Heterogeneous Peer Effects and Rank Concerns: Theory and Evidence

Michela M. Tincani[1]

June 29, 2015

[1]Tincani: University College London, 30 Gordon Street, London, WC1H0BE, UK, m.tincani@ucl.ac.uk. I am particularly grateful to Orazio Attanasio, Ed Hopkins, Aureo de Paula, Petra Todd and Ken Wolpin for their feedback and encouragement. I would also like to thank Alberto Bisin, Denis Chetverikov, Martin Cripps, Mariacristina De Nardi, Steven Durlauf, Jan Eeckhout, David Figlio, Tatiana Kornienko, Dennis Kristenses, Lars Nesheim, Imran Rasul, Nikita Roketskiy, Bryony Reich, Martin Weidner, Daniel Wilhelm and participants at various seminars and conferences for insightful comments and discussions. I thank Julia Schmieder for excellent research assistance, and Rodrigo Astroza for helping me understand how earthquake damage propagates. Research funding from the Centre for Microdata Methods and Practice and from the European Research Council's grant number IHKDC-249612 is gratefully acknowledged. I am grateful to the Chilean Ministry of Education and to the *Agencia de Calidad de la Educación* for access to some of the data used in this research. The views reported here are those of the author and are not necessarily reflective of views at the Ministry and at the *Agencia*

Abstract

Much evidence exists of heterogeneous and non-linear ability peer effects in test scores. However, little is known about the mechanisms that generate them and whether this evidence can be used to improve the organization of classrooms. This paper is the first to study student rank concerns as a mechanism behind ability peer effects. First, it uses a theoretical model where students care about their achievement relative to that of their peers to derive predictions on the shape of peer effects. Second, it proposes a new method to identify heterogeneous and non-linear peer effects. Third, it tests the theoretical predictions in a novel empirical setting that uses the Chilean 2010 earthquake as a natural experiment. The results indicate that rank concerns generate peer effects among Chilean 8^{th} graders. An important implication is that educators can exploit the incentives generated by academic competition when choosing classroom assignment rules.

KEYWORDS: Peer Effects in Education, Model of Social Interactions, Rank Concerns, Testing Theoretical Predictions, Natural Experiment.

1 Introduction

Peers can have very important effects on the development of one's human capital, and the study of peer effects is a cornerstone in the Economics of Education literature. One of the most important goals of peer effect research in education, dating at least to the Coleman report (Coleman 1966), is to design classroom allocation rules that improve student outcomes. As a first step towards this goal, most of the existing research has been concerned with identifying and quantifying peer effects.

Estimates from linear-in-means models of the impact of average peer ability on student outcomes vary greatly, and many studies find no effects (Angrist 2014). However, several studies that relax the assumptions of the linear-in-means model, for example, by allowing higher moments of the peer ability distribution to matter or by allowing for heterogeneous impacts, find larger and significant peer effects in the classroom (Sacerdote 2014).[1] While this suggests that policies that regroup students across classrooms may generate social gains, we still lack an understanding of the mechanisms behind peer effects (Sacerdote 2011, Epple and Romano 2011). As highlighted by the work in Carrell, Sacerdote, and West (2013), this limits our ability to use peer effect estimates to improve the organization of classrooms.

This paper proposes and tests a new mechanism of social interactions in the classroom that can help us understand some of the existing evidence, as well as serve as a framework for future research on peer regrouping policies. First, it develops a theoretical model that has implications for the shape of peer effects. Second, it proposes a new method to identify heterogeneous and non-linear peer effects. Third, it tests the theoretical predictions in a new empirical setting that uses the Chilean 2010 earthquake as a natural experiment. In doing so, it empirically distinguishes the proposed mechanism from alternative ones.

