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 Theory

A.1 Homogeneous Demand

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

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

Define $n_H = \frac{K-w}{r}$ and $n_L = \frac{K}{r}$ to be the additional capacity increase that schools can afford under high and low technologies, 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 the main results, we prove the following result, which applies to both low (L) and high-saturation (H) treatments.

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 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. □

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 $\hat{\mu}_1^*(p) > 0$. □

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 $i = 1, 2$. □

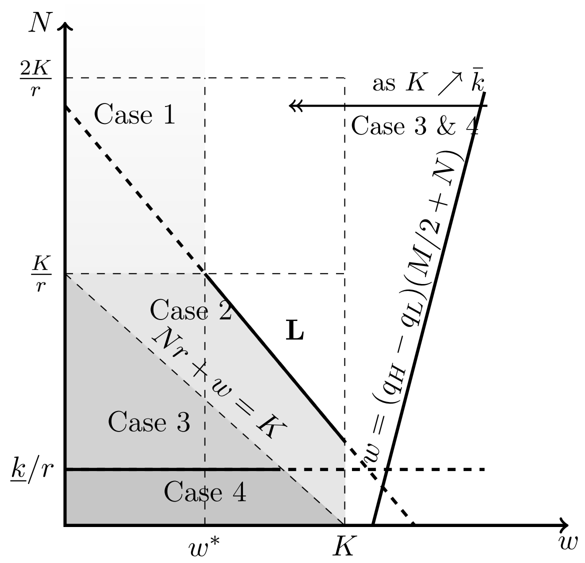
Lemma A6. *For each $i \in \{1, 2\}$, $\bar{p}_i > p_i^*$, and there exists no p, p' with $p_i^* < p < p' < q_i$ such that $\mu_i^*([p, p']) = 0$.*

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^*. □

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 1 (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: $K \le Nr$ (or equivalently $n_L \le N$): 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 $\Pi_H \ge \Pi_L$, implying $w \le w^*$.

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 $x_1 = N$.*

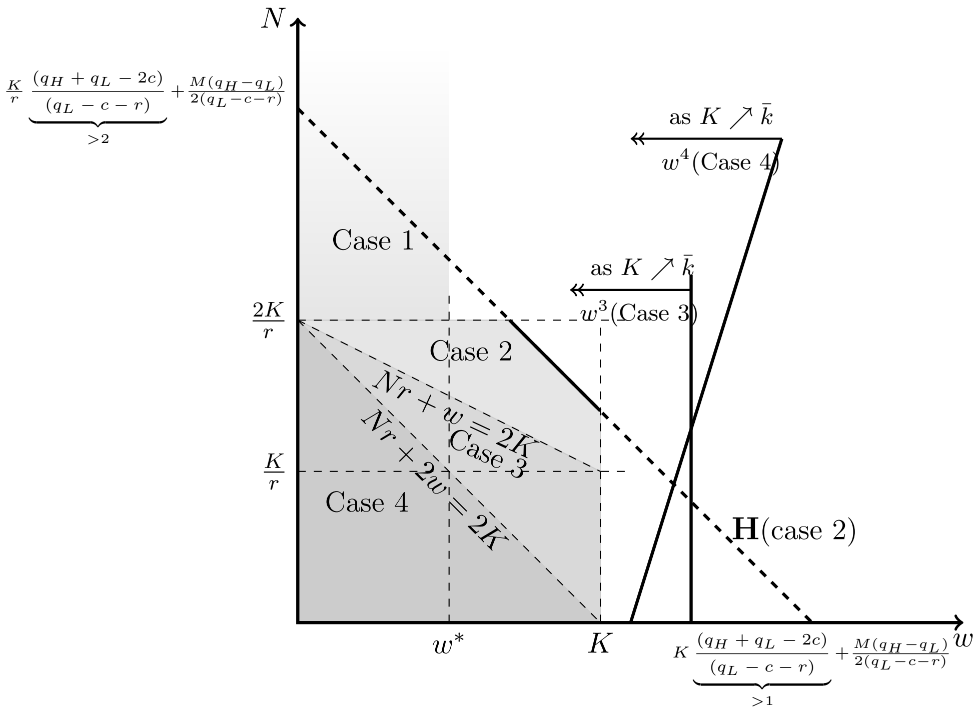
Proof. Suppose for a contradiction that $x_1 = N + e$ where $e > 0$. This time school 1 randomly picks a price over the range $[p_1^*, q_H]$ with a probability measure μ_1 and school 2 randomly picks a price over the range $[p_2^*, q_L]$ with a probability measure μ_2. Schools' profit functions are given 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 1 (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 1 (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.

A.2 Generalization of the Model

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 $D(p) = T(1-\frac{p}{q})$. We adopt the rationing rule of Kreps and Scheinkman (1983), henceforth KS. 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. 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 modification of the example in the main text to 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.[1] 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:[2]

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.[3] But, once school 2's capacity increases to 6, there is no pure strategy NE.[4] The threat of mixed strategy equilibrium prices forces schools to not expand their capacities beyond their Cournot optimal capacities.[5]

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

[1] 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.

[2] 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.

[3] 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).

[4] 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 $\$4 \times 6 = 24$.

[5] 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.

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

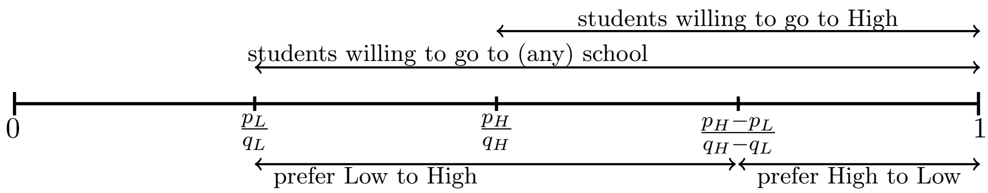
$$B(y) = \frac{1-y}{2}.$$

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 1: 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

$$x_H < 1 - \frac{p_H - p_L}{q_H - q_L} \tag{19}$$

$$1 \leq x_H + 2x_L \tag{20}$$

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).

