: Distribution of Welfare Changes in the Three Counterfactuals (in %)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Percentile				
	5	10	25	50	75	90	95
Counterfactual #1	-0.26	-0.001	0	0	0.15	0.48	0.77
Counterfactual #2	-0.74	-0.35	0	0.19	0.87	2.05	3.30
Counterfactual #3 ($m = 1$)	-0.36	-0.15	0	0.28	0.75	1.45	2.07
Counterfactual #3 ($m = 2$)	-0.23	0.16	0.73	1.75	3.65	6.18	8.10

This table reports the percent welfare changes from switching from the current set of values for h_{jt} to the counterfactual set at the percentiles listed in the column headings. In each of the counterfactuals, the social planner chooses h_{jt} to maximize social welfare. In the first counterfactual, welfare is calculated using the demand estimates, and poor and rich types sort based on the demand estimates. In the second counterfactual, the social planner places the same weight on school quality for poor students as for rich students, but poor and rich students still sort based on the demand estimates. In the third set of counterfactuals, the social planner places the same weight on school quality for poor students as for rich students, and both types' coefficient on school quality in their demand function is replaced by $m \times \widehat{\delta_{rich}}$.

Appendix A: Descriptive Evidence on the Determinants of School Choice

In this appendix, I provide descriptive evidence that poor students are less responsive to school quality when they make enrollment decisions. I regress measures of information about school quality and self-reported reasons parents choose a school on the probability that a household is rich according to the classification in Section 6. In Appendix Table A9, I estimate five associations using the household survey. First, I regress an indicator variable that is equal to 1 if a child ever changes schools over the course of the survey on the probability that the child is a rich type. Here, I restrict the sample to children who are always enrolled in school between 2004 and 2007. Column 1 shows that there is a strong positive and statistically significant relationship: moving from having a 0 probability of being a rich type to a probability of 1 increases the likelihood that a child changes schools by 20 percentage points. In column 2, I regress an indicator variable for whether a child's parents know her teacher's name on the probability that a child is a rich type. Again, there is a strong and statistically significant relationship: moving from a 0 probability of being a rich type to a probability of 1 increases the likelihood that parents know the teacher's name by 16 percentage points. In column 3, I regress an indicator variable equal to 1 if parents report choosing a child's school based on distance on the probability the child is a rich type. Here, the sample size drops substantially since parents were only asked why they chose a given school in 2004. I find that rich types are less likely to choose a school based on distance by 16 percentage points, though this effect is not significant. In column 4, I run a similar regression with an indicator variable for choosing a school based on quality as the outcome. Column 4 indicates that rich types are significantly more likely to report choosing a school based on quality. Finally, in column 5, I regress a school's estimated type-specific value-added for a household on parents' ranking of that school's quality (from 1-5), the household's probability of being a rich type, and the interaction of these two variables. To account for extreme outliers in the value-added estimation, I trim the top and bottom 5% of the value-added estimates. The interaction term, though marginally significant, indicates that moving from having a 0 probability of being a rich type to a probability of 1 approximately doubles the association between a parent's rankings of school quality and the school's type-specific value-added.

Taken together, the associations reported in Appendix Table A9 show that poor types both report caring less about quality when they make enrollment decisions and have less information with which to make these decisions. This is consistent with estimates of the determinants of school choice, which show that $\delta_{high} > \delta_{low}$.

Appendix B: Mathematical Appendix

In this appendix, I provide proofs for propositions 2.1 and 2.2. Additionally, I prove proposition A1 which shows that, when δ_{poor} is not sufficiently low, there is no pure strategy equilibrium in the simplified model with two private schools.

Proposition 2.1. *For N=1, there is a unique equilibrium where the single private school chooses* $h^* = \max(1-(-u_o)^{1/2}, 0)$.

Proof. When there is one private school, it maximizes its share when it minimizes the share lost to the public school. The school will discontinuously receive all the rich types as long as $-(1-h)^2 \geq u_o$. It can never receive all the poor types, so it will always choose its location to receive all the rich types at $h \geq 1-(-u_o)^{1/2}$. Then, the school minimizes the loss of the poor types subject to this constraint at $h^* = \max(1-(-u_o)^{1/2}, 0)$. □

Proposition 2.2. *For N=2, if δ_{poor} is sufficiently small, the unique equilibrium is* $(h_1, h_2) = (1, 1)$.

Proof. The proof proceeds in two steps. First, I show that $(1,1)$ is an equilibrium for a sufficiently low δ_{poor}. Then I show that no other equilibrium is possible.

