Journal of Public Economics 111 (2014) 29–45

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Encouraging classroom peer interactions: Evidence from Chinese migrant schools☆

Tao Li [a,*], Li Han [b], Linxiu Zhang [c], Scott Rozelle [d]

[a] University of Macau, Faculty of Social Sciences, Av. Padre Tomas Pereira, Taipa, Macau
[b] Hong Kong University of Science and Technology, Hong Kong
[c] Chinese Academy of Sciences, China
[d] Stanford University, United States

ARTICLE INFO

Article history:
Received 27 December 2012
Received in revised form 20 December 2013
Accepted 27 December 2013
Available online 9 January 2014

ABSTRACT

In a randomized trial conducted with primary school students in China, we find that pairing high and low achieving classmates as benchmates and offering them group incentives for learning improved low achiever test scores by approximately 0.265 standard deviations without harming the high achievers. Offering only low achievers incentives for learning in a separate trial had no effect. Pure peer effects at the benchmate level are not sufficiently powerful to explain the differences between these two results. We interpret our evidence as suggesting that group incentives can increase the effectiveness of peer effects.

1. Introduction

The educational outcomes of low achieving students may improve if they are moved from a low achieving peer group to a higher achieving one through school integration policies such as detracking, school busing or housing vouchers (see the reviews of the peer effects literature by Epple and Romano, 2011; Sacerdote, 2011). Alternatively, their educational outcomes may improve if they experience more positive interactions with higher achieving peers in their *current* peer group — a hypothesis that has so far received little attention.

There are several reasons why we want to study this hypothesis. First, stimulating positive interactions between students from different backgrounds is arguably one of the ultimate goals of school integration policies. Mixing together a diversity of students within schools and classrooms is only one means for facilitating this purpose. Furthermore, altering peer groups in this way is both expensive and time-consuming, and yet there is no guarantee that it will automatically produce the desired peer interactions for low achieving students, who usually come from disadvantaged social backgrounds. Carrell et al. (2011) demonstrated a case in which randomly assigned higher achieving peers failed to benefit low achievers, and the most plausible explanation appeared to be the lack of interactions between these two groups. The lack of interactions between white and black students in officially integrated schools is also well documented (e.g., Echenique and Fryer, 2007). Finally, even if school integration policies successfully induce high and low achieving students to interact with each other, enhancing the quality of these interactions is still beneficial.

The way classrooms are typically managed in China offers us a convenient opportunity to study the above hypothesis. Traditional pair and row seating is the predominant classroom layout in China. Benchmate pairs typically sit next to each other throughout a semester and frequently interact with each other on a daily basis. By strategically reshuffling benchmates, student-level peer interactions could be influenced without the need to alter school or classroom composition.

Our *peer incentive experiment*, the focus of this paper, was designed to estimate the effects of enhancing benchmate-level peer interactions between high and low achieving classmates on the academic performance of low achievers. The experiment was implemented in 44 classes from 11 migrant primary schools in Beijing. Based on baseline test scores, we randomly assigned half of the bottom twenty students to the treatment group, and the other half to the control group. The treatment included an opportunity component as well as an incentive component. The opportunity component was that each treated student was randomly assigned to one of the top ten performing classmates as a benchmate for a semester. The incentive component was that the top three benchmate pairs (that is the three benchmate pairs in which the treated students made the largest test-score gains over a semester) in each class would get a monetary reward. The purpose of such a tournament-based group incentive was to encourage benchmates to interact with each other in a way that would contribute to the weaker partner's academic performance.

By comparing the treatment with the control students in the same classes (i.e., a within-class evaluation design), we found a robust effect

☆ We would like to thank Weili Ding, Esther Duflo, Gigi Foster, Caroline Hoxby, Michael Kremer, Victor Lavy, Steven Lehrer, Patrick McEwan, Albert Park, Sujata Visaria and Hongliang Zhang, two anonymous reviewers and the seminar participants at Harvard, Stanford, Peking University, NEUDC, Oxford and Chinese University of Hong Kong for their helpful comments on the drafts of this paper. Financial support from Ford Foundation and the University of Macau (Grant SRG009-FSH13-LT) is gratefully acknowledged.

* Corresponding author.
E-mail address: litao@post.harvard.edu (T. Li).

0047-2727/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jpubeco.2013.12.014

of approximately 0.265 standard deviations (s.d.) in the low achievers' evaluation test scores using various estimation strategies. This finding is significant because it clearly demonstrates that policymakers can make peer effects more effective than they would otherwise be without the time or expense associated with manipulating classroom or school composition.

Behind any economic intervention that employs a group incentive scheme there are many potential mechanisms that could be at play. In our peer incentive experiment, there are at least three potential mechanisms to consider: treated low achievers improved their scores (1) because of peer interactions stimulated by group incentives (2) because of their own desire to win rewards from their scores, or (3) simply because of having better opportunities to interact with a high-achieving classmate. Unpacking our main effects has both theoretical and practical implications.

For this purpose, we additionally ran a separate *individual incentive experiment* that studied the effects of offering low achievers exactly the same level of incentives for improving their scores (but did not match them up with a high-achieving peer) in 47 classes from 12 different migrant primary schools in Beijing. We also followed the peer effects literature (e.g., Sacerdote, 2001; Zimmerman, 2003) and estimated conventional, reduced-form *pure benchmate effects* (i.e., pure peer effects at the benchmate level) by exploiting two exogenous changes to benchmate composition in the peer incentive classes. Neither the individual incentive nor the pure benchmate effects are statistically distinguishable from zero. The evidence supports a straightforward interpretation of our primary finding: in our peer incentive experiment it is the group incentive, rather than either of the two alternative mechanisms, that made the peer effects more effective than they would have been otherwise.