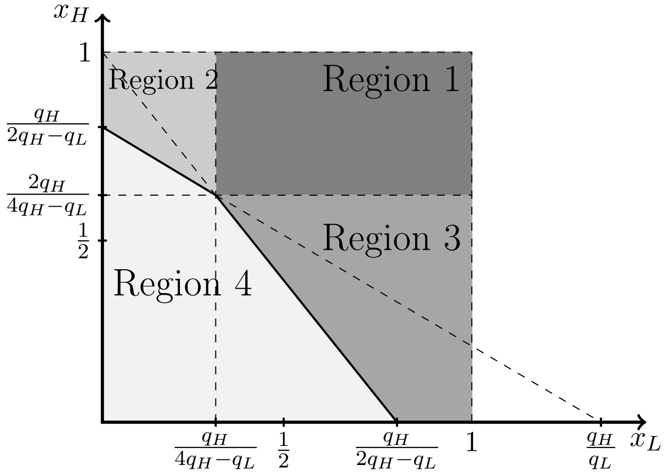
Summary of the Equilibrium: The equilibrium prices can be summarized in the following picture where

Region 1: Parameters satisfy $x_H \geq \frac{2q_H}{4q_H - q_L}$ and $x_L \geq \frac{q_H}{4q_H - q_L}$. 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 $q_L x_L + (2q_H - q_L)x_H \geq q_H$. 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 $q_H x_H + (2q_H - q_L)x_L \geq q_H$. 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 low and high saturation treatment, 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 2. *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

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

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

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

<u>High saturation treatment with (H, H) Equilibrium</u>: 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

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

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

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

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

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

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

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 $\Pi^H_{Low} - \Pi^H_{(H,L)} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} = \frac{KMq_L}{2r} + \widehat{\Pi}^H_{(H,L)} - \Pi^L_{Low} \leq 0$.

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

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.

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

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. □

B Weighting of average treatment effects with unequal selection probabilities

B.1 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 fairly simply 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:

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

B.2 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.

C Sampling, Surveys and Data

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. Note that while the ballot randomization included all 855 schools, the final analysis sample has 852 schools (unbeknownst to us 1 school had closed down by the time of the ballot and the other 2 were actually refusals that were mis-recorded by field staff). Appendix Figure C1 summarizes the number of villages and schools in each experimental group.

Power Calculations

We use 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).

Survey Instruments

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 C2.

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 C3 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.[6] 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.[7] 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 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.

[6] 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.

[7] 97 percent of schools (394/408) had positive enrollment in grade 4.

Data Definitions

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.

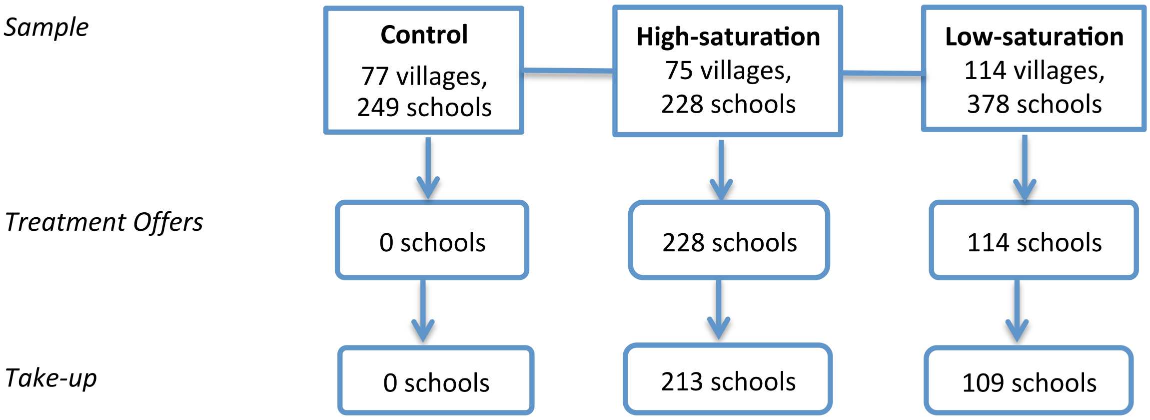
Variable	**Description**	**Survey Source**
Group A: Village Level		
Grant per capita	Grant amount per private school going child in treatment villages. For L villages, this is Rs 50,000/total private enrollment, and for H villages, this equals (50,000*# of private schools in village)/total private enrollment. 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	
Enrollment	School enrollment in all grades, verified through school registers. Coded as 0 after school closure.	School
Fees	Monthly tuition fees charged by the school averaged across all grades.	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.	School
Collected Revenues	Self-reported measure on total monthly fee collections from all enrolled students. Coded as 0 after school closure unless otherwise specified.	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.	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
Fixed Costs	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 0 after school closure.	School

Continued on next page

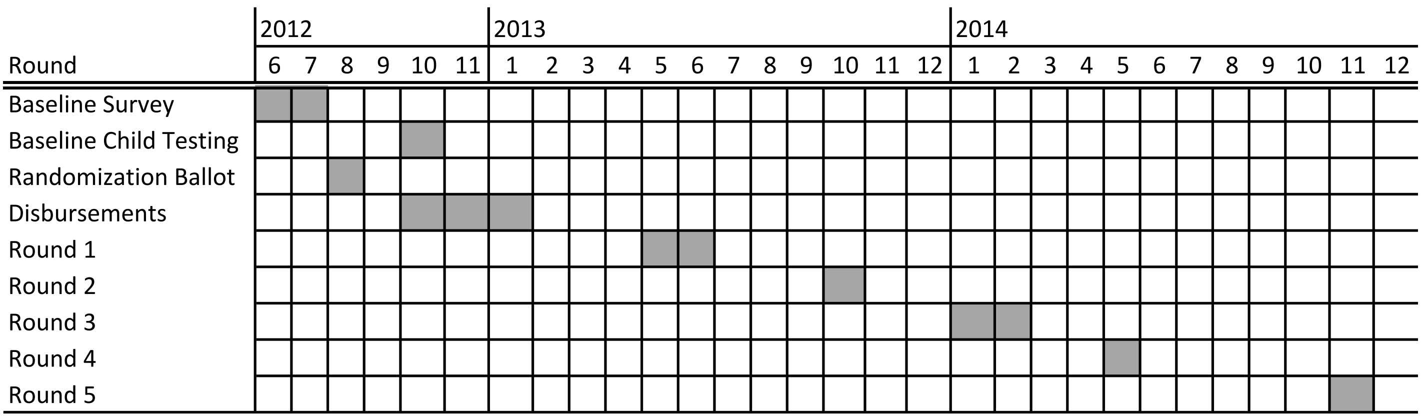
Continued from previous page

Variable	**Description**	**Survey Source**
Variable Costs	Sum of spending on teacher salaries, non-teaching staff salaries, rent and utilities for a given month. Coded as 0 after school closure.	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.	School owner
Household borrowing: Loan value	Value of total borrowing in PKR by the owner household from any source for any purpose.	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

Appendix Figure C1: Sample Details



Appendix Figure C2: Project Timeline



Notes: Rounds 1-3 correspond to the first school year 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 C3: 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. See Appendix C for details.

D Balance and Attrition

In this section, we discuss and address issues of experimental balance and attrition in detail.