Step 1: $(1,1)$ *is an equilibrium.* If school 1 chooses $h_1 = 1$, the only possible best responses for 2 are $h_2 = 0$ or $h_2 = 1$ since, if $h_2 < 1$, 2 will lose all high types and 2 will gain more and more low types the closer it places to 0. Then, it is school 2's best response to locate at 1 if

$$\frac{1}{2} + \frac{e^{-\delta_{poor}}}{e^{-\delta_{poor}u_o} + 2e^{-\delta_{poor}}} > \frac{1}{1 + e^{-\delta_{poor}u_o} + e^{-\delta_{poor}}}.$$

We can see that when δ_{poor} is sufficiently low, this will be the case since the derivative of the left-hand side (LHS) with respect to δ_{poor} is always negative:

$$\begin{aligned}\frac{\partial LHS}{\partial \delta_{poor}} &= \frac{-e^{-\delta_{poor}}(e^{-\delta_{poor}u_o} + 2e^{-\delta_{poor}}) - e^{-\delta_{poor}}(-u_o e^{-\delta_{poor}u_o} - 2e^{-\delta_{poor}})}{(e^{-\delta_{poor}u_o} + 2e^{-\delta_{poor}})^2} \\ &= \frac{(u_o - 1)e^{-\delta_{poor}u_o}}{(e^{-\delta_{poor}u_o} + 2e^{-\delta_{poor}})^2} < 0 \end{aligned} \tag{16}$$

and the derivative of the right-hand side (RHS) with respect to δ_{poor} is always positive:

$$\frac{\partial RHS}{\partial \delta_{poor}} = \frac{-(-u_o e^{-\delta_{poor}u_o} - e^{-\delta_{poor}})}{(e^{-\delta_{poor}u_o} + e^{-\delta_{poor}} + 1)^2} > 0.$$

This shows that there is a single crossing in the profit functions from placing at 1 and 0, implying that there exists a δ^*_{poor} such that, for all $\delta_{poor} < \delta^*_{poor}$, 2's best response is to choose $h_2 = 1$. Since the best response functions of 1 and 2 are symmetric, for δ_{poor} sufficiently small $(1,1)$ is an equilibrium.

Step 2: $(1,1)$ *is unique.* If 1 chooses $h_1 \neq 1$, 2's best response function is $max(h_1 + \epsilon, 1 - (-u_o)^{1/2})$, $\epsilon > 0$, since 2 can take all the rich types and split the poor types by deviating ϵ above h_1 as long as the rich types prefer 2 to their outside option. Since 1 and 2 have symmetric best response functions, it is clear that for both schools to play their best response, at least one school must place at 1. In step 1, we showed that if one school locates at 1, the other school's best response is to also locate at 1 if δ_{poor} is sufficiently low. Therefore $(1,1)$ is the unique equilibrium. □

Proposition A1. *When* $N = 2$ *and* $\delta^*_{poor} < \delta_{poor} < \delta_{high}$, *where* δ^*_{poor} *is defined in the same way as above, there is no pure strategy equilibrium.*

Proof. If $\delta_{poor} > \delta^*_{poor}$, this implies that the school 2's best response to school 1's choice of 1 will be to choose 0. However, if school 2 chooses 0, it is no longer school 1's best response to choose 1 since school 1 can choose $h_1 = 1 - (-u_o)^{1/2}$, retaining all the rich types and gaining some of the poor types. However, if 1 chooses any h_1 besides 1, it will no longer be in school 2's best interest to choose $h_2 = 0$. Instead, 2 will choose a location $h_2 = \max(1 - (-u_o)^{1/2}, h_1 + \epsilon)$, $\epsilon > 0$, if $h_1 \neq 1$ and 0 otherwise. Since schools 1's and 2's best response functions are symmetric, we can see that there is no set of locations (h_1, h_2) such that both schools are playing their best responses, and there is no pure strategy equilibrium. □

Appendix C: Estimation of the Determinants of School Choice

In the first subsection of this appendix, I discuss how I estimate the parameters $\{\mathbf{\Gamma_z^{indiv}}, \delta_{poor}, \delta_{rich}, \zeta_{jn}\}$ from equation (5). In the second subsection, I discuss how I estimate the effect of fees, $\mathbf{\Gamma_z^{school}}$, on the utility of poor types.

Estimation of $\{\mathbf{\Gamma_z^{indiv}}, \delta_{poor}, \delta_{rich}, \zeta_{jn}\}$

I estimate the parameters $\{\mathbf{\Gamma_z^{indiv}}, \delta_{poor}, \delta_{rich}, \zeta_{jn}\}$ that maximize the log likelihood function

$$\mathcal{L} = \sum_{ijt} \mathbb{1}_{ijt} log(p_{ijt}), \tag{17}$$

where $\mathbb{1}_{ijt}$ is an indicator variable equal to 1 if a student i attends a school j in year t and p_{ijt} is the probability i attends j in year t given by equation (9). In practice, I do this using the Artelys Knitro package in matlab to minimize the negative log likelihood function. To reduce computational time, I provide the derivatives of equation (17) with respect to $\theta = \{\mathbf{\Gamma_z^{indiv}}, \delta_{poor}, \delta_{rich}, \zeta_{jn}\}$. For notational simplicity, let X_{ijt} also include $VA_{j,rich,n}$ and $VA_{j,poor,n}$. Then, the derivative of equation (17) with respect to the vector θ is

$$\sum_{ijt} \mathbb{1}_{ijt} \frac{1}{p_{ijt}} \Big(P(type_i = rich) \frac{\partial p_{ij,rich,t}}{\partial \theta} + (1 - P(type_i = rich)) \frac{\partial p_{ij,poor,t}}{\partial \theta} \Big), \tag{18}$$

where p_{ijzt} is the probability that a student i chooses j in year t conditional on that student being type z, and

$$\frac{\partial p_{ijzt}}{\partial \theta} = p_{ijzt} \Big(X_{ijzt} - \frac{\sum_k X_{ikt} e^{\theta X_{ikt}}}{\sum_k e^{\theta X_{ikt}}} \Big)$$

for the elements of θ that are in the utility function for type z and 0 for the remaining elements of θ. To ensure that I find the global maximum of equation (17), I estimate $\{\mathbf{\Gamma_z^{indiv}}, \delta_{poor}, \delta_{rich}, \zeta_{jn}\}$ with 20 randomly chosen start points and choose the parameter estimates that produce the largest value for the log likelihood function.