At the class level, we randomly assigned 35 extra classes to be control classes. By comparing students from the experimental classes to their counterparts in the control classes with similar baseline test scores (i.e., an across-class evaluation design), we found a small and statistically insignificant spillover effect for the untreated students in the experimental classes, including the high achievers in the peer incentive classes. Our results suggest that encouraging peer interactions inside a given peer group may be a less controversial way to make use of peer effects because it brings about efficiency gains.

We nevertheless acknowledge that this paper has several important limitations. Because of the small number of schools involved in our study, we cannot discuss the effects that result from an entire school being treated. One could imagine that the culture of the school could change in a general way. A larger study involving school-level treatment would improve the external validity of our research. Another line of future research would be evaluating the long-term effects of our peer incentive treatment. The effects reported here were short term (one semester only); the long-term effects, if any, are unclear.

The rest of the paper is organized as follows. Section 2 reviews the literature. Section 3 presents a conceptual framework. Section 4 describes our programs and data. Section 5 describes the evaluation design and reports results from the peer and individual incentive experiments. Section 6 reports the estimation strategies and the results of the pure benchmate effects. Section 7 concludes the paper. Details of program implementation and some extra robustness checks are in the Appendix.

2. Related literatures

To the best of our knowledge, benchmate pairs are the smallest set of peer groups that the classroom peer effects literature has ever studied. Benchmate interactions are entirely voluntary. There are no assigned tutoring sessions. Teachers are not involved in the daily interaction process. These two characteristics make benchmate interactions fundamentally different from cooperative learning intensively studied by educational psychologists (Johnson and Johnson, 1997) or group studying and peer tutoring studied by other economists (Angrist et al., 2009; Blimpo, 2010).

Nearly all the previous empirical literature on educational peer effects focuses almost exclusively on the task of establishing whether peer effects exist by exploiting exogenous changes in peer group composition.[1] The reduced-form peer effects estimated in this way cannot be used to distinguish among externalities from different channels (Manski, 1993). With objectives similar to the work in this paper, several authors have recently tried to estimate peer effects emanating from different student behavior, such as student efforts (Cooley, 2009), the choice of college major (Giorgi et al., 2009) and classroom infractions (Kinsler, 2010). None of these papers, however, have explicitly studied peer interactions. To our knowledge only a few papers have attempted to do so. Relying on surveys and administrative data, Stinebrickner and Stinebrickner (2006, 2008) found that college roommate peer effects are most likely to arise through roommates influencing each other's time-use rather than through their interacting on academic matters. The paper by Carrell et al. (2011) is in spirit closer to ours. They found that high and low achieving peers may be reluctant to interact in schools, which might contribute to the poor academic performance of low achievers. However, unlike our experiment, their study was not designed to provide causal evidence of the effect of peer interactions on educational outcomes.

In another literature it has been well established that cash incentives are effective in stimulating peer interactions in workplaces (Hamilton et al., 2003; Boning et al., 2007; Chan et al., 2010). As far as we know, however, there is no parallel study on peer interactions in schools, except for two papers by Babcock et al. (2010), Babcock and Hartman (2011) that we will discuss below. This lack is a bit surprising because peer effects are considered to be a central input into the education production process (Epple and Romano, 2011). The absence of the use of cash incentives to encourage peer interactions cannot be explained by a lack of interest in using cash incentives in education. The use of cash incentives to solicit other types of socially desirable behavior in education has flourished in recent years, such as conditional cash transfer programs surveyed by Rawlings and Rubio (2005), teacher merit pay programs surveyed by Podgursky and Springer (2007), and randomized trials encouraging college student workout behaviors by Babcock et al. (2010) and Babcock and Hartman (2011). The latter two papers are similar in spirit to ours in that they examined the effects of cash incentives on randomly-assigned or self-selected student peer groups. Importantly, the two Babcock studies did not examine academic outcomes.

The segment of the literature that examines the use of cash incentives in school that is probably most relevant to our study is the set of studies that examine pay-for-grades programs (which we call individual incentive experiments in this paper). Over the past decade, a large number of such programs have been implemented around the world. Despite much enthusiasm, the estimated program effects on actual learning are still mixed (for a review, see Slavin, 2010). Evidence of the effect of pay-for-grades programs on secondary school students, usually those preparing for important high school exit exams, tend to be positive and significant (Mauldon et al., 2000; Spencer et al., 2005; Angrist and Lavy, 2009; Jackson, 2010). In contrast, evidence for primary school students is noisier (Kremer et al., 2009; Bettinger, 2012; Fryer, 2011). After analyzing the effects from the largest pay-for-grades experiment conducted in 261 American public schools, Fryer (2011) suggested that individual incentives tied to student test scores were not effective because students did not know how to improve learning on their own. Fryer's conclusion underscores the need to compare pay-for-grades programs with programs that not only pay for grades but

[1] An incomplete list of recent contributions include the following: Foster (2006), Ding and Lehrer (2007), Figlio (2007), Lyle (2007), Carrell et al. (2009), Carrell and Hoekstra (2010), Burke and Sass (2011), Gibbons and Telhaj (2011), Lavy and Schlosser (2011), Imberman et al. (2012), Lavy et al. (2012), etc.