ONLINE APPENDIX

Upping the Ante: The Equilibrium Effects of Unconditional Grants to Private Schools

T. Andrabi, J. Das, A.I. Khwaja, S. Ozyurt, and N. Singh

Contents

A Sample, Data, and Additional Results

A.1 Sample

Sampling Frame

Villages: Our sampling frame includes any village with at least two non-public schools, i.e. private or NGO, in rural areas of Faisalabad district in the Punjab province. The data come from the National Education Census (NEC) 2005 and are verified and updated during field visits in 2012. There are 334 eligible villages in Faisalabad, comprising 42 percent of all villages in the district; 266 villages are chosen from this eligible set to be part of the study based on power calculations.

Schools: Our intervention focuses on the impact of untied funding to non-public schools. The underlying assumption here is that a school owner or manager has discretion over spending in their own school. If instead the school is part of a network of schools and is centrally managed, as is the case for certain NGO schools in the area, then it is often unclear how money is allocated across schools in the network. Therefore, we decided to exclude schools in our sample where we could not obtain guarantees from officials that the money would be spent only on the randomly selected schools. In practice, this was a minor concern since it only excluded 5 schools (less than 1 percent of non-public schools) across all 266 villages from participation in the study. The final set of eligible schools for participation in the study was 880.

Study Sample

All eligible schools that consented to participate across the 266 villages are included in the final randomization sample for the study. This includes 822 private and 33 NGO schools, for a total of 855 schools; there were 25 eligible schools (about 3 percent) that refused to participate in either the ballot or the surveys. The reasons for refusals were: impending school closure, lack of trust, survey burden, etc. Appendix Figure A1 summarizes the number of villages and schools in each experimental group.

Power Calculations

We used longitudinal LEAPS data for power calculations and were able to compare power under various randomization designs. Given high auto-correlation in school revenues, we chose a stratified randomization design, which lowers the likelihood of imbalance across treatment arms and increases precision since experimental groups are more comparable within strata than across strata (Bruhn and McKenzie, 2009). The sample size was chosen so that the experiment had 90 percent power to detect a 20 percent increase in revenue for H schools, and 78 percent power for the same percentage increase in revenue for L^t schools (both at 5% significance level).

A.2 Data

We use data from a range of surveys over the project period. We outline the content and the respondents of the different surveys below. For the exact timing of the surveys, please refer to Appendix Figure A2.

Survey Instruments

Village Listing: This survey collects identifying data such as school names and contact numbers for all public and private schools in our sampling frame.

School Survey Long: This survey is administered twice, once at baseline in summer 2012 and again after treatment in the first follow-up round in May 2013. It contains two modules: the first module collects detailed information on school characteristics, operations and priorities; and the second module collects household and financial information from school owners. The preferred respondent for the first module is the operational head of the school, i.e. the individual managing day-to-day operations; if this individual was absent the day of the survey, either the school owner, the principal or the head teacher could complete the survey. For the second module, the preferred respondent was either the legal owner or the financial decision-maker of the school. In practice, the positions of operational head or school owner are often filled by the same individual.

School Survey Short: This survey is administered quarterly between October 2013 and December 2014, for a total of four rounds of data. Unlike the long school survey, this survey focuses on our key outcome variables: enrollment, fees, revenues and costs. The preferred respondent is the operational head of the school, followed by the school owner or the head teacher. Please consult Appendix Figure A3 to see the availability of outcomes across rounds.

Child Tests and Questionnaire: We test and collect data from children in our sample schools twice, once at baseline and once after treatment in follow-up round 3. Tests in Urdu, English and Mathematics are administered in both rounds; these tests were previously used and validated for the LEAPS project (Andrabi et al., 2002). Baseline child tests are only administered to a randomly selected half of the sample (426 schools) in November 2012. Testing is completed in 408 schools for over 5000 children, primarily in grade 4.[1] If a school had zero enrollment in grade 4 however, then the preference ordering of grades to test was grade 3, 5, and then 6.[2] A follow-up round of testing was conducted for the full sample in January 2014. We tested two grades between 3 and 6 at each school to ensure that zero enrollment in any one grade still provided us with some test scores from every school. From a roster of 20,201 enrolled children in this round, we tested 18,376 children (the rest were absent). For children tested at baseline, we test them again in whichever grade they are in as long as they were enrolled at the same school. We also test any new children that join the baseline test cohort. In the follow-up round, children also complete a short survey, which collects family and household information (assets, parental education, etc.), information on study habits, and self-reports on school enrollment.

Teacher Rosters: This survey collects teacher roster information from all teachers at a school. Data include variables such as teacher qualifications, salary, residence, tenure at school and in the profession. It was administered thrice during the project period, bundled with other surveys. The first collection was combined with baseline child testing in November 2012, and hence data was collected from only half of the sample. Two follow-up rounds with the full sample took place in May 2013 (round 1) and November 2014 (round 5).

Investment Plans: These data are collected once from the treatment schools as part of the disbursement activities during September-December 2012. The plans required school

[1] The remaining schools had either closed down (2), refused surveying (10) or had zero enrollment in the tested grades at the time of surveying (6). The number of enrolled children is 5611, of which 5018 children are tested; the remaining 11% are absent.

[2] 97 percent of schools (394/408) had positive enrollment in grade 4.

owners to write down their planned investments and the expected increase in revenues from these investments— whether through increases in enrollment or fees. School owners also submitted a desired disbursement schedule for the funds based on the timing of their investments.

Variable Description

The table below lists, defines and provides the data source for key variables in our empirical analysis. Group A are variables measured at the village level; Group B at the school level; and Group C at the teacher level.

Variable	**Description**	**Survey Source**
Group A: Village Level		
Grant per capita	Grant amount per private school going child in treatment villages. For L villages, this is Rs 50,000/total private enrollment, and for H villages, this equals (50,000*# of private schools in village)/total private enrollment. Control schools are assigned a value of 0.	School
Group B: School Level		
Closure	An indicator variable taking the value '1' if a school closed during the study period.	School
Refusal	An indicator variable taking the value '1' if a school refused a given survey. A closed school is assigned a missing value.	
Enrollment	School enrollment in all grades, verified through school registers. Coded as 0 after school closure, unless otherwise stated.	School
Fees	Monthly tuition fees charged by the school averaged across all grades. Coded as missing after school closure, unless otherwise stated.	School
Posted Revenues	Sum of revenues across all grades obtained by multiplying enrollment in each grade by the monthly fee charged for that grade. Coded as 0 after school closure, unless otherwise stated.	School
Collected Revenues	Self-reported measure on total monthly fee collections from all enrolled students. Coded as 0 after school closure, unless otherwise stated.	School
Test Scores	Child test scores in English, Math and Urdu, are averaged across enrolled children to generate school-level test scores in these subjects. Tests are graded using item response theory (IRT), which appropriately adjusts for the difficulty of each question and allows for comparison across years in standard deviation units. Variable is missing for closed schools, unless otherwise missing	Child tests
Stayer	A stayer is a child who self-reports being at the same school for at least 18 months in round 3.	Child survey

Continued on next page

Continued from previous page

Variable	**Description**	**Survey Source**
Fixed Expenditures	Sum of spending on infrastructure (construction/rental of a new building, additional classroom, furniture and fixtures), educational materials, and other miscellaneous items in a given year. Data is collected at the item level, e.g. furniture, equipment, textbooks etc. Coded as missing after school closure, unless otherwise stated.	School
Variable Expenditures	Sum of spending on teacher salaries, non-teaching staff salaries, rent and utilities for a given month. Coded as missing after school closure, unless otherwise stated.	School
Sources of school funding (Y/N)	Indicator variables for whether school items were purchased through (i) self-financing- school fees or owner's own household income, or (ii) credit- loans from a bank or MFI	School
Household borrowing (Y/N)	Indicator variables for borrowing behavior of the school owner's household: whether household ever borrowed from any sources, formal sources (e.g. bank, MFI) and informal (e.g. family, friend, pawnshop, moneylender) sources. Coded as missing after school closure, unless otherwise stated.	School owner
Household borrowing: Loan value	Value of total borrowing in PKR by the owner household from any source for any purpose. Coded as missing after school closure, unless otherwise stated.	School owner
Group C: Teacher Level		
Teacher salaries	Monthly salary collected for each teacher present during survey.	Teacher roster
Teacher start date	YYYY-MM at which the teacher started her tenure at the school. This allows us to tag a teacher as a newly arrived or an existing teacher relative to treatment date.	Teacher roster

A.3 Weighting with unequal selection probabilities

Saturation Weights

Our experimental design is a two-stage randomization. First, villages are assigned to one of three groups: Pure Control; High-saturation, H; and Low-saturation, L; based on power calculations, $\frac{3}{7}$ of the villages are assigned to the L arm, and $\frac{2}{7}$ each to the H arm and the control group. Second, in the L arm, one school in each village is further randomly selected to receive a grant offer; meanwhile, all schools in H and no school in control villages receive grant offers. This design is slightly different from randomization saturation designs that have been recently used to measure spillover effects (see Crépon et al., 2013; Baird et al., 2016) since the proportion of schools that receive grant offers is not randomly assigned within L villages. Instead, since we are interested in examining what happens when a single school is made the grant offer, the proportion of schools within L villages assigned to treatment depends on village size at the time of treatment; this changes the probability of selection into treatment for all schools in these villages. For instance, if a L village has 2 schools, then probability of treatment is 0.5 for a given school, whereas if the village has 5 schools, the selection probability reduces to 0.20.

While this consideration does not affect the estimates for the H arm, the impact for schools in the L arm need to adjust for this differential selection probability. This can be done by constructing appropriate weights for schools in the L villages. Not doing so would overweight treated schools in small villages and untreated schools in large villages. Following the terminology in Baird et al. (2016), we refer to the weights given below as saturation weights, s_g where g represents the treatment group:

- $s_{high} = s_{control} = 1$
- $s_{lowtreated} = B$, where B is the number of private schools in the village
- $s_{lowuntreated} = \frac{B}{B-1}$

To see why weighting is necessary, consider this example. Assume we are interested in the following unweighted simple difference regression: $Y_{ij} = \alpha + \beta T_{ij} + \epsilon_{ij}$, where i indexes a school in village j; T_{ij} is a treatment indicator that takes value 1 for a treated school in L villages and 0 for all control schools. That is, we are only interested in the difference in outcomes between low-treated and control schools. Without weighting, our treatment effect is the usual $\beta = [E(TT')]^{-1}E(TY)$.

If instead we were to account for the differential probability of selection of the low-treated schools, we would weight these observations by B and control observations by 1. This weighting transforms the simple difference regression as follows: $\tilde{Y_{ij}} = \tilde{\alpha} + \beta_0 \tilde{T_{ij}} + \tilde{\epsilon_{ij}}$, and our $\beta_0 = [E(\tilde{T}\tilde{T}')]^{-1}E(\tilde{T}\tilde{Y})$, where $\tilde{T}$ and $\tilde{Y}$ are obtained by multiplying through by $\sqrt{B_j}$ where B_j is the weight assigned to the low-treated observation based on village size. Note that the bias from not weighting is therefore more severe as village size increases. However, since our empirical village size distribution is quite tight (varying only between 1 and 9 private schools), in practice, weighting does not make much of a difference to our results.

While we must account for weights to address the endogenous sampling at the school level in the low-saturation treatment, we do not need weights to account for the unequal probability of village level assignment in the first stage since this assignment is independent of village characteristics. Nevertheless, if we were to do so, our results are nearly identical. The weights in this case would be as follows:

- $s_{high} = s_{control} = \frac{7}{2}$
- $s_{lowtreated} = \frac{7}{3}B$
- $s_{lowuntreated} = \frac{7}{3}\frac{B}{B-1}$

Tracking Weights

In addition to the saturation weights, tracking weights are required to account for the randomized intensive tracking procedure used in round 5. These weights are only used for regressions containing data from round 5; regressions using data from rounds 1-4 only require saturation weights. We implemented this randomized tracking procedure in order to address attrition concerns, which we expected to be more severe two years after treatment. We describe below the details of the procedure and specify the tracking weights for round 5 data.

In round 5, 60 schools do not complete surveys despite being operational. We randomly select half of these schools to be intensively tracked, i.e. our enumerators make multiple visits to these schools to track down the respondent, and, if necessary, survey the respondents over the phone or at non-school premises. These efforts increase our round 5 survey completion rate from 88 to 94 percent. To account for the additional attention received by this tracked subsample, we assign a weight of 2 if the school was selected to be part of the intensively tracked subsample, and 0 if it was not.

A.4 Private and Social Returns Calculations

In this section, we describe our internal rate of return (IRR) and welfare calculations. We caution that these estimates are speculative in nature and are intended to provide qualitative insights as to the trade-offs between adopting a low vs. high-saturation financing approach.

IRR and Loan-loss guarantee

In this exercise, we aim to establish whether lenders would be willing to lend to schools in this sector. We conduct two types of IRR calculations and then assess whether schools would be able to pay back a Rs.50,000 loan at 15% interest rate. We calculate: (i) Returns over a 2 year period with resale of assets at 50% value at the end of the term; and (ii) Returns over a 5 year period with no resale of assets. We use the three different estimates of collected revenues and variable expenditures shown in Appendix Table B19 that treat closure differently to do these calculations[3]; since there are no effects on variable expenditures for L^t schools, we use zeros in those calculations instead of the point estimates from the table.[4] In addition, we also consider fixed expenditures

[3] In particular, we either code closed schools as having: (i) missing revenues and expenditures ('closed-as-missing'); or (ii) zero revenues and zero expenditures ('closed-as-0'); or (iii) imputed values for revenues and expenditures based on trends in open schools in the control group ('closed-as-imputed'). In the third approach, we regress an outcome variable on baseline covariates, strata, and round fixed effects for open schools in the control villages and use the coefficients from the regression to predict and "fill in" the outcome variables for closed schools in the H, L^u and control groups.

[4] If instead, we used the exact point estimates for variable expenditures for the L^t schools, we would get slightly lower IRR for our closed-as-0 approach and even larger IRR for the other two approaches (as the point estimates for variable expenditures are negative).

for assets purchased in year 1, making the same adjustments for closure each time.[5] The table below shows the range of IRR calculations. To calculate the IRR for our pooled treatment, we simply take a weighted average of the IRR for the other two treatments.

Table: Internal Rate of Return (IRR) Estimates

	Closed-as-missing		Closed-as-0		Closed-as-imputed	
	(1) 2-year	(2) 5-year	(3) 2-year	(4) 5-year	(5) 2-year	(6) 5-year
Low-treated	92%	114%	147%	166%	84%	106%
High	10%	30%	27%	48%	26%	47%
Pooled Treatment	37%	58%	67%	87%	45%	67%

These IRR figures compare very favorably to the prevailing market interest rates in Pakistan, which range from 15-20%, suggesting that this may be a profitable lending sector. If we look at the H and L approaches separately, the IRR estimates for the L approach are always larger, though even the H approach gives estimates just at or above market rates. By our most conservative estimates, if we were to offer our grant as a Rs.50,000 loan at 15% interest rate, it would take an L^t school 1 year to pay off the loan and an H school 3 years to pay off their loan.

These rates of return along with the lower default rates may lead the lender to prefer the L over the H approach, unless the fixed cost of visiting three villages (versus one) is much higher. A social planner who cares about child test scores may however prefer the H approach. To incentivize the H approach, the social planner could subsidize the lender based on the expected losses from defaults in a manner that makes the lender indifferent between the L and H approaches.

We calculate this subsidy amount as follows. We first note that closure rates are differential across the L^t and H groups by 7 percentage points. The closure rate in L^t group is 1% and 8% for the H group. If we conservatively assume that closed schools would default on their loans completely, then we can estimate the expected loss that would make a lender indifferent. The expected loss for a given school in L^t group is Rs.575, while it is Rs.5,800 for a H school. Therefore, for every Rs.150K given out in loans, the social planner would need to subsidize the lender by Rs.15,675 to make them indifferent between the two approaches.

Welfare Calculations

While a complete welfare calculus is beyond the scope of this paper, we document changes for four beneficiary groups from our intervention: school owners, teachers, parents and children. We compare gains from a *total* grant of PKR 150K under two different financial saturations— the L arm where we give PKR 50K to one school in three villages ('the L policy'), and the H arm where each school in one village receives PKR 50K ('the H policy').

[5] Our estimates for fixed expenditures (top coded at the 99th percentile) are as follows. For closed-as-missing, we have we have Rs.34,851 for H schools and Rs.28,867 for L^t schools. For closed-as-0, we have Rs.34,455 for H schools and Rs.31,184 for L^t schools. Finally, for closed-as-imputed, we have we have Rs.34,221 for H schools and Rs.28,653 for L^t schools. P-values are always below 0.05 in all these specifications.

School Owners: We consider profits, estimated as actual collected revenues minus variable expenditures via a seemingly unrelated regression approach, as the gains for school owners. Appendix Table B19 provides these estimates for H and L^t schools under different assumptions regarding school closure; we either code closed schools as missing, or as zero, or use imputed data. The L policy would garner profits ranging from Rs.206K (closed-as-imputed) to Rs.243K (closed-as-0), whereas the H policy would give profits that are lower by an order of magnitude and ranging from Rs.62K (closed-as-missing) to Rs.84K (closed-as-0).

Teachers: We use changes in the teacher wage bill to understand how the intervention affected the teacher market. Recall from Table 8 that we do not observe significant overall changes in number of teachers employed by schools, but do observe teacher churn in the H arm. Under the assumption that this churn arises simply from switches in employment status for teachers, we can use these estimates of wage gains to compute changes in teacher welfare. As with gains for school owners, we can estimate a range of effects depending on how we treat school closure. Following the same approach of varying assumptions on school closure, we find that the average wage bill under the H policy ranges from Rs.81K (closed-as-0) to Rs.99K (closed-as-missing). In contrast, gains for teachers under the L policy are much lower ranging from Rs.-30K (closed-as-missing) to Rs.8K (closed-as-0). These numbers suggest that gains are partially transferred from school owners to teachers in the H relative to the L policy.

Parents: Calculating consumer surplus requires some strong assumptions on the demand function. These assumptions include: (i) the demand curve can be approximated as linear; and (ii) there is an upper bound to demand at zero price because of the reasonable assumption of 'closed' markets in our context. If one is willing to make such assumptions, one can use the following approach to estimate consumer surplus.

Since quality does not change in the L arm, our treatment effects arise from a movement along the demand curve. We derive this linear demand curve using two points from our experiment— the baseline price-enrollment (PQ) combination of (238, 164), denoted by (P_0, Q_0) in the figure, and the L^t PQ-combination, denoted by (P_L, Q_L). For P_L, we use posted fees, weighted by the number of children (the H effect is Rs.13, p-value 0.243, and the L^t effect is -Rs1.5, p-value 0.894). The choice of Q_L ranges from 11 to 22 children depending on how we treat school closures. These two points allow us to generate linear demand curves and calculate consumer surplus gains. In particular, we can calculate the baseline consumer surplus, the triangle CS_0, and the additional surplus gain in L^t from movement down the demand curve. This additional surplus is calculated as the difference in areas of the two triangles generated by the baseline and L^t PQ-combinations. For a total 150K in grants across three villages, the annual increase in CS ranges from Rs.9200-Rs.9500. The increase in consumer surplus in L^t is driven by the fee reduction faced by the inframarginal children, as the newly enrolled, 'marginal,' children are at the cusp of indifference before the intervention.

For the H arm, we see test score gains accompanied by fee increases. This implies a movement of the demand curve. Given our earlier assumption of an upper bound on demand arising from closed markets, an increase in quality pivots our baseline demand curve outward. We use our H estimates to obtain a new, pivoted demand curve, where fees increase by Rs.13 and enrollment by 7-9 children depending on how school closures are treated. The annual gains in consumer surplus from a grant investment of RS.150K range from Rs.5900 to Rs.10,300.

As we see, our results on consumer surplus are quite sensitive to the point estimates we use and there are scenarios where each of the H or L policy performs better than the other. This suggests that once we account for uncertainty, these differences will be indistinguishable. Therefore, the main takeaway from these calculations is that unlike the gains for schools owners and teachers, consumer surplus under both approaches is qualitatively similar.

Children: We measure benefit to children in terms of test score gains. Conceptually, there are two types of children we need to consider: (i) children that remain at their baseline schools, and (ii) children that newly enroll at the school.

We know from Appendix Table B11 that the H schools shows test score increases for children who remain enrolled at their baseline schools. Because welfare requires us to care about the average *child* in the village, rather than the average *school*, we run the same regression as in Appendix Table B11, Column 1, at the child level. We find that the average child in H schools shows a gain of 0.22sd (p-value=0.017), and an average child in L^t schools gains 0.10sd (p-value=0.33), and this difference is significant at p-value of 0.24. In particular, considering a total baseline enrollment of 492 children from 3 schools, our H child test score gains suggest a total increase of 108.2sd under the H policy. In comparison, the total gain under the L policy is substantially lower at 49.2sd, even if we conservatively take the (statistically insignificant) 0.10sd coefficient at face value.

For newly enrolled children, we rely on our previous work, Andrabi et al. (2011), showing test score gains of 0.33sd for children who switch from public to private schools.[6] In H villages, this leads to a total test score gain of 8.9 standard deviations as each of the three schools gains 9 children (0.33sd*9*3). For the L^t sample, each school gains 22 children (Table 5, Column 1), which means a total increase of 66 children across 3 villages, and a total test score increase of 21.8sd (0.33*22*3). These calculations rely on estimates of new enrollment where closed schools are coded to have 0 enrollment, as in Table 5. If we instead used enrollment figures with closed-as-missing, we obtain gains of 6.9sd (0.33sd*7*3) under the H policy, and 12.9sd (0.33*13*3) from the L policy. Using imputed values for closed schools instead gives us gains of 6.9sd and 10.9sd under the H and L policy, respectively.

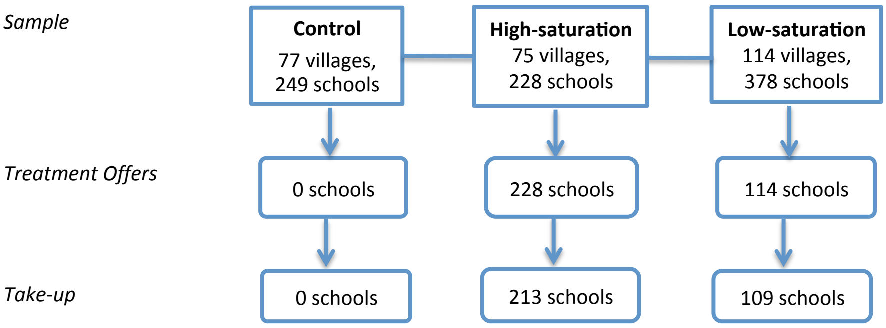
Summing the gains for already and newly enrolled children, we obtain a total sd gain of between 115.1-117.1 for H and 60.1-71 for L approaches.

These calculations assume that test score gains accrue to children across all grades, which may be reasonable given that fee increases are observed across grades (see Appendix Table B9). Using the same method, if we instead restrict to the tested children in grades 3-5, we obtain a total increase of 30.4sd in H compared with a 16.9sd increase in L^t.

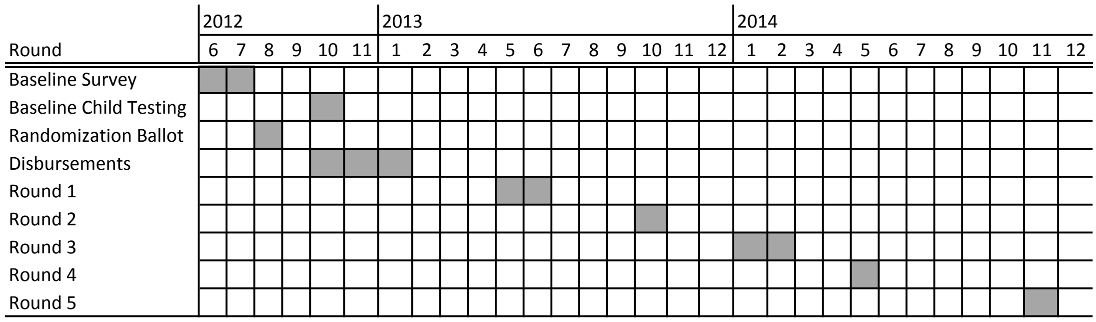
[6] Our current study was not designed to estimate the effects for newly enrolled children since it would have been enormously expensive to test all enrolled children in each public and private school in the village, and identifying marginal movers for testing at baseline is a difficult, if not impossible, task.