D.1 Balance

As noted in the main text of the paper, our randomization is always balanced in distributional tests across the village and school level. While there is no mean imbalance at the village level in univariate comparisons, we do detect mean imbalance in a few comparisons between the L^t schools and schools in H and control. This imbalance is primarily driven by the skewness (heavy right tail) of some of our covariates. To see this, recall that our randomization is stratified by village size and average revenue and takes place in two stages, first at the village level and then at the school level. While stratification helps in reducing the ex-ante probability of imbalance at the village level, it does not automatically guarantee the same for school level regressions. Instead, the source of imbalance for the L^t group is related to distributional skewness and the sample sizes we realize as a result of our design. Because only 1 school in a low-saturation village is offered a grant, there are 114 L^t schools in comparison with 228 H and 249 control schools. The smaller sample size for the L^t group increases the likelihood that the distributional overlap for a given covariate between the L^t group and the H or control group may have uneven mass, especially in the tails of the distribution. It is therefore reassuring that though we may have mean imbalance in comparisons with the L^t group, the Kolmogorov-Smirnov (K-S) tests in Appendix Table D1 show that we cannot reject that the covariate distributions are the same for comparisons between L^t and other groups. Nevertheless, we conduct two types of additional analyses, presented below, to address any concerns arising from the detected imbalance.

First, we conduct simulations to see whether we still observe mean covariate imbalance when we randomly select data from 1 school in the control or H arm to compare with our L^t sample. The thought experiment here is as follows: Assume we only had money to survey 1 school in each experimental group, but the treatment condition remained the same (i.e. all schools are treated in H; 1 school in L; and no schools in control). Our school level balance regressions would now only use data from the surveyed schools. Since these sample sizes are more comparable, the likelihood of imbalance is now lower. Indeed, when we run 1000 simulations of this procedure, we find no imbalance on average using this approach between either L^t and control, or L^t and H schools. This approach can also be applied to estimate our treatment effects, and we find that our key results are quite similar in magnitudes though we lose some precision due to the smaller sample sizes. This exercise lends support to the idea that the mean imbalance at the school level does not reflect a randomization failure but rather issues of covariate overlap in group distributions.

Second, we assess the robustness of our results by trimming the right tails, top 2%, of the imbalanced variables and re-running the balance and main outcome regressions. The previous analysis provides justification for undertaking these approaches as a way to understand our treatment effects. Appendix Table D2 shows our balance regressions with trimmed baseline variables. There is no average imbalance for enrollment or fees in comparisons between L^t versus control; we observe some imbalance at the 10% level for H vs L^t schools for fees. However, observing 3 out of 32 imbalanced tests at the 10% level may occur by random chance. Our outcome regressions using trimmed baseline data in Appendix Tables D3 are also nearly identical to the tables in the main text. Together, these tests reveal that the limited imbalance we detect does not pose any noteworthy concerns for our results.

D.2 Attrition

Even though we have high survey completion rates throughout the study, we do observe some differential response rates between the L^t and control schools (see Appendix Table D4). It is not surprising that treated schools, especially in the L arm, may be more willing to participate in surveys given the sizable nature of the cash grant they received. Here, we check robustness of our results to this (small) differential attrition using predicted attrition weights. The procedure is as follows: We calculate the probability of refusal (in any follow-up round) given treatment variables and a set of covariates using a probit model, and use the predicted values to construct weights.[8] The weight is the inverse probability of response $(1 - prob(attrition))^{-1}$, and is simply multiplied to the existing saturation weight. This procedure gives greater weight to those observations that are more likely to refuse in a subsequent round.

Appendix Table D5 shows our key regressions using attrition weights. Given the low levels of attrition, our results, unsurprisingly, are similar in magnitudes and significance to tables in the main text.

[8] The probit model reveals that only our treatment variable has any predictive power for attrition.

Table D1: Randomization Balance

	(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	2.5	0.011	0.010	0.001	0.95	1.00	1.00	
			[0.95]	[0.95]	[0.99]				
Number of private schools	266	3.3	0.021	0.162	-0.141	1.00	1.00	0.99	
			[0.85]	[0.16]	[0.18]				
Private enrollment	266	523.5	-23.549	11.202	-34.750	0.28	0.86	0.30	
			[0.51]	[0.71]	[0.29]				
Average monthly fee (PKR)	266	232.1	12.668	-12.855	25.523	0.46	0.85	0.57	
			[0.41]	[0.20]	[0.07]				
Average test score	133	-0.222	-0.013	0.031	-0.044	0.27	0.51	0.35	
			[0.88]	[0.75]	[0.57]				
Overall Effect: p-value			0.95	0.96	0.99				

Panel B: Private school level variables									
			Tests of difference				K-S Test p-values		
		Control							
	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	0.18	0.69	0.90
			[0.66]	[0.07]	[0.91]	[0.17]			
Monthly fee (PKR)	851	238.1	24.1	-32.3	-10.7	56.4	0.94	0.42	0.24
			[0.15]	[0.02]	[0.35]	[0.00]			
Annual expenses (PKR)	837	78860.9	21,559.2	-16,659.5	-5,747.2	38,218.7	0.58	0.88	0.57
			[0.13]	[0.15]	[0.60]	[0.01]			
Monthly expenses (PKR)	848	25387.0	2,692.7	-2,373.7	2,280.1	5,066.3	0.81	0.82	0.94
			[0.32]	[0.43]	[0.28]	[0.16]			
Infrastucture index (PCA)	835	-0.008	0.073	0.308	-0.074	-0.235	0.22	0.40	0.27
			[0.64]	[0.17]	[0.56]	[0.33]			
School age (in years)	852	8.3	0.028	0.296	0.220	-0.268	0.98	0.73	0.61
			[0.96]	[0.69]	[0.70]	[0.72]			
Number of teachers	851	8.2	0.015	-0.408	0.242	0.423	1.00	0.95	0.81
			[0.97]	[0.39]	[0.48]	[0.37]			
Average test score	401	-0.210	-0.054	0.160	-0.052	-0.214	0.55	0.39	0.11
			[0.53]	[0.18]	[0.61]	[0.05]			
Overall Effect: p-value			0.83	0.28	0.24	0.33			

Notes: * p<0.1, ** p<0.05, *** p<0.01

a) This table shows randomization checks at the village and private school level, Panel A and B respectively, for key variables in our study. Across both panels, column 1 shows number of observations and col 2 shows the control mean. Panel A, cols 3-5 and Panel B, 3-6 show tests of differences-- regression coefficients and p-values in square brackets-- between experimental groups. Panel A, cols 6-8, and Panel B, cols 7-9 show p-values from Kolmogorov-Smirnov (K-S) tests of equality of distributions. In the bottom row, we report p-value from a test asking whether variables jointly predict treatment status for each 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 C. There are fewer observations for test scores since half of the sample was tested at baseline.