Finally, I estimate the standard errors using the fact that in general, for maximum likelihood estimation, $\sqrt{C}(\hat{\theta} - \theta^*) \rightarrow \mathcal{N}(0, I^{-1})$, where the information matrix $I(\theta)$ is given by the expectation of the outer-products of the first derivatives (given by equation (18)) of the log likelihood function and C is the number of observations (here, children). Therefore, the covariance matrix is:

$$\frac{1}{C}\left(\sum_i \frac{\partial \mathcal{L}}{\partial \theta}\frac{\partial \mathcal{L}'}{\partial \theta}\right)^{-1}.$$

Estimation of $\boldsymbol{\Gamma}_{\mathbf{z}}^{\mathbf{school}}$

To construct the instrument for school fees, I regress teacher salaries in private schools on teacher characteristics as follows:

$$salary_{ijt} = \boldsymbol{\Upsilon} Z_{it} + \eta_j + \alpha_t + \epsilon_{ijt},$$

where $salary_{ijt}$ is the salary of teacher i in school j in year t, η_j is a school fixed effect, α_t is a year fixed effect, and Z_{it} are teacher characteristics consisting of fixed effects for qualifications, experience, training, and age. I regress salaries on these characteristics to ensure that differences in the cost of teachers are not explained by differences in teacher quality, which could be related to ξ_{jn}. Then, I predict the residual:

$$\widehat{salary}_{ijt} = \widehat{\eta}_j + \widehat{\epsilon_{ijt}}.$$

For each village v, I create the leave-one out average measure $cost_v = \sum_{m \in T, m \neq v} \widehat{salary}_{ijt}$, where T is the set of villages in the same sub-district as v.

In equation form, the final instrument is then calculated from the regression

$$cost_v \times I_j^{private} = \rho_0 + \rho_1 I_j^{private} + \rho_2 cost_v + \mu_j, \tag{19}$$

where $I_j^{private}$ is an indicator variable equal to 1 if a school is private and the final instrument is the estimate of the residual, $\hat{\mu}_j$. Under the assumption that $\mu_j \perp \xi_j$, the moment conditions are given by

$$\Phi = \begin{pmatrix} \xi'_{jn}\hat{\mu}_j \\ \xi'_{jn}\hat{\mu}_j^2 \\ \xi'_{jn}1 \end{pmatrix}$$

where $\xi_{jn} = \zeta_{jn} - \boldsymbol{\Gamma}_{\mathbf{z}}^{\mathbf{school}} X_{jn}^{school} - c$, and c is a constant. To estimate c and $\boldsymbol{\Gamma}_{\mathbf{z}}^{\mathbf{school}}$, I again use Knitro. I first solve for the parameters that minimize $\hat{\Phi}'\hat{\Phi}$. Given these parameters, I estimate the optimal weighting matrix, C (for details, see Cragg (1983) and Hansen (1982)). Having estimated C, I re-estimate the parameters, minimizing $\hat{\Phi}' C \hat{\Phi}$. I calculate the standard errors using the standard "sandwich formula" for GMM.

Appendix D: Equilibrium Choice of Horizontal Quality Estimation

In Section 5, I estimate h_{jt} for every private school j in every year t and β using a two stage procedure. In the first stage, to estimate the h_{jt} that satisfy equation (11), I solve

$$\min_{h_{jt}} O_{jt}(h_{jt}) \quad \text{for each } jt,$$

where

$$\begin{aligned} O_{jt}(h_{jt}) =& \Big(\sum_{it} P(type_i = rich)\delta_{rich}(1-h_{jt})p_{ij,rich,t}(1-p_{ij,rich,t}) \\ &+ (1-P(type_i = rich))\delta_{poor}h_{jt}(p^2_{ij,poor,t} - p_{ij,poor,t})\Big)^2. \end{aligned}$$

In practice, this is implemented using the Artelys Knitro package in matlab.

To minimize computational time, I parallelize the loop through jt. Additionally, I provide the solver with the first derivative of the objective function, which is given by

$$\begin{aligned} \frac{\partial O_{jt}(h_{jt})}{\partial h_{jt}} =& 2O_{jt}(h_{jt})\Big(\sum_{it} -P(type_i = rich)\delta_{rich}p_{ij,rich,t}(1-p_{ij,rich,t}) \\ &+ \delta_{rich}(1-h_{jt})(\frac{\partial \rho_{ij,rich,t}}{h_{jt}} - 2p_{ij,rich,t}\frac{\partial \rho_{ij,rich,t}}{h_{jt}}) \\ &+ (1-P(type_i = rich))\delta_{poor}(p^2_{ij,poor,t} - p_{ij,poor,t}) \\ &+ (1-P(type_i = rich))\delta_{poor}h_{jt}(2p_{ij,poor,t}\frac{\partial \rho_{ij,rich,t}}{h_{jt}} - \frac{\partial \rho_{ij,rich,t}}{h_{jt}})\Big), \end{aligned}$$

where $\frac{\partial \rho_{ij,rich,t}}{h_{jt}} = \frac{\partial p_{ij,rich,t}}{h_{jt}}\frac{1}{2\beta}$ and $\frac{\partial \rho_{ij,poor,t}}{h_{jt}} = \frac{\partial p_{ij,poor,t}}{h_{jt}}\frac{1}{2\beta}$, and $\frac{\partial p_{ij,rich,t}}{h_{jt}}$ and $\frac{\partial p_{ij,poor,t}}{h_{jt}}$ are given by equation (10). Additionally, for some schools, the objective function is extremely small, leading to numerical instability. To address this problem, I scale the objective function and its derivative by $\frac{1}{O_{jt}(0)}$.

The estimation procedure for β in the second stage is straightforward. Following equation (12), I simply regress $\widehat{VA_{j,high,n}} - \widehat{VA_{j,low,n}}$ on $(2h_{jt} - 1)$, restricting the constant to be zero.

Appendix E: Counterfactual Estimation

This appendix provides additional details on the estimation of the social planners' choice of h_{jt} in the three counterfactuals.

Counterfactual 1: Removing the Spence Distortion. Define H to be set of h_{jt} belonging to private schools. The social planner chooses a vector of horizontal quality, $\mathbf{h^s_{jt}} \in H$ to solve

$$\mathbf{h^s_{jt}} = \max_{\mathbf{h_{jt}} \in H} W_1,$$

where W_1 is given by equation (14). To construct counterfactual values of $VA_{j,t,poor}$ and $VA_{j,t,rich}$ for each school, I need to identify v. I estimate v using the relationships given by equations (3) and (4). Given the estimates of h_{jt} from Section 5.4, each of these equations can be used to solve for an estimate of v, giving two different estimates. My final measure of v_{jt}, $\overline{v_{jt}}$, is then the average of the estimates from the two equations. Then, in the counterfactual, $VA_{j,t,poor} = \overline{v_{jt}} - \hat{\beta}(h_{jt})^2$ and $VA_{j,t,rich} = \overline{v_{jt}} - \hat{\beta}(1-h_{jt})^2$, where $\hat{\beta}$ is the estimate of β from Section 5.4. Then, for a given choice of h_{jt}, I can recalculate a school's $VA_{j,t,rich}$ and $VA_{j,t,poor}$ and solve for W_1.

In practice, I solve the social planner's maximization problem in Matlab using Kintro. To minimize computational time, I solve for each village in each year separately, since the village-year problems are separable. I also provide the solver with the first derivative of the social planner's problem. This is given by

$$\sum_i \frac{-2\beta(1-P(type_i = rich))}{\alpha_{poor}} \sum_{jt \in H} h_{jt} \frac{e^{\delta_{poor} VA_{j,poor,n} + \mathbf{\Gamma^{indiv}_{poor}} X^{indiv}_{ijt} + \zeta_{jn}}}{\sum_k e^{\delta_{poor} VA_{k,poor,n} + \mathbf{\Gamma^{indiv}_{poor}} X^{indiv}_{ijt} + \zeta_{jn}}}$$
$$+ \frac{2\beta P(type_i = rich)}{\alpha_{rich}} \sum_{jt \in H} (1-h_{jt}) \frac{e^{\delta_{rich} VA_{j,rich,n} + \mathbf{\Gamma^{indiv}_{rich}} X^{indiv}_{ijt} + \zeta_{jn}}}{\sum_k e^{\delta_{rich} VA_{k,rich,n} + \mathbf{\Gamma^{indiv}_{rich}} X^{indiv}_{ijt} + \zeta_{jn}}}.$$

Finally, I use 20 randomly chosen start points for each village-year to ensure that the optimal solution is found.

Counterfactual 2: The Social Planner Equally Weights School Quality for Rich and Poor. The procedure is the same as above except that the social planner chooses

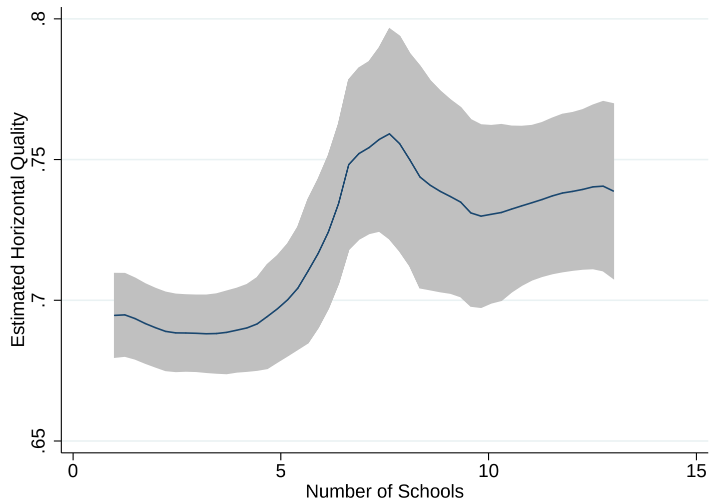
$\mathbf{h^s_{jt}} \in H$ to maximize W_2. The derivative of the objective function is now

$$\sum_i \Big(\frac{-2\beta(1-P(type_i = rich))}{\alpha_{poor}} \sum_{jt\in H} \Big(h_{jt} \frac{e^{\delta_{poor} VA_{j,poor,n} + \mathbf{\Gamma^{indiv}_{poor}} X^{indiv}_{ijt} + \zeta_{jn}}}{\sum_k e^{\delta_{poor} VA_{j,poor,n} + \mathbf{\Gamma^{indiv}_{poor}} X^{indiv}_{ijt} + \zeta_{jn}}}$$
$$+ \frac{(1-P(type_i = rich))}{\alpha_{poor}} \Big(-2\beta h_{jt} p_{ij,poor,t} \delta_{poor} + 2p^2_{ij,poor,t} h_{jt}\beta) VA_{j,poor,n} - 2h_{jt}\beta p_{ij,poor,t}$$
$$+ \sum_{k\neq j\in H} 2h_{jt}\beta p_{ik,low,t} VA_{k,poor,n}\delta_{low} \sum_{l\in H} p_{il,low,t} \Big)\Big)$$
$$+ \frac{2\beta P(type_i = rich)}{\alpha_{rich}} \sum_{jt\in H} (1-h_{jt}) \frac{e^{\delta_{rich} VA_{j,rich,n} + \mathbf{\Gamma^{indiv}_{rich}} X^{indiv}_{ijt} + \zeta_{jn}}}{\sum_k e^{\delta_{rich} VA_{k,rich,n} + \mathbf{\Gamma^{indiv}_{rich}} X^{indiv}_{ijt} + \zeta_{jn}}} \Big).$$

Counterfactual 3: Improved Sorting The procedure for identifying the socially optimal values of h_{jt} is the same as in counterfactual #1, except that δ_{poor} and δ_{rich} are now replaced with $m \times \widehat{\delta_{rich}}$. To get the change in welfare, I calculate welfare under the original allocation using a W_2-style formulation but modifying W_2 so that rich types can also value school quality with a different coefficient than the coefficient they use to sort (and using $\widehat{\delta_{poor}}$ and $\widehat{\delta_{rich}}$ to determine sorting). This modification is exactly the same as the modification of W_1 to allow poor types' value for school quality to be different than the coefficient they use for sorting in W_2. For the new allocation, which allows for improved sorting, both welfare measures will deliver the same value, and I calculate welfare where $\delta_{low} = \delta_{high} = m \times \widehat{\delta_{rich}}$. This ensures that welfare is calculated under the same preferences for both the original and counterfactual allocations.

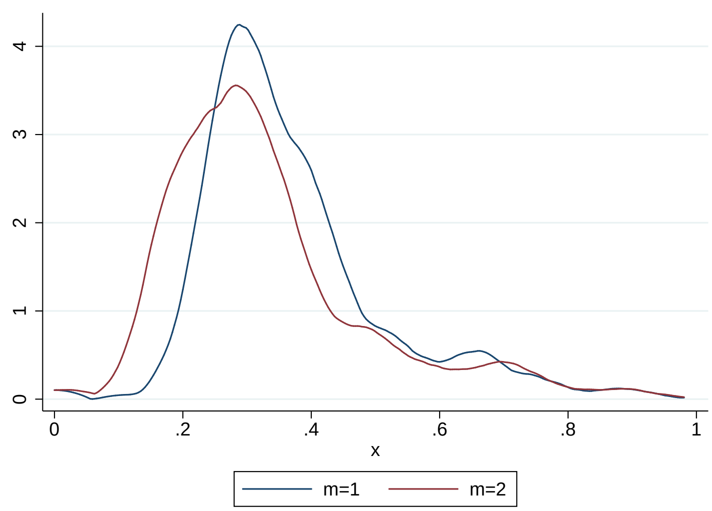
Appendix Figures

Figure A1: Relationship Between Horizontal Quality and the Number of Private Schools in the Market



This figure plots the relationship between the estimates of h_{jt} for private schools in the data and the number of private schools in the market in a village-year.

Figure A2: Distribution of Socially Optimal Horizontal Quality Under $m = 1$ and $m = 2$



This figure shows kernel density plots of private schools' socially optimal h_{jt} in counterfactual #3, where $\delta_{poor} = \delta_{rich} = m \times \widehat{\delta_{rich}}$.

Appendix Tables

Table A1: Summary Statistics for Public and Private Schools

	(1)	(2)	(3)	(4)	(5)	(6)
		Private			Public	
	Mean	SD	N	Mean	SD	N
Fee (Rupees)	1,360	963	1,166	11	155	1,928
Maximum Grade Offered	7.549	1.975	1,166	5.935	1.987	1,925
Student-Teacher Ratio	21.172	13.554	1,168	38.718	33.974	1,924
Has Library	0.392	0.488	1,168	0.224	0.417	1,930
Has Computer	0.266	0.442	1,168	0.010	0.101	1,930
Has Sports	0.349	0.477	1,168	0.110	0.313	1,930
Has Hall	0.195	0.397	1,168	0.069	0.253	1,930
Has Wall	0.962	0.190	1,168	0.658	0.474	1,930
Has Fans	0.942	0.233	1,164	0.476	0.500	1,926
Has Electricity	0.959	0.199	1,167	0.542	0.498	1,930
Number Permanent Classrooms	4.235	4.143	1,166	3.386	3.042	1,928
Number of Semi-Permanent Classrooms	1.854	2.990	1,168	0.664	1.526	1,929
Number of Staff Rooms	0.531	0.532	1,168	0.265	0.476	1,928
Number of Stores	0.428	0.571	1,168	0.269	0.623	1,929
Number of Toilets	0.668	0.851	1,168	0.315	0.744	1,929
Number of Blackboards	7.031	4.457	1,168	5.295	4.151	1,930

This table reports summary statistics for private and public schools in the LEAPS survey from 2004-2007. An observation is a school-year.

Table A2: Summary Statistics for Tested Students in Public and Private Schools

	(1)	(2)	(3)	(4)	(5)	(6)
		Private			Public	
	Mean	SD	N	Mean	SD	N
Female	0.451	0.498	14,202	0.450	0.498	28,499
Age	10.338	1.820	14,200	10.582	1.819	28,496
Grade	4.120	1.027	14,202	4.133	0.988	28,499
Mother Some Primary	0.481	0.500	14,202	0.299	0.458	28,499
Father Some Primary	0.748	0.434	14,202	0.579	0.494	28,499
Beds	0.998	0.044	14,202	0.996	0.060	28,499
Radio	0.632	0.482	14,202	0.541	0.498	28,499
TV	0.761	0.427	14,202	0.589	0.492	28,499
Refrigerator	0.600	0.490	14,202	0.322	0.467	28,499
Bicycle	0.746	0.435	14,202	0.713	0.452	28,499
Plough	0.250	0.433	14,202	0.222	0.416	28,499
Small Ag. Tools	0.697	0.460	14,202	0.720	0.449	28,499
Tables	0.952	0.213	14,202	0.855	0.352	28,499
Chairs	0.952	0.215	14,202	0.846	0.361	28,499
Fans	0.974	0.161	14,202	0.924	0.265	28,499
Tractor	0.159	0.366	14,201	0.115	0.319	28,499
Cattle	0.529	0.499	14,201	0.601	0.490	28,499
Goats	0.527	0.499	14,201	0.663	0.473	28,499
Chicken	0.573	0.495	14,201	0.649	0.477	28,499
Watches	0.972	0.166	14,201	0.958	0.202	28,499
Motor Rickshaw	0.040	0.196	14,201	0.039	0.193	28,499
Motorcycle	0.295	0.456	14,200	0.169	0.375	28,499
Car	0.124	0.329	14,201	0.048	0.215	28,499
Telephone	0.577	0.494	14,201	0.364	0.481	28,499
Tubewell	0.233	0.423	14,201	0.158	0.364	28,499
Math	0.376	0.829	14,202	-0.044	0.972	28,499
Urdu	0.429	0.857	14,202	-0.099	0.972	28,499
English	0.537	0.753	14,202	-0.205	0.939	28,499
Yearly Gains in Math	0.395	0.641	6,828	0.390	0.717	14,402
Yearly Gains in Urdu	0.432	0.581	6,828	0.443	0.670	14,402
Yearly Gains in English	0.350	0.580	6,828	0.391	0.694	14,402

This table reports summary statistics at the student-year level for the sample of tested students in private and government schools in the LEAPS survey from 2004-2007.

Table A3: Summary Statistics for Household Sample of Children Aged 5-15

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Private			Public			Not Enrolled	
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Female	0.434	0.496	2,892	0.479	0.500	6,851	0.593	0.491	4,169
Age	9.537	2.865	2,892	9.967	2.821	6,851	10.732	3.413	4,169
Distance to Current School (km)	0.484	0.687	2,892	0.724	0.927	6,851	–	–	–
Tables	0.717	0.451	2,892	0.623	0.485	6,851	0.475	0.499	4,169
Chairs	0.914	0.280	2,892	0.791	0.407	6,848	0.584	0.493	4,168
Fans	0.733	0.443	2,892	0.689	0.463	6,851	0.610	0.488	4,169
Sewing Machine	0.864	0.343	2,892	0.752	0.432	6,851	0.575	0.494	4,169
Air Cooler	0.153	0.360	2,892	0.074	0.262	6,851	0.042	0.201	4,169
Air Conditioner	0.268	0.443	2,892	0.247	0.431	6,851	0.220	0.414	4,169
Refrigerator	0.368	0.482	2,892	0.182	0.386	6,851	0.107	0.309	4,169
Radio	0.554	0.497	2,892	0.476	0.499	6,851	0.390	0.488	4,169
TV	0.524	0.500	2,892	0.419	0.493	6,851	0.318	0.466	4,169
VCR	0.112	0.315	2,892	0.051	0.220	6,851	0.036	0.186	4,169
Watches	0.908	0.290	2,892	0.907	0.290	6,851	0.849	0.358	4,169
Guns	0.129	0.336	2,892	0.071	0.256	6,851	0.052	0.222	4,169
Plough	0.173	0.378	2,892	0.134	0.340	6,851	0.105	0.306	4,169
Thresher	0.175	0.380	2,892	0.091	0.288	6,851	0.055	0.227	4,169
Tractor	0.092	0.289	2,892	0.059	0.236	6,851	0.042	0.201	4,169
Tubewell	0.221	0.415	2,892	0.166	0.372	6,851	0.133	0.339	4,169
Agricultural Machinery	0.263	0.440	2,892	0.239	0.427	6,851	0.206	0.404	4,169
Agricultural Hand Tools	0.660	0.474	2,892	0.657	0.475	6,851	0.618	0.486	4,169
Motorcycle	0.283	0.451	2,891	0.231	0.421	6,851	0.212	0.409	4,169
Car	0.084	0.278	2,889	0.032	0.176	6,851	0.030	0.169	4,169
Bicycle	0.615	0.487	2,891	0.635	0.482	6,851	0.590	0.492	4,169
Cows	0.654	0.476	2,891	0.715	0.451	6,851	0.710	0.454	4,169
Goats	0.485	0.500	2,891	0.570	0.495	6,851	0.602	0.490	4,169
Chickens	0.388	0.487	2,891	0.424	0.494	6,851	0.395	0.489	4,169