In the model I propose, peer effects arise because of standard technological spillovers operating through the mean of peer ability, and because students have rank concerns. While students in various countries have been found to care about their rank

[1] For example, in a widely cited work using data from North Carolina, Hoxby and Weingarth (2005) calculate that increasing by ten percentage points the fraction of low-achievers in a classroom increases low-achievers' test scores by 18.5 percent of a standard deviation, while increasing by the same amount the fraction of high-achievers increases high-achievers' test scores by a staggering 40 percent of a standard deviation. For comparison, increasing teacher quality by one standard deviation or reducing class size by ten students increases student test scores by 10 percent of a standard deviation (Rivkin, Hanushek, and Kain 2005).

(Tran and Zeckhauser 2012, Azmat and Iriberri 2010), this is the first paper to study rank concerns as a mechanism underlying peer effects. In the model, achievement is produced through costly effort. Students are heterogeneous in terms of ability to study, which reduces the cost of exerting effort. Intuitively, how much effort students exert to improve their rank depends on the ability composition of their peers. The model is based on the theory of conspicuous consumption in Hopkins and Kornienko (2004). The main model prediction is that making peer ability more dispersed benefits some students and harms others depending on their ability and on the type of rank preference. The model has also the implication that achievement is monotone decreasing in cost of effort.

To test the main model prediction, an exogenous change in peer ability variance is needed. The key identification problem with observational data is that classrooms with different student compositions are different also in other unobserved ways that make it impossible to separate peer effects from other confounding effects (Manski 1993).[2] The key identification idea is that, as I document, the impacts of the Chilean 2010 earthquake were heterogeneous even across students in the same classroom, because the earthquake affected households differently depending on their distance from the rupture. To the extent that damage to a student's home affects a student's ability to study, classrooms that have different distributions of damages have also different distributions of students' ability to study. Therefore, the effect of the variance of ability to study on achievement can be obtained by comparing classrooms that are identical except for the variance in damages. Importantly, I show that these comparisons are free of the typical confounding factors that would arise if we used variation in any other determinant of a student's ability to study. For example, while in classrooms with different variances in initial achievement teachers may teach differently (Duflo, Dupas, and Kremer 2011), thus confounding the peer effect estimates, I do not find evidence that teachers adapt their teaching to the variance

[2]To overcome the problem of correlated effects, some studies have used data with random allocation of students to dorms. See, for example, Sacerdote (2001), Zimmerman (2003), Stinebrickner and Stinebrickner (2006), Kremer and Levy (2008), and Garlick (2014). In contrast, as noted in the survey by Epple and Romano (2011), very few experiments with random or quasi-random allocation to classrooms exist (Duflo, Dupas, and Kremer 2011, Whitmore 2005, Kang 2007). Notice that to test the model's implications it is sufficient to identify contextual/exogenous effects in the terminology of Manski (1993), i.e. the effect on own outcomes of pre-determined classmate characteristics (ability to study). In particular, it is not necessary to identify endogenous peer effects (the effect of peer outcomes on own outcomes), which generate a simultaneity problem known as the reflection problem.

in damages. Natural disasters have been used before to identify peer effects.[3] This study differs from previous work because it uses a continuous rather than a binary measure of exposure.

In terms of data construction, I use results from the structural engineering literature to build a measure of damage to each student's home caused by the 2010 Chilean earthquake. The measure is based on seismic intensity according to the Medvedev-Sponheuer-Karnik scale. I then merge this dataset with four waves of a large administrative dataset with information on students, teachers, classrooms and schools (*Sistema de Medición de la Calidad de la Educación*, SIMCE 2005, 2007, 2009, 2011). The resulting dataset contains two cohorts, observed before and after the earthquake, of $110,822$ students divided into $3,712$ classrooms.

I build an econometric model that exploits the natural experiment to estimate the heterogeneous impact on achievement of the variance in ability to study. For this purpose, I combine a semiparametric single-index model with a kernel-weighted difference-in-differences estimator. This model has several desirable features. The semiparametric approach imposes minimal assumptions on the technology of test score production. This allows me to test the main model's prediction (by detecting any pattern of heterogeneity in the peer effects across students), as well as to test the additional model prediction of monotonicity of the production technology. Moreover, the difference-in-differences approach accounts for an artifact introduced by the natural experiment; the variance in damages is determined by the geographic dispersion of the students in the class, which could be correlated with unobserved student and/or classroom characteristics that could confound the peer effect estimates. For this reason, I use the pre-earthquake cohort of students (2005-2009), who were not affected by the earthquake, to estimate and difference out any potentially confounding correlation between geographic dispersion and outcomes.