Table D2: Randomization Balance, Trimmed Sample

	(1)	(2)	(3)	(4)	(5)	(6)
		Control	Tests of difference			
Private school level variables	N	Mean	H-C=0	L^t-C =0	L^u-C =0	H-L^t =0
Enrollment	836	154.1	-5.7	-13.8	-2.0	8.1
			[0.39]	[0.14]	[0.77]	[0.35]
Monthly fee (PKR)	834	221.6	2.5	-20.3	-8.4	22.8
			[0.81]	[0.13]	[0.38]	[0.07]
Annual expenses (PKR)	821	65441.7	5,875.8	-5,477.6	-4,902.8	11,353.3
			[0.53]	[0.60]	[0.57]	[0.32]
Monthly expenses (PKR)	832	22293.5	1,061.4	-2,774.9	2,720.2	3,836.4
			[0.49]	[0.14]	[0.10]	[0.05]
Infrastucture index (PCA)	819	-0.141	0.077	0.133	-0.012	-0.056
			[0.41]	[0.31]	[0.88]	[0.69]
School age (No of years)	836	7.9	-0.191	0.615	0.171	-0.806
	.	.	[0.69]	[0.40]	[0.74]	[0.25]
Number of teachers	834	7.7	-0.045	-0.290	0.316	0.245
			[0.88]	[0.44]	[0.31]	[0.47]
Average test score	393	-0.242	-0.020	0.074	-0.029	-0.095
			[0.81]	[0.48]	[0.75]	[0.34]
Overall Effect: p-value			0.85	0.47	0.94	0.38

Notes: * p<0.1, ** p<0.05, *** p<0.01

a) This table reproduces Table D1, Panel B, using trimmed data to assess whether mean imbalance in Table D1, Panel B, is driven by large values in the right tails. The trimming procedure makes the top 2% of baseline values missing for each variable. Column 1 shows the number of observations, and col 2 shows the control mean. The remaining columns show tests of difference -- regression coefficients and p-values in square brackets-- between groups. In the bottom row, we report p-values from a test asking whether variables jointly predict treatment status for each group.

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

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

Table D3: Main Outcomes, Trimmed Sample

	(1) Enrollment	(2) Fees	(3) Score
High	10.50*	13.20*	0.154*
	(5.73)	(7.20)	(0.08)
Low Treated	24.01***	-1.49	0.005
	(7.39)	(7.42)	(0.10)
Low Untreated	-2.16	-1.58	0.033
	(5.44)	(6.10)	(0.07)
Baseline	0.78***	0.75***	0.473***
	(0.04)	(0.04)	(0.09)
R-Squared	0.52	0.58	0.19
Observations	3985	2272	720
# Schools (Rounds)	821 (5)	786 (3)	720 (1)
Mean Depvar	154.13	221.58	-0.24
Test pval (H=0)	0.07	0.07	0.07
Test pval ($L^t = 0$)	0.00	0.84	0.96
Test pval ($L^t = $ H)	0.07	0.06	0.13

Notes: * p<0.1, ** p<0.05, *** p<0.01

a) This table reproduces our results using baseline variables trimmed at the top 2% as controls; the trimming procedure drops the top 2% of the baseline measure of the dependent variable from the regression. Columns 1-3 show impacts on enrollment, fees and test-scores.

b) Regressions are weighted to adjust for sampling and tracking as necessary and include strata and round fixed effects, with clustered standard errors at the village level. The number of observations may vary across columns as data are pooled across rounds and not all outcomes are measured in every round. We thus also report the number of schools and round for each regression; any variation in the number of schools arises from attrition or missing values for some variables.

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 D4: Differential Attrition

	(1)	(2)	(3)	(4)	(5)
	Control	High	Low Treated	Low Untreated	N
Panel A: Differential 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: Differential Baseline Characteristics for Attriters (At least once refused) by Treatment Status					
Enrollment	191.4	8.4	6.4	-33.0*	79
		(44.68)	(28.77)	(18.74)	
Monthly fee (PKR)	257.5	-28.5	-47.5	37.2	79
		(60.78)	(42.46)	(50.90)	
Annual fixed expenses (PKR)	103745.0	55017.7	20106.0	-49684.0	77
		(90071.94)	(26347.19)	(39480.86)	
Monthly variable costs (PKR)	31768.8	7830.1	44448.2	-4501.2	79
		(19060.95)	(31225.62)	(9184.26)	
Infrastructure index	0.062	0.536	1.140	-0.192	78
		(0.39)	(0.74)	(0.36)	
School age (No of years)	8.8	6.3*	-3.47	0.59	79
		(3.64)	(2.79)	(2.62)	
Number of teachers	9.7	1.01	-0.61	-0.81	79
		(2.59)	(0.94)	(0.79)	

Notes: * p<0.1, ** p<0.05, *** p<0.01

a) This table examines differential attrition, defined as refusal to participate in follow-up surveying, across experimental groups, and assesses whether attriters have systematically different baseline characteristics across groups. Panel A tests for differential attrition in each follow-up round (1-5) 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. 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).

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 D5: Main Outcomes, using Attrition-predicted Weights

	(1) Enrollment	(2) Fees	(3) Score
High	8.71	25.69***	0.17*
	(5.55)	(7.88)	(0.09)
Low Treated	16.73**	5.47	-0.04
	(7.19)	(7.86)	(0.11)
Low Untreated	0.91	6.30	0.06
	(5.27)	(6.40)	(0.07)
Baseline	0.77***	0.82***	0.37***
	(0.04)	(0.04)	(0.11)
R-Squared	0.62	0.71	0.16
Observations	3878	2230	706
# Schools (Rounds)	797 (5)	769 (3)	706 (1)
Mean Depvar	163.64	238.13	-0.21
Test pval (H=0)	0.12	0.00	0.05
Test pval ($L^t = 0$)	0.02	0.49	0.72
Test pval ($L^t =$ H)	0.24	0.01	0.05

Notes: * p<0.1, ** p<0.05, *** p<0.01
a) This table checks whether our results are robust to accounting for differential attrition using the inverse probability weighting technique. In addition to using saturation or tracking weights, we now weight all regressions with attrition-predicted weights. This procedure is described in detail in Appendix D. Cols 1-3 show impacts on enrollment, fees, and test scores with these weights.
b) Regressions are weighted to adjust for sampling, tracking where necessary, and attrition, and include strata and round fixed effects, with standard errors clustered at the village level.
The number of observations may vary across columns as data are pooled across rounds and not all outcomes are measured in every round. We thus also report the number of schools and rounds for each regression; any variation in the number of schools arises from attrition or missing values for some variables.
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.

E Additional Results

This section includes additional tables referenced in the main text.