This table reports summary statistics at the child-year level for children aged 5-15 in the 1,740 surveyed households in the LEAPS survey from 2004-2007.

Table A4: Association Between Number of Private Schools and Village Characteristics

	(1) 1981 Population	(2) 1998 Population	(3) % Pop. Change	(4) Mean Assets	(5) Percent Own Land	(6) Gini Coefficient
num_pri_{vt}	248.879**	373.904*	0.009	0.017	-0.001	0.001
	(108.864)	(210.127)	(0.025)	(0.036)	(0.010)	(0.008)
Year FE	Y	Y	Y	Y	Y	Y
Village FE	N	N	N	Y	Y	Y
Number of observations	448	448	436	448	336	448
Clusters	112	112	109	112	112	112
Adjusted R^2	0.146	0.119	-0.005	0.964	0.841	0.265

This table reports the association between the number of private schools in a village and village-level measures of population in 1981, population in 1998, change in population between 1981 and 1998, wealth, and inequality. In all columns, an observation is a village-year. In columns 1-3, the outcome data comes from the 1981 and 1998 Punjab population censuses. In column 4, mean assets is generated by conducting a principal components analysis of indicator variables for asset ownership and predicting the first component. The outcome, a proxy for wealth, is the village-year mean of this asset measure. In column 5, the outcome is the percent of surveyed households who reported owning land. There are fewer observations for this outcome because the survey did not include information about land ownership in round 2. In column 6, the outcome is the Gini coefficient for the village-year, generated using the wealth measure from the principal components analysis. Standard errors are clustered at the village level. *, **, and *** denote 10, 5, and 1% significance respectively.

Table A5: Effect of Private School Entry on Inequality in Test Scores as Measured by the Gap Between the 90th and 10th Percentile

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	90-10 Gap in Math		90-10 Gap in English		90-10 Gap in Urdu		90-10 Gap in Mean	
	Private	Public	Private	Public	Private	Public	Private	Public
num_pri_{vt}	0.168***	0.026	0.230***	0.102	0.098***	0.075	0.216***	0.064
	(0.058)	(0.049)	(0.087)	(0.066)	(0.037)	(0.064)	(0.073)	(0.054)
Village FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Number of observations	430	447	430	447	426	447	430	447
Clusters	110	112	110	112	110	112	110	112
Adjusted R^2	0.365	0.407	0.307	0.359	0.246	0.256	0.333	0.374

This table reports estimates of the effect of the number of private schools in the village on the gap in test scores between the 90th and 10th percentile student in the public and private sectors separately. An observation is the regression is a village-year. The regressions use data from the LEAPS tested sample. Standard errors are clustered at the village level. The number of clusters differs by sector because not all villages have a private school in operation during the study period. *, **, and *** denote 10, 5, and 1% significance respectively.

Table A6: Out of Sample Validation of Type-Specific School Value-Addeds

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Math Score		English Score		Urdu Score		Mean Score	
	Poor Type	Rich Type	Poor Type	Rich Type	Poor Type	Rich Type	Poor Type	Rich Type
$Math\,VA_{poor}$	0.881***	-0.149						
	(0.083)	(0.092)						
$Math\,VA_{rich}$	0.019	1.139***						
	(0.064)	(0.079)						
$English\,VA_{poor}$			0.896***	-0.088				
			(0.069)	(0.078)				
$English\,VA_{rich}$			0.014	1.071***				
			(0.050)	(0.064)				
$Urdu\,VA_{poor}$					0.967***	-0.089		
					(0.075)	(0.086)		
$Urdu\,VA_{rich}$					0.041	1.028***		
					(0.059)	(0.107)		
$Mean\,VA_{poor}$							0.996***	-0.128
							(0.091)	(0.107)
$Mean\,VA_{rich}$							0.001	1.207***
							(0.061)	(0.089)
Grade × Lagged Test Score Controls	Y	Y	Y	Y	Y	Y	Y	Y
Number of observations	1,242	708	1,242	708	1,242	708	1,242	708
Clusters	105	101	105	101	105	101	105	101
Adjusted R^2	0.628	0.598	0.640	0.665	0.670	0.690	0.717	0.721

This table reports the coefficients for regressions of the test scores of rich and poor students who change schools on their new school's value-addeds for rich and poor types, controlling for the grade-specific effects of lagged test scores and lagged test scores squared, year fixed effects interacted with whether the village had the report card program, and district fixed effects. The type-specific value-addeds were calculated leaving out all students who switched schools. Standard errors are clustered at the village level. *, **, and *** denote 10, 5, and 1% significance respectively.

Table A7: Tests for Pre-Trends in the Effect of Number of Private Schools on Student Outcomes

	(1) Math	(2) English	(3) Urdu	(4) Mean
$\mathbb{1}_{rich} \times num\_pri_{v,t+1}$	0.025	-0.014	-0.015*	-0.001
	(0.021)	(0.013)	(0.008)	(0.010)
$num\_pri_{v,t+1}$	-0.025	0.001	0.002	-0.007
	(0.025)	(0.010)	(0.006)	(0.013)
Peer Controls	Y	Y	Y	Y
Student-Teacher Ratio Controls	Y	Y	Y	Y
Number of Private Schools Controls	Y	Y	Y	Y
Lagged Test Score Controls	Y	Y	Y	Y
School by Type FE	Y	Y	Y	Y
Grade by Type FE	Y	Y	Y	Y
Year by Type FE	Y	Y	Y	Y
Number of observations	2,622	2,622	2,622	2,622
Clusters	104	104	104	104
Adjusted R^2	0.564	0.625	0.631	0.713

This table reports estimates of the effect of the forward lagged number of private schools in the village on test scores. The regressions use data from the LEAPS tested sample. Lagged test score controls consist of the relevant lagged test score and its square interacted with grade by type fixed effects. Number of private schools controls consist of num_pri_{vt} and its interaction with $\mathbb{1}_{rich}$. The peer controls consist of the school-level mean and variance of lagged test scores in year t, as well as the percent of rich types in a school, all of which are allowed to have different effects for rich and poor types. The student-teacher ratio controls consist of a control for the school's student-teacher ratio in year t, which is allowed to have different effects for rich and poor types. Observations are at the student-year level. Standard errors are clustered at the village level. *, **, and *** denote 10, 5, and 1% significance respectively.

Table A8: Effect of the Number of Private Schools on Student Outcomes for Students Who Always Attend Private Schools

	(1) Math	(2) English	(3) Urdu	(4) Mean
$\mathbb{1}_{rich} \times num\_pri_{vt}$	0.136*	0.074	0.067	0.096*
	(0.070)	(0.059)	(0.060)	(0.053)
num_pri_{vt}	-0.124***	-0.061	-0.046	-0.076*
	(0.043)	(0.064)	(0.044)	(0.039)
Peer Controls	Y	Y	Y	Y
Lagged Test Score Controls	Y	Y	Y	Y
School by Type FE	Y	Y	Y	Y
Grade by Type FE	Y	Y	Y	Y
Year by Type FE	Y	Y	Y	Y
Number of observations	5,910	5,910	5,910	5,910
Clusters	108	108	108	108
Adjusted R^2	0.592	0.610	0.650	0.712

This table reports estimates of the effect of the number of private schools in the village on test scores for students who always attend private schools during the sample period and who are observed in school prior to the exit or entry event. The regressions use data from the LEAPS tested sample. Lagged test score controls consist of the relevant lagged test score and its square interacted with grade by type fixed effects. The peer controls consist of school-level mean and variance of lagged test scores in year t, as well as the percent of rich types in a school, all of which are allowed to have different effects for rich and poor types. Standard errors are clustered at the village level. *, **, and *** denote 10, 5, and 1% significance respectively.

Table A9: Knowledge of Educational Quality and Determinants of School Choice

	(1) Changed Schools	(2) Knows Teacher's Name	(3) Chose School for Distance	(4) Chose School for Quality	(5) Mean School VA
$P(type_i = rich)$	0.197***	0.156***	-0.163	0.377***	0.090
	(0.058)	(0.047)	(0.102)	(0.097)	(0.072)
$rank_{ij}$					0.035***
					(0.009)
$P(type_i = rich) \times rank_{ij}$					0.034*
					(0.020)
Mean	0.347	0.532	0.427	0.210	0.010
Observation Level	Child	Child-Year	Child	Child	Parent-School-Year
Number of observations	5,621	13,645	2,873	2,873	21,270
Clusters	1,696	1,687	1,155	1,155	6845
Adjusted R^2	0.003	0.002	0.002	0.017	0.030

This table reports descriptive statistics on rich and poor types' knowledge of educational markets and the determinants of their enrollment decisions in the household survey data. Column 1 regresses an indicator variable for changing schools at least once over the course of the study period on the probability of being a rich type for children who were always enrolled in school; each observation is a child. Column 2 regresses an indicator variable for whether a parent knows a child's teacher's name on the probability of being a rich type; an observation is a child-year. Column 3 regresses an indicator variable for if a parent reports distance is the main reason they choose their child's school on the probability of being a rich type. An observation is a child, since the question was only asked in round 1. Column 4 regresses an indicator variable for if a parent reports quality is the main reason they chose their child's school on the probability of being a rich type. As before, an observation is a child. All standard errors for columns 1-4 are clustered at the household level. Column 5 regresses a school's expected value-added for a household on parents' assessment of the school's quality, the household's probability of being a rich type, and their interaction. Each observation is at the parent-school-year level, and standard errors are clustered at the school level. *, **, and *** denote 10, 5, and 1% significance respectively.