As a preliminary data analysis, I provide the first evaluation of the impact of the Chilean earthquake on student test scores. Using difference-in-differences value-added test score regression models, I estimate that being exposed to the earthquake reduced test scores by 0.05 standard deviations (sd) (p-value<0.001). Moreover, every USD 100 in earthquake damages caused a reduction of 0.016 sd in test scores (p-value<0.001).

[3] In the educational peer effects literature see, for example, Cipollone and Rosolia (2007), Imberman, Kugler, and Sacerdote (2012), and Sacerdote (2008).

The main empirical finding is that increasing the variance of peer ability to study benefits low-ability students, harms middle-ability students, and it harms high-ability students in Spanish classes while it benefits those in Mathematics classes. While these rich empirical patterns are hard to rationalize with standard models of peer effects, the parsimonious theoretical model can explain them in a simple and intuitive way. Low-ability students exert more effort in classrooms with larger ability variance because there are more students close to their ability level, therefore, surpassing the student next up in the ability distribution is less costly. As a consequence, middle-ability students face stronger competition from below and this gives them an incentive to exert more effort. However, they also have an incentive to exert less effort, because in classrooms with larger ability variance there are fewer students close to their ability level, therefore, surpassing the student next up in the ability distribution is more costly. The model predicts that the incentive to decrease effort prevails for middle-ability students, yielding lower test scores as observed in the data. Also high-ability students face two opposing incentives, and the model predicts that test scores increase in Mathematics and decrease in Spanish, as observed in the data, whenever the rank preference is stronger in Mathematics than in Spanish. Statistical tests do not reject any of the theoretical model's implications.

The implications for the estimation of peer effects are far-reaching. When there are rank concerns, peer effects operate through the entire distribution of ability. Commonly used empirical models that focus on the mean of peer ability may fail in out of sample predictions, like in Carrell, Sacerdote, and West (2013), and they may fail to detect peer effects when these are present.[4] This is especially relevant when peer effects are assumed to imply clustering of outcomes around the mean, an assumption often made in variance contrast methods (Glaeser, Sacerdote, and Scheinkman 1996, Graham 2008). Rank concerns generate peer effects without necessarily implying outcome clustering.

There are also important implications for ability tracking, the most important peer regrouping policy. A common concern is that assigning students of similar ability to

[4] In a related paper, Tincani (2014), I show that all the (puzzling) results of the peer regrouping experiment in Carrell, Sacerdote, and West (2013) can be rationalized by the model presented in this paper. The model can also explain similar results from the recent peer regrouping experiment in Booij, Leuven, and Oosterbeek (2014). To the best of my knowledge, these are the only two studies with peer regrouping experiments at the classroom/study group level, and that generate peer effects that are believed to be due to peer-to-peer interactions rather than peer-to-teacher interactions.

the same classroom harms low-ability students, who are tracked with other low-ability students, unless teachers teach more productively in more homogeneous classrooms. This concern is founded when only the mean of peer ability matters for peer-to-peer interactions, which is not the case if students have rank concerns.[5] In fact, under the type of rank concerns for which I find evidence in the data, students in tracked classrooms have a stronger incentive to exert effort. Intuitively, competing is easier amongst equals. Therefore, this paper uncovers a new channel of operation of tracking policies. The practical implication of this result is that educators can exploit the student incentives generated by classroom assignment policies to motivate students. It would be helpful for future research to collect measures of student rank concerns and use them to investigate the optimal organization of classrooms.[6]

The rest of the paper is organized as follows. Section 2 reviews the most relevant literature, section 3 presents the theoretical model, and it is followed by section 4 that describes the data and the context in which the theory is tested. Section 5 introduces the main empirical framework, and section 6 presents the estimation results and links them to the theoretical predictions. Robustness and alternative mechanisms are discussed in section 7, and section 8 concludes.