Table E1: 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
High	-0.007	0.002	-0.010	0.020	-0.033	1,063.0
	(0.01)	(0.01)	(0.05)	(0.02)	(0.05)	(15,092.8)
Low Treated	-0.0004	-0.006	-0.039	0.010	-0.053	17,384.2
	(0.01)	(0.01)	(0.05)	(0.02)	(0.05)	(29,982.8)
Low Untreated	-0.002	-0.011	-0.005	0.035*	-0.055	13,611.9
	(0.01)	(0.01)	(0.04)	(0.02)	(0.04)	(21,581.8)
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
Observations	795	795	784	784	784	784
# Schools (Rounds)	795 (1)	795 (1)	784 (1)	784 (1)	784 (1)	784 (1)
Mean Depvar	0.99	0.02	0.23	0.02	0.21	44,782.7
Test pval (H=0)	0.48	0.88	0.83	0.23	0.47	0.94
Test pval (L^t= 0)	0.97	0.68	0.45	0.64	0.27	0.56
Test pval (L^t=H)	0.53	0.56	0.60	0.65	0.69	0.60

Notes: * p<0.10, ** p<0.05, *** p<0.001

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. Data for columns 1-2 are from the school survey and from the school owner survey for cols 3-6. The dependent variables in col 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. Col 3 reports whether the household of the school owner has ever borrowed any money for any reason. Cols 4-5 disaggregate this household borrowing into formal and informal sources. Col 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 and round fixed effects, with standard errors clustered at the village level. The number of observations and unique schools are the same since we use one round of data. Observations may vary across columns due to attrition and missing values. The mean of the dependent variable is the follow-up control mean.

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 E2: Enrollment by Grades

	(1) Lower than 1	(2) 1 to 3	(3) 4 to 5	(4) 6 to 8	(5) 9 to 12
High	3.11	2.49	1.57	1.82	1.36
	(2.15)	(2.05)	(1.11)	(1.55)	(1.15)
Low Treated	6.51**	8.81***	2.85**	4.33**	3.73
	(2.52)	(2.57)	(1.27)	(2.04)	(2.45)
Low Untreated	1.31	1.78	1.32	0.63	-1.29
	(1.95)	(1.83)	(1.06)	(1.48)	(1.29)
Baseline	0.59***	0.73***	0.70***	0.62***	0.78***
	(0.06)	(0.05)	(0.03)	(0.04)	(0.10)
R-Squared	0.38	0.54	0.59	0.57	0.65
Observations	3,334	3,420	3,420	3,420	3,420
# Schools (Rounds)	852 (4)	855 (4)	855 (4)	855 (4)	855 (4)
Mean Depvar	49.89	53.68	28.15	23.10	8.22
Test pval (H=0)	0.15	0.22	0.16	0.24	0.24
Test pval (L^t= 0)	0.01	0.00	0.03	0.03	0.13
Test pval (L^t=H)	0.17	0.01	0.28	0.20	0.39

Notes: * p<0.1, ** p<0.05, *** p<0.01

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 are assigned 0 enrollment.

b) Regressions are weighted to adjust for sampling and include strata and round fixed effects, with standard errors clustered at village level. We report the number of observations and the unique number of schools and rounds in each regression; the number of unique schools may be fewer than the full sample due to attrition or missing values for some variables. 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 E3: 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
Observations	765	711
# Schools (Rounds)	765 (1)	711 (1)
Mean Depvar	14.69	0.07
Test pval (H=0)	0.62	0.10
Test pval (L^t= 0)	0.30	0.02
Test pval (L^t=H)	0.56	0.21

Notes: * p<0.1, ** p<0.05, *** p<0.01

a) This table examines changes in child enrollment status. The dependent variables are from tested children in round 3. Col 1 is the number of children enrolled in grade 4, and col 2 is the fraction of those 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 fewer than 18 months before are considered new.

b) Regressions are weighted to adjust for sampling and include strata and round fixed effects, with standard errors clustered at village level. The number of observations and schools is the same in this table since we survey children just once. Observations may be lower than the full sample due to missing values for some variables. The mean of the dependent variable is its baseline value or the follow-up control mean.

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 E4: Monthly Tuition Fees by Grades

	(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
Observations	2,277	2,278	2,240	1,485	360
# Schools (Rounds)	789 (3)	789 (3)	773 (3)	542 (3)	144 (3)
Mean Depvar	169.89	207.82	237.43	319.88	425.94
Test pval (H=0)	0.17	0.08	0.12	0.04	0.04
Test pval (L^t= 0)	0.37	0.61	0.32	0.14	0.26
Test pval (L^t=H)	0.08	0.05	0.04	0.00	0.53

Notes: * p<0.1, ** p<0.05, *** p<0.01

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. We report the number of observations and the unique number of schools and rounds in each regression. Higher grades have fewer school observations because fewer schools offer those grade levels and hence post tuition fees. These observations are subsequently coded as missing. In contrast, in Table E2, enrollment in higher grades is coded as 0 if a school does not offer those grades. The pattern of results in Table E2 stay the same if we restrict its sample to the sample in this table. 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 E5: School Test Scores, Different Controls

	No controls				Additional controls			
	(1) Math	(2) Eng	(3) Urdu	(4) Avg	(5) Math	(6) Eng	(7) Urdu	(8) Avg
High	0.155	0.181*	0.115	0.150	0.157*	0.185*	0.108	0.151*
	(0.105)	(0.102)	(0.092)	(0.096)	(0.093)	(0.094)	(0.088)	(0.088)
Low Treated	-0.066	0.108	-0.059	-0.006	-0.0832	0.069	-0.087	-0.038
	(0.122)	(0.114)	(0.114)	(0.111)	(0.106)	(0.104)	(0.102)	(0.0981)
Low Untreated	0.021	0.055	0.007	0.028	0.005	0.046	-0.024	0.007
	(0.091)	(0.091)	(0.081)	(0.083)	(0.078)	(0.082)	(0.077)	(0.074)
Baseline					0.373***	0.457***	0.312***	0.433***
					(0.077)	(0.064)	(0.01)	(0.086)
R-Squared	0.08	0.06	0.08	0.08	0.27	0.20	0.21	0.24
Observations	732	732	732	732	722	722	722	722
# Schools (Rounds)	732 (1)	732 (1)	732 (1)	732 (1)	722 (1)	722 (1)	722 (1)	722 (1)
Mean Depvar	-0.21	-0.18	-0.24	-0.21	-0.21	-0.18	-0.24	-0.21
Test pval (H=0)	0.14	0.08	0.21	0.12	0.09	0.05	0.22	0.08
Test pval (L^t= 0)	0.59	0.34	0.60	0.96	0.43	0.51	0.40	0.70
Test pval (L^t=H)	0.07	0.52	0.13	0.16	0.02	0.27	0.05	0.05