A.5 Figures

Appendix Figure A1: Sample Details



Appendix Figure A2: Project Timeline



Notes: Rounds 1-3 correspond to the first school year after treatment and rounds 4 and 5 refer to the second school year after treatment. A school year in this region is typically from April-March, with a three month break for summer between June-August.

Appendix Figure A3: Data Availability by Survey Rounds

Outcome	Baseline	Round 1	Round 2	Round 3	Round 4	Round 5
Enrollment	✓	✓	✓	✓	✓	✓
Fees	✓	✓	✓	✓	✓	✓
Posted Revenues	✓	✓	✓		✓	
Collected Revenues			✓	✓	✓	✓
Expenditures	✓	✓				✓
Test Scores*	✓			✓		
Teacher variables*	✓	✓				✓

Notes: This table shows data availablity in each round for key outcomes. Different modules are administered in different rounds based on cost and attrition concerns. Variables with a star marking are only collected for half of the sample at baseline.

B Additional Tables

This section includes additional tables referenced in the main text.

Table B1: Randomization Balance, By Treatment Saturation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Village level variables									
		Control	Tests of difference			K-S Test p-values			
	N	Mean	H-C=0	L-C=0	H-L=0	H=C	L=C	H=L	
Number of public schools	266	2.451	0.011	0.010	0.001	0.953	1.000	1.000	
			[0.946]	[0.949]	[0.993]				
Number of private schools	266	3.331	0.021	0.162	-0.141	1.000	1.000	0.989	
			[0.851]	[0.164]	[0.176]				
Private enrollment	266	523.519	-23.5	11.2	-34.8	0.277	0.859	0.295	
			[0.512]	[0.714]	[0.287]				
Average fee (PKR)	266	232.069	12.7	-12.9	25.5	0.463	0.850	0.571	
			[0.413]	[0.204]	[0.067]				
Average test score	133	-0.222	-0.013	0.031	-0.044	0.267	0.505	0.346	
			[0.883]	[0.755]	[0.574]				
Overall Effect: p-value			0.95	0.96	0.99				
Panel B: Private school level variables									
		Control	Tests of difference				K-S Test p-values		
	N	Mean	H-C=0	L^t-C =0	L^u-C =0	H-L^t =0	H=C	L^t=C	H=L^t
Enrollment	851	163.6	-3.9	-18.9	0.9	15.0	0.184	0.693	0.903
			[0.658]	[0.073]	[0.913]	[0.168]			
Fees, monthly (PKR)	851	238.1	24.1	-32.3	-10.7	56.4	0.942	0.418	0.238
			[0.149]	[0.024]	[0.353]	[0.002]			
Fixed expenditures, annual (PKR)	837	78,861	21,559	-16,659	-5,747	38,219	0.575	0.883	0.570
			[0.134]	[0.154]	[0.601]	[0.013]			
Variable expenditures, annual (PKR)	848	304,644	32,312	-28,484	27,361	60,796	0.806	0.825	0.941
			[0.322]	[0.428]	[0.282]	[0.157]			
Infrastucture index (PCA)	835	-0.008	0.073	0.308	-0.074	-0.235	0.220	0.400	0.267
			[0.636]	[0.173]	[0.557]	[0.332]			
School age (No of years)	852	8.277	0.028	0.296	0.220	-0.268	0.975	0.726	0.606
			[0.961]	[0.691]	[0.699]	[0.721]			
Number of teachers	851	8.157	0.015	-0.408	0.242	0.423	1.000	0.950	0.813
			[0.965]	[0.394]	[0.484]	[0.366]			
Average test score	401	-0.210	-0.054	0.160	-0.052	-0.214	0.549	0.385	0.112
			[0.528]	[0.184]	[0.615]	[0.046]			
Overall Effect: p-value			0.48	0.37	0.78	0.16			

Notes:

a) This table shows randomization balance checks at the village and school level, Panels A and B respectively, for key variables in our study. Across both panels, column 1 shows the number of observations and column 2 shows the control mean. Panel A, columns 3-5, and Panel B, columns 3-6 show tests of differences-- regression coefficients and p-values in square brackets-- between experimental groups. Panel A, columns 6-8, and Panel B, columns 7-9 show p-values from Kolmogorov-Smirnov (K-S) tests of equality of distributions. The overall effect in each panel reports p-values from tests that ask whether variables jointly predict treatment status for a given group.

b) All regressions include strata fixed effects. Panel A regressions have robust standard errors. Panel B regressions are weighted to adjust for sampling and have clustered errors at the village level.

c) All variables are defined in Appendix A. There are fewer observations for test scores since half of the sample was tested at baseline.

Table B2: Attrition, By Treatment Saturation

	(1) Control Mean	(2) High	(3) Low Treated	(4) Low Untreated	(5) N
Panel A: Survey Attrition					
Round 1	0.059	-0.032	-0.044	-0.035	824
		(0.02)	(0.02)	(0.02)	
Round 2	0.052	-0.028	-0.045	-0.031	806
		(0.02)	(0.02)	(0.02)	
Round 3	0.087	-0.063	-0.079	-0.038	798
		(0.02)	(0.02)	(0.02)	
Round 4	0.054	-0.030	-0.054	-0.029	781
		(0.02)	(0.02)	(0.02)	
Round 5	0.126	-0.084	-0.106	-0.030	758
		(0.02)	(0.02)	(0.03)	
Always refused	0.033	-0.007	-0.033	-0.025	758
		(0.02)	(0.01)	(0.01)	
Panel B: Baseline Characteristics for Attriters (At least once refused) by Treatment Status					
Enrollment	191.417	8.422	6.449	-33.0	79
		(44.68)	(28.77)	(18.74)	
Fees, monthly (PKR)	257.485	-28.527	-47.518	37.183	79
		(60.78)	(42.46)	(50.90)	
Fixed expenditures, annual (PKR)	103,745	55,018	20,106	-49,684	77
		(90071.94)	(26347.19)	(39480.86)	
Variable expenditures, annual (PKR)	381,225	93,961	533,379	-54,014	79
		(228731.40)	(374707.38)	(110211.16)	
Infrastructure Index	0.062	0.536	1.140	-0.192	78
		(0.39)	(0.74)	(0.36)	
School age (No of years)	8.806	6.34	-3.469	0.585	79
		(3.64)	(2.79)	(2.62)	
Number of teachers	9.667	1.006	-0.611	-0.808	79
		(2.59)	(0.94)	(0.79)	
Average Test score (SD)	-0.099	-1.17	0.61	-0.023	33
		(0.10)	(0.31)	(0.23)	

Notes:

a) This table examines differential attrition, defined as refusal to participate in follow-up surveying, across all experimental groups. Column 1 gives the control group mean; columns 2-4 show regression coefficients for each of the experimental groups for a given variable; and column 5 shows the number of observations. Panel A tests for differential attrition in each follow-up round and across all rounds. Only 14 schools refuse surveying in every follow-up round. Panel B restricts to attriters to look for any differences in baseline characteristics by treatment group. Since doing this exercise on 14 schools would not be informative, we conservatively define an attriter to be any school that refuses surveying at least once after treatment (79 schools). We only have test scores for a fraction of the attriters since we only tested half the sample at baseline.

b) All regressions include strata fixed effects and are weighted to adjust for sampling, with clustered standard errors at the village level. The number of observations in Panel A is declining over time because closed schools are coded as missing in these regressions.

Table B3: Addressing Attrition and Balance Concerns, Pooled Treatment Effects

	Expenditures		Revenues					
	(1) Fixed	(2) Variable	(3) Posted	(4) Collected	(5) Enrollment	(6) Fees	(7) Test score	(8) Closure†
Panel A: Year 1 Effects, Attrition Re-weighted								
Treatment	30,133.846	3,861.724	70,027.843	58,745.633	11.041	10.299	0.008	
	(8,126.784)	(19,317.148)	(37,533.931)	(43,093.669)	(3.919)	(6.351)	(0.070)	
N Schools (Rounds)	748 (1)	759 (1)	789 (2)	786 (2)	790 (3)	766 (2)	706 (1)	
Test pval (T=0)	0.000	0.842	0.063	0.174	0.005	0.106	0.914	
Panel B: Year 1 Effects, Using Post Double Selection Lasso Controls								
Treatment	28,411.117	34,227.462	90,955.716	61,980.902	14.917	9.057	0.018	-0.044
	(6,787.568)	(24,226.675)	(33,260.114)	(36,834.556)	(4.015)	(5.532)	(0.068)	(0.016)
N Schools (Rounds)	781 (1)	792 (1)	830 (2)	827 (2)	831 (3)	798 (2)	732 (1)	855 (1)
Test pval (T=0)	0.000	0.158	0.006	0.092	0.000	0.102	0.794	0.006
Panel C: Year 2 Effects, Attrition Re-weighted								
Treatment (T)	1,198.555	7,308.817	93,093.585	64,952.916	16.035	7.974		
	(7,076.352)	(17,460.130)	(48,549.005)	(44,518.288)	(6.466)	(6.896)		
N Schools (Rounds)	664 (1)	663 (1)	784 (1)	788 (2)	788 (2)	724 (1)		
Test pval (T=0)	0.866	0.676	0.056	0.146	0.014	0.249		
Panel D: Year 2 Effects, Using Post Double Selection Lasso Controls								
Treatment	909.459	8,505.642	134,783.541	84,747.999	21.147	7.056		-0.051
	(5,871.390)	(16,048.610)	(51,722.183)	(35,185.705)	(6.072)	(6.219)		(0.022)
N Schools (Rounds)	687 (1)	685 (1)	826 (1)	830 (2)	830 (2)	751 (1)		855 (1)
Test pval (T=0)	0.877	0.596	0.009	0.016	0.000	0.257		0.017
Panel E: Average Effects Across Years 1 and 2, Attrition Re-weighted								
Treatment (T)	16,623.001	6,015.325	77,702.811	61,893.997	13.064	9.560	0.008	
	(5,361.526)	(16,161.857)	(40,180.901)	(43,113.551)	(4.708)	(5.944)	(0.070)	
N Schools (Rounds)	767 (2)	768 (2)	794 (3)	793 (4)	797 (5)	769 (3)	706 (1)	
Test pval (T=0)	0.002	0.710	0.054	0.152	0.006	0.109	0.914	
Panel F: Average Effects Across Years 1 and 2, Using Post Double Selection Lasso Controls								
Treatment (T)	14,361.879	12,992.360	100,070.552	69,199.284	17.016	8.955	0.018	-0.051
	(5,160.214)	(15,809.417)	(37,058.822)	(38,097.838)	(4.626)	(4.998)	(0.068)	(0.022)
N Schools (Rounds)	800 (2)	802 (2)	837 (3)	836 (4)	840 (5)	802 (3)	732 (1)	855 (1)
Test pval (T=0)	0.005	0.411	0.007	0.069	0.000	0.073	0.794	0.017

Notes:
a) This table addresses concerns regarding attrition and balance for the pooled treatment effects from Table 2. In Panels A, C, and E, we show our effects for year 1, 2, and averaged across both years, respectively, after weighting to account for differential attrition using the inverse probability weighting technique. In Panels B, D, and E, we show effects after using the post double-selection Lasso technique to control for baseline variables and any imbalance therein. Treatment refers to any H or L^t schools offered a grant; the comparison group includes L^u and control schools. The dependent variables are as follows: annual fixed expenditures (column 1); annual variable expenditures (column 2); annual posted revenues (column 3); annual collected revenues (column 4); enrollment (column 5); average monthly fees charged to students (column 6); average test score for all children in a given school across English, Math and Urdu (column 7); and closure (column 8). Data are pooled across rounds, except for column 8, where closure is coded based on the last follow-up round in the year(s) under consideration. For columns 1-2 and 6-7, if a school closes down, the variable is coded is missing, and for columns 3-5, the variable is coded as 0. Since closure is never missing in the data, there is no need to apply the attrition-reweighting technique for it in Panel A, C, or E.
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition or because the variable was not collected in a given round. We also control for baseline value of dependent variable, wherever available; in the case of column 4, we just use the baseline posted revenues as a control.
c) For each panel, we show the p-values from the test of a zero average impact of the treatment (T=0).
† Data is based on the last follow-up round in the year(s) under consideration.

Table B4: Fixed and Variable Expenditures

	Fixed expenditures (annual)			Variable expenditures (annual)		
	(1) Raw	(2) Top Coded 1%	(3) Trim Top 1%	(4) Raw	(5) Top Coded 1%	(6) Trim Top 1%
			Panel A: Year 1 Effects			
High	35,188.3	34,850.9	34,320.3	28,371.4	31,108.1	34,092.2
	(10,064.5)	(9,787.3)	(8,857.9)	(22,189.0)	(22,672.2)	(22,299.0)
Low Treated	28,518.6	28,866.7	33,562.3	-16,544.0	-19,023.3	-24,118.2
	(11,993.6)	(11,912.4)	(11,325.7)	(25,656.9)	(20,811.2)	(18,030.6)
Low Untreated	4,667.3	3,180.6	3,313.3	-540.8	375.2	-1,559.6
	(10,407.1)	(9,967.4)	(8,736.0)	(17,753.4)	(17,518.6)	(16,663.3)
R-Squared	0.11	0.11	0.11	0.72	0.72	0.60
N Schools (Rounds)	767 (1)	767 (1)	753 (1)	788 (1)	788 (1)	780 (1)
Test pval (H=0)	0.001	0.000	0.000	0.202	0.171	0.128
Test pval (L^t=0)	0.018	0.016	0.003	0.520	0.362	0.182
Test pval (L^t=H)	0.582	0.613	0.948	0.135	0.047	0.009
			Panel B: Year 2 Effects			
High	2,215.7	1,474.6	-2,071.2	49,244.6	43,906.9	24,928.9
	(7,909.4)	(7,609.9)	(6,777.1)	(30,604.6)	(27,313.5)	(22,393.0)
Low Treated	3,002.8	3,147.4	5,588.7	-11,669.0	-13,209.5	-14,331.3
	(9,822.6)	(9,761.2)	(9,574.3)	(22,481.7)	(21,249.6)	(20,453.5)
Low Untreated	3,915.0	1,847.6	-737.7	-7,297.0	-540.2	4,666.0
	(8,870.1)	(8,178.2)	(7,192.0)	(22,294.1)	(20,865.9)	(20,358.4)
R-Squared	0.05	0.05	0.06	0.64	0.61	0.56
N Schools (Rounds)	676 (1)	676 (1)	662 (1)	682 (1)	682 (1)	675 (1)
Test pval (H=0)	0.780	0.847	0.760	0.109	0.109	0.267
Test pval (L^t=0)	0.760	0.747	0.560	0.604	0.535	0.484
Test pval (L^t=H)	0.933	0.855	0.369	0.029	0.030	0.076

Notes:
a) This table repeats Table 3 to examine impacts on annual fixed and variable expenditures in each year separately. Panel A shows year 1 impacts and Panel B shows year 2 impacts. As before, fixed expenditures are measured on an annual basis and include spending on infrastructure or educational materials and supplies; variable expenditures include expenses incurred on a monthly basis—teaching and non-teaching staff salaries, utilities and rent— and we annualize the variable for ease of comparison. Data on expenditures are collected twice, once in each of the two years after treatment. Columns 1-3 focus on fixed expenditures, and columns 4-6 on variable expenditures. Top coding of the data assigns the value at the 99th percentile to the top 1% of data (columns 2 and 5). Trimming top 1% of the data assigns a missing value to data above the 99th percentile (columns 3 and 6). Both the top coding and trimming procedures are applied to each round of the data separately. Variables are coded as missing once a school closes down.
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition.
c) For each panel, we show p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table B5: School Revenues

	Posted Revenues (annual)			Collected Revenues (annual)		
	(1) Raw	(2) Top Coded 1%	(3) Trim Top 1%	(4) Raw	(5) Top Coded 1%	(6) Trim Top 1%
Panel A: Year 1 Effects						
High	67,187.7	59,115.8	49,509.7	51,101.1	56,005.4	48,546.5
	(35,908.0)	(28,079.6)	(23,859.7)	(33,586.1)	(25,323.1)	(21,148.5)
Low Treated	123,138.5	100,291.4	86,318.7	97,756.5	76,130.4	59,241.0
	(56,458.9)	(45,831.4)	(44,251.5)	(55,214.2)	(35,059.1)	(32,514.5)
Low Untreated	2,202.8	-1,159.0	4,973.2	14,354.9	15,640.8	16,716.9
	(33,655.7)	(26,970.1)	(22,002.7)	(30,332.8)	(24,088.4)	(19,396.8)
R-Squared	0.69	0.67	0.61	0.59	0.66	0.57
N Schools (Rounds)	825 (2)	825 (2)	814 (2)	822 (2)	822 (2)	810 (2)
Test pval (H=0)	0.062	0.036	0.039	0.129	0.028	0.022
Test pval (L^t=0)	0.030	0.030	0.052	0.078	0.031	0.070
Test pval (L^t=H)	0.351	0.393	0.409	0.434	0.601	0.756
Panel B: Year 2 Effects						
High	63,149.1	62,819.1	72,790.9	53,768.6	54,972.3	36,828.2
	(64,522.7)	(43,944.6)	(38,437.5)	(58,342.8)	(36,530.9)	(29,115.6)
Low Treated	137,664.3	135,800.7	124,694.0	91,027.3	90,869.5	69,893.1
	(70,351.5)	(59,724.7)	(53,462.2)	(57,747.3)	(44,978.6)	(38,970.8)
Low Untreated	-24,868.4	-22,347.9	1,376.5	-13,102.0	-13,015.5	135.1
	(38,970.4)	(36,880.3)	(30,022.2)	(33,770.4)	(31,721.8)	(24,994.3)
R-Squared	0.58	0.63	0.53	0.53	0.60	0.50
N Schools (Rounds)	821 (1)	821 (1)	809 (1)	825 (2)	825 (2)	813 (2)
Test pval (H=0)	0.329	0.154	0.059	0.358	0.134	0.207
Test pval (L^t=0)	0.051	0.024	0.020	0.116	0.044	0.074
Test pval (L^t=H)	0.400	0.278	0.385	0.630	0.495	0.418

Notes:
a) This table repeats Table 4 to examine impacts on annual school revenues in each year separately. Panel A shows year 1 impacts and Panel B shows year 2 impacts. Columns 1-3 consider posted revenues, defined as the sum of revenues expected from each grade based on enrollment and posted ffectfees. Columns 4-6 consider collected revenues, defined as revenues actually collected from students at the school. Across all columns, schools are coded to have zero revenues once they close down and in all subsequent rounds. Top coding of the data assigns the value at the 99th percentile to the top 1% of data (columns 2 and 5). Trimming top 1% of the data assigns a missing value to data above the 99th percentile (columns 3 and 6). Both the top coding and trimming procedures are applied to each round of the data separately. Across all columns, schools are coded to have zero revenues once they close down.
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. Across all columns, we use baseline posted revenues as the baseline control as we did not measure collected revenues at baseline. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition or because the variable was not collected in a given round.
c) For each panel, we show p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table B6: School Enrollment and Fees

	(1) Enrollment	(2) Posted Fees	(3) Collected Fees	(4) Closure†
	Panel A: Year 1 Effects			
High	8.86	17.68	19.45	-0.02
	(5.38)	(7.63)	(8.09)	(0.03)
Low Treated	18.83	1.93	7.71	-0.06
	(7.00)	(7.93)	(11.42)	(0.02)
Low Untreated	-0.31	0.07	16.62	-0.01
	(5.09)	(6.24)	(7.50)	(0.02)
R-Squared	0.69	0.71	0.56	0.03
N Schools (Rounds)	827 (3)	796 (2)	776 (2)	855 (1)
Test pval (H=0)	0.101	0.021	0.017	0.487
Test pval (L^t=0)	0.008	0.808	0.500	0.010
Test pval (L^t=H)	0.148	0.059	0.307	0.081
	Panel B: Year 2 Effects			
High	9.12	21.04	22.46	-0.02
	(7.99)	(10.27)	(12.91)	(0.03)
Low Treated	26.02	-2.51	-2.86	-0.09
	(10.01)	(9.43)	(10.69)	(0.03)
Low Untreated	1.00	-0.38	3.73	-0.03
	(7.23)	(9.13)	(8.69)	(0.03)
R-Squared	0.53	0.73	0.51	0.05
N Schools (Rounds)	826 (2)	749 (1)	754 (2)	855 (1)
Test pval (H=0)	0.255	0.041	0.083	0.597
Test pval (L^t=0)	0.010	0.791	0.789	0.007
Test pval (L^t=H)	0.100	0.014	0.060	0.045

Notes:
a) This table repeats Table 5 to examine impacts on enrollment and fees in each year separately. Panel A show year 1 effects and Panel B shows year 2 effects. The dependent variables are as follows: enrollment (column 1); posted tuition fees, averaged across all grades taught at the school (column 2); collected tuition fees, generated by dividing collected revenues by enrollment in the given round (column 3); and closure (column 4). Whereas columns 1-3 pool data over the two years, closure is measured at the end of year 2. Once a school closes down, enrollment is coded as 0 and posted and collected fees are coded as missing.
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. Columns 1-2 use the baseline of the dependent variable as a control, and column 3 uses posted fees as the baseline control since collected revenues are not measured at baseline. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition or because the variable was not collected in a given round.
c) For each panel, we show p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.
† Data is based on the last follow-up round in year 2.

Table B7: Enrollment by Grades (Average Effects across Years 1 and 2)

	(1) Lower than 1	(2) 1 to 3	(3) 4 to 5	(4) 6 to 8	(5) 9 to 12
High	2.46	1.03	1.06	1.57	1.56
	(2.18)	(1.85)	(0.97)	(1.52)	(1.19)
Low Treated	5.30	6.10	1.49	3.31	3.56
	(2.52)	(2.41)	(1.19)	(2.03)	(2.49)
Low Untreated	0.47	-0.05	0.43	-0.09	-1.28
	(1.94)	(1.62)	(0.97)	(1.47)	(1.32)
R-Squared	0.40	0.57	0.63	0.59	0.66
# Schools (Rounds)	835 (4)	838 (4)	838 (4)	838 (4)	838 (4)
Depvar Mean	49.89	53.68	28.15	23.10	8.22
Test pval (H=0)	0.259	0.580	0.274	0.302	0.189
Test pval (L^t= 0)	0.036	0.012	0.211	0.104	0.154
Test pval (L^t=H)	0.260	0.042	0.704	0.367	0.462

Notes:
a) This table disaggregates school enrollment into grade bins to examine the source of enrollment gains over the two years of the study. Data from rounds 1-4 are used since grade-wise enrollment was not collected in round 5. All grades in closed schools and grades where there may be no enrolled students in open schools are assigned 0 enrollment.
b) Regressions are weighted to adjust for sampling and include strata and round fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition. The mean of the dependent variable is its baseline value.
c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table B8: Enrollment Decomposition Using Year 1 Child Data

	(1) Enrollment	(2) % New
High	0.348	0.025
	(0.702)	(0.015)
Low Treated	0.776	0.056
	(0.740)	(0.024)
Low Untreated	-0.382	0.024
	(0.706)	(0.017)
Baseline	0.641	
	(0.048)	
R-Squared	0.61	0.04
# Schools (Rounds)	765 (1)	711 (1)
Depvar Mean	14.69	0.07
Test pval (H=0)	0.620	0.097
Test pval (L^t= 0)	0.295	0.023
Test pval (L^t=H)	0.559	0.208

Notes:
a) This table examines changes in child enrollment status in the first year after treatment. The dependent variables are from tested children in round 3. Column 1 is the number of children enrolled in grade 4. Column 2 is the fraction of the grade 4 children who newly enroll in the school after treatment. Enrollment status is determined based on child self-reports; any child who reports joining the school in the 18 months prior to round 3 is considered new.
b) Regressions are weighted to adjust for sampling and include strata fixed effects, with standard errors clustered at village level. The number of schools may vary across columns due to attrition. The mean of the dependent variable is its baseline value in column 1 or the follow-up control mean in column 2.
c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table B9: Posted Tuition Fees by Grades (Average Effects across Years 1 and 2)

	(1) Lower than 1	(2) 1 to 3	(3) 4 to 5	(4) 6 to 8	(5) 9 to 12
High	14.43	21.22	19.38	36.87	142.64
	(10.49)	(12.12)	(12.54)	(17.75)	(66.98)
Low Treated	-4.85	-3.22	-8.05	-18.75	88.64
	(5.39)	(6.39)	(8.04)	(12.58)	(78.69)
Low Untreated	2.33	4.23	-1.06	-2.44	-68.85
	(4.59)	(6.21)	(6.54)	(11.24)	(54.93)
Baseline	0.83	0.75	0.79	0.67	0.47
	(0.05)	(0.05)	(0.04)	(0.06)	(0.13)
R-Squared	0.64	0.60	0.59	0.57	0.48
# Schools (Rounds)	789 (3)	789 (3)	773 (3)	542 (3)	144 (3)
Depvar Mean	169.89	207.82	237.43	319.88	425.94
Test pval (H=0)	0.170	0.081	0.124	0.039	0.036
Test pval ($L^t = 0$)	0.369	0.615	0.318	0.137	0.263
Test pval (L^t=H)	0.082	0.047	0.043	0.003	0.530

Notes:
a) This table averages monthly tuition fees by grade bins to assess whether fee changes occur in specific grades. Fees for closed schools or schools that do not offer certain grade levels are coded as missing.
b) Regressions are weighted to adjust for sampling and include strata and round fixed effects, with standard errors clustered at village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. Higher grades have fewer school observations because fewer schools offer those grade levels and hence post tuition fees, and these observations are subsequently coded as missing. In contrast, enrollment in higher grades is coded as 0 if a school does not offer those grades; if we restrict to the same sample as in this table, the pattern of enrollment results by grades stays the same. The mean of the dependent variable in all regressions is its baseline value.
c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools."

Table B10: Addressing Attrition and Balance Concerns, by Treatment Saturation

	Expenditures (Top Coded 1%)		Revenues (Top Coded 1%)					
	(1) Fixed	(2) Variable	(3) Posted	(4) Collected	(5) Enrollment	(6) Fees	(7) Test score	(8) Closure†
Panel A: Average Effects across Years 1 and 2, Attrition Re-weighted								
High	20,146.293	32,559.591	63,209.698	64,248.117	8.709	25.685	0.169	
	(5,855.537)	(21,444.531)	(25,240.393)	(23,775.902)	(5.552)	(7.884)	(0.087)	
Low Treated	17,314.817	-16,886.897	77,128.409	63,058.133	16.733	5.472	-0.038	
	(7,239.409)	(18,998.287)	(39,183.949)	(36,170.642)	(7.190)	(7.862)	(0.105)	
Low Untreated	1,224.638	-2,438.691	12,313.601	19,615.156	0.910	6.298	0.055	
	(6,335.265)	(16,150.625)	(22,445.382)	(21,828.791)	(5.267)	(6.403)	(0.073)	
R-Squared	0.11	0.67	0.66	0.62	0.62	0.71	0.16	
N Schools (Rounds)	767 (2)	768 (2)	794 (3)	793 (4)	797 (5)	769 (3)	706 (1)	
Test pval (H=0)	0.001	0.130	0.013	0.007	0.118	0.001	0.053	
Test pval (L^t=0)	0.017	0.375	0.050	0.082	0.021	0.487	0.720	
Test pval (L^t=H)	0.691	0.036	0.740	0.976	0.243	0.015	0.054	
Panel B: Average Effects across Years 1 and 2, Using Post Double Selection Lasso Controls								
High	15,528.993	39,370.239	78,226.389	51,275.541	7.181	20.806	0.149	-0.014
	(6,296.553)	(19,057.035)	(36,686.848)	(25,402.537)	(5.670)	(7.120)	(0.084)	(0.035)
Low Treated	13,447.899	9,780.557	91,724.468	72,350.527	22.960	2.233	0.004	-0.096
	(7,301.677)	(17,649.506)	(40,474.870)	(31,924.428)	(6.694)	(6.876)	(0.099)	(0.030)
Low Untreated	-1,700.195	15,908.029	27,133.990	22,480.622	1.250	0.222	0.060	-0.022
	(6,271.418)	(15,696.194)	(25,090.081)	(23,251.376)	(5.424)	(5.953)	(0.071)	(0.029)
N Schools (Rounds)	800 (2)	802 (2)	837 (3)	836 (4)	840 (5)	802 (3)	732 (1)	855 (1)
Test pval (H=0)	0.014	0.039	0.033	0.044	0.205	0.003	0.076	0.687
Test pval (L^t=0)	0.066	0.579	0.023	0.023	0.001	0.745	0.969	0.001
Test pval (L^t=H)	0.771	0.113	0.763	0.550	0.013	0.010	0.138	0.016

a) This table addresses concerns regarding attrition and balance for our effects by treatment saturation. Panel A shows estimates after weighting to account for differential attrition using the inverse probability weighting technique; Panel B shows estimates using the post double-selection Lasso technique to control for baseline variables and any imbalance therein. The dependent variables are as follows: annual fixed expenditures (column 1); annual variable expenditures (column 2); annual posted revenues (column 3); annual collected revenues (column 4); enrollment (column 5); average monthly fees charged to students (column 6); average test score for all children in a given school across English, Math and Urdu (column 7); and closure (column 8). For columns 1-4, we use the top-coded (at the 99pctl for each round) version of our dependent variables. For columns 1-2 and 6-7, if a school closes down, the variable is coded is missing, and for columns 3-5, the variable is coded as 0. We pool available data for these regressions, except for column 8, which is measured at the end of year 2 (as indicated by the † symbol). Since closure is never missing in the data, there is no need to apply the attrition-reweighting technique for it in Panel A.
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition or because the variable was not collected in a given round.
c) For each panel, we report p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.
† Data is based on the last follow-up round in year 2.

Table B11: Test Scores, Stayers Only (Year 1)

	(1) Average	(2) Math	(3) English	(4) Urdu
Panel A: Effects By Treatment Saturation				
High	0.132	0.150	0.191	0.120
	(0.0767)	(0.0927)	(0.0976)	(0.0848)
Low Treated	-0.0335	-0.114	0.0541	-0.0903
	(0.0890)	(0.115)	(0.111)	(0.111)
Low Untreated	0.0163	0.0307	0.0550	0.0146
	(0.0631)	(0.0774)	(0.0836)	(0.0706)
R-Squared	0.17	0.17	0.13	0.15
N Schools (Rounds)	720 (1)	720 (1)	720 (1)	720 (1)
Test pval (H=0)	0.086	0.108	0.052	0.159
Test pval (L^t= 0)	0.707	0.322	0.625	0.417
Test pval (L^t=H)	0.063	0.018	0.211	0.056
Panel B: Pooled Treatment Effects				
Treatment	0.0140	-0.0413	0.0681	-0.0266
	(0.0601)	(0.0790)	(0.0750)	(0.0766)
R-Squared	0.16	0.16	0.13	0.14
N Schools (Rounds)	720 (1)	720 (1)	720 (1)	720 (1)

Notes:
a) This table examines whether our school test score results are driven by compositional changes. As before, school test scores are generated by averaging child average (across all subjects) test scores for a given school. We repeat all of the regressions in Table 6, but only include all children who report being at the same school for at least 1.5 years ('stayers'). For this sample, Panel A shows results by treatment saturation and Panel B shows pooled treatment effects.
b) Regressions are weighted to adjust for sampling and include strata fixed effects. Standard errors are clustered at the village level. We include a dummy variable for the untested sample at baseline across all columns and replace the baseline test score with a constant. Since the choice of the testing sample at baseline was random, this procedure allows us to control for baseline test scores of stayers wherever available. The number of schools is lower than the full sample due to attrition, closure, and zero enrollment in some schools in the tested grades.
c) For Panel A, we show p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools, and for Panel B, we show p-values from the test of a zero average impact of treatment (T=0).

Table B12: School Infrastructure

	Infrastructure Spending	Number purchased		Facility present (Y/N)			Upgradation
	(1) Amount	(2) Desks	(3) Chairs	(4) Computers	(5) Library	(6) Sports	(7) # Rooms
Panel A: Year 2 Effects							
High	125.38	0.53	1.34	0.06	0.00	0.07	0.21
	(7,575.44)	(1.54)	(1.02)	(0.06)	(0.03)	(0.03)	(0.36)
Low Treated	-2,560.79	-1.58	0.77	0.10	-0.01	0.02	-0.31
	(8,621.92)	(1.60)	(0.58)	(0.06)	(0.04)	(0.04)	(0.35)
Low Untreated	145.61	-1.86	0.29	-0.03	0.02	0.03	-0.08
	(7,930.17)	(1.46)	(0.43)	(0.05)	(0.03)	(0.03)	(0.31)
R-squared	0.05	0.08	0.04	0.16	0.04	0.12	0.57
N Schools (Rounds)	678 (1)	652 (1)	685 (1)	689 (1)	689 (1)	689 (1)	689 (1)
Test pval (H=0)	0.987	0.731	0.190	0.311	0.959	0.050	0.555
Test pval (L^t=0)	0.767	0.324	0.190	0.129	0.865	0.572	0.378
Test pval (L^t=H)	0.752	0.184	0.615	0.528	0.831	0.260	0.165
Panel B: Average Effects across Years 1 and 2							
High	13,882.37	3.60	2.74	0.14	0.06	0.09	0.45
	(5,806.93)	(1.13)	(0.79)	(0.04)	(0.03)	(0.03)	(0.24)
Low Treated	7,993.58	3.89	3.55	0.13	-0.01	-0.01	0.02
	(6,045.80)	(1.46)	(1.50)	(0.05)	(0.04)	(0.03)	(0.29)
Low Untreated	-859.19	-0.09	0.72	0.01	-0.01	0.02	0.03
	(5,925.98)	(1.08)	(0.70)	(0.04)	(0.03)	(0.02)	(0.23)
R-squared	0.06	0.10	0.08	0.16	0.16	0.10	0.57
N Schools (Rounds)	789 (2)	791 (2)	796 (2)	801 (2)	801 (2)	801 (2)	801 (2)
Test pval (H=0)	0.018	0.002	0.001	0.001	0.033	0.006	0.062
Test pval (L^t=0)	0.187	0.008	0.018	0.010	0.688	0.826	0.952
Test pval (L^t=H)	0.337	0.854	0.626	0.917	0.032	0.008	0.161

a) This table repeats Table 7 to examine outcomes relating to school infrastructure. Panel A shows year 2 effects and Panel B shows average effects across years 1 and 2. Column 1 reports annual infrastructure expenditures, which comprises the largest component of annual fixed expenditures and includes spending on furniture, fixtures, or upgradation of classroom facilities. Columns 2-3 consider the number of new desks and chairs purchased in the last year; columns 4-6 are dummy variables for the presence of particular school facilities; and column 7 measures the number of rooms upgraded from temporary to permanent or semi-permanent classrooms. Variables are coded as missing once a school closes down.
b) Regressions are weighted to adjust for sampling and tracking wherever necessary and include strata and round fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition.
c) For each panel, we show p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table B13: Teacher Compensation and Composition

	School level			Teacher level salaries (monthly)		
	(1) Wage bill (Annual)	(2) Teachers (number)	(3) New Teachers (number)	(4) All	(5) New	(6) Existing
Panel A: Year 1 Effects						
High	29,361.94	0.13	0.33	407.63	396.81	419.42
	(19,512.03)	(0.30)	(0.18)	(220.81)	(282.32)	(246.42)
Low Treated	-9,224.02	-0.12	0.09	-98.84	-44.89	-119.46
	(23,785.37)	(0.33)	(0.22)	(251.22)	(413.02)	(234.30)
Low Untreated	1,722.03	0.26	0.40	155.44	62.42	188.17
	(16,707.71)	(0.29)	(0.19)	(183.66)	(236.55)	(189.50)
R-Squared	0.66	0.57	0.15	0.19	0.24	0.18
N Schools (Rounds)	788 (1)	794 (1)	794 (1)	794 (1)	556 (1)	783 (1)
N Teachers				6243	1442	4797
Test pval (H=0)	0.134	0.659	0.066	0.066	0.161	0.090
Test pval (L^t=0)	0.698	0.726	0.685	0.694	0.914	0.611
Test pval (L^t=H)	0.142	0.442	0.260	0.074	0.295	0.069
Panel B: Year 2 Effects						
High	38,351.67	0.46	0.53	610.02	658.16	555.48
	(21,933.39)	(0.35)	(0.30)	(318.00)	(321.44)	(363.68)
Low Treated	-13,310.47	-0.05	0.05	-281.99	-119.27	-420.30
	(17,785.76)	(0.36)	(0.36)	(320.06)	(436.62)	(290.61)
Low Untreated	468.63	0.26	0.07	208.91	88.27	298.29
	(15,897.07)	(0.31)	(0.27)	(249.17)	(273.56)	(261.95)
R-Squared	0.61	0.49	0.14	0.21	0.23	0.22
N Schools (Rounds)	682 (1)	696 (1)	696 (1)	690 (1)	644 (1)	678 (1)
N Teachers				5482	2461	3021
Test pval (H=0)	0.082	0.185	0.079	0.056	0.042	0.128
Test pval (L^t=0)	0.455	0.886	0.897	0.379	0.785	0.149
Test pval (L^t=H)	0.022	0.199	0.180	0.030	0.123	0.023

Notes:
a) This table repeats Table 8 to consider the impacts on teacher compensation and composition in each year separately. Panel A show year 1 effects and Panel B shows year 2 effects. The dependent variables are as follows: (annualized) wage bill, the largest component of variable expenditures (column 1); total number of teachers (column 2); number of new teachers (column 3); and monthly salaries for all, new, and existing teachers (columns 4-6). Column 1 data come from school surveys, whereas data for columns 2-6 come from teacher rosters collected at each school. In columns 2-3, we collapse data from these rosters at the school level to understand changes in teacher composition; columns 4-6 are measured at the teacher level. For columns 5 and 6, whether a teacher is new or existing is determined by their start date at the school relative to baseline. Variables are coded as missing once a school closes down.
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition.
c) For each panel, we show p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table B14: Baseline Correlates of Closure

	(1)	(2)	(3)	(4)	(5)	(6)
		Closed Schools				
Baseline variable	*Open Control (mean)*	*Control (mean)*	*High*	*Low-treated*	*Low- untreated*	*N*
Enrollment	174.724	108.676	-43.746	-34.574	-28.540	94
			(18.59)	(15.06)	(16.94)	
Fees, monthly (PKR)	235.809	239.141	5.722	-96.878	-42.413	95
			(79.45)	(41.81)	(32.99)	
Fixed expenditures, annual (PKR)	78,311	56,192	-27,955	17,643	-38,854	92
			(31043.36)	(21633.93)	(31801.69)	
Variable expenditures, annual (PKR)	303,505	172,759	-39,515	-61,404	-42,731	95
			(30426.71)	(26323.68)	(25926.72)	
Infrastructure Index	-0.144	0.500	1.027	-0.615	0.059	93
			(1.11)	(0.61)	(0.58)	
School age (No of years)	8.542	5.765	1.369	1.967	-0.708	95
			(1.57)	(1.80)	(1.70)	
Number of teachers	8.500	5.176	-0.749	-0.641	0.191	95
			(0.72)	(0.62)	(0.65)	
Average Test score (SD)	-0.228	0.124	-0.164	-0.174	-0.820	40
			(0.42)	(0.70)	(0.45)	

Notes:

a) This table shows the baseline correlates of school closure. Column 1 and 2 report the baseline mean for schools that remain open or close down, respectively, in the control group two years after treatment. Columns 3-5 report the regression estimate for how these baseline variables vary by treatment status relative to control. Column 6 reports the observations, which are different due to either missing values at baseline or our decision to only test children in half the schools at random at baseline.

b) Regressions are weighted to adjust for sampling and include strata fixed effects. Standard errors are clustered at the village level.

Table B15: Main Outcomes, Accounting for Differential Closure

	Expenditures Top Coded 1%		Revenues Top Coded 1%				
	(1) Fixed	(2) Variable	(3) Posted	(4) Collected	(5) Enrollment	(6) Fees	(7) Test Score
Panel A: Using imputations for closed schools in H, C and L^u							
High	18,667.920	30,718.215	62,124.976	58,667.176	6.672	18.017	0.136
	(5,667.679)	(19,180.223)	(29,340.763)	(28,321.977)	(4.869)	(7.802)	(0.083)
Low Treated	15,391.701	-15,940.882	86,160.870	54,458.801	10.552	0.600	-0.021
	(7,082.862)	(17,806.230)	(47,564.681)	(39,121.093)	(7.332)	(7.447)	(0.100)
Low Untreated	3,419.507	852.545	-18,386.481	-10,596.479	-2.678	-0.357	0.032
	(6,319.213)	(14,496.586)	(27,119.455)	(25,981.818)	(4.857)	(6.258)	(0.068)
R-Squared	0.11	0.68	0.67	0.64	0.64	0.72	0.17
N Schools (Rounds)	810 (2)	821 (2)	826 (3)	821 (4)	830 (5)	824 (3)	768 (1)
Test pval (H=0)	0.001	0.110	0.035	0.039	0.172	0.022	0.102
Test pval (L^t=0)	0.031	0.371	0.071	0.165	0.151	0.936	0.833
Test pval (L^t=H)	0.640	0.033	0.634	0.922	0.593	0.027	0.119
Panel B: Using predicted probabilities to increase closure in the L^t group							
High	19,717.561	35,820.627	59,897.659	55,365.320	8.937	18.716	0.154
	(5,877.814)	(22,208.296)	(31,304.285)	(29,120.460)	(6.099)	(7.883)	(0.085)
Low Treated	16,647.952	-14,846.444	101,063.604	75,647.215	18.510	1.118	0.025
	(7,372.689)	(18,963.221)	(47,233.877)	(39,156.463)	(7.790)	(7.583)	(0.099)
Low Untreated	2,894.259	-333.941	-8,514.624	993.052	0.165	0.037	0.034
	(6,377.546)	(16,207.296)	(28,296.806)	(26,667.181)	(5.546)	(6.498)	(0.070)
R-Squared	0.11	0.67	0.65	0.62	0.62	0.72	0.19
N Schools (Rounds)	781 (2)	792 (2)	832 (5)	831 (4)	836 (5)	795 (3)	720 (1)
Test pval (H=0)	0.001	0.108	0.057	0.058	0.144	0.018	0.071
Test pval (L^t=0)	0.025	0.434	0.033	0.054	0.018	0.883	0.804
Test pval (L^t=H)	0.667	0.034	0.422	0.647	0.223	0.028	0.192

a) This table shows the sensitivity of our results to differential closure during the experiment. We use two strategies to consider how effects would change if there were no differential closure across treatments. In Panel A, we impute data for closed (H, L^u, and control) schools by using coefficients from a regression of open schools in the control group that predict outcomes using baseline variables and round and strata fixed effects. On the other hand, in Panel B, we use baseline variables to generate predicted closure probabilities and drop L^t schools in the top 5% of predicted probabilities to eliminate the closure difference across all groups. The baseline covariates used in both approaches include enrollment, fees, test score, school age, number of teachers, an infrastructure index, and fixed and variable expenditures. The dependent variables are as follows: annual fixed expenditures (column 1); annual variable expenditures (column 2); annual posted revenues (column 3); annual collected revenues (column 4); enrollment (column 5); average monthly fees charged to students (column 6); and average test score for all children in a given school across English, Math, and Urdu (column 7). For columns 1-4, we use the top-coded (at the 99pctl for each round) version of our dependent variables. For columns 1-2 and 6-7, if a school closes down, the variable is coded is missing, and for columns 3-5, the variable is coded as 0. Data are pooled across all available rounds.
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of stacked rounds of data in parantheses. The number of schools may vary across columns due to attrition or because the variable was not collected in a given round.
c) For each panel, we report p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table B16: Credit Behavior (Year 1)

	School funding sources (Y/N)		HH borrowing (Y/N)			HH loan value
	(1) Self-financed	(2) Credit	(3) Any	(4) Formal	(5) Informal	(6) Any
			Panel A: Pooled Treatment			
Treatment	-0.001	0.003	-0.025	-0.008	-0.012	3129.459
	(0.01)	(0.01)	(0.04)	(0.02)	(0.03)	(21524.23)
Baseline	0.079	-0.017	0.078	0.210	-0.000	0.063
	(0.09)	(0.01)	(0.04)	(0.05)	(0.04)	(0.03)
R-Squared	0.03	0.02	0.04	0.13	0.02	0.03
N Schools (Rounds)	795 (1)	795 (1)	784 (1)	784 (1)	784 (1)	784 (1)
Test pval (T=0)	0.85	0.71	0.48	0.64	0.71	0.88
			Panel B: By Treatment Saturation			
High	-0.007	0.002	-0.010	0.020	-0.033	1063.026
	(0.01)	(0.01)	(0.05)	(0.02)	(0.05)	(15092.81)
Low Treated	-0.000	-0.006	-0.039	0.010	-0.053	17384.174
	(0.01)	(0.01)	(0.05)	(0.02)	(0.05)	(29982.80)
Low Untreated	-0.002	-0.011	-0.005	0.035	-0.055	13611.930
	(0.01)	(0.01)	(0.04)	(0.02)	(0.04)	(21581.81)
Baseline	0.078	-0.017	0.080	0.208	0.003	0.064
	(0.09)	(0.01)	(0.04)	(0.05)	(0.04)	(0.03)
R-Squared	0.03	0.02	0.04	0.14	0.02	0.03
N Schools (Rounds)	795 (1)	795 (1)	784 (1)	784 (1)	784 (1)	784 (1)
Test pval (H=0)	0.481	0.882	0.829	0.229	0.467	0.944
Test pval (L^t= 0)	0.970	0.679	0.454	0.640	0.272	0.563
Test pval (L^t=H)	0.529	0.564	0.598	0.647	0.687	0.605

Notes:
a) This table looks at credit behavior of school owners in year 1 to understand whether the treatment simply acted as a substitute for other types of credit. Panel A considers the pooled treatment specification; and Panel B separates effects by treatment saturation. Data for columns 1-2 are from the school survey, and for columns 3-6 from the school owner survey. The dependent variables in columns 1-2 are indicators for whether a school reports financing school expenditures through fees or owner income or through a formal or informal financial institution, respectively. Column 3 reports whether the household of the school owner has ever borrowed any money for any reason. Columns 4-5 disaggregate this household borrowing into formal and informal sources. Column 6 examines total borrowing by the owner's household for any reason; if the owner household did not borrow, the loan value is coded as 0. Schools that closed or refused surveying are coded as missing for credit behavior.
b) Regressions are weighted to adjust for sampling and include strata fixed effects, with standard errors clustered at the village level. The number of schools may vary across columns due to attrition.
c) For Panel A, we show p-values from the test of a zero average impact of treatment (T=0), and for Panel B, we show p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table B17: Main Outcomes, controlling for Grant size per capita

	Expenditures Top Coded 1%		Revenues Top Coded 1%						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Fixed	Variable	Posted	Collected	Enrollment	Fees	Test score	Closure	Teacher Wage bill
High	27473.657	85376.276	16425.280	7615.056	-2.714	10.764	0.227	-0.066	80544.506
	(9255.720)	(37392.112)	(55024.482)	(49298.416)	(10.605)	(12.677)	(0.165)	(0.053)	(31555.745)
Low Treated	19105.495	1379.429	97997.666	67971.906	18.050	-2.128	-0.004	-0.102	5182.574
	(8068.281)	(20658.418)	(48149.096)	(39119.314)	(8.345)	(8.197)	(0.110)	(0.033)	(19151.278)
Low Untreated	5137.001	14278.245	-21558.137	-13284.489	-3.310	-2.431	0.055	-0.043	14908.747
	(6790.888)	(18000.504)	(33202.803)	(30069.266)	(6.245)	(7.383)	(0.083)	(0.032)	(14866.258)
Grant per capita	-20.544	-131.921	114.511	125.605	0.031	0.022	-0.000	0.000	-126.961
	(17.259)	(57.633)	(88.001)	(78.779)	(0.020)	(0.024)	(0.000)	(0.000)	(52.140)
R-Squared	0.11	0.67	0.65	0.62	0.62	0.72	0.17	0.05	0.63
# Schools (Rounds)	786 (2)	797 (2)	832 (3)	831 (4)	836 (5)	800 (3)	725 (1)	855 (1)	797 (2)
Depvar Mean	77876.64	299322.06	463848.68	362731.75	163.64	238.13	-0.21	-0.21	-0.21
Test pval (H=0)	0.003	0.023	0.766	0.877	0.798	0.397	0.171	0.215	0.011
Test pval ($L^t = 0$)	0.019	0.947	0.043	0.083	0.031	0.795	0.973	0.002	0.787
Test pval (L^t=H)	0.269	0.008	0.196	0.272	0.030	0.211	0.101	0.433	0.011

Notes:
a) This table repeats our main results with an additional village level control variable, grant amount per capita. This control variable captures whether our differential results for the H and L^t schools are driven by total resources provided to a village. It is constructed by adding the total amount of funding received by treatment villages, which is 50,000 PKR for L villages, and a multiple of 50,000 PKR based on the number of private schools in H villages. The dependent variables are as follows: annual fixed expenditures (column 1); annual variable expenditures (column 2); annual posted revenues (column 3); annual collected revenues (column 4); enrollment (column 5); average monthly fees charged to students (column 6); average test score for all children in a given school across English, Math, and Urdu (column 7); closure (column 8); and teacher wage bill (column 9). For columns 1-4, we use the top-coded (at the 99pctl for each round) version of our dependent variables. For columns 1-2 and 6-7, if a school closes down, the variable is coded is missing, and for columns 3-5, the variable is coded as 0. We pool all available data for these regressions, except for closure, which is measured at the end of year 2 (as indicated by the † symbol).
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects, with standard errors clustered at village level. The number of schools may vary across columns due to attrition or because the variable was not collected in a given round.
c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.
† Data is based on the last follow-up round in year 2.

Table B18: Main Outcomes, Interacted with Baseline Availability of Bank Account

	Expenditures Top Coded 1%		Revenues Top Coded 1%					
	(1) Fixed	(2) Variable	(3) Posted	(4) Collected	(5) Enrollment	(6) Fees	(7) Test score	(8) Closure
High	22693.11	46504.06	50501.10	59859.31	6.93	18.28	0.12	-0.02
	(6912.47)	(25883.56)	(38805.19)	(38468.26)	(7.36)	(10.09)	(0.10)	(0.04)
Low Treated	16632.75	4918.03	125834.37	101864.70	21.85	-1.76	0.02	-0.08
	(9324.69)	(23965.87)	(65534.61)	(53785.75)	(10.35)	(10.14)	(0.13)	(0.04)
Low Untreated	7080.63	987.27	-11071.39	-3708.17	-0.49	0.75	0.00	-0.05
	(7864.86)	(20382.75)	(34443.97)	(32151.87)	(6.86)	(8.14)	(0.08)	(0.03)
High*NoBankAct	-10317.17	-39747.69	33866.30	-16334.23	7.55	2.09	0.11	0.02
	(13218.76)	(49389.79)	(60125.76)	(58014.51)	(10.72)	(15.12)	(0.16)	(0.06)
Low Treated*NoBankAct	534.48	-57977.39	-41904.05	-57886.87	0.05	6.98	-0.13	-0.01
	(14403.22)	(39490.30)	(77825.28)	(63672.79)	(14.41)	(14.93)	(0.22)	(0.07)
Low Untreated*NoBankAct	-16797.43	-7031.23	9024.36	18625.38	2.93	-2.91	0.09	0.07
	(13062.49)	(36300.18)	(60128.18)	(54893.42)	(11.63)	(13.60)	(0.15)	(0.07)
HH does not have bank act	-4968.14	6159.47	-9912.64	11552.79	-1.13	-0.77	-0.10	-0.03
	(9286.36)	(23226.75)	(41208.82)	(34708.30)	(7.42)	(10.01)	(0.11)	(0.05)
R-Squared	0.11	0.68	0.65	0.62	0.62	0.72	0.17	0.05
# Schools (Rounds)	786 (2)	797 (2)	832 (3)	831 (4)	836 (5)	800 (3)	725 (1)	855 (1)
Depvar Mean	77876.64	299322.06	463848.68	362731.75	163.64	238.13	-0.21	-0.21

Notes:
a) This table examines whether our results are driven by baseline access to bank accounts in school owner households. We reproduce our key results after adding an interaction with a dummy variable for whether the owner's household does not have a bank account with our treatment indicators. The primary coefficients of interest are the three interaction terms with the treatment groups, which tell us whether treated schools where the owner did not have access to a bank account at baseline benefited more from treatment. The dependent variables are as follows: annual fixed expenditures (column 1); annual variable expenditures (column 2); annual posted revenues (column 3); annual collected revenues (column 4); enrollment (column 5); average monthly fees charged to students (column 6); average test score for all children in a given school across English, Math, and Urdu (column 7); and closure (column 8). For columns 1-4, we use the top-coded (at the 99pctl for each round) version of our dependent variables. For columns 1-2 and 6-7, if a school closes down, the variable is coded is missing, and for columns 3-5, the variable is coded as 0. We pool available data for these regressions, except for closure, which is measured at the end of year 2 (as indicated by the † symbol).
b) Regressions are weighted to adjust for sampling and tracking and include strata and round fixed effects, with standard errors clustered at village level. The number of schools may vary across columns due to attrition or because the variable was not collected in a given round.
† Data is based on the last follow-up round in year 2.

Table B19: Profits, by Treatment Saturation

	(1) Collected Revenues Top coded 1%	(2) Variable expenditures Top coded 1%	(3) Profits (1)-(2)
Panel A: Closed schools as missing			
High	56,437.356	35,840.649	20,596.71
	(29,891.931)	(22,203.899)	(27,809.13)
Low Treated	58,306.034	-14,456.076	72,762.11
	(41,558.970)	(18,586.535)	(39,489.56)
Test pval (H=0)	0.060	0.108	0.459
Test pval (L^t= 0)	0.162	0.437	0.065
Test pval (L^t=H)	0.967	0.032	0.194
Panel B: Closed schools as 0			
High	55,625.198	27,501.188	28,124.01
	(29,061.826)	(20,174.033)	(26,268.74)
Low Treated	83,396.433	2,423.772	80,972.66
	(39,089.283)	(17,695.007)	(37,002.37)
Test pval (H=0)	0.057	0.174	0.284
Test pval (L^t= 0)	0.034	0.891	0.029
Test pval (L^t=H)	0.530	0.245	0.180
Panel C: Closed schools with imputed values			
High	58,667.176	31,121.539	27,545.64
	(28,321.977)	(20,337.311)	(26,233.01)
Low Treated	54,458.801	-14,243.563	68,702.36
	(39,121.093)	(17,875.309)	(37,368.25)
Test pval (H=0)	0.039	0.127	0.294
Test pval (L^t= 0)	0.165	0.426	0.066
Test pval (L^t=H)	0.922	0.042	0.285

Notes:
a) This table reports the separate treatment impacts on operating profits and its components, with the three panels showing the sensitivity of our results to how closed schools are treated. Once a school closes, we either code variables as missing (Panel A), or as 0 (Panel B), or use imputed values (Panel C). In the third approach, we impute data by using coefficients from a regression of open schools in the control group to predict outcomes using baseline variables and round and strata fixed effects for H, L^u, and control schools. Columns 1 and 2 show impacts on annual collected revenues and annual variable expenditures, the two components of profits; we use the top coded (at the 99pctl) versions of these variables. Column 3 shows the impacts on profits— the coefficients and standard errors are calculated using a seemingly unrelated regression approach that uses the estimates from columns 1 and 2. For each panel, we also show p-values from tests of difference for our main comparisons.

Table B20: Distributional Effects, Socioeconomic Status and Gender

		Disadvantaged children			Female children		
	(1) Mean HH Assets	(2) Fraction	(3) Number	(4) Number New	(5) Fraction	(6) Number	(7) Number New
High	-0.002	-0.006	-0.417	0.329	0.030	-0.009	0.401
	(0.007)	(0.023)	(0.865)	(0.181)	(0.019)	(0.844)	(0.269)
Low Treated	-0.006	0.033	0.100	0.519	0.027	-0.398	0.219
	(0.008)	(0.030)	(1.090)	(0.247)	(0.022)	(1.020)	(0.208)
Low Untreated	-0.004	0.008	-0.249	0.018	0.032	0.258	0.027
	(0.007)	(0.023)	(0.911)	(0.157)	(0.018)	(0.901)	(0.165)
R-Squared	0.09	0.10	0.17	0.10	0.05	0.16	0.06
# Schools (Rounds)	725 (1)	725 (1)	725 (1)	725 (1)	725 (1)	725 (1)	725 (1)
Control Mean	0.56	0.36	8.85	0.66	0.44	12.11	0.85
Test pval (H=0)	0.75	0.81	0.63	0.07	0.11	0.99	0.14
Test pval (L^t= 0)	0.47	0.27	0.93	0.04	0.22	0.70	0.29
Test pval (L^t=H)	0.62	0.17	0.55	0.43	0.87	0.69	0.54

Notes:
a) This table considers the distributional impacts of our treatments. These data come from child surveys conducted along with child testing in the first year after treatment. For columns 1-4, we leverage a measure of average household assets generated from child reports on whether 24 different assets are present in their households. Column 1 looks at the mean household assets of children enrolled at the school in the follow-up round. Columns 2 and 3 looks at the fraction and number of disadvantaged children at the schools, where a child is defined as disadvantaged if their household assets are below the 25th percentile of the sample. Column 4 considers the number of new disadvantaged children. Columns 5-7 consider the fraction, number, and number new of female children at the school level.
b) Regressions are weighted to adjust for sampling and include strata fixed effects, with standard standard errors clustered at village level. As we do not have baseline data for these variables, we report the control mean of the dependent variable in the bottom panel.
c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

C Theory

Our theoretical exercise consists of two parts. First, we introduce credit constrained firms and quality into a canonical Kreps and Scheinkman (1983) framework. Schools in our model are willing to increase their capacities or qualities, but are credit constrained beyond their initial capital; the unconditional grants alleviate these constraints. Second, we introduce comparative static exercises through the provision of unconditional grants and study the equilibrium with varying degrees of financial saturation.

Setup

Two identical private schools, indexed by $i = 1, 2$, choose whether to invest in capacity, $x_i \geq 0$, or quality, q_t, where $t \in \{H, L\}$ is high or low quality.[7] High quality is conceptualized as investments that may allow schools to offer better quality/test scores and charge higher prices, such as specialty infrastructure (e.g. library or sports facility) or higher-quality teachers. On the other hand, capacity investment, such as basic hard infrastructure (desks, chairs) or basic renovations/upgrades, allow schools to retain or increase enrollment but do not change existing students' willingness to pay.[8]

SCHOOLS: Each school i maximizes $\Pi_i = (p_i - c)x_i^e + K_i - rx_i - w_t$ subject to $rx_i + w_t \leq K_i$ and $x_i^e \leq x_i$, where x_i^e is the enrollment, p_i is the price of school i per seat, c is the constant marginal cost for a seat, r is the fixed cost for a seat, w_t is the fixed cost for quality type, and K_i is the amount of fixed capital the school has. Schools face the same marginal and fixed costs for investments. The fixed cost for low quality is normalized to 0, and so w is the fixed cost of delivering high quality.[9]

STUDENTS: There are T students each of whom demands only one seat. Each student j has a taste parameter for quality θ_j and maximizes utility $U(\theta_j, q_t, p_i) = \theta_j q_t - p_i$ by choosing a school with quality q_t and fee p_i. The value of the outside option is zero for all students, and students choose to go to school as long as $U \geq 0$.

TIMING: The investment game has three stages. In the first stage, schools simultaneously choose their capacity and quality. After observing these choices, schools simultaneously choose their prices in the second stage. Demand is realized after the prices are revealed. Standard allocation rules are assumed. In particular, we assume (i) the school offering the higher surplus to students serves the entire market up to its capacity and the residual demand is met by the other school; (ii) if schools set the same price and quality, market demand is split in proportion to their capacities as long as their capacities are not met; (iii) if schools choose different qualities but offer the same surplus, then the school offering the higher quality serves the entire market up to its

[7] The model can be extended to allow for school heterogeneity but doing so does not generate results that are qualitatively different from those generated in this basic version. See the section at the end of this Appendix for further discussion on this issue.

[8] The model's use of capacity is more general than just desks, chairs and classrooms. The following interpretation is also consistent with the model's use of capacity: Fixing the fees, capacity investment is all kind of investments that make students (who are not attending a private school) willing to go to a private school. Therefore, the model's use of capacity includes some forces that horizontal differentiation would generate. Nevertheless, one could incorporate horizontal differentiation along with vertical differentiation in the model (see the numerical examples at the end of this Appendix). Such inclusions do not necessarily change the main message of the theory, and are thus omitted for simplicity.

[9] Alternative parameterizations for the profit function will naturally lead to different set of equilibrium outcomes. However, our main results, which are concerned about the comparative static comparison between the high- and low-saturation treatments, will remain unaffected as long as parameterizations do not vary with the treatment arm. This point is discussed further at the end of this Appendix.

capacity and the residual demand is met by the other school.

Equilibrium Analysis

Consider first the baseline scenario, where it is given that each school produces low quality and has $M/2$ students. Therefore, $M < T$ refers to the covered market and $N = T - M$ is the size of the uncovered market. It is straightforward to show that at baseline schools charge the same price $p = q_L$, extract full consumer surplus and earn positive profits. Schools do not lower prices since they cannot meet the additional demand. Next, we examine the subgame perfect Nash equilibrium (or equilibrium in short) of the investment game under the two comparative static exercises: The L arm where only one school receives a grant $K > 0$, and the H arm where both schools receive the same grant K. Grants received are common knowledge among all schools.

C.1 Theory with Homogeneous Demand

Here we assume that students are homogeneous with $\theta = 1$. When schools receive additional financing, they have to trade off increasing capacity at the risk of price competition versus increasing quality at a (possibly) higher cost. Two key parameters that influence the investment strategies of the schools are the cost of quality, w, and the size of the uncovered market, N. When both w and N are very low, schools prefer to invest in quality in both treatment arms. For sufficiently high values of w, schools in both treatments prefer to invest in capacity as long as N is quite large. As N decreases, schools will invest in capacity as long as increasing revenues through new students is more rewarding than increasing revenues among existing students through higher quality and prices, but spend less of their grants to escape from price competition. At a threshold level of N, at least one of the schools switches to quality investment instead of a partial expansion in capacity. This threshold for N decreases as w increases, suggesting a negative relationship between the two. We formally prove these claims for both treatment arms and characterize the $wN-$space where quality investment by at least one school is consistent with equilibrium.

Because the schools are credit constrained, they cannot afford high quality if its cost is greater than the grant size. Therefore, we are concerned with the part of the $wN-$space where quality investment is feasible, i.e., $w \leq K$. We also parametrize the size of the grant, K to be neither 'too small' nor 'too large.' In particular, we assume that K is large enough such that investing in quality is not always the optimal action but small enough so that rate of return of each investment is positive.[10]

[10] We suppose that $\underline{k} < K < \bar{k}$ where $\underline{k} = \frac{Mr}{2}\left(\frac{q_H - q_L}{q_l - c}\right)$ and $\bar{k} = \frac{M}{2}(q_H - q_L)$. If the inequality $\underline{k} < K$ does not hold, then the revenue from capacity investment, $\frac{K}{r}(q_L - c)$, is lower than revenue from quality (only) investment, $\frac{M}{2}(q_H - q_L)$, and thus, quality investment is always optimal. The rate of return from capacity investment is positive because we assume $q_L - c - r > 0$. Finally, $K < \bar{k}$ implies that rate of return from quality (only) investment is also positive. The last assumption is not crucial for the qualitative nature of our results, but eliminates a significant number of additional constraints one needs to consider. The figures in the proof of Theorem 0 indicate how $\mathbf{E_L}$ and $\mathbf{E_H}$ sets would change if we relax this assumption.

Figure C1(a): Low-saturation Treatment

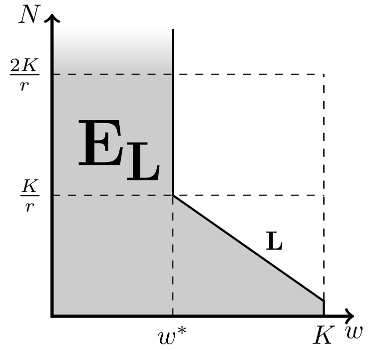
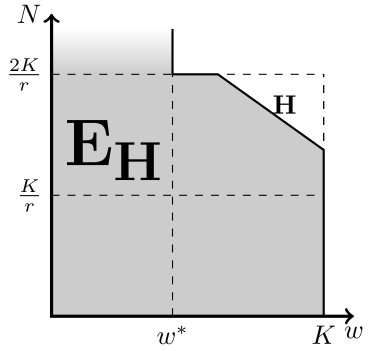


Figure C1(b): High-saturation Treatment



Theorem 0. *The shaded regions* $\mathbf{E_L}$ *and* $\mathbf{E_H}$ *in Figure C1 represent the set of parameters in* $wN-$*space where there exists an equilibrium of the investment game in the low and high-saturation treatment, respectively, such that (at least one) treated school invests on quality.*

The area E_H indicates parameters (w and N) that sustain equilibrium in which at least one of the schools invests in quality. A reader may think that there may also exist equilibria in E_H where none of the schools will invest in quality. This is not correct. That is, in region E_H there is no equilibrium where both schools invest in capacity.

Moreover, in the high-saturation treatment, if parameters are in region E_L, which is a subset of E_H, then the only equilibrium is such that both schools invest in quality. In this region, there is no equilibrium where only one school invests in quality and other invests in capacity. Here is the reason: When parameters are in region E_L, a school prefers investing in quality even though it has the monopoly power to cover the entire uncovered market. When there are two schools with the cash grant to invest in capacity, the schools will certainly prefer to invest in quality for all parameters in E_L. Both schools investing in quality remains to be the only equilibrium as we slowly move outside of this region, i.e., as N gradually increases. This is true because schools will continue to prefer investing in quality rather than capacity as the size of the uncovered market is not as high as the level that triggers deviation to capacity investment. However, schools' incentive to invest in quality will decrease gradually as N and w increase, and so at some point (before we hit the boundary of E_H), we will start observing other (or multiple) equilibria, i.e., one school invests in quality and other invests in capacity.

Suppose that the size of the uncovered market is sufficiently large such that the treated L school cannot cover it even if it spends the entire grant in capacity, i.e., $K/r \leq N$. If this school increases capacity, then the gain in profits is equal to the return on each new student times the number of new students, $(q_L - c)\frac{K}{r}$. If it invests in quality instead, then the gain in profits is equal to the increase in return on existing students from the higher price times the number of existing students plus the return from higher quality to each new student times the number of new students, $(q_h - q_L)\frac{M}{2} + (q_h - c)\frac{K-w}{r}$. Therefore, investing in capacity is more profitable if the former term is greater than the latter, yielding the condition $w > w^*$ where $w^* = r\left(\frac{q_H - q_L}{q_H - c}\right)\left(\frac{M}{2} + \frac{K}{r}\right)$. However, if the size of the uncovered market is smaller, in particular $N < K/r$, then spending the entire grant in capacity implies that the treated school must steal some students from the rival school, resulting in a price war and lower payoffs. In order to escape from lower payoffs, treated school will partially invest in capacity. The line **L** indicates the parameters w and N that equate the treated school's profit from quality investment to its profit from partial capacity investment.[11]

On the other hand, schools will never engage in a price war in the H treatment arm as long as the uncovered market size is large enough, so that schools cannot cover it even if both spend the entire grant on capacity, i.e. $2K/r \leq N$. Therefore, for these values of N, equilibrium predictions will be no different than the L arm. However, when N is less than $2K/r$, spending the entire grant on additional capacity implies that schools must steal some students from the rival school, resulting again in a price war. The constraint indicating the indifference between profit from quality investment and profit from partial capacity investment, the line **H** in Figure $C1(b)$, is much farther out because now both schools can invest in capacity, and hence price competition is likely even for higher values of the uncovered market size, N.[12] The next result is self evident from the last two figures and thus provided with no formal proof.

Corollary 1. *If the treated school in the low-saturation treatment invests on quality, then there must exist an equilibrium in the high-saturation treatment that at least one school invests on quality. However, the converse is not always true.*

Proof of Theorem 0: Let schools choose $x_1, x_2 \geq 0$ and $q_1, q_2 \in \{q_H, q_L\}$ in the first stage of the investment game and p_1, p_2 in the second stage. Define s_i to be school i's surplus, that is $s_i = q_i - p_i$. Therefore, school i's profit function is:

$$\Pi_i = \begin{cases} (p_i - c)(x_i + \frac{M}{2}) - rx_i - w_t + K, & \text{if } [s_i > s_j] \textit{ or } [s_i = s_j \textit{ and } q_i > q_j] \\ (p_i - c)(N - x_j + \frac{M}{2}) - rx_i - w_t + K, & \text{if } [s_i < s_j] \textit{ or } [s_i = s_j \textit{ and } q_i < q_j] \\ (p_i - c)\frac{(M/2 + x_i)T}{M + x_i + x_j} - rx_i - w_t + K, & \text{if } s_i = s_j \textit{ and } q_i = q_j \end{cases}$$

Define $n_H = \frac{K-w}{r}$ and $n_L = \frac{K}{r}$ to be the additional capacity increase that schools can afford together with high and low quality investment, respectively. Note that feasibility requires that $x_i \leq n_L$ and $x_i \leq n_H$ whenever $q_i = q_H$. One can easily verify that if the schools' capacity choices x_1 and x_2 are such that $x_1 + x_2 \leq N$, then in the pricing stage, school i picks $p_i = q_i$. Let μ be a probability density function with support $[\underline{p}, \bar{p}]$. Then for notational simplicity, we use $\hat{\mu}(p)$ for any $p \in [\underline{p}, \bar{p}]$ to denote $\mu(\{p\})$. Before proving Theorem 0, we prove the following result, which applies to both L and H treatment arms.

Proposition A. *Suppose that the schools' quality choices are $q_1, q_2 \in \{q_H, q_L\}$ and capacity choices are $x_1, x_2 \geq 0$ with $x_1, x_2 \leq N + \frac{M}{2}$ and $x_1 + x_2 > N$. Then, in the (second) pricing stage, there exists no pure strategy equilibrium. However, there exists a mixed strategy equilibrium (μ_1^*, μ_2^*), where for $i = 1, 2$, μ_i^* is*

(i) a probability density function with support $[p_i^*, q_i]$, satisfying $c < p_i^* < q_i$, and

(ii) atomless except possibly at q_i, that is $\hat{\mu}_i^*(p) = 0$ for all $p \in [p_i^*, q_i)$.

(iii) Furthermore, $\hat{\mu}_1^*(p_1)\hat{\mu}_2^*(p_2) = 0$ for all $p_1 \in [p_1^*, q_1]$ and $p_2 \in [p_2^*, q_2]$ satisfying $q_1 - p_1 = q_2 - p_2$.

Proof of Proposition A. Because no school alone can cover the entire market, i.e., $x_i < N + \frac{M}{2}$, $p_1 = p_2 = c$ cannot be an equilibrium outcome. Likewise, given that the schools compete in a Bertrand fashion and total capacity, $M + x_1 + x_2$, is greater than

[11] More formally, **L** represents the line $(q_H - c)\left(\frac{M}{2} + \frac{K-w}{r}\right) = (q_L - c)\left(\frac{M}{2} + N\right) + K - Nr$.

[12] More formally, **H** represents the line $(q_H - c)(\frac{M}{2} + \frac{K-w}{r}) = (q_L - c)(\frac{M}{2} + N - \frac{K}{r}) - Nr$.

total demand, $M+N$, showing that there is no pure strategy equilibrium is straightforward, and left to the readers.

However, by Theorem 5 of Dasgupta and Maskin (1986), the game has a mixed-strategy equilibrium: The discontinuities in the profit functions $\Pi_i(p_1, p_2)$ are restricted to the price couples where both schools offer the same surplus, that is $\{(p_1, p_2) \in [c, q_H]^2 | q_1 - p_1 = q_2 - p_2\}$. Lowering its price from a position $c < q_1 - p_1 = q_2 - p_2 \leq q_H$, a school discontinuously increases its profit. Hence, $\Pi_i(p_1, p_2)$ is weakly lower semi-continuous. $\Pi_i(p_1, p_2)$ is also clearly bounded. Finally, $\Pi_1 + \Pi_2$ is upper semi-continuous because discontinuous shifts in students from one school to another occur where either both schools derive the same profit per student (when $q_1 = q_2$) or the total profit stays the same or jumps per student because the higher quality school steals the student from the low quality school and charges higher price (when $q_1 \neq q_2$). Thus, by Theorem 5 of Dasgupta and Maskin (1986), the game has a mixed-strategy equilibrium.

Suppose that (μ_1^*, μ_2^*) is a mixed-strategy equilibrium of the pricing stage. Let $\bar{p}_i$ be the supremum of the support of μ_i^*, so $\bar{p}_i = \inf\{p \in [c, q_i] | p \in supp(\mu_i^*)\}$. Likewise, let p_i^* be the infimum of the support of μ_i^*. Define $s(p_i, q_i)$ to be the surplus that school i offers, so $s(p_i, q_i) = q_i - p_i$. We will prove the remaining claims of the proposition through a series of Lemmata.

Lemma A1. *$s(p_1^*, q_1) = s(p_2^*, q_2)$ and $p_i^* > c$ for $i = 1, 2$.*

Proof. Note that the claim turns into the condition $p_1^* = p_2^* > c$ when $q_1 = q_2$. To show $s(p_1^*, q_1) = s(p_2^*, q_2)$, suppose for a contradiction that $s(p_1^*, q_1) \neq s(p_2^*, q_2)$. Assume, without loss of generality, that $s(p_1^*, q_1) > s(p_2^*, q_2)$. For any $p_1 \geq p_1^*$ in the support of μ_1^* satisfying $s(p_1^*, q_1) \geq s(p_1, q_1) > s(p_2^*, q_2)$, player 1 can increase its expected profit by deviating to a price $p_1' = p_1 + \epsilon$ satisfying $s(p_1', q_1) > s(p_2^*, q_2)$. This is true because by slightly increasing its price from p_1 to p_1' school 1 keeps its expected enrollment the same. This opportunity of a profitable deviation contradicts with the optimality of equilibrium. The case for $s(p_1^*, q_1) < s(p_2^*, q_2)$ is symmetric. Thus, we must have $s(p_1^*, q_1) = s(p_2^*, q_2)$.

Showing that $p_i^* > c$ for $i = 1, 2$ is straightforward: Suppose for a contradiction that $p_i = c$ for some i, so school i is making zero profit per student it enrolls. However, because no school can cover the entire market, i.e., $x_j < \frac{M}{2} + N$, school i can get positive residual demand and positive profit by picking a price strictly above c, contradicting the optimality of equilibrium. □

Definition 1. *Let $[a_i, b_i)$ be a non-empty subset of $[c, q_i]$ for $i = 1, 2$. Then $[a_1, b_1)$ and $[a_2, b_2)$ are called surplus-equivalent if $s(a_1, q_1) = s(a_2, q_2)$ and $s(b_1, q_1) = s(b_2, q_2)$.*

Lemma A2. *Let $[a_i, b_i)$ be a non-empty subset of $[c, q_i]$ for $i = 1, 2$. If $[a_1, b_1)$ and $[a_2, b_2)$ are surplus equivalent, then $\mu_1^*([a_1, b_1)) = 0$ if and only if $\mu_2^*([a_2, b_2)) = 0$.*

Proof. Take any two such intervals and suppose, without loss of generality, $\mu_1^*([a_1, b_1)) = 0$. That is, $[a_1, b_1)$ is not in the support of μ_1^*. Therefore, for any $p \in [a_2, b_2)$, player 2's expected enrollment does not change by moving to a higher price within this set $[a_2, b_2)$. However, player 2 receives a higher profit simply because it is charging a higher price per student. Hence, optimality of equilibrium implies that player 2 should never name a price in the interval $[a_2, b_2)$, implying that $\mu_2^*([a_2, b_2)) = 0$. □

Lemma A3. *If $p_i \in (c, q_i]$ for $i = 1, 2$ with $s(p_1, q_1) = s(p_2, q_2)$, then $\hat{\mu}_1^*(p_1)\hat{\mu}_1^*(p_2) = 0$.*

Proof. Suppose for a contradiction that there exists some p_1 and p_2 as in the premises of this claim such that $\hat{\mu}_1^*(p_1)\hat{\mu}_1^*(p_2) > 0$. Because $\hat{\mu}_1^*(p_1) > 0$, player 2 can enjoy the discrete chance of price-undercutting his opponent. That is, there exists sufficiently small $\epsilon > 0$ such that player 2 gets strictly higher profit by naming price $p_2 - \epsilon$ rather than price p_2. This contradicts the optimality of the equilibrium. $\square$

Lemma A4. *Equilibrium strategies must be atomless except possibly at $\bar{p}_i$. More formally, suppose that $s(\bar{p}_i, q_i) \geq s(\bar{p}_j, q_j)$ where $i, j \in \{1, 2\}$ and $j \neq i$, then for any $k \in \{1, 2\}$ and $p \in [c, q_H]$, satisfying $p \neq \bar{p}_j$, it must be the case that $\hat{\mu}^*_k(p) = 0$.*

Proof. Suppose without loss of generality that $k = 1$ and suppose for a contradiction that $\hat{\mu}^*_1(p) > 0$ for some $p \in [c, q_H] \setminus \{\bar{p}_j\}$. Therefore, there must exist sufficiently small $\epsilon > 0$ and $\delta > 0$ such that for all $p_2 \in I \equiv [q_2 - s(p, q_1), q_2 - s(p, q_1) + \epsilon)$ player 2 prefers to name a price $p_2 - \delta$ instead of p_2 and enjoy the discrete chance of price-undercutting his opponent. Therefore, the optimality of the equilibrium strategies suggests that $\mu^*_2(I) = 0$. Because the intervals $[p, p + \epsilon)$ and I are surplus-equivalent, Lemma A2 implies that we must have $\mu^*_1([p, p + \epsilon)) = 0$, contradicting $\hat{\mu}^*_1(p) > 0$. $\square$

Lemma A5. $s(\bar{p}_1, q_1) = s(\bar{p}_2, q_2) = 0$, *and thus $\bar{p}_i = q_i$ for $i = 1, 2$.*

Proof. To show $s(\bar{p}_1, q_1) = s(\bar{p}_2, q_2)$ suppose for a contradiction that $s(\bar{p}_1, q_1) \neq s(\bar{p}_2, q_2)$. Suppose, without loss of generality, that $s(\bar{p}_2, q_2) > s(\bar{p}_1, q_1)$. Therefore, by Lemma A4 we have $\mu^*_2([\bar{p}_2, \tilde{p}_2)) = 0$ where $\tilde{p}_2 \equiv q_2 - s(\bar{p}_1, q_1)$, and by Lemma A2 $\mu^*_1([\tilde{p}_1, \bar{p}_1)) = 0$ where $\tilde{p}_1 \equiv q_1 - s(\bar{p}_2, q_2))$. In fact, there must exist some small $\epsilon > 0$ such that $\mu^*_1([\tilde{p}_1 - \epsilon, \bar{p}_1)) = 0$. The last claim is true because player 1 prefers to deviate from any $p \in [\tilde{p}_1 - \epsilon, \tilde{p}_1]$ to price $\bar{p}_1$ since the change in profit, $\Pi_1(p, p_2) - \Pi_1(\bar{p}_1, p_2)$ is equal to $(p - c)\mu^*([p, \tilde{p}_1])x_1 - (\bar{p}_1 - c)(T - x_2) < 0$ as ϵ converges zero. Because the sets $[\bar{p}_2 - \epsilon, \tilde{p}_2)$ and $[\tilde{p}_1 - \epsilon, \bar{p}_1)$ are surplus-equivalent and $\mu^*_1([\tilde{p}_1 - \epsilon, \bar{p}_1)) = 0$, Lemma A2 implies that $\mu^*_2([\bar{p}_2 - \epsilon, \tilde{p}_2)) = 0$, contradicting that $\bar{p}_2$ is the supremum of the support of μ^*_2. Thus, $s(\bar{p}_1, q_1) = s(\bar{p}_2, q_2)$ must hold.

To show that $s(\bar{p}_i, q_i) = 0$ for $i = 1, 2$, assume for a contradiction that $s(\bar{p}_1, q_1) = s(\bar{p}_2, q_2) > 0$. By Lemma A3 we know that $\hat{\mu}^*_1(\bar{p}_1)\hat{\mu}^*_1(\bar{p}_2) = 0$. Suppose, without loss of generality, that $\hat{\mu}^*_1(\bar{p}_1) = 0$. Therefore, player 2 can profitable deviate from price $\bar{p}_2$ to price q_2: the deviation does not change player 2's expected enrollment, but it increases expected profit simply because player 2 is charging a higher price per student it enrolls. This contradicts with the optimality of the equilibrium, and so we must have $s(\bar{p}_i, q_i) = 0$ for $i = 1, 2$. $\square$

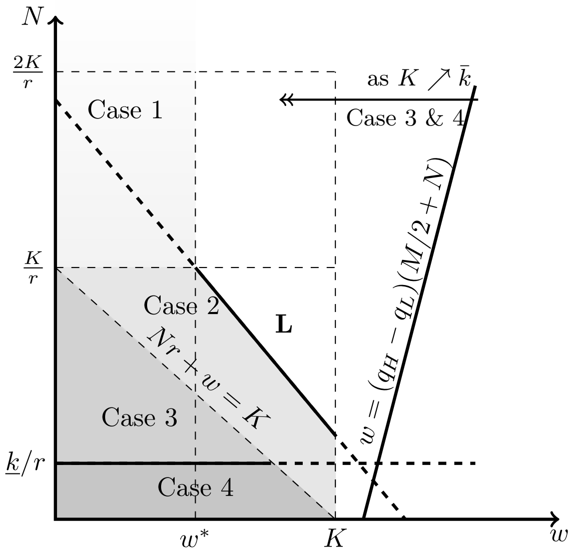
Lemma A6. *For each $i \in \{1, 2\}$, $\bar{p}_i > p^*_i$, and there exists no p, p' with $p^*_i < p < p' < q_i$ such that $\mu^*_i([p, p']) = 0$.*

Proof. If $\bar{p}_i = p^*_i$ for some i, that is player i is playing a pure strategy, then player j can profitably deviate from q_j by price undercutting its opponent, contradicting the optimality of equilibrium.

Next, suppose for a contradiction that there exists p, p' with $p^*_i < p < p' < q_i$ such that $\mu^*_i([p, p']) = 0$. By Lemma A2, there exists p_j, p'_j that are surplus equivalent to p, p', respectively, and $\mu^*_j([p_j, p'_j]) = 0$. Then the optimality of equilibrium and Lemma A4 implies that there exists some $\epsilon > 0$ such that $\mu^*_i([p - \epsilon, p']) = 0$. This is true because instead of picking a price in $[p - \epsilon, p]$, school i would keep expected enrollment the same and increase its profit by picking a higher price p'. Repeating the same arguments will eventually yield the conclusion that we have $\mu^*_i([p^*_i, p']) = 0$, contradicting the assumption that p^*_i is the infimum of the support of μ^*_i. $\square$

For the rest of the proofs, we use Π_t to denote the profit of a school that picks quality $t \in \{H, L\}$. Let Π_H^{Dev} denote the deviation profit of a school that deviates from high to low quality (once the other school's actions are fixed). Similarly, Π_L^{Dev} denotes the deviation profit of a school that deviates from low to high quality.

Proof of Theorem 0 (Low-Saturation Treatment). Suppose that (only) school 1 receives the grant. Because the schools are symmetric, this does not affect our analysis. There are four exhaustive cases we must consider for the low-saturation treatment and all these cases are summarized in the following figure:



Case 1: $K \leq Nr$ (or equivalently $n_L \leq N$): There would be no price competition among the schools whether school 1 invests in capacity or quality. Therefore, $\Pi_H = (q_H - c)\left(\frac{M}{2} + \frac{K-w}{r}\right)$ and $\Pi_L = (q_L - c)\left(\frac{M}{2} + \frac{K}{r}\right)$. Thus, there is an equilibrium where school 1 invests in quality if and only if $\Pi_H \geq \Pi_L$, implying $w \leq w^*$.

Case 2: $K - w \leq Nr < K$ (or equivalently $n_H \leq N < n_L$): If school 1 invests in quality, then $\Pi_H = (q_H - c)\left(\frac{M}{2} + \frac{K-w}{r}\right)$. But if it invests in capacity, then its optimal choice would be $x_1 = N$ (as we formally prove below) and profit would be $\Pi_L = (q_L - c)\left(\frac{M}{2} + N\right) + K - Nr$.

Claim: *If school 1 invests in capacity, then its optimal capacity choice x_1 is such that $x_1 = N$.*

Proof. Suppose for a contradiction that $x_1 = N + e$ where $e > 0$. In the mixed strategy equilibrium of the pricing stage, each school i randomly picks a price over the range $[p_i^*, q_L]$ with a probability measure μ_i. School 1's profit functions are given by $\Pi_1(q_L, \mu_2) = (q_L - c)\left[\frac{\hat{\mu}_2(M/2+x_1)(M+N)}{M+x_1} + (1-\hat{\mu}_2)\left(\frac{M}{2} + N\right)\right] + K - rx_1$, where $\hat{\mu}_2 = \hat{\mu}_2(q_L)$, and $\Pi_1(p_1^*, \mu_2) = (p_1^* - c)(x_1 + M/2) + K - rx_1$. However, school 2's profit functions are $\Pi_2(q_L, \mu_1) = (q_L - c)\left[\frac{\hat{\mu}_1(M/2)(M+N)}{M+x_1} + (1-\hat{\mu}_1)\left(\frac{M}{2} + N - x_1\right)\right]$, where $\hat{\mu}_1 = \hat{\mu}_1(q_L)$ and $\Pi_2(p_2^*, \mu_1) = (p_2^* - c)(\frac{M}{2})$.

In equilibrium both schools offer the same surplus, and so $p_1^* = p_2^*$ holds. Moreover, because each school i is indifferent between q_L and p_i^* we must have $\Pi_1(q_L, \mu_2) = \Pi_1(p_1^*, \mu_2)$ and $\Pi_2(q_L, \mu_1) = \Pi_2(p_2^*, \mu_1)$. We can solve these equalities for $\hat{\mu}_1$ and $\hat{\mu}_2$. However, we know that in equilibrium we must have $\hat{\mu}_1\hat{\mu}_2 = 0$. If $\hat{\mu}_2 = 0$, then it is easy to see that $\Pi_1(q_L, \mu_2)$ decreases with x_1 (or e), and thus optimal capacity should be $x_1 = N$. However, $\hat{\mu}_1 = 0$ yields $\hat{\mu}_2 = -\frac{4(e+N)(e+M+N)}{M^2} < 0$, contradicting with the optimality of equilibrium because we should have $\hat{\mu}_2 \geq 0$. Thus, school 1's optimal capacity is $x_1 = N$. □

Therefore, school 1 selects high quality if and only if $\Pi_H \geq \Pi_L$, which implies

$$(q_L - c - r)N + (q_H - c)\frac{w}{r} \leq \frac{M}{2}(q_H - q_L) + (q_h - c - r)\frac{K}{r}.$$

The last condition gives us the line **L**. Drawing the line **L** on $wN-$space implies that the $N-$intercept is greater than K/r and the $w-$intercept is greater than K whenever $K < \bar{k}$. Moreover, when $w = w^*$, N takes the value K/r and when $w = K$, N takes a value which is less than K/r because $K > \underline{k}$.

Case 3: $\frac{Mr}{2}\frac{(q_H - q_L)}{(q_L - c)} \leq Nr < K - w$ (or equivalently $\underline{k}/r \leq N < n_H$)

Claim: *If school 1 invests in quality, then its optimal capacity choice x_1 is such that $x_1 = N$.*

Proof. Suppose for a contradiction that $x_1 = N + e$ where $e > 0$. This time school 1 randomly picks a price over the range $[p_1^*, q_H]$ with a probability measure μ_1 and school 2 randomly picks a price over the range $[p_2^*, q_L]$ with a probability measure μ_2. Schools' profit functions are given by $\Pi_1(q_H, \mu_2) = (q_H - c)\left[\hat{\mu}_2\left(\frac{M}{2} + x_1\right) + (1 - \hat{\mu}_2)(M/2 + N)\right] + K - rx_1 - w$ and $\Pi_1(p_1^*, \mu_2) = (p_1^* - c)(x_1 + \frac{M}{2}) + K - rx_1 - w$ for school 1 and $\Pi_2(q_L, \mu_1) = (q_L - c)(\frac{M}{2} + N - x_1)$ and $\Pi_2(p_2^*, \mu_1) = (p_2^* - c)(\frac{M}{2})$ for school 2.

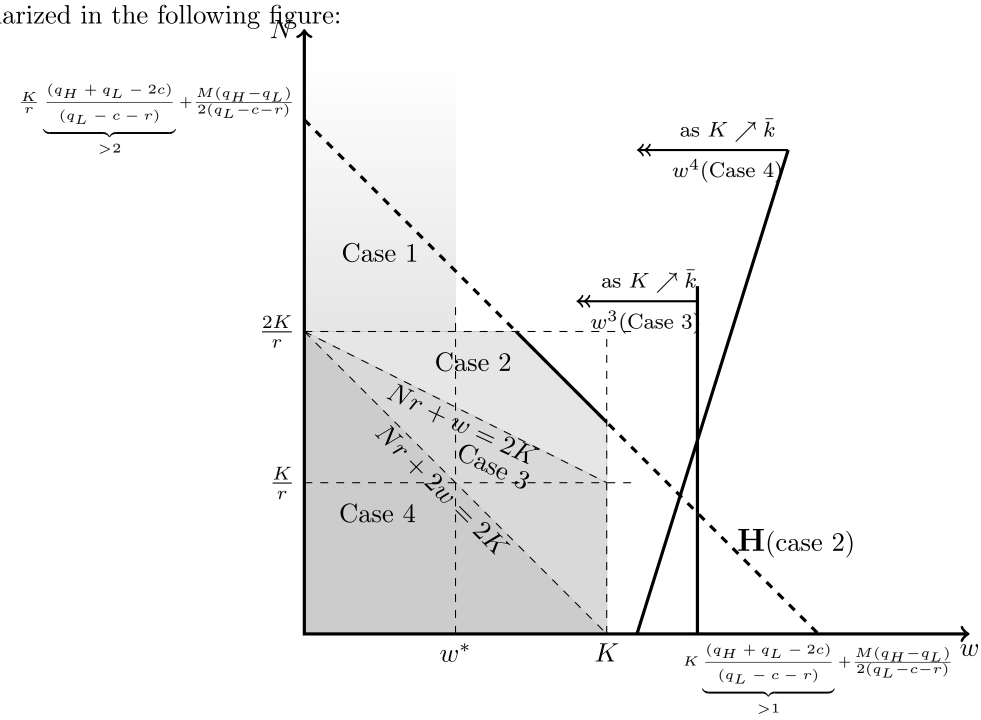
This time equilibrium prices must satisfy $q_H - p_1^* = q_L - p_2^*$. Solving this equality along with $\Pi_1(q_L, \mu_2) = \Pi_1(p_1^*, \mu_2)$, and $\Pi_2(q_L, \mu_1) = \Pi_2(p_2^*, \mu_1)$ implies that either $\hat{\mu}_2 = 0$, and thus $\Pi_1(q_L, \mu_2)$ decreases with x_1 and the optimal capacity should be $x_1 = N$, or $\hat{\mu}_1 = 0$ and $\hat{\mu}_2 \geq 0$. However, solving for $\hat{\mu}_2$ implies that $\hat{\mu}_2 = \frac{q_H - q_L}{q_H - c} - \frac{2(q_L - c)(e+N)}{M(q_H - c)}$ which is less than zero for all $e > 0$ whenever $\underline{k}r \leq N$. This contradicts with the optimality of the equilibrium, and thus school 1's optimal capacity is $x_1 = N$. □

Therefore, school 1's profit is $\Pi_H = (q_H - c)(\frac{M}{2} + N) + K - w - Nr$ if it invests in quality and $\Pi_L = (q_L - c)(\frac{M}{2} + N) + K - Nr$ if it invests in capacity. Therefore, investing in quality is optimal if and only if $w \leq (q_H - q_L)(\frac{M}{2} + N)$ which holds for all N and w as long as $K < \bar{k}$.

Case 4: $Nr < \frac{Mr}{2}\frac{(q_H - q_L)}{(q_L - c)}$ (or equivalently $Nr < \underline{k}$): In this case, school 1 prefers to select $x_1 > N$ and start a price war. This is true because the profit maximizing capacity (derived from the profit function Π_H calculated in the previous case) is greater than N, and so price competition ensues. Therefore, school 1's profit function is strictly greater than $(q_H - c)(\frac{M}{2} + N) + K - w - Nr$ if it invests in quality. However, if school 1 invests in capacity, then as we proved in the second case school 1 prefers to choose its capacity as N, and thus its profit would be $\Pi_L = (q_L - c)(\frac{M}{2} + N) + K - Nr$. Therefore, school

1 prefers to invest in quality as long as the first term is greater than or equal to Π_L, implying that $w \leq (q_H - q_L)(\frac{M}{2} + N)$ which is less than K because $K < \bar{k}$.

Proof of Theorem 0 (High-Saturation Treatment). There are four exhaustive cases we must consider for the high-saturation treatment and all these cases are summarized in the following figure:



Case 1: Suppose that $2K \leq Nr$ (or equivalently, $2n_L \leq N$): Because the uncovered market is large, price competition never occurs in this case. Therefore, $\Pi_H = (q_H - c)(\frac{M}{2} + \frac{K-w}{r})$ and $\Pi_L = (q_L - c)(\frac{M}{2} + \frac{K}{r})$. Moreover, $\Pi_H^{Dev} = (q_L - c)(\frac{M}{2} + \frac{K}{r})$ and $\Pi_L^{Dev} = (q_H - c)(\frac{M}{2} + \frac{K-w}{r})$.

To have an equilibrium where one school invests in high quality and the other invests in low quality, we must have $\Pi_H \geq \Pi_H^{Dev} = \Pi_L$ and $\Pi_L \geq \Pi_L^{Dev} = \Pi_H$ implying that $w = w^*$, which is less than K because $\underline{k} < K$. To have an equilibrium where both schools pick the high quality, we must have $\Pi_H \geq \Pi_H^{Dev}$, implying $w \leq w^*$. Hence, there exists an equilibrium where at least one school invests in quality if and only if $w \leq w^*$.

Case 2: Suppose that $2K - w \leq Nr < 2K$ (or equivalently, $n_L + n_H \leq N < 2n_L$): Because we still gave $n_H + n_H \leq N$, there exists an equilibrium where (H, H) is an equilibrium outcome for all values of $w \leq w^*$. Now, consider an equilibrium where only one school, say school 1, invests in high quality, and so (H, L) is the outcome. In this case $n_L + n_H \leq N$ and no price competition occurs, so $\Pi_H = (q_H - c)(\frac{M}{2} + \frac{K-w}{r})$ and $\Pi_L = (q_L - c)(\frac{M}{2} + \frac{K}{r})$. Moreover, $\Pi_L^{Dev} = (q_H - c)(\frac{M}{2} + \frac{K-w}{r})$ because the other school has picked n_H and $2n_H < N$. However, if school 1 deviates to low quality and picks quantity higher than n_L, price competition ensues. First we prove that it is not optimal for school 1 to pick a large capacity if it deviates to L.

Claim: *Consider an equilibrium strategy where both schools invest in capacity only and $x_2 = n_L$. Then school 1's optimal capacity choice x_1 is such that $x_1 = N - n_L$.*

Proof. Suppose for a contradiction that $x_1 = N - n_L + e$ where $e > 0$. In the mixed strategy equilibrium each school i randomly picks a price over the range $[p_i^*, q_L]$ with a probability measure μ_i and we have

$$\Pi_1(q_L, \mu_2) = (q_L - c)\left[\frac{\hat{\mu}_2(M/2 + x_1)(M+N)}{M + x_1 + x_2} + (1 - \hat{\mu}_2)\left(\frac{M}{2} + N - x_2\right)\right] + K - rx_1 \quad (1)$$

and

$$\Pi_1(p_1^*, \mu_2) = (p_1^* - c)(x_1 + M/2) + K - rx_1 \quad (2)$$

where $\hat{\mu}_2 = \mu_2(\{q_L\})$. Moreover,

$$\Pi_2(q_L, \mu_1) = (q_L - c)\left[\frac{\hat{\mu}_1(M/2 + x_2)(M+N)}{M + x_1 + x_2} + (1 - \hat{\mu}_1)\left(\frac{M}{2} + N - x_1\right)\right] + K - rx_2 \quad (3)$$

and

$$\Pi_2(p_2^*, \mu_1) = (p_2^* - c)(M/2 + x_2) + K - rx_2 \quad (4)$$

where $\hat{\mu}_1 = \mu_1(\{q_L\})$. In equilibrium we have $p_1^* = p_2^*$, $\Pi_1(q_L, \mu_2) = \Pi_1(p_1^*, \mu_2)$, and $\Pi_2(q_L, \mu_1) = \Pi_2(p_2^*, \mu_1)$. Moreover, if $\hat{\mu}_2 = 0$, then $\Pi_1(q_L, \mu_2)$ decreases with x_1, and thus the optimal capacity should be $x_1 = N - x_2$. Therefore, we must have $\hat{\mu}_1 = 0$. Solving for $\hat{\mu}_2 \geq 0$, and then solving $\partial\Pi_1(q_L, \mu_2)/\partial e = 0$ implies

$$e = \frac{K}{r} - \frac{N}{2} - \frac{Mr + 2K}{4(q_L - c)}.$$

Because $N \geq (2K - w)/r$, e is less than or equal to $-\frac{K-w}{r} - \frac{Mr+2K}{4(q_L-c)}$, which is negative because $K < w$, contradicting with the initial assumption that $e > 0$. □

Therefore, if school 1 deviates to low quality, then its payoff is $\Pi_H^{Dev} = (q_L - c)(\frac{M}{2} + N - \frac{K}{r}) - Nr$. Thus, there is an equilibrium with one school investing in quality and other investing in capacity if and only if $\Pi_L \geq \Pi_L^{Dev}$ and $\Pi_H \geq \Pi_H^{Dev}$, which implies the following two inequalities: $w \geq w^*$ and

$$w \leq \frac{Mr(q_H - q_L)}{2(q_H - c)} + \frac{(q_H + q_L - 2c)K}{q_H - c} - \frac{Nr(q_L - c - r)}{q_H - c}.$$

The last condition gives us the line **H**. Drawing the line **H** on $wN-$space implies that the N-intercept is greater than $2K/r$ because $\frac{q_H + q_L - 2c}{q_L - c - r} > 2$ and the w-intercept is bigger than K because $\frac{q_H + q_L - 2c}{q_H - c} > 1$. However, when $w = K$, **H** gives the value of $\frac{M(q_H - q_L)}{2(q_L - c - r)} + \frac{K(q_L - c)}{r(q_L - c - r)}$ for N which is strictly greater than K/r. However, it is less than or greater than $2K/r$ depending on whether $\frac{Mr(q_H - q_L)}{2(q_L - c - 2r)}$ is greater or less than K/r. That is, for sufficiently small values of K, **H** lies above $2K/r$. However, it is easy to verify that **H** always lies above K/r.

Case 3: Suppose that $2K - 2w \leq Nr < 2K - w$ (or equivalently, $2n_H \leq N < n_L + n_H$): Note that for all values of $w \leq w^*$ there exists an equilibrium where (H, H) is an equilibrium outcome. This is true because Π_H is the same as the one we calculated in Case 1 in the proof of Theorem 0 (Low-saturation Treatment) but Π_H^{Dev} is much less.

If (H, L) is an equilibrium outcome, then the optimal capacity for school 2 is $x_2 = N - x_1$. The reason for this is that if it ever starts a price war (i.e., a mixing equilibrium), then school 2 will only get the residual demand when it picks the price of q_L, implying that its payoff will be a decreasing function of x_2 as long as $x_2 > N - x_1$. On the other hand, because schools' profits increase with their capacity, as long as there is no price competition, the school 1's optimal capacity choice will be $x_1 = n_H = \frac{K-w}{r}$. Thus, in an equilibrium where (H, L) is the outcome, the profit functions are $\Pi_H = (q_H - c)\left(\frac{M}{2} + \frac{K-w}{r}\right)$ and $\Pi_L = (q_L - c)\left(\frac{M}{2} + N - \frac{K-w}{r}\right) + K - r\left(N - \frac{K-w}{r}\right)$. If school 2 deviates to high quality, then its deviation payoff is $\Pi_L^{Dev} = (q_H - c)\left(\frac{M}{2} + N - \frac{K-w}{r}\right)$ because $2n_H \leq N$. Now we prove that it is not optimal for school 1 to deviate to L and pick a large capacity that will ensue price competition.

Claim: *Consider an equilibrium strategy where both schools invest in capacity only and $x_2 = N - n_H$. Then school 1's optimal capacity choice x_1 is such that $x_1 = n_H$.*

Proof. Suppose for a contradiction that $x_1 = n_H + e$ where $e > 0$. In the mixed strategy equilibrium schools' profit functions are given by Equations 1-4 of Case 2. Once again, solving $p_1^* = p_2^*$, $\Pi_1(q_L, \mu_2) = \Pi_1(p_1^*, \mu_2)$, and $\Pi_2(q_L, \mu_1) = \Pi_2(p_2^*, \mu_1)$ imply that if $\hat{\mu}_2 = 0$, then $\Pi_1(q_L, \mu_2)$ decreases with x_1, and so the optimal capacity should be $x_1 = N - x_2$. Therefore, we must have $\hat{\mu}_1 = 0$. Solving for $\hat{\mu}_2 \geq 0$, and then solving $\partial \Pi_1(q_L, \mu_2)/\partial e = 0$ implies

$$e = \underbrace{\frac{N(q_L - c - r)}{2(q_L - c)} + \frac{w(2q_L - 2c - r)}{2r(q_L - c)}}_{e_1} - \frac{K(2q_L - 2c - r)}{r(q_l - c)} - \frac{Mr}{4(q_L - c)}.$$

which is strictly less than zero because $e_1 \leq \left(\frac{w}{2r} + \frac{N}{2}\right)\frac{(2q_L - 2c - r)}{(q_L - c)}$ and it is less than $\frac{K}{r}\frac{(2q_L - 2c - r)}{(q_L - c)}$ because we are in the region where $w + Nr < 2K$. However, $e < 0$ contradicts with our initial assumption. □

Therefore, $x_1 = n_H$ is the optimal choice for school 1 if it deviates to low quality, and thus we have $\Pi_H^{Dev} = (q_L - c)\left(\frac{M}{2} + \frac{K-w}{r}\right) + w$. To have an equilibrium outcome (H, L) we must have $\Pi_q \geq \Pi_q^{Dev}$ for each $q \in \{H, L\}$. Equivalently,

$$(q_L - c - r)N + \frac{w}{r}(q_H + q_L - 2c - r) \geq (q_H - q_L)\left(\frac{M}{2} + \frac{K}{r}\right) - 2K$$

and

$$(q_H - q_L)\left(\frac{M}{2} + \frac{K}{r}\right) \geq \frac{w}{r}(q_H - q_L + r).$$

It is easy to verify that the first inequality holds for all $w \geq w^*$ and $N \geq 0$. The second inequality implies $w \leq \frac{(q_H - q_L)r}{(q_H - q_L + r)}\left(\frac{M}{2} + \frac{K}{r}\right) \equiv w^3$ which is strictly higher than K whenever $K \leq \bar{k}$.

Case 4: Suppose that $Nr < 2K - 2w$ (or equivalently, $N < 2n_H$): We will prove, for all parameters in this range, that there exists an equilibrium where both schools invest in quality and $x_1 = x_2 = N/2$. For this purpose, we first show that school 1's best response is to pick $x_1 = N/2$ in equilibrium where both schools invest in quality and $x_2 = N/2$. Suppose for a contradiction that school 1 picks $x_1 = N/2 + e$ where $e > 0$. Then in the mixed strategy equilibrium of the pricing stage, each school i randomly picks a price over the range $[p_i^*, q_H]$ with a probability measure μ_i and the profit functions are given by

$$\Pi_1(q_H,\mu_2) = (q_H-c)\left[\frac{\hat{\mu}_2(\frac{M}{2}+x_1)(M+N)}{M+x_1+x_2} + (1-\hat{\mu}_2)\left(\frac{M}{2}+N-x_2\right)\right] + K - rx_1 - w$$

where $\hat{\mu}_2 = \mu_2(\{q_H\})$ and $\Pi_1(p_1^*,\mu_2) = (p_1^*-c)(x_1+M/2) - +K - rx_1 - w$. On the other hand,

$$\Pi_2(q_H,\mu_1) = (q_H-c)\left[\frac{\hat{\mu}_1(\frac{M}{2}+x_2)(M+N)}{M+x_1+x_2} + (1-\hat{\mu}_1)\left(\frac{M}{2}+N-x_1\right)\right] + K - rx_2 - w$$

where $\hat{\mu}_1 = \mu_1(\{q_H\})$ and $\Pi_2(p_2^*,\mu_1) = (p_2^*-c)(x_2+M/2) + K - rx_2 - w$.

Once again, solving $p_1^* = p_2^*$, $\Pi_1(q_H,\mu_2) = \Pi_1(p_1^*,\mu_2)$, and $\Pi_2(q_H,\mu_1) = \Pi_2(p_2^*,\mu_1)$ imply that if $\hat{\mu}_2 = 0$, then $\Pi_1(q_L,\mu_2)$ decreases with x_1, and so the optimal capacity should be $x_1 = N - x_2$. Therefore, we must have $\hat{\mu}_1 = 0$. Solving for $\mu_2 \geq 0$ yields $\hat{\mu}_2 = -\frac{4e(e+M+N)}{(M+N)^2}$ which is clearly negative for all values of $e > 0$, yielding the desired contradiction. Therefore, school 1's optimal capacity choice is $x_1 = N - x_2 = N/2$.

In equilibrium with (H,H) and $x_i = N/2$ for $i = 1,2$, profit function is $\Pi_H = (q_H - c)\left(\frac{M+N}{2}\right) + K - w - \frac{Nr}{2}$. However, if a school deviates to low quality, then its optimal capacity choice would still be $N/2$ because entering into price war is advantageous for the opponent, making profit of the deviating school a decreasing function of its own capacity (beyond $N/2$). Therefore, $\Pi_H^{Dev} = (q_L - c)\left(\frac{M+N}{2}\right) + K - \frac{Nr}{2}$. Thus, no deviation implies that $w \leq (q_H - q_L)\left(\frac{M+N}{2}\right) \equiv w^4$ which holds for all $w \leq \bar{k}$ and $N \geq 0$. That is, for all the parameters of interest, (H,H) is an equilibrium outcome.

C.2 Generalization of the Model

In this section, we suppose that each of T students has a taste parameter for quality θ_j that is uniformly distributed over $[0, 1]$ and rest of the model is exactly the same as before. Therefore, if the schools have quality q and price p, then demand is $D(p) = T(1-\frac{p}{q})$. In what follows, we first characterize the second stage equilibrium prices (given the schools' quality and capacity choices), and thus calculate the schools' equilibrium payoffs as a function of their quality and capacity. We do not need to characterize equilibrium prices when the schools' qualities are the same because they are given by KS. For that reason, we will only provide the equilibrium prices when schools' qualities are different. After the second stage equilibrium characterization, we prove, for a reasonable set of parameters, that if the treated school in the L arm invests in quality, then at least one of the schools in the H arm must invest in quality (Theorem 1). We prove this result formally for the case $w = K$, which significantly reduces the number of cases we need to consider. Therefore, even when the cost of quality investment is very high, quality investment in the H arm is optimal if it is optimal in the L arm. There is no reason to suspect that our result would be altered if the cost of quality investment is less than the grant amount, and thus we omit the formal proof for $w < K$. To build intuition, consider the following example with 10 consumers, A to J, who value low quality in descending order:

Consumers	A	B	C	D	E	F	G	H	I	J
Value for low quality	10	9	8	7	6	5	4	3	2	1

where A values low quality at \$10 and J at \$1. Following KS, the rationing rule allocates consumers to schools in order of maximal surplus.[13] Fix the capacity of the first school at 2 and let the capacity of the second school increase from 1 to 6. As School 2's capacity increases from 1 to 5, equilibrium prices in the second stage drop from \$8 to \$4 as summarized in the next table:[14]

Capacity of School 2	1	2	3	4	5	6
NE prices	8	7	6	5	4	mixed

The reason for the existence of pure strategy equilibrium prices is provided by Proposition 1 of KS that the schools' unique equilibrium price is the market clearing price whenever both schools' capacity is less than or equal to their Cournot best response capacities.[15] But, once school 2's capacity increases to 6, there is no pure strategy NE.[16] The threat of mixed strategy equilibrium prices forces schools to not expand their capacities beyond their Cournot optimal capacities.[17]

[13]Suppose that both schools have a capacity of 2 and school 1 charges \$7 and School 2 charges \$9. Then, the rationing rule implies that consumers A and B will choose School 1 since they obtain a higher surplus by doing so and consumer C is rationed out of the market.

[14]For example, the equilibrium price is \$8 when School 2 capacity is 1 because if school 1 charges more than \$8, given the rationing rule, A derives maximal surplus from choosing school 2 and School 1's enrollment declines to 1. A lower price also decreases profits since additional demand cannot be met through existing capacity.

[15]Given that school 1's capacity is 2, school 2's Cournot best response capacity is both 4 and 5 (if only integer values are allowed).

[16]Now $p = \$3$ is no longer a NE, since school 2 can increase profits by charging \$4 and serving 5 students rather than charging \$3 and enrolling 6 students. But, \$4 is not a NE either, since $\$4 - \epsilon$ will allow 6 students to enroll for a profit just below $\$4 \times 6 = 24$.

[17]In our example, suppose now that schools can also offer high quality, which doubles consumer valuation (A values low quality at \$10 but high quality at \$20). Now, when School 1 has a capacity of 2 and school 2 has a capacity of 6, in an equilibrium where school 2 chooses high quality, school 1 charges \$3 and caters to consumers G and H and school 2 charges \$9 and caters to consumers A through F.

Equilibrium Prices when Qualities are the Same

Following this basic intuition, when both schools' qualities are the same in the first stage, we are in the KS world, where the schools' optimal capacity choices will be equal to their Cournot quantity choices in the absence of credit constraint. However, if schools are credit constrained, then they will choose their capacities according to their capital up to their Cournot capacity.

In the Cournot version of our model, when schools' quantities are x_1 and x_2, the market price is $P(x_1 + x_2) = q(1 - x_1 + x_2)$. Therefore, the best response function for school with no capacity cost is

$$B(y) = \underset{0 \leq x \leq 1-y}{\arg\max}\{xTP(x + y)\}$$

which implies that

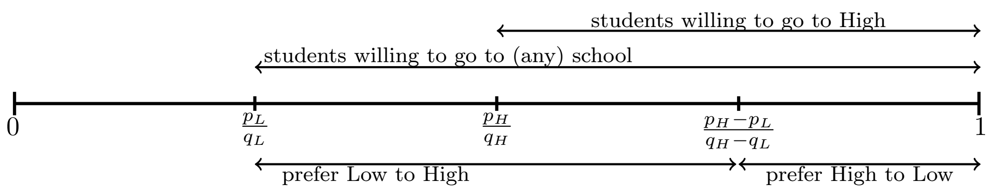
$$B(y) = \frac{1 - y}{2}.$$

According to Proposition 1 of KS, if $x_i \leq B(x_j)$ for $i, j = 1, 2$ and $i \neq j$, then a subgame equilibrium is for each school to name price $P(x_1 + x_2)$ with probability one. The equilibrium revenues are $x_i P(x_1 + x_2)$ for school i. However, if $x_i \geq x_j$ and $x_i > B(x_j)$, then the price equilibrium is randomized (price war) and school i's expected revenue is $R(x_j) = B(x_j)P(B(x_j) + x_j)$, i.e., school i cannot fully utilize its capacity, and school j's profit is somewhere between $[\frac{x_j}{x_i} R(x_j), R(x_j)]$.

Equilibrium Prices when Qualities are Different

Suppose that one school has quality q_H and the other school has quality q_L. Let x_H and x_L denote these schools' capacity choices and p_H and p_L be their prices, where $\frac{p_L}{q_L} \leq \frac{p_H}{q_H}$. The next figure summarizes students' preferences as a function of their taste parameter $\theta \in [0, 1]$.

Figure: Student's preferences over the space of taste parameter



Therefore, demand for the high quality school is $D_H = 1 - \frac{p_H - p_L}{q_H - q_L}$ and enrollment is $e_H = \min\left(x_H, 1 - \frac{p_H - p_L}{q_H - q_L}\right)$. Demand for the low quality school is

$$D_L = \begin{cases} \frac{p_H - p_L}{q_H - q_L} - \frac{p_L}{q_L}, & \text{if } x_H \geq 1 - \frac{p_H - p_L}{q_H - q_L} \\ 1 - \frac{p_L}{q_L} - x_H, & \text{otherwise,} \end{cases}$$

and enrollment of the low quality school is $e_L = \min\left(x_L, \max\left(\frac{p_H - p_L}{q_H - q_L} - \frac{p_L}{q_L},\ 1 - \frac{p_L}{q_L} - x_H\right)\right)$.

Best response prices: Next, we find the best response functions for the schools given their first stage choices, q_H, q_L, x_H and x_L. The high quality school's profit is $p_H e_H$

which takes its maximum value at $p_H = \frac{q_H - q_L + p_L}{2}$. Therefore, the best response price for the high quality school is $P_H(p_L) = \frac{q_H - q_L + p_L}{2}$ whenever the school's capacity does not fall short of the demand at these prices, i.e. $p_L \leq (q_H - q_L)(2x_H - 1)$. Otherwise, i.e. $p_L > (q_H - q_L)(2x_H - 1)$, we have $P_H(p_L) = p_L + (1 - x_H)(q_H - q_L)$. To sum,

$$P_H(p_L) = \begin{cases} \frac{q_H - q_L + p_L}{2}, & \text{if } p_L \leq (q_H - q_L)(2x_H - 1) \\ p_L + (1 - x_H)(q_H - q_L), & \text{otherwise.} \end{cases}$$

Now, given x_H, x_L and p_H, we find the best response price for the low quality school, p_L. We know that if $x_H \geq 1 - \frac{p_H - p_L}{q_H - q_L}$, then the enrollment is $e_L = \min\left(x_L, \frac{p_H - p_L}{q_H - q_L} - \frac{p_L}{q_L}\right)$. However, if $x_H < 1 - \frac{p_H - p_L}{q_H - q_L}$, then the enrollment is $e_L = \min\left(x_L, 1 - \frac{p_L}{q_L} - x_H\right)$. Therefore, the profit functions are as follows:

1) $x_H \geq 1 - \frac{p_H - p_L}{q_H - q_L}$

   (i) If $x_L < \frac{p_H - p_L}{q_H - q_L} - \frac{p_L}{q_L}$, then $e_L = x_L$, and so $\Pi_L = p_L x_L$.

   (ii) If $x_L \geq \frac{p_H - p_L}{q_H - q_L} - \frac{p_L}{q_L}$, then $e_L = \frac{p_H - p_L}{q_H - q_L} - \frac{p_L}{q_L}$, and so $\Pi_L = p_L\left(\frac{p_H - p_L}{q_H - q_L} - \frac{p_L}{q_L}\right)$.

2) $x_H < 1 - \frac{p_H - p_L}{q_H - q_L}$

   (i) If $x_L < 1 - \frac{p_L}{q_L} - x_H$, then $e_L = x_L$, and so $\Pi_L = p_L x_L$.

   (ii) If $x_L \geq 1 - \frac{p_L}{q_L} - x_H$, then $e_L = 1 - \frac{p_L}{q_L} - x_H$, and so $\Pi_L = p_L\left(1 - \frac{p_L}{q_L} - x_H\right)$.

Profit maximizing p_L's yield the following best response function:

$$P_L(p_H) = \begin{cases} \frac{p_H q_L}{2q_H}, & \text{if } x_H \geq 1 - \frac{p_H - p_L}{q_H - q_L} \text{ and } p_H \leq 2x_L(q_H - q_L) \\ \frac{p_H q_L - x_L q_L (q_H - q_L)}{q_H}, & \text{if } x_H \geq 1 - \frac{p_H - p_L}{q_H - q_L} \text{ and } p_H > 2x_L(q_H - q_L) \\ \frac{(1 - x_H) q_L}{2}, & \text{if } x_H < 1 - \frac{p_H - p_L}{q_H - q_L} \text{ and } x_H + 2x_L \geq 1 \\ q_L(1 - x_L - x_H), & \text{if } x_H < 1 - \frac{p_H - p_L}{q_H - q_L} \text{ and } x_H + 2x_L < 1 \end{cases}$$

Finding Optimal Prices: Solving the best response functions simultaneously implies working out the following eight cases:

Case 1: Consider the parameters satisfying

$$p_L \leq (q_H - q_L)(2x_H - 1) \tag{5}$$

so that the best response function for the high quality school is $P_H(p_L) = \frac{q_H - q_L + p_L}{2}$. We need to consider the following four sub-cases:

Case 1.1: Consider the parameters satisfying

$$x_H \geq 1 - \frac{p_H - p_L}{q_H - q_L} \tag{6}$$

$$p_H \leq 2x_L(q_H - q_L) \tag{7}$$

so that the best response function for the low quality school is $P_L(p_H) = \frac{p_H q_L}{2q_H}$. Solving the best response functions simultaneously yields

$$p_L = \frac{q_L(q_H - q_L)}{4q_H - q_L}$$
$$p_H = \frac{2q_H(q_H - q_L)}{4q_H - q_L}$$

Therefore, the inequalities (5) and (6) yield $x_H \geq \frac{2q_H}{4q_H - q_L}$ and equation (7) yields $x_L \geq \frac{q_H}{4q_H - q_L}$.

Case 1.2: Consider the parameters satisfying

$$x_H \geq 1 - \frac{p_H - p_L}{q_H - q_L} \tag{8}$$
$$p_H > 2x_L(q_H - q_L) \tag{9}$$

so that the best response function for the low quality school is $P_L(p_H) = \frac{p_H q_L - x_L q_L(q_H - q_L)}{q_H}$. Solving them simultaneously yields

$$p_L = \frac{q_L(q_H - q_L)(1 - 2x_L)}{2q_H - q_L}$$
$$p_H = \frac{(q_H - q_L)(q_H - q_L x_L)}{2q_H - q_L}$$

Therefore, the inequalities (5) and (8) yield $q_H \leq q_L x_L + (2q_H - q_L)x_H$ and equation (9) yields $x_L < \frac{q_H}{4q_H - q_L}$.

Case 1.3: Consider the parameters satisfying

$$x_H < 1 - \frac{p_H - p_L}{q_H - q_L} \tag{10}$$
$$1 \leq x_H + 2x_L \tag{11}$$

so that the best response function for the low quality school is $P_L(p_H) = \frac{(1-x_H)q_L}{2}$. Solving them simultaneously yields

$$p_L = \frac{(1 - x_H)q_L}{2}$$
$$p_H = \frac{q_H - q_L}{2} + \frac{q_L(1 - x_H)}{4}$$

The inequality (10) yields $x_H < \frac{2q_H - q_L}{4q_H - 3q_L}$ and the inequality (5) yields $x_H \geq \frac{2q_H - q_L}{4q_H - 3q_L}$, which cannot be satisfied simultaneously. Therefore, there cannot exist an equilibrium for the parameter values satisfying inequalities (5), (10) and (11).

Case 1.4: Consider the parameters satisfying

$$x_H < 1 - \frac{p_H - p_L}{q_H - q_L} \tag{12}$$
$$1 > x_H + 2x_L \tag{13}$$

so that the best response function for the low quality school is $P_L(p_H) = q_L(1-x_H-x_L)$. Solving them simultaneously yields

$$p_L = q_L(1 - x_H - x_L)$$
$$p_H = \frac{q_H - q_L(x_L + x_H)}{2}$$

The inequality (12) yields $x_H < \frac{q_H - q_L x_L}{2q_H - q_L}$ and the inequality (5) yields $x_H \geq \frac{q_H - q_L x_L}{2q_H - q_L}$, which cannot be satisfied simultaneously. Therefore, there cannot exist an equilibrium for the parameter values satisfying inequalities (12), (13) and (5).

Case 2: Now, consider the parameters satisfying

$$p_L > (q_H - q_L)(2x_H - 1) \tag{14}$$

so that the best response function for the high quality school is $P_H(p_L) = p_L + (1 - x_H)(q_H - q_L)$. We need to consider the following four sub-cases:

Case 2.1: Consider the parameters satisfying

$$x_H \geq 1 - \frac{p_H - p_L}{q_H - q_L} \tag{15}$$
$$p_H \leq 2x_L(q_H - q_L) \tag{16}$$

so that the best response function for the low quality school is $P_L(p_H) = \frac{p_H q_L}{2q_H}$. Solving the best response functions simultaneously yields

$$p_L = \frac{q_L(q_H - q_L)(1 - x_H)}{2q_H - q_L}$$
$$p_H = \frac{2q_H(q_H - q_L)(1 - x_H)}{2q_H - q_L}$$

Therefore, the inequalities (14), (15), and (16) yield $x_H < \frac{2q_H}{4q_H - q_L}$, $x_H \geq x_H$, and $q_H x_H + (2q_H - q_L)x_L \geq q_H$ respectively.

Case 2.2: Consider the parameters satisfying

$$x_H \geq 1 - \frac{p_H - p_L}{q_H - q_L} \tag{17}$$
$$p_H > 2x_L(q_H - q_L) \tag{18}$$

so that the best response function for the low quality school is $P_L(p_H) = \frac{p_H q_L - x_L q_L(q_H - q_L)}{q_H}$. Solving them simultaneously yields

$$p_L = q_L(1 - x_H - x_L)$$
$$p_H = (1 - x_H)q_H - x_L q_L$$

Therefore, the inequalities (14), (17), and (18) yield $q_H > x_L q_L + x_H(2q_H - q_L)$, $x_H \geq x_H$, and $q_H x_H + (2q_H - q_L)x_L < q_H$ respectively.