2 Literature Review

This is the first paper to study and find evidence of rank concerns as a mechanism underlying peer effects.[7] Relatively few studies explore the mechanisms behind peer effects. Lavy and Schlosser (2011) and Lavy, Paserman, and Schlosser (2012)

[5] Duflo, Dupas, and Kremer (2011) run an experiment where Kenyan first-graders are randomly allocated to tracked and non-tracked classrooms and find beneficial impacts of tracking on students of all ability levels. They attribute those positive impacts to teachers and find supporting evidence for this. While rank concerns are unlikely to be driving their results, the linear-in-means model of peer effects that they adopt rules out *ex ante* that peer-to-peer interactions can generate positive effects in the low-ability track.

[6] Future research could further analyze how the incentives generated by peer composition interact with those provided through the grading system or through merit fellowships and financial awards. See Dubey and Geanakoplos (2010) for a theoretical analysis of the grading system incentives. A large number of studies analyze students' response to merit and financial incentives (Angrist and Lavy 2009, Kremer, Miguel, and Thornton 2009, Fryer 2010, Levitt, List, and Sadoff 2011, Cotton, Hickman, and Price 2014).

[7] The idea that rank concerns could generate peer effects dates back to at least Jencks and Mayer (1990) who, however, do not explore it. Related to this idea are the works of Murphy and Weinhardt (2014) and Elsner and Isphording (2015), who empirically analyze the importance of past class rank for future performance in school.

use teacher and student surveys to understand how gender variation and proportion of low-ability students impact class outcomes. Using a different approach, Blume, Brock, Durlauf, and Jayaraman (2014) and Fruehwirth (2013) provide microfoundations to the widely used linear-in-means peer effect specification, proving that it can be rationalized by a desire to conform. De Giorgi and Pellizzari (2013) develop and test behavioral models that can rationalize observed outcome clustering within classrooms at Bocconi University.[8] This paper uses a different and novel approach: it first presents a plausible mechanism of interactions in the classroom, and then derives testable implications for the shape of the peer effects. By testing the model and ruling out alternative mechanisms, this is one of the first papers to investigate how classroom composition affects student incentives.[9] As emphasized in the survey in Epple and Romano (2011), modeling "the way in which students, teachers, and principals are affected by the incentives created by differing administrative assignments of students to peer groups" is a necessary first step to identify the optimal design of classrooms.

This paper focuses on the dispersion of ability in the classroom. Previous research has found that ability dispersion plays an important role in determining student achievement. See, for example, Carrell, Sacerdote, and West (2013), Booij, Leuven, and Oosterbeek (2014), Lyle (2009), Duflo, Dupas, and Kremer (2011), Ding and Lehrer (2007), Hoxby and Weingarth (2005), Vigdor and Nechyba (2007).[10] Interestingly, linear-in-means models appear better suited to capture peer effects in social behaviors such as crime and smoking (Sacerdote 2014). Through the lens of this paper's finding, this can indicate that a desire to conform might be a more plausible explanation for this kind of social behaviors than for test scores.

By uncovering a new possible channel of operation of tracking policies, this paper is related to the literature on ability tracking. With the exception of Garlick (2014), who studies tracking in university dorms, all studies of ability tracking that use ran-

[8] See also Calvó-Armengol, Patacchini, and Zenou (2009), who provide microfoundations to the Katz-Bonacich centrality measure in a network.

[9] To the best of my knowledge, only Todd and Wolpin (2014) analyze how student incentives are affected by the ability composition of one's peers without requiring that the resulting reduced-form peer-effect specification be linear-in-means. Fu and Mehta (2015) analyze how parental effort decisions are affected by tracking, and assume that peer spill-overs are of the linear-in-means type. The structural models in these papers, however, are not tested, rather, they are estