

*Supplementary Material to:*

“Peer Effects, Teacher Incentives, and the Impact of Tracking:

Evidence from a Randomized Evaluation in Kenya”

by Esther Duflo, Pascaline Dupas, and Michael Kremer

This document corresponds to the Proofs Appendix the paper refers to.

## Appendix : Proofs

### *Proof of Proposition 2*

Consider the convex case. Since the distribution of pre-test scores is assumed to be symmetric and strictly quasi-concave, the peak of the distribution must be at the median. To see that  $x^*$  must be above the median, suppose that  $x^*$  were less than the median. Denote the distance between  $x^*$  and the median as  $D$ . Now consider an alternative  $x^*$ , denoted  $x^{*'}$ , equal to the median plus  $D$ . By symmetry of the distribution, the total number of students at any distance from  $x^{*'}$  equals the total number of students at any distance from  $x^*$ . Furthermore, the distribution of  $h(x_i - x^*)$  is identical to the distribution of  $h(x_i - x^{*'})$ . That, the distribution of the teachers' impact on students' scores is the same. However, the distribution of students within range  $\theta$  of  $x^{*'}$  first order stochastically dominates the distribution of students within a range  $\theta$  of  $x^*$ . Thus, by convexity of  $P(\cdot)$ , the teacher would be better off with the target teaching level  $x^{*'}$ , since she then improves the scores of higher-scoring students.

Since the distribution is symmetric, quasi-concave, if  $P(\cdot)$  is linear, maximizing teacher payoff means maximizing  $\int h(x^* - x_i)f(x_i)dx_i$ . In this case, the median maximizes teacher payoffs. This implies that if  $P(\cdot)$  is convex, the first order condition can only be satisfied for  $x^*$  greater than the median.

Arguments for the linear and concave case are analogous. ■

### *Proof of Proposition 3*

Consider first the case in which  $f(\cdot)$  is increasing in peer test scores. A uniform marginal increase in peer baseline achievement will lead to an increase in the focus teaching level.

Students with  $x > x^*$  and  $x < x^* + \theta$  will be closer to the target teaching level. They will thus benefit not only from the direct impact of higher-achieving peers but also from the indirect impact on teachers' choice of target instruction level. Students whose initial test scores were above  $x^* + \theta$  are still too far from the target level of instruction, but still benefit from the increase in test scores (note that in the case where the teacher reward is a convex function of student test scores, there may not be any student above  $x^* + \theta$ , as  $x^*$  may have been chosen to be within  $\theta$  of the top of the distribution).

Students with scores between  $x^* - \theta$  and  $x^*$  benefit from the higher achievement of their peers and from any increase in teacher effort associated with the higher peer achievement. On the other hand, these students now are further away from the new target teaching level. The overall effect is ambiguous.

Students with scores less than  $x^* - \theta$  were not in range of the teacher's instruction prior to the increase in test scores, and are not advantaged or disadvantaged by the change in the target teaching level. However, they benefit from the higher-achievement of their peers.

In the case where  $f(\cdot)$  is a constant (no direct peer effects), the proof follows from the discussion of the indirect effects. ■

### ***Proof of Proposition 5***

Consider first the case of convex payoffs. Suppose that  $D_U = D_L$ . In that case, both the lower track teacher and the upper track teacher would have the same number of students within any distance, by the symmetry of the original distribution.

The first order necessary condition for an optimum is that marginally increasing  $D$  reduces the contribution to the  $P$  function from students to the left of  $x^*$  by the same amount it increases

the contribution to the  $P$  function from students to the right of  $x^*$ . This necessary condition cannot be satisfied simultaneously for both the low achievement class and high achievement class if the target teaching levels in each class are symmetric around the median. To see this note that, by symmetry, increasing  $D$  will have the same effect on the total number of students within distance  $\theta$  of  $x^*$  for both sections. That is, the distribution of  $h(x_i - x^*)$  will be affected in exactly the same way. However, increasing  $D$  in the lower track will move  $x^*$  away from the higher-scoring students of the class, while it moves  $x^*$  closer to the higher-scoring students in the upper track. This implies that if the third derivative is non-negative (i.e. the degree of convexity is non-decreasing), there are more gains from increasing  $D$  in the higher track than in the lower track. Hence, if  $x^*_L$  is chosen optimally such that the gains from increasing  $x^*_L$  equal the losses from doing so, it will always be the case that at a symmetric  $x^*_U$  (i.e. such that  $D_U = D_L$ ), increasing  $x^*_U$  will increase teacher payoffs.

So far we have shown that choosing  $D_U = D_L$  cannot be optimal. To complete the proof for the convex case, note that choosing  $D_U < D_L$  will not be optimal either, since this moving to  $D_U = D_L$  will always increase payoff of the teacher in the upper track..

Arguments are analogous for the linear and concave cases. Under linearity, the median student will be equidistant from the target teaching level in the lower and upper section. Under concavity, they will be closer in the top section. ■

### ***Proof of Proposition 6***

Consider the convex case. When choosing the optimal effort level, the teacher equalizes the marginal benefit of effort,  $g'(e)$  times  $P'(y) \equiv \int P'(y_i) f(x_i) dx_i$  to the marginal cost of effort,  $c'(e)$ . The argument of this proof is as follows: Since  $P'(y)$  for the teacher in the upper section

is higher than  $P'(y)$  for the teacher in the lower section, it must be that the teacher in the upper section exerts more effort since  $g(\cdot)$  is concave, while  $c(\cdot)$  is strictly convex.

It remains to be shown that  $P'(y)$  for the teacher in the upper section is higher than  $P'(y)$  for the teacher in the lower section. Recall that by Proposition 5, under a convex payoff function, we always have  $D_U > D_L$ , so the teacher in the upper section chooses the target level of instruction to be further away from the median student than the teacher in the lower section.

Note that by convexity of  $P(\cdot)$ , if the overall payoff  $P(y)$  in the upper section exceeds overall payoff in the lower section, then the marginal payoff  $P'(y)$  is also higher in the upper section than in the lower section. Furthermore, even if the upper track teacher chose  $D_U = D_L$ , the average payoff would be higher in the upper section. Since the teacher in the upper section maximizes  $P(y)$  by choosing  $D_U > D_L$  (while having the option of  $D_U = D_L$ ) it must be that his payoff is even higher than at  $D_U = D_L$ . But this implies that  $P'(y)$  for the teacher in the upper section is higher than  $P'(y)$  for the teacher in the lower section.

The proofs for the linear and concave cases follow a similar logic.

The second result (that for high enough  $\lambda$ , the difference between effort levels of contract teachers assigned to the high- and low-achievement classes will become arbitrarily small) is due to the assumption that the cost of effort becomes arbitrarily high as a maximum effort level  $\bar{e}$  is approached. ■