Notes: * p<0.10, ** p<0.05, *** p<0.001
a) This table conducts robustness checks on our school test score results. School test scores are generated by averaging child average (across all subjects) test scores for a given school. Columns 1-4 are the same regressions as Table 4, Columns 1-4, but without any baseline controls. Columns 5-8 repeat these regressions with additional controls, which include the baseline score, percentage of students in specific grades and percentage female. Test scores are averaged across all children in a given school separately for each round, and child composition is hence different across rounds.
b) Regressions are weighted to adjust for sampling and include strata fixed effects, with standard errors clustered at village level. We include a dummy variable for the untested sample at baseline across all columns and replace the baseline score with a constant. Since the testing sample was chosen randomly at baseline, this procedure allows us to control for baseline test scores wherever available. The number of observations and the unique number of schools are the same since test scores are only collected once after treatment. The number of schools is lower than the full sample due to attrition and zero enrollment in some schools in the tested grades. The mean of the dependent variable is the test score for those schools tested at random at baseline.
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 E6: Test Scores, Stayers Only

	School level				Child level
	(1)	(2)	(3)	(4)	(5)
	Math	Eng	Urdu	Avg	Avg
High	0.150	0.191*	0.120	0.132*	0.235**
	(0.093)	(0.098)	(0.085)	(0.077)	(0.094)
Low Treated	-0.114	0.054	-0.090	-0.034	0.095
	(0.115)	(0.111)	(0.111)	(0.089)	(0.108)
Low Untreated	0.031	0.055	0.015	0.016	0.002
	(0.077)	(0.084)	(0.071)	(0.063)	(0.083)
Baseline Score	0.279**	0.429***	0.365***	0.337***	0.637***
	(0.135)	(0.118)	(0.109)	(0.098)	(0.049)
R-Squared	0.17	0.13	0.15	0.17	0.21
Observations	720	720	720	720	11,676
# Schools (Rounds)	720 (1)	720 (1)	720 (1)	720 (1)	711 (1)
Mean Depvar	-0.21	-0.21	-0.21	-0.21	-0.18
Test pval (H=0)	0.11	0.05	0.16	0.09	0.01
Test pval (L^t= 0)	0.32	0.62	0.42	0.71	0.38
Test pval (L^t=H)	0.02	0.21	0.06	0.06	0.19

Notes: * p<0.10, ** p<0.05, *** p<0.001

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 4, but only include all children who report being at the same school for at least 1.5 years.

b) Regressions are weighted to adjust for sampling and include strata fixed effects, with standard errors clustered at village level. We include a dummy variable for the untested sample at baseline across all columns and replace the baseline score with a constant. Since the testing sample was chosen randomly at baseline, this procedure allows us to control for baseline test scores wherever available. The number of observations and the unique number of schools are the same since test scores are only collected once after treatment. The number of schools is lower than the full sample due to attrition and zero enrollment in some schools in the tested grades. The mean of the dependent variable is the test score for those tested at random at baseline.

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 E7: Main Outcomes, Interacted with Baseline Availability of Bank Account

	(1) Enrollment	(2) Fees	(3) Score
High	6.93	18.28*	0.118
	(7.36)	(10.09)	(0.10)
Low Treated	21.85**	-1.76	0.021
	(10.35)	(10.14)	(0.13)
Low Untreated	-0.49	0.75	0.005
	(6.86)	(8.14)	(0.08)
High*NoBankAct	7.55	2.09	0.110
	(10.72)	(15.12)	(0.16)
Low Treated*NoBankAct	0.05	6.98	-0.133
	(14.41)	(14.93)	(0.22)
Low Untreated*NoBankAct	2.93	-2.91	0.091
	(11.63)	(13.60)	(0.15)
HH does not have bank act	-1.13	-0.77	-0.102
	(7.42)	(10.01)	(0.11)
Baseline	0.75***	0.83***	0.35***
	(0.05)	(0.04)	(0.11)
R-Squared	0.62	0.72	0.17
Observations	4,059	2,312	725
# Schools (Rounds)	836 (5)	800 (3)	725 (1)
Mean Depvar	163.64	238.13	-0.21

Notes: * $p<0.1$, ** $p<0.05$, *** $p<0.01$

a) This table examines whether our results are driven by baseline access to bank accounts in school owner households. Cols 1-3 reproduce our key results adding an interaction with a dummy variable for whether the owner's household does not have a bank account with 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.

b) Regressions are weighted to adjust for sampling and tracking and include strata and round fixed effects, with standard standard errors clustered at village level. The number of observations may vary across columns as data are pooled across rounds and not all outcomes are measured in every round. We thus also report the number of schools and rounds for each regression, and any remaining variation in the number of schools arises from attrition or missing values for variables. The mean of the dependent variable is its baseline value or the follow-up control mean.

Table E8: Main Outcomes, controlling for Grant size per capita

	(1) Enrollment	(2) Fees	(3) Score
High	-2.714	10.764	0.227
	(10.605)	(12.677)	(0.165)
Low Treated	18.050**	-2.128	-0.004
	(8.345)	(8.197)	(0.110)
Low Untreated	-3.310	-2.431	0.055
	(6.245)	(7.383)	(0.083)
Grant per capita	0.031	0.022	-0.0002
	(0.020)	(0.024)	(0.0004)
Baseline	0.760***	0.826***	0.359***
	(0.047)	(0.037)	(0.114)
R-Squared	0.62	0.72	0.17
Observations	4,059	2,312	725
# Schools (Rounds)	836 (5)	800 (3)	725 (1)
Mean Depvar	163.64	238.13	-0.21
Test pval (H=0)	0.80	0.40	0.17
Test pval (L^t= 0)	0.03	0.80	0.97
Test pval (L^t=H)	0.03	0.21	0.10

Notes: * p<0.1, ** p<0.05, *** p<0.01

a) This table repeats our main results with an additional village level control variable, grant amount per capita. This control variable captures whether our results 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 low-saturation villages and a multiple of 50,000 PKR based on the number of private schools in high-saturation villages.

b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata fixed effects, with standard errors clustered at village level. The number of observations may vary across columns as data are pooled across rounds and not all outcomes are measured in every round. We thus also report the number of schools and round for each regression. Any remaining variation in the number of schools arises from attrition or missing values for some variables. 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 E9: School Infrastructure (Year 2)