Case 2.3: Consider the parameters satisfying

$$x_H < 1 - \frac{p_H - p_L}{q_H - q_L} \tag{19}$$
$$1 \le x_H + 2x_L \tag{20}$$

so that the best response function for the low quality school is $P_L(p_H) = \frac{(1-x_H)q_L}{2}$. Solving them simultaneously yields

$$p_L = \frac{(1-x_H)q_L}{2}$$
$$p_H = (1-x_H)(q_H - \frac{q_L}{2})$$

The inequality (19) yields $x_H < x_H$ implying that there cannot exist an equilibrium for the parameter values satisfying inequalities (14), (19), and (20).

Case 2.4: Consider the parameters satisfying

$$x_H < 1 - \frac{p_H - p_L}{q_H - q_L} \tag{21}$$
$$1 > x_H + 2x_L \tag{22}$$

so that the best response function for the low quality school is $P_L(p_H) = q_L(1-x_H-x_L)$. Solving them simultaneously yields

$$p_L = q_L(1 - x_H - x_L)$$
$$p_H = (1-x_H)q_H - q_L x_L$$

The inequality (21) yields $x_H < x_H$ implying that there cannot exist an equilibrium for the parameter values satisfying inequalities (14), (21), and (22).

Summary of the Equilibrium: The equilibrium prices can be summarized in the following picture where

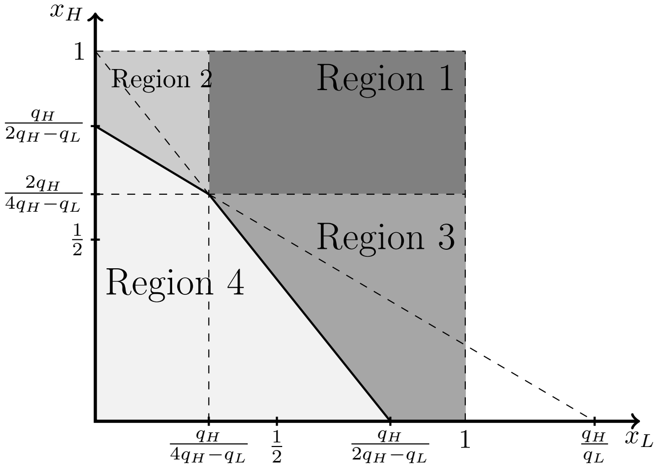
Region 1: Parameters satisfy $x_H \ge \frac{2q_H}{4q_H - q_L}$ and $x_L \ge \frac{q_H}{4q_H - q_L}$. Equilibrium prices are $p_L = \frac{q_L(q_H - q_L)}{4q_H - q_L}$ and $p_H = \frac{2q_H(q_H - q_L)}{4q_H - q_L}$. Therefore, enrollment and revenue (per student) of the high quality school are $e_H = \frac{2q_H}{4q_H - q_L}$ and $\Pi_H = \frac{4q_H^2(q_H - q_L)}{(4q_H - q_L)^2}$. Note that this is not the profit function of the high quality school, and so the cost of choosing capacity x_H and high quality are excluded.

Region 2: Parameters satisfy $x_L < \frac{q_H}{4q_H - q_L}$ and $q_L x_L + (2q_H - q_L)x_H \ge q_H$. Equilibrium prices are $p_L = \frac{q_L(q_H - q_L)(1-2x_L)}{2q_H - q_L}$ and $p_H = \frac{(q_H - q_L)(q_H - q_L x_L)}{2q_H - q_L}$. Therefore, enrollment and revenue (per student) of the high quality school are $e_H = \frac{q_H - q_L x_L}{2q_H - q_L}$ and $\Pi_H = (q_H - q_L)\frac{(q_H - q_L x_L)^2}{(2q_H - q_L)^2}$.

Region 3: Parameters satisfy $x_H < \frac{2q_H}{4q_H - q_L}$ and $q_H x_H + (2q_H - q_L)x_L \ge q_H$. Equilibrium prices are $p_L = \frac{q_L(q_H - q_L)(1-x_H)}{2q_H - q_L}$ and $p_H = \frac{2q_H(q_H - q_L)(1-x_H)}{2q_H - q_L}$. Therefore, enrollment and revenue of the high quality school are $e_H = x_H$ and

$\Pi_H = \frac{2q_H(q_H-q_L)(1-x_H)x_H}{2q_H-q_L}$. Moreover, the profit of the low quality school is $\Pi_L = p_L\left(\frac{p_H-p_L}{q_H-q_L} - \frac{p_L}{q_L}\right) = \frac{q_H q_L(q_H-q_L)(1-x_H)^2}{(2q_H-q_L)^2}$.

Region 4: Parameters satisfy $q_H x_H + (2q_H - q_L)x_L < q_H$ and $q_L x_L + (2q_H - q_L)x_H < q_H$. Equilibrium prices are $p_L = q_L(1 - x_H - x_L)$ and $p_H = (1 - x_H)q_H - x_L q_L$. Enrollment and revenue of the high quality school are $e_H = x_H$ and $\Pi_H = x_H[(1 - x_H)q_H - x_L q_L]$. Enrollment and revenue of the low quality school are $e_L = x_L$ and $\Pi_L = p_L x_L = q_L(1 - x_H - x_L)x_L$.



The First Stage Equilibrium: Quality and Capacity

Now we consider the first stage equilibrium strategies. In the baseline, we still assume that schools do not have enough capital to adopt high quality, and thus both schools are of low quality. Moreover, the schools' initial capacity is $x_1 = x_2 = \frac{M}{2}$. Therefore, the baseline market price is $P(M) = q_L(1 - M)$. We make the following two assumptions regarding the size of the covered market, M:

Assumption 1: $2 \leq TM$. That is, total private school enrollment is at least 2.

Assumption 2: $\frac{M}{2} \leq \frac{1}{3}\left(1 - \frac{r}{q_L}\right)$.

Assumption 3: $\frac{K}{Tr} + \frac{M}{2} \leq \frac{2q_H}{4q_H-q_L}$.

If the second assumption does not hold, then the treated school in the L arm would prefer not to increase its capacity. This assumption implies that schools do not have enough capital to pick their Cournot optimal capacities at baseline. If the third assumption does not hold, then the treated school can increase its capacity to the level where it can cover more than half of the market. We impose these three assumptions simply because parameters that do not satisfy them seem irrelevant for our sample. We also like to note the following observations that help us to pin down what the equilibrium prices will be when schools' quality choices are different.

Observation 1: $x_1 = x_2 = \frac{M}{2}$ satisfy the constraint $q_H x_1 + x_2(2q_H - q_L) < q_H$ if assumption 2 holds.

Observation 2: $\frac{2q_H}{4q_H - q_L} > \frac{1}{2}$, and so $\frac{M}{2} < \frac{2q_H}{4q_H - q_L}$.

Therefore, the schools would be in Region 4 with their baseline capacities. If school 1 receives a grant and invests in quality and capacity, then the schools either stay in Region 4, i.e. school 1 picks its quality such that x_H, x_L satisfies the constraints of Region 4, or move to Region 2. However, the next result shows that schools will always stay in Region 4, both in the H and L arms, if the schools' quality choices are different.

Lemma 1. *Both in the low and high saturation treatments, if schools' quality choices are different, then their equilibrium capacities x_L and x_H must be such that both $q_H x_H + x_L(2q_H - q_L) < q_H$ and $q_L x_L + x_H(2q_H - q_L) < q_H$ hold.*

Proof. Whether it is the low or high saturation treatment, suppose that school 1 receives the grant and invests in higher quality while school 2 remains in low quality. We know by assumption 3 that school 1's final capacity will never be above $2q_H/(4q_H - q_L)$. Therefore, schools' equilibrium capacities x_H and x_L will be in Region 4 or in Region 3. Next, we show that school 2 will never pick its capacity high enough to move Region 3 even if it can afford it.

School 2's profit, if it picks x such that $x + \frac{M}{2}$ and x_H remains in Region 4, is

$$\Pi_L = Tq_L(x + \frac{M}{2})(1 - x_H - \frac{M}{2} - x) + K - Trx.$$

The first order conditions imply that the optimal (additional) capacity is $\frac{1 - x_H - r/q_L}{2} - \frac{M}{2}$ or less if the grant is not large enough to cover this additional capacity. On the other hand, the capacity school 2 needs to move to Region 3, x_L, must satisfy $x_L \geq \frac{q_H(1-x_H)}{2q_H - q_L}$, which is strictly higher $x + \frac{M}{2}$. Therefore, given school 1's choice, school 2's optimal capacity will be such that schools remain in Region 4.

On the other hand, if school 2 could pick the capacity required to move into Region 3, the profit maximizing capacity would be $\frac{q_H(1-x_H)}{2q_H - q_L}$ because school 2's profit does not depend on its capacity beyond this level. Therefore, the profit under this capacity level would be

$$\Pi^3 = \frac{Tq_H(1 - x_H)}{2q_H - q_L}\left(\frac{q_L(q_H - q_L)(1 - x_H)}{2q_H - q_L} - r\right) - Tr\frac{M}{2}.$$

However, if school 2 picks x and remains in Region 4, then its profit would be

$$\Pi^4 = \frac{Tq_L}{2}\left(1 - x_H - \frac{r}{q_L}\right)^2 - Tr\frac{M}{2}.$$

The difference yields

$$\Pi^3 - \Pi^4 = -\frac{T(2q_H r + q_L^2(1 - x_H) - q_L r)^2}{4q_L(2q_H - q_L)^2} < 0$$

implying that school 2 prefers to choose a lower capacity and remain in Region 4 even if it can choose a higher capacity. □

Theorem 1. *If the treated school in the low saturation treatment invests in quality, then there must exist an equilibrium in the high saturation treatment where at least one school invests in quality. However, the converse is not always true.*

Proof. We prove our claim for $w = K$.

Low saturation treatment: If school 1 invests in quality its profit is

$$\Pi^H_{Low} = \frac{TM}{2}\left[\left(1-\frac{M}{2}\right)q_H - \frac{M}{2}q_L\right]$$

However, if school 1 invests in capacity, then its optimal capacity choice is $x^l = \frac{1}{2}\left(1-\frac{3M}{2}-\frac{r}{q_L}\right)$ and profit is

$$\Pi^L_{Low} = \begin{cases} K + T\left[\frac{(2-M)^2}{16}q_L - \frac{(2-3M)}{4}r + \frac{r^2}{4q_L}\right], & \text{if } x^l \le \min\left(\frac{K}{Tr}, B(\frac{M}{2})\right) \\ Tq_L\left(\frac{K}{Tr}+\frac{M}{2}\right)\left(1-M-\frac{K}{Tr}\right), & \text{if } \frac{K}{Tr} < x^l \le B(\frac{M}{2}) \\ Tq_L\left(B(\frac{M}{2})+\frac{M}{2}\right)\left(1-M-B(\frac{M}{2})\right) + K - TrB(\frac{M}{2}), & \text{if } B(\frac{M}{2}) < \min\left(x^l, \frac{K}{Tr}\right) \end{cases}$$

High saturation treatment with (H,L) Equilibrium: We are trying to create an equilibrium where at least one school invests in high quality. In an equilibrium where only one school invests in quality, the low quality school's optimal capacity choice is $x^l = \frac{1}{2}\left(1-\frac{3M}{2}-\frac{r}{q_L}\right)$ and profit is

$$\Pi^L_{(H,L)} = \begin{cases} K + T\left[\frac{(2-M)^2}{16}q_L - \frac{(2-3M)}{4}r + \frac{r^2}{4q_L}\right], & \text{if } x^l \le \min\left(\frac{K}{Tr}, B(\frac{M}{2})\right) \\ Tq_L\left(\frac{K}{Tr}+\frac{M}{2}\right)\left(1-\frac{K}{Tr}-M\right), & \text{if } \frac{K}{Tr} < x^l \le B(\frac{M}{2}) \\ Tq_L\left(B(\frac{M}{2})+\frac{M}{2}\right)\left(1-B(\frac{M}{2})-M\right) + K - TrB(\frac{M}{2}), & \text{if } B(\frac{M}{2}) < \min\left(x^l, \frac{K}{Tr}\right) \end{cases}$$

On the other hand, the high quality school's equilibrium profit is

$$\Pi^H_{(H,L)} = \frac{TM}{2}\left[\left(1-\frac{M}{2}\right)q_H - x_L q_L\right]$$

where

$$x_L = \begin{cases} \frac{M}{2} + x^l, & \text{if } x^l \le \min\left(\frac{K}{Tr}, B(\frac{M}{2})\right) \\ \frac{M}{2} + \frac{K}{Tr}, & \text{if } \frac{K}{Tr} < x^l \le B(\frac{M}{2}) \\ \frac{M}{2} + B(\frac{M}{2}), & \text{if } B(\frac{M}{2}) < \min\left(x^l, \frac{K}{Tr}\right) \end{cases}$$

Deviation payoffs from (H,L)*:* If the low type deviates to high quality, then we are back in KS world, and thus its (highest) deviation payoff will be

$$\widehat{\Pi}^L_{(H,L)} = \frac{TM}{2}(1-M)q_H.$$

However, if the high quality school deviates to low quality, then we are again in KS world. Thus, given that the other school's capacity is x_L, deviating school's optimal capacity is $\widehat{x} = \frac{1}{2}\left(1-M-x_L-\frac{r}{q_L}\right)$ and optimal profit is

$$\widehat{\Pi}^H_{(H,L)} = \begin{cases} K + T\left[\frac{(1-x_L)^2}{4}q_L - \frac{(1-x_L-M)}{2}r + \frac{r^2}{4q_L}\right], & \text{if } \widehat{x} \le \min\left(\frac{K}{Tr}, B(x_L)\right) \\ Tq_L\left(\frac{K}{Tr}+\frac{M}{2}\right)\left(1-\frac{M}{2}-x_L-\frac{K}{Tr}\right), & \text{if } \frac{K}{Tr} < \widehat{x} \le B(x_L) \\ Tq_L\left(B(x_L)+\frac{M}{2}\right)\left(1-\frac{M}{2}-x_L-B(x_L)\right) + K - TrB(x_L), & \text{if } B(x_L) < \min\left(\widehat{x}, \frac{K}{Tr}\right) \end{cases}$$

High saturation treatment with (H,H) Equilibrium: Because $w = K$, schools cannot increase their capacities. Moreover, we are in KS world, and so the equilibrium payoff is

$$\Pi_{(H,H)} = \frac{TM}{2}(1-M)q_H.$$

Deviation payoffs from (H,H): If a school deviates then the payoff is identical with the equilibrium of (H,L). Therefore, the deviating school's optimal capacity is $x^l = \frac{1}{2}\left(1 - \frac{3M}{2} - \frac{r}{q_L}\right)$ and profit is

$$\widehat{\Pi}_{(H,H)} = \begin{cases} K + T\left[\frac{(2-M)^2}{16}q_L - \frac{(2-3M)}{4}r + \frac{r^2}{4q_L}\right], & \text{if } x^l \leq \min\left(\frac{K}{Tr}, B(\frac{M}{2})\right) \\ Tq_L\left(\frac{K}{Tr} + \frac{M}{2}\right)\left(1 - \frac{K}{Tr} - M\right), & \text{if } \frac{K}{Tr} < x^l \leq B(\frac{M}{2}) \\ Tq_L\left(B(\frac{M}{2}) + \frac{M}{2}\right)\left(1 - B(\frac{M}{2}) - M\right) + K - TrB(\frac{M}{2}), & \text{if } B(\frac{M}{2}) < \min\left(x^l, \frac{K}{Tr}\right) \end{cases}$$

Note the following:

Claim 1. *If $x^l < \min\left(\frac{K}{Tr}, B(\frac{M}{2})\right)$, then $\widehat{x} < \min\left(\frac{K}{Tr}, B(x_L)\right)$.*

Proof. Assume that x^l satisfies the above inequality. Then $x_L = \frac{M}{2} + x^l$, $B(x_L) = B(\frac{M}{2}) - \frac{x^l}{2}$, and $\widehat{x} = \frac{x^l}{2}$, which is less than $\frac{K}{Tr}$. Moreover, $\widehat{x} < B(x_L)$ because $x^l < B(\frac{M}{2})$, and thus the desired result. □

Claim 2. *If $\frac{K}{Tr} < x^l \leq B(\frac{M}{2})$, then either $\widehat{x} < \min\left(\frac{K}{Tr}, B(x_L)\right)$ or $\frac{K}{Tr} < \widehat{x} \leq B(x_L)$.*

Proof. In this case $x_L = \frac{M}{2} + \frac{K}{Tr}$, $B(x_L) = B(\frac{M}{2}) - \frac{K}{2Tr}$, and $\widehat{x} = x^l - \frac{K}{2Tr}$. Therefore, we have $\widehat{x} \leq B(x_L)$ because $x^l < B(\frac{M}{2})$. However, $\widehat{x}$ may be greater or less than $\frac{K}{Tr}$, hence the desired result. □

Claim 3. *If $B(\frac{M}{2}) < \min\left(\frac{K}{Tr}, x^l\right)$, then $B(x_L) < \min\left(\frac{K}{Tr}, \widehat{x}\right)$.*

Proof. In this case $x_L = \frac{M}{2} + B(\frac{M}{2})$, $B(x_L) = \frac{1}{2}B(\frac{M}{2})$, and $\widehat{x} = x^l - \frac{1}{2}B(\frac{M}{2})$, Therefore, we have $\widehat{x} > B(x_L)$ and $B(x_L) < B(\frac{M}{2}) < \frac{K}{Tr}$, and thus the desired result. □

Lemma 1. *Suppose that $x^l \leq \min\left(\frac{K}{Tr}, B(\frac{M}{2})\right)$ and $\widehat{x} \leq \min\left(\frac{K}{Tr}, B(x_L)\right)$. If the treated school in the low saturation treatment invests in quality, then there is an equilibrium in the high saturation treatment such that at least one school invests in quality.*

Proof. For the given parameter values we know that the optimal capacity of the low quality school in low saturation treatment is x^l, and thus $x_L = \frac{M}{2} + x^l$ and $\widehat{x} = \frac{x^l}{2}$. Assume that the treated school in the low saturation treatment invests in quality. Then we must have

$$\Pi^H_{Low} \geq \Pi^L_{Low}$$

or equivalently, $\frac{TM}{2}\left[\left(1 - \frac{M}{2}\right)q_H - \frac{M}{2}q_L\right] \geq K + T\left[\frac{(2-M)^2}{16}q_L - \frac{(2-3M)}{4}r + \frac{r^2}{4q_L}\right]$. We need to show that either (H,L) or (H,H) is an equilibrium outcome. Equivalently, we need to prove that either the inequalities in (1) or (2) below hold:

(1) Both the low and high quality schools do not deviate from (H,L), i.e.,

$$\Pi^L_{(H,L)} \geq \widehat{\Pi}^L_{(H,L)} \quad \text{and} \quad \Pi^H_{(H,L)} \geq \widehat{\Pi}^H_{(H,L)}.$$

Equivalently, $K + T\left[\frac{(2-M)^2}{16}q_L - \frac{(2-3M)}{4}r + \frac{r^2}{4q_L}\right] \geq \frac{TM}{2}(1-M)q_H$ and $\frac{TM}{2}\left[\left(1 - \frac{M}{2}\right)q_H - x_L q_L\right] \geq K + T\left[\frac{(1-x_L)^2}{4}q_L - \frac{(1-x_L-M)}{2}r + \frac{r^2}{4q_L}\right]$ hold.

(2) Alternatively, the schools do not deviate from (H,H), that is

$$\Pi_{(H,H)} \geq \widehat{\Pi}_{(H,H)}$$

or equivalently, $\frac{TM}{2}(1-M)q_H \geq K + T\left[\frac{(2-M)^2}{16}q_L - \frac{(2-3M)}{4}r + \frac{r^2}{4q_L}\right]$.

Note that if $\Pi^L_{(H,L)} < \widehat{\Pi}^L_{(H,L)}$, then the inequality in (2) holds, and so we have an equilibrium where both schools pick high quality. Inversely, if the inequality in (2) does not hold, then $\Pi^L_{(H,L)} \geq \widehat{\Pi}^L_{(H,L)}$, i.e., the low quality school does not deviate from (H,L). If we show that the high quality school also doesn't deviate from (H,L), then we complete our proof. Because $\Pi^H_{Low} \geq \Pi^L_{Low}$, showing $\Pi^H_{(H,L)} - \Pi^H_{Low} \geq \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low}$ would prove that the second inequality in (1) holds as well. Therefore, we will prove that $\Pi^H_{Low} - \Pi^H_{(H,L)} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} = \frac{TMq_L}{2}x^l + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} \leq 0$.

$$\begin{aligned}
\frac{TMq_L}{2}x^l + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} &= \frac{Tq_L}{4}x^l\left[\frac{r}{q_L} - 2 + 3M + x^l\right] \\
&= \frac{Tq_L}{4}x^l\left[\frac{r}{2q_L} - \frac{3}{2} + \frac{3M}{2}\right] \text{ since } x^l = \frac{1}{2}\left(1 - \frac{3M}{2} - \frac{r}{q_L}\right) \\
&\leq \frac{Tq_L}{4}x^l\left[-\frac{r}{2q_L} - \frac{1}{2}\right] \text{ since } \frac{3M}{2} \leq 1 - \frac{r}{q_L} \text{ by Assumption 2} \\
&< 0.
\end{aligned}$$

Thus, either (H,L) or (H,H) is an equilibrium outcome. □

Lemma 2. *Suppose that* $\frac{K}{Tr} < x^l \leq B(\frac{M}{2})$ *and* $\widehat{x} \leq \min\left(\frac{K}{Tr}, B(x_L)\right)$. *If the treated school in low saturation treatment invests in quality, then there is an equilibrium in the high saturation treatment such that at least one school invests in quality.*

Proof. For the given parameter values we know that the optimal capacity of the low quality school is x^l is greater than $\frac{K}{TR}$, and thus $x_L = \frac{M}{2} + \frac{K}{Tr}$. Moreover, because $\widehat{x} < \min\left(\frac{K}{Tr}, B(x_L)\right)$ holds, we have $x^l < \frac{3K}{2Tr}$. Assume that the treated school in the low saturation treatment invests in quality. Then we must have

$$\Pi^H_{Low} \geq \Pi^L_{Low}$$

or equivalently, $\frac{TM}{2}\left[\left(1-\frac{M}{2}\right)q_H - \frac{M}{2}q_L\right] \geq Tq_L\left(\frac{K}{Tr} + \frac{M}{2}\right)\left(1 - M - \frac{K}{Tr}\right)$. Then we need to show that either (H,L) or (H,H) is an equilibrium. Equivalently, we need to show that either the inequalities in (1) or (2) below hold:

(1) Both the low and high quality schools do not deviate from (H,L), i.e.,

$$\Pi^L_{(H,L)} \geq \widehat{\Pi}^L_{(H,L)} \quad \text{and} \quad \Pi^H_{(H,L)} \geq \widehat{\Pi}^H_{(H,L)}.$$

Equivalently, $Tq_L\left(\frac{K}{Tr} + \frac{M}{2}\right)\left(1 - M - \frac{K}{Tr}\right) \geq \frac{TM}{2}(1-M)q_H$ and $\frac{TM}{2}\left[\left(1-\frac{M}{2}\right)q_H - x_Lq_L\right] \geq K + T\left[\frac{(1-x_L)^2}{4}q_L - \frac{(1-x_L-M)}{2}r + \frac{r^2}{4q_L}\right]$ hold.

(2) Alternatively, the schools do not deviate from (H,H), that is

$$\Pi_{(H,H)} \geq \widehat{\Pi}_{(H,H)}$$

or equivalently, $\frac{TM}{2}(1-M)q_H \geq Tq_L\left(\frac{K}{Tr} + \frac{M}{2}\right)\left(1 - M - \frac{K}{Tr}\right)$.