	Spending	Number purchased		Facility present (Y/N)			Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Amount (PKR)	Desks	Chairs	Computers	Library	Sports	# Rooms Upgraded
High	606.00	0.56	1.16	0.06	-0.00	0.05*	0.24
	(6537.56)	(1.39)	(0.83)	(0.05)	(0.03)	(0.03)	(0.37)
Low Treated	353.44	-0.92	0.84	0.14**	0.00	0.02	0.31
	(7911.96)	(1.44)	(0.54)	(0.06)	(0.03)	(0.03)	(0.36)
Low Untreated	1497.67	-1.46	0.28	-0.02	0.02	0.02	0.08
	(7029.37)	(1.28)	(0.38)	(0.04)	(0.03)	(0.03)	(0.33)
Baseline	0.04	0.08**	0.01	0.31***	0.02	0.07*	0.74***
	(0.03)	(0.04)	(0.02)	(0.05)	(0.03)	(0.04)	(0.05)
R-Squared	0.05	0.08	0.04	0.16	0.04	0.11	0.51
Observations	770	746	780	784	784	784	784
# Schools (Rounds)	770 (1)	746 (1)	780 (1)	784 (1)	784 (1)	784 (1)	784 (1)
Mean Depvar	57258.48	14.59	10.92	0.39	0.35	0.19	6.36
Test pval (H=0)	0.93	0.68	0.16	0.26	1.00	0.06	0.52
Test pval (L^t= 0)	0.96	0.53	0.12	0.03	0.95	0.46	0.39
Test pval (L^t=H)	0.97	0.32	0.74	0.21	0.95	0.44	0.86

Notes: * p<0.10, ** p<0.05, *** p<0.01

a) This table examines outcomes relating to school infrastructure using data from round 5 only. Column 1 is the annual fixed expenditure on infrastructure– e.g. furniture, fixtures, or facilities. Columns 2-3 refer to the number of desks and chairs purchased. Columns 4-6 are dummy variables for the presence of particular school facilities. Finally, column 7 measures the number of rooms upgraded from temporary to permanent or semi-permanent classrooms. Closed schools take on a value of 0 in all columns.

b) Regressions are weighted to adjust for sampling and include strata fixed effects, with standard errors clustered at the village level. The number of observations and unique schools are the same since we only use one round of data. 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 E10: Revenues, excluding Closed schools

	Overall Posted (monthly)			Overall Collected (monthly)		
	(1)	(2)	(3)	(4)	(5)	(6)
	Full	Top Coded 1%	Trim Top 1%	Full	Top Coded 1%	Trim Top 1%
High	5,471.4	4,872.2*	4,543.6**	4,748.8	4,775.2**	3,593.5*
	(3,432.9)	(2,498.8)	(2,094.2)	(3,482.7)	(2,425.1)	(1,871.3)
Low Treated	8,589.9*	7,287.7*	6,271.1*	5,600.5	4,747.5	3,191.9
	(4,988.8)	(4,032.3)	(3,742.7)	(4,804.2)	(3,349.9)	(2,964.8)
Low Untreated	-1,239.5	-1,434.3	-405.0	-119.6	-298.1	6.9
	(2,843.0)	(2,378.4)	(1,847.0)	(2,753.9)	(2,364.5)	(1,765.4)
Baseline Posted Revenues	1.0***	1.0***	0.9***	0.8***	0.9***	0.7***
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
R-Squared	0.66	0.67	0.61	0.57	0.64	0.56
Observations	2,312	2,312	2,276	2,948	2,948	2,900
# Schools (Rounds)	800 (3)	800 (3)	788 (3)	781 (4)	781 (4)	770 (4)
Mean Depvar	40,181.0	38,654.1	36,199.2	30,865.0	30,208.8	27,653.0
Test pval (H=0)	0.11	0.05	0.03	0.17	0.05	0.06
Test pval (L^t= 0)	0.09	0.07	0.10	0.24	0.16	0.28
Test pval (L^t=H)	0.57	0.57	0.65	0.87	0.99	0.89

Notes: * p<0.1, ** p<0.05, *** p<0.01

a) This table repeats Table 2, Columns 2-7, to look at monthly posted and collected revenues dropping schools once they close down. Columns 1-3 consider posted revenues, defined as the sum of revenues expected from each grade based on enrollment and posted fees. Cols 4-6 consider collected revenues, defined as self-reported revenues actually collected from all students at the school. Top coding of the data assigns the value at the 99th percentile to the top 1% of data. Trimming top 1% of data assigns a missing value to data above the 99th pctl. Both top coding and trimming are applied to each round of data separately.

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 observations may vary across columns as data are pooled across rounds and not all outcomes are measured in every round. We thus also report the unique number of schools and rounds in each regression. Any remaining variation in the number of schools arises from missing values for some variables. The mean of the dependent variable is its baseline value or the follow-up control mean.

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.

F Private and Social Returns Calculations

In this section, we describe our calculations from Section 4 in the main text as well as show IRR calculations. Note that this exercise is necessarily suggestive since 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.

Note that for these calculations, we take all point estimates seriously even if they are not statistically significant or precise. We use these estimates to 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, and the H arm where each school in one village receives PKR 50K.

We now proceed by looking at each beneficiary group separately.

F.1 Welfare Calculations

Summary of calculations: We reproduce the table from the main text below for reference.

	In Rupees			Standard Deviations
Group	Owners	Teachers	Parents	Children
L^t	10,918	-2,514	4,080	61.1
H	5,295	8,662	7,560	117.2

School Owners: We consider net collected revenues, subtracting variable costs from actual collected revenues, as the monthly gains for school owners. Closed schools are considered missing in these calculations (different from Table 2) because we are interested in the gains for school owners rather than the average impact on schools. That is, we implicitly assume that owners who close down their school make (on the margin) a similar amount to what they did before closing the school. Imputing a zero revenue value would be a less plausible and more extreme assumption.

Using Table Table E10, col 5, monthly collected revenues for L^t are Rs.4,748 and Rs.4,775 for H schools. Variable costs are computed using estimates from Table 5, col 6– the cumulative effect is divided by 24 (12 months per year over 2 years of the intervention) for a monthly increase of Rs.1,109 for L^t and Rs.3,010 for H schools. Thus, we have a monthly profit of Rs.3,639 for L^t and Rs.1,765 for H schools. Multiplying by 3 gives us the owner estimates in table above.

Teachers: We use changes in the teacher wage bill to understand how the intervention affected the teacher market. Recall from Table 7 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. We see that the average monthly wage bill in H increases by Rs.2,742 relative to control and decreases by Rs.838 for the L^t schools (Table 7, Column 2). We simply multiply these coefficients by 3, and find that teachers in H increase their overall income by Rs.8,226, while teachers in L^t over three villages decrease their overall income by Rs.2,514.

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.

Since quality does not change in the L arm, our treatment effects arise from a movement along the demand curve, as in Appendix Figure F1, Panel A. 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). Since collected fees drop by Rs 8 (Table 3, Col 9) and enrollment increases by 12 children (Table 3, col 5), the L^t PQ-combination is P_L=230Rs and Q_L=176. Hence, our linear demand curve is $Q = 521 - 1.5P$.

From Appendix Figure F1, Panel A, 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, represented by the dotted quadrilateral region. This additional surplus is calculated as the difference in areas of the two triangles generated by the baseline and L^t PQ-combinations and equals Rs.1,360. For a total 150K in grants across three villages, the increase in CS is therefore Rs.4,080. The increase in consumer surplus in L^t is largely driven by the fee reduction faced by the inframarginal children; the newly enrolled, 'marginal,' children were just at the cusp of indifference before the intervention and so their gains are quite small.

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, as in Appendix Figure F1, Panel B. We use our H estimates to obtain this new demand curve. Since collected fees increase by Rs 29 and enrollment by 9 children, our pivoted linear demand curve is $Q = 521 - 1.3P$. The consumer surplus from this new demand curve is Rs.11,485; relative to the baseline consumer surplus, this represents an increased surplus of Rs.2,520 per school. The total consumer surplus increase from grant investment of RS.150K is thus Rs 7,560.

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.

As seen in Appendix Table E6, the H arm dramatically improves test scores for already enrolled children. In particular, considering a total baseline enrollment of 492 children from 3 schools, our H child test score gains of 0.22 sd (Table 4, Col 5) suggest a total increase of 108.2 sd. In comparison, the total gain in L^t is substantially lower at 49.2sd, even if we take the (statistically insignificant) 0.1sd coefficient at face value.

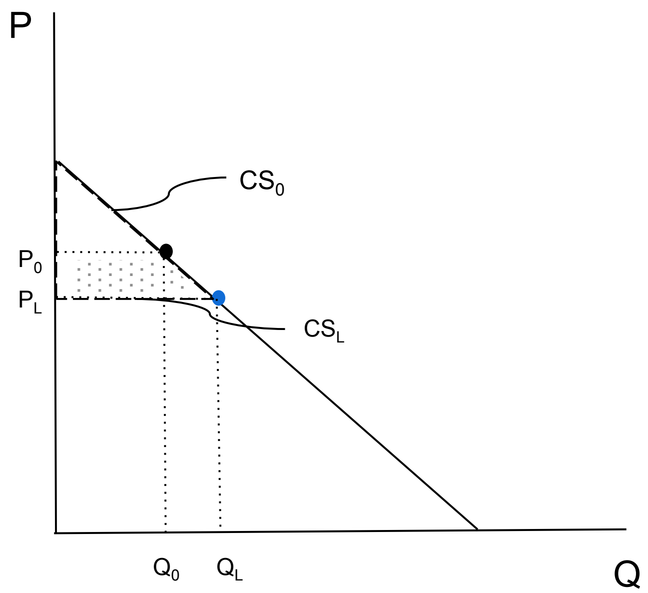
For newly enrolled children, we rely on our previous work, Andrabi et al. (2017), showing test score gains of 0.33sd for children who switch from public to private schools.[9] In H villages, this leads to a total test score gain of 8.9 standard deviations as each of the three schools gains 9 children (0.33sd*9*3). For the L^t sample, each school gains 12 children (Table 3, Col 5), which means a total increase of 36 children across 3 villages, and a total test score increase of 11.9sd (0.33*36).

Summing the gains for already and newly enrolled children, we obtain a total sd gain of 117.2 for H and 61.1 for L approaches.

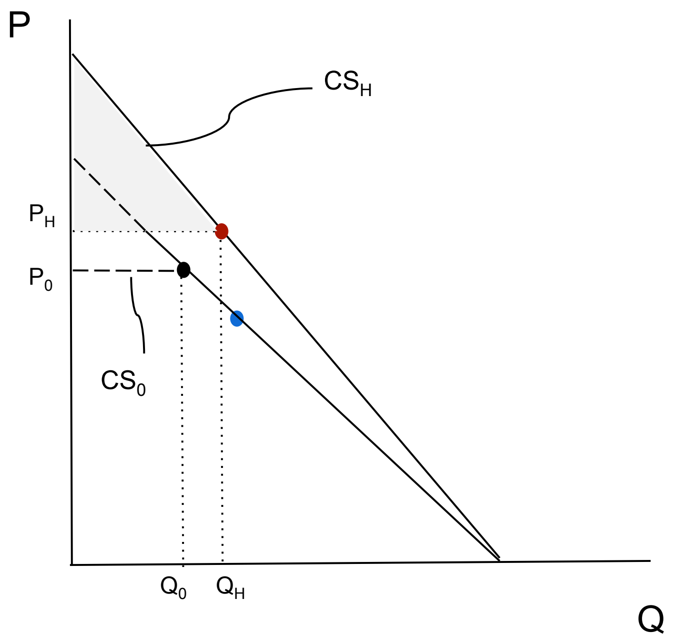
These calculations assume that test score gains accrue to children across all grades, which may be reasonable given that fee increases are observed across grades (Appendix Table E4). Using the same method, if we instead restrict to the tested children in grades 3-5, we obtain a total increase of 31sd in H compared with a 18.2sd increase in L^t.

[9]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.

Appendix Figure F1: Consumer Surplus



Panel A: Consumer surplus at baseline, CS_0, and in L^t from movement along demand curve



Panel B: Consumer surplus in H after a pivot of the demand curve

F.2 IRR and Loan-loss guarantee

The welfare calculations show the tension between private and social returns posed by the two financing treatments. We will now compute the internal rate of return (IRR) directly, and see 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 based on the IRR. We begin by calculating: (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 still use the same estimates of collected revenues and costs as for the welfare calculations, but now also consider fixed costs for assets purchased in year 1 (Table 5, Col 1). With these assumptions, we find returns between 61-83% for L^t and between 12-32% for H schools.

These rates of return are above or just around market interest rates in Pakistan, which range from 15-20%, suggesting that this may be a profitable lending sector. If we were to offer our grant as a RS 50,000 loan at 15% interest rate, it would take a L^t school 1.5 years to pay off the loan and a H school 4 years to pay off their loan.

The higher rates of return coupled with the lower chance of default (Table 3, Col 4) may lead the lender to prefer L over the H approach, unless the fixed costs 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 7pp (Table 3, col 4). The closure rate in L^t group is 1% and 8% for the H group. If we 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.613, while it is Rs.6400 for a H school. For every Rs.150K given out in loans, the social planner would need to subsidize the lender by Rs.17,363 over a two year period of the loan to make them indifferent between the two approaches. This subsidy compares favorably to the annual consumer surplus gain estimated to be Rs.41,760 higher ([Rs.7,560-Rs.4,080]*12) in the H arm as compared to the L arm.

References

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

Andrabi, T., Das, J., and Khwaja, A. I. (2017). Report cards: The impact of providing school and child test scores on educational markets. *American Economic Review*, 107(6):1535–63.

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.