Note that if $\Pi^L_{(H,L)} < \widehat{\Pi}^L_{(H,L)}$, then the inequality in (2) holds, and so we have an equilibrium where both schools pick high quality. Inversely, if the inequality in (2) does not hold, then $\Pi^L_{(H,L)} \geq \widehat{\Pi}^L_{(H,L)}$, i.e., the low quality school does not deviate from (H,L). If we show that the high quality school also doesn't deviate from (H,L), then we complete our proof. Because $\Pi^H_{Low} \geq \Pi^L_{Low}$, showing $\Pi^H_{(H,L)} - \Pi^H_{Low} \geq \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low}$ would prove that the second inequality in (1) holds as well. Therefore, we will prove that $\Pi^H_{Low} - \Pi^H_{(H,L)} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} = \frac{KMq_L}{2r} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} \leq 0$.

$$\begin{aligned}
\frac{KMq_L}{2r} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} &= \underbrace{\frac{T}{16q_L}(2r-(2-3M)q_L)^2}_{=Tq_L(x^l)^2} + \underbrace{\frac{3K}{4r}(2r-(2-3M)q_L)}_{-\frac{3Kq_Lx^l}{r}} + \frac{5K^2q_L}{4r^2T} \\
&= \frac{Kq_L}{r}\left(\frac{Tr}{K}(x^l)^2 - 3x^l + \frac{5K}{4Tr}\right) \\
&\leq \frac{Kq_L}{r}\left(\frac{Tr}{K}(x^l)^2 - 3x^l + \frac{5}{4}x^l\right) \text{ since } \frac{K}{Tr} < x^l \\
&= \frac{Kq_L}{r}\left(\frac{Tr}{K}(x^l)^2 - \frac{7}{4}x^l\right) \\
&\leq \frac{Kq_L}{r}\left(\frac{3}{2x^l}(x^l)^2 - \frac{7}{4}x^l\right) \text{ since } x^l < \frac{3K}{2Tr} \\
&< 0.
\end{aligned}$$

Thus, either (H,L) or (H,H) is an equilibrium outcome. □

Lemma 3. *Suppose that $\frac{K}{Tr} < x^l \leq B(\frac{M}{2})$ and $\frac{K}{Tr} < \widehat{x} \leq B(x_L)$. If the treated school in the low saturation treatment invests in quality, then there is an equilibrium in the high saturation treatment such that at least one school invests in quality.*

Proof. Assume that the treated school in the low saturation treatment invests in quality. Then we must have

$$\Pi^H_{Low} \geq \Pi^L_{Low}$$

or equivalently, $\frac{TM}{2}\left[\left(1-\frac{M}{2}\right)q_H - \frac{M}{2}q_L\right] \geq Tq_L\left(\frac{K}{Tr}+\frac{M}{2}\right)\left(1-M-\frac{K}{Tr}\right)$. Then we need to show that either (H,L) or (H,H) is an equilibrium. Equivalently, we need to show that either the inequalities in (1) or (2) below hold:

(1) Both the low and high quality schools do not deviate from (H,L), i.e.,

$$\Pi^L_{(H,L)} \geq \widehat{\Pi}^L_{(H,L)} \quad \text{and} \quad \Pi^H_{(H,L)} \geq \widehat{\Pi}^H_{(H,L)}.$$

Equivalently, $Tq_L\left(\frac{K}{Tr}+\frac{M}{2}\right)\left(1-M-\frac{K}{Tr}\right) \geq \frac{TM}{2}(1-M)q_H$ and $\frac{TM}{2}\left[\left(1-\frac{M}{2}\right)q_H - x_Lq_L\right] \geq Tq_L\left(\frac{K}{Tr}+\frac{M}{2}\right)\left(1-M-x_L-\frac{K}{Tr}\right)$ hold.

(2) Alternatively, the schools do not deviate from (H,H), that is

$$\Pi_{(H,H)} \geq \widehat{\Pi}_{(H,H)}$$

or equivalently, $\frac{TM}{2}(1-M)q_H \geq Tq_L\left(\frac{K}{Tr}+\frac{M}{2}\right)\left(1-M-\frac{K}{Tr}\right)$.

Same as before if we show that the high quality school doesn't deviate from (H,L), i.e., $\Pi^H_{Low} - \Pi^H_{(H,L)} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} = \frac{KMq_L}{2r} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} \leq 0$, then we complete our proof.

$$\begin{aligned}\frac{KMq_L}{2r} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} &= \frac{KMq_L}{2r} + Tq_L\left(\frac{K}{Tr} + \frac{M}{2}\right)\left(-\frac{K}{Tr}\right)\\ &= \frac{Kq_L}{r}\left(\frac{M}{2} - \frac{K}{Tr} - \frac{M}{2}\right)\\ &< 0.\end{aligned}$$

Thus, either (H,L) or (H,H) is an equilibrium outcome. □

Lemma 4. *Suppose that $B(\frac{M}{2}) < \min\left\{\frac{K}{Tr}, x^l\right\}$ and $B(x_L) < \min\left\{\frac{K}{Tr}, \widehat{x}\right\}$. If the treated school in the low saturation treatment invests in quality, then there is an equilibrium in the high saturation treatment such that at least one school invests in quality.*

Proof. For the given parameter values $B(\frac{M}{2}) = \frac{1}{2} - \frac{M}{4}$, $x_L = \frac{M}{2} + B(\frac{M}{2})$, and $B(x_L) = \frac{1}{2}B(\frac{M}{2})$. Assume that the treated school in the low saturation treatment invests in quality. Then we must have $\Pi^H_{Low} \geq \Pi^L_{Low}$ or equivalently, $\frac{TM}{2}\left[\left(1 - \frac{M}{2}\right)q_H - \frac{M}{2}q_L\right] \geq Tq_L\left(B(\frac{M}{2}) + \frac{M}{2}\right)\left(1 - M - B(\frac{M}{2})\right) + K - TrB(\frac{M}{2})$. Then we need to show that either (H,L) or (H,H) is an equilibrium. Equivalently, we need to show that either the inequalities in (1) or (2) below hold:

(1) Both the low and high quality schools do not deviate from (H,L), i.e., $\Pi^L_{(H,L)} \geq \widehat{\Pi}^L_{(H,L)}$ and $\Pi^H_{(H,L)} \geq \widehat{\Pi}^H_{(H,L)}$. Equivalently, $Tq_L\left(B(\frac{M}{2}) + \frac{M}{2}\right)\left(1 - M - B(\frac{M}{2})\right) + K - TrB(\frac{M}{2}) \geq \frac{TM}{2}(1-M)q_H$ and $\frac{TM}{2}\left[\left(1 - \frac{M}{2}\right)q_H - x_Lq_L\right] \geq Tq_L\left(B(x_L) + \frac{M}{2}\right)(1 - M - x_LB(x_L)) + K - TrB(x_L)$ hold.

(2) Alternatively, the schools do not deviate from (H,H), that is $\Pi_{(H,H)} \geq \widehat{\Pi}_{(H,H)}$ or equivalently, $\frac{TM}{2}(1-M)q_H \geq Tq_L\left(B(\frac{M}{2}) + \frac{M}{2}\right)\left(1 - M - B(\frac{M}{2})\right) + K - TrB(\frac{M}{2})$.

Same as before if we show that the high quality school doesn't deviate from (H,L), i.e., $\Pi^H_{Low} - \Pi^H_{(H,L)} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} = \frac{TMq_L}{2}B(\frac{M}{2}) + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} \leq 0$, then we complete our proof.

$$\begin{aligned}\frac{TMq_L}{2}B(\frac{M}{2}) + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} &= \frac{TMq_LB(\frac{M}{2})}{2} + \frac{TrB(\frac{M}{2})}{2} + \frac{Tq_LB(\frac{M}{2})}{2}\left[\frac{M}{2} + \frac{B(\frac{M}{2})}{2} - 1\right]\\ &= \frac{TB(\frac{M}{2})}{2}\left[r + q_L\left(\frac{11M}{8} - \frac{3}{4}\right)\right]\\ &< 0 \text{ since } \frac{M}{2} < \frac{1}{3}\left(1 - \frac{r}{q_L}\right) \text{ by Assumption 2.}\end{aligned}$$

Thus, either (H,L) or (H,H) is an equilibrium outcome. □

Finally, the converse of the claim is not necessarily true because $\Pi^H_{Low} - \Pi^H_{(H,L)} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low}$ is strictly negative. That is, there are many parameters in which at least one school invests in quality in the high saturation treatment, but the treated school invests only in capacity in the low saturation treatment. □

C.3 Potential Extensions

We build a model with two ex-ante identical schools, each representing an "average school" in our sample. The main focus of our exercise is to answer a question that has a comparative static flavor: given that an average school in L treatment (or simply

L school) invests in quality, would the same (average) school invest in quality if it were in H treatment. Theorem 0 provides an affirmative answer to this question by characterizing the set of equilibrium in both treatments, whereas Theorem 1 skips this step (as equilibrium characterization under downward sloping demand curve is much more involved) and directly proves our main result.

Our simple workhorse model ignores the cases where schools are (initially) heterogeneous in different dimensions. The assumptions that schools are initially symmetric or not horizontally differentiated, and that schools cannot horizontally differentiate themselves are strong ones. In this section, we elaborate why these missing ingredients do not necessarily change the main message of our model, and discuss that the story/prediction of the model is in fact far more robust than it seems. At the end of this section, we also provide three numerical examples that clarify some of our discussions below.

Initial School Heterogeneity (Capacity)

The model can easily be extended to the case where one school is smaller than the other in terms of initial capacity. In this case, the small school (only when it is sufficiently small) would be more willing to start a price war and steal students from the bigger school (on the equilibrium path). The reason for this is simple: Reduced fees for all students means a very big loss for the bigger school, and so the bigger school becomes more accommodating (not so aggressive) in a price war. Thus, the small school uses this in its advantage and steals some students from its rival without making a huge discount on its fee.

Therefore, the equilibrium outcomes for a given set of parameters (as characterized in Theorem 0) would certainly be different when one school is very big, and the other is very small. However, our comparison between two treatments will still survive. That is, if L school invests in quality, then the same school would have incentive to invest in quality if it were in H treatment. This is true because if L school invests in quality, then the size of the uncovered market that the school is facing should be so small, or starting a price war and stealing some students should not be as profitable as investing in quality and raising fees. However, these (competitive market) forces will work exactly in the same direction (potentially stronger) as before when both schools receive the grant. Thus, L school would also invest in quality if it were one of the schools in H arm.

School Closures

Although our model does not allow school closures or entry, it can easily be modified to accommodate these features. As we suggest by the previous arguments and by the numerical example at the end of this section, small schools would be more eager to start a price war and steal students from the bigger school in L treatment when the small school is sufficiently small. However, the same school will not be successful in increasing its enrollment (or stealing students from the big school) in the H treatment because the big school can push back the competition by investing in higher quality. Therefore, if low enrollment is the major reason for school closures, then in the L treatment, smaller schools are more likely to increase their enrollment, and thus less likely to shut down.

Initial School Heterogeneity (Horizontal Differentiation)

The model can also be extended to the case where initially schools are horizontally differentiated, and so each has some monopoly power. This modification would potentially lead to two uncovered markets; one for each school. One may think each market as consisting of students who live close to one school and those who are closer to the rival school. Therefore, each school would be more advantaged than its rival in its own uncovered market. Horizontal differentiation would make student stealing and price competition a very costly strategy. Therefore, in comparison to the case with no horizontal differentiation, schools would have more incentive to escape from price war, and so more incentive to invest in quality.

Nevertheless, this modification would not change our main message: If L school invests in quality, which means that the size of the schools' own uncovered market was so small or stealing new students from the other school was less profitable than investing in quality and raising fees. These forces will work exactly in the same direction and potentially stronger when both schools receive the grant and have the opportunity to make investments. Thus, L school should also invest in quality if it was one of the schools in H treatment.

Initial School Heterogeneity and Ability to Horizontally Differentiate

The model and the results can be extended to the case where schools have different initial sizes and are horizontally differentiated, and furthermore, can invest in (i) capacity (ii) quality, and (iii) horizontal differentiation. Schools' ability to further differentiate themselves horizontally would increase their incentives to increase their capacities (as differentiation may increase their monopoly power and demands further), reduce their incentives to start a price war (stealing from other schools gets more difficult), and so may decrease their incentives to invest in quality.

Nevertheless, our comparison between the two treatment arms should survive: If L invests in quality, then it means that horizontal differentiation and creating extra demand, and so increasing capacity is not as profitable as raising fees and per student profitability (potentially because the school is already in its limits to market saturation or quality investment is less costly, or rate of return of quality investment is higher). But then the same school would invest in quality if it was one of the schools in H treatment because now competition makes profitability in enrollment arm less likely as the other school will have incentive to differentiate itself horizontally or vertically (so stealing students gets even more difficult).

Variable vs Fixed Costs of Investments

The model assumes that cost of quality investment is fixed cost. What if it was (partially or fully) variable cost? The model would be extended to cover these cases and it does not necessarily change our main message. That is, if L school invests in quality, then the same school would invest in quality if it were in H treatment. Please see the numerical example at the end for an alternative explanation.

In our current setup, if cost of quality investment was fully variable cost, such as higher teachers' salary, but capacity is fully fixed cost, then credit constrained schools should choose their quality levels optimally before receiving the cash grants, and so our treatments should have no impact on their behavior on quality dimension in L and H arms.

Therefore, if quality investment is a fully variable cost, then we need to make a further assumption to answer why schools do not invest in quality before intervention but do so after receiving the cash grant. One potential explanation would be that the schools are insurance (in addition to or instead of credit) constrained. Then consider the following modification to the model: After making their investment decisions (quality or capacity) schools receive a demand shock, which determines whether the demand will be low or high. In case of a high demand, both investments yield positive profit. In case of a low demand, schools make negative profits and shut down if they make investments that increase their variable cost (such as quality investment) and do not have enough cash to cover their losses. In this environment, schools would be choosing their ex-ante (i.e., before treatment) quality levels sub-optimally simply because they cannot raise enough cash to buffer their ex-post (i.e., after demand realization) losses.

If the L school invests in quality in this environment, then this means that its (expected) return is higher than making capacity increase (potentially because uncovered market was so small or risks of negative profits were not so high, or returns from quality investment were so high). However, these forces will work exactly in the same direction when both schools receive the grant and have the opportunity to make investments. Thus, the same school in L treatment would also invest in quality if it were one of the schools in H treatment.

Finally, if quality investment were to require components that are partially variable and partially fixed costs, then this modification would not change the players' incentives so much from our current model, where schools cannot invest on quality because they are credit constrained and cash grant alleviates these constraints. Therefore, introducing quality investment as a variable cost does not necessarily change the main message of our theorems.

C.4 Numerical Examples

Initial School Heterogeneity in capacity

In this example, we show that when schools are heterogeneous in initial capacities and one school is sufficiently smaller than the other, then in equilibrium the smaller school would invest in capacity even though it triggers price war. Moreover, the predictions of our simple model still hold in this case, that is if (smaller) school prefers to invest in high quality in L arm, then it will prefer to invest in high quality in H arm. Put differently, (smaller) schools may prefer to invest in capacity in L treatment but in quality in H treatment. Suppose for now that

1. School 1's initial capacity is 20.

2. School 2's initial capacity is 100.

3. Size of the uncovered market, $N = 0$.

4. Size of cash grant, $K = 8$

5. Cost per unit of extra capacity, $r = 1$.

6. Willingness to pay for low quality, $q_L = 2$.

7. Willingness to pay for high quality is a free variable, q_H, and

8. Cost of high quality is also a free variable w.

L treatment: Suppose that school 1 receives the cash grant. In equilibrium, if school 1 invests only in capacity, it is optimal for this school to invest all its money to capacity and buy 8 new chairs. However, because the size of the uncovered market is zero, i.e., $N = 0$, this means price war: School 1 and 2 pick their prices randomly over the interval $[1.84, 2]$ and school 1's profit is $\frac{1288}{25} = 51.52$.[18]

However, if school 1 invests in quality and pay w for high quality, then in equilibrium its optimal decision would be to invest the remaining $K - w$ for extra seats, which again means price competition. In this case, school 1's profit would be

$$(20 + x)p(x) \tag{23}$$

where $x = K - w$ is the number of additional chairs school 1 can afford after paying the cost of high quality and $p()$ is the lower boundary for the price schools pick in the price war. Note that $p(.)$ inversely depends on x and it is always less than q_H.

Therefore, in L treatment schools 1's optimal decision (whether to invest in quality or capacity) is determined whether the profit in (23) is bigger or less than 51.52, namely

$$(20 + x)p(x) > (<)\ 51.52 \tag{24}$$

H treatment: In this case, regardless of what school 2 does, school 1's profit if it invests only in capacity can be at most 51.52. On the other hand, if school 1 invests in quality, then its profit will be

$$(20 + x^H)p^H + 8 - x^H - w \tag{25}$$

where p^H is a decreasing function of x^H.

When $w = K$, that is schools can't invest in extra chair if they invest in quality, then $x = x^H = 0$, and thus both profits in (23) and (25) are the same. Thus, our theorem holds. Namely, if school 1 invests in quality in L arm then it invests in quality in H arm.

When $w < K$, for example consider the case where $w = 5$, equilibrium calculations (which we skip here) show that in the subgame where school 2 does not invest in quality we have $x^H = x$ and $p^H = p$ for all values of $q_H > q_L$. Therefore, the profit of school 1 in (23) and (25) are the same. On the other hand, in a subgame where school 2 invests in quality, it is easy to prove that school 1 also prefers to invest in quality. Thus, our theorem also holds when $w < K$.

Initial School Heterogeneity Horizontal Differentiation

In this example, we show that our results would still hold even when schools are initially heterogeneous in the sense that they are horizontally differentiated. More specifically, we provide an example where schools choose capacity investment in L treatment, whereas one of the schools prefers to invest in quality in H treatment to avoid price competition. For simplicity, we model initial heterogeneity as each school having different (local) uncovered markets. Suppose for now that

1. Each school's initial capacity is 10.
2. Size of the uncovered market for School 1, $N_1 = 2$.

[18] If school 1 buys 0 chairs and simply keeps the money with it, then its profit would be 48, which is less than 51.52.

3. Size of the uncovered market for School 2, $N_2 = 4$.

4. Size of cash grant, $K = 8$

5. Cost per unit of extra capacity, $r = 1$.

6. Cost of high quality $w = 8$.

7. Willingness to pay for low quality, $q_L = 10$.

8. Willingness to pay for high quality, $q_H = 12.9$.

9. If a student goes to a school not in his/her district valuations reduce by 2.[19]

L treatment:

1. School 1 gets the grant: In equilibrium, school 1 invests in capacity. This is true because the profit in capacity investment (130) is higher than profit in quality investment (129).

   (i) School 1 invests in capacity: In equilibrium, school 1 buys 6 extra chairs and keep the remaining money for profit of 130.[20]

   (ii) School 1 invests in quality: In equilibrium school 1's profit is 129.

2. School 2 gets the grant: In equilibrium, school 2 invests in capacity. This is true because the profit in capacity investment (144) is higher than profit in quality investment (129).

   (i) School 2 invests in capacity: In equilibrium, school 2 buys 4 more chairs and keeps the price at 10, the total profit is $14 \times 10 + 4 = 144$.

   (ii) School 2 invests in quality: In equilibrium school 2's profit is 129.

H treatment: The only equilibrium is that school 1 invests in quality (with profit 129) and school 2 buys 4 chairs with profit of 144. This is true because

- If school 1 deviates and invests in capacity, the optimal action is to buy only 2 chairs for profit of $12 \times 10 + 6 = 126$. Investing more than 2 chairs means price competition with school 2, yielding a profit at most $12 \times 8 + 2 = 98$.
- If school 2 deviates and invests in quality, its profit will be 129, which is less than 144. Thus no school has incentive to deviate.
- Moreover, there is no equilibrium where both schools invest in capacity in H treatment: This is true because

[19] Therefore, schools can attract all uncovered students, but have to reduce down their prices to attract students of other district.

[20] If school 1 invests 6 chairs (total capacity of school 1 is then 16), then no need for price competition, but price must reduce to 8 to attract 4 more students from district 2 (schools can't price discriminate students), and so profit is $16 \times 8 + 2 = 130$. However, if school 1 invests only 2 chairs (total capacity of school 1 is then 12), then no need for price competition and no need to reduce the price, and so profit is $12 \times 10 + 6 = 126$. Finally, if school 1 invests more than 6 chairs (total capacity of school 1 is, for example, 18), then equilibrium fee is in mixed strategies but the expected profit is no more than 130

- If that was an equilibrium, then School 1 buys 2 seats and keeps its fee as 10. Beyond this 2 chairs, if school 1 adds one more chair, the marginal benefit is 1 student x price to attract that student (8) = 8. However, the marginal cost of adding one more chair is total number of students (=12) x fee forgone to attract one more student (=2), which is equal to 24, much higher than the marginal benefit. Thus, in an equilibrium where both school invests in capacity, school 1 invests 2 additional chairs and keep 6 of the cash grant, with a total profit of 126.
- However, if school 1 invests in quality instead, the profit would be $10 \times 12.9 = 129$, which is higher than 126. Hence, there cannot exist an equilibrium where both schools invest in capacity

Variable Cost for Quality Investment

In this example, we show that our results would still hold even when quality investment is a variable cost. We still suppose that capacity investment is a fixed cost, however quality investment increases the marginal cost of each enrolled student. As before, schools can't invest in capacity because they are cash constrained. Differently, now we assume that schools collect profits in two rounds. There is no strategic component in the second round. The investment decisions are all made in the first round. Schools can't invest in quality without the cash grant because (1) they can increase their fees (for high quality) with a lag; say a year later, and (2) quality investment increases the marginal cost in such a way that schools make negative profit for the first year and schools cannot roll over negative profit next year without any credit. Therefore, cash grant helps schools elevate the burden on their balance sheet and survive to the next round. Therefore, suppose for now that

1. Each school's initial capacity is 10.

2. Size of the uncovered market, $N = 4$.

3. Size of cash grant, $K = 5$.

4. Cost per unit of extra capacity, $r = 1$.

5. Willingness to pay for low quality, $q_L = 2$.

6. Willingness to pay for high quality, q_H, is a free variable.

7. Investment in capacity is a fixed cost whereas investment in quality increases the marginal cost per student to $c = 2.5$.

L treatment:

If a school invests on capacity in equilibrium, it picks 4 chairs and receives profit $2 \times 14 + 1 = 29$ for period 1 and $2 \times 14 = 28$ in period 2, with a total profit of 57.

However, if the same school invests in quality, the first period profit would be

$$\Pi_1 = (2 - c)10 + 5 = 0$$

and the second period profit is

$$\Pi_2 = (q_H - c)10.$$

Thus, the treated school invests in quality whenever

$$\Pi_1 + \Pi_2 \geq 57.$$

H treatment: Suppose that both schools invest in quality in H treatment. it means that each school's first period profit is

$$\Pi_1 = (2 - c)10 + 5 = 0$$

and the second period profit is

$$\Pi_2 = (q_H - c)10.$$

However, if one of the schools deviates and invests on capacity, it picks 4 chairs and receives profit $2 \times 14 + 1 = 29$ for period 1 and $2 \times 14 = 28$ in period 2, with a total profit of 57.

Therefore, if a school invests in quality in L treatment, (that is $\Pi_1 + \Pi_2 \geq 57$), then both schools investing in capacity in H treatment is also an equilibrium.

In fact, if $\Pi_1 + \Pi_2 \geq 57$ holds, then there is no equilibrium in H treatment where both schools invest in capacity: Suppose that both schools invest in capacity in H treatment. It means, each school buys 2 chairs (to avoid price competition) and receives profit of $2 \times 12 + 3 = 27$ for period 1 and $2 \times 12 = 24$ in period 2, with a total profit of 51.

However, if one of the schools deviates and invests in quality, the first period profit would be

$$\Pi_1 = (2 - c)10 + 5 = 0$$

and the second period profit is

$$\Pi_2 = (q_H - c)10.$$

Therefore, if a school invests in quality in L treatment, (that is $\Pi_1 + \Pi_2 \geq 57$), then there is no equilibrium where both schools invest in capacity in H treatment (because $\Pi_1 + \Pi_2 > 51$).

The example, and the analysis, would be more involved if the first period profit Π_1 was positive because schools that invest in quality would have incentive to invest in capacity as well. However, this simple example shows that if L school invests in quality, this is because (1) the parents' willingness to pay, i.e., q_H, is much higher than the marginal cost of high quality or (2) the size of the uncovered market is small enough. But then the same forces and additional competition for uncovered students will (more strongly) push schools to invest in quality in H treatment.

References

Andrabi, T., Das, J., and Khwaja, A. I. (2002). Test Feasibility Survey PAKISTAN : Education Sector.

Andrabi, T., Das, J., Khwaja, A. I., and Zajonc, T. (2011). Do value-added estimates add value? accounting for learning dynamics. *American Economic Journal: Applied Economics*, 3(3):29–54.

Baird, S., Bohren, J. A., Mcintosh, C., and Ozler, B. (2016). Designing Experiments to Measure Spillover Effects. *Policy Research Working Paper No. 6824*.

Bruhn, M. and McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1(4):200–232.

Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., and Zamora, P. (2013). Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. *Quarterly Journal of Economics*, 128(2):531–580.

Dasgupta, P. and Maskin, E. (1986). DasguptaMaskin1986.pdf. *The Review of Economic Studies*, 53(1):1–26.

Kreps, D. M. and Scheinkman, J. a. (1983). Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes. *The Bell Journal of Economics*, 14(2):326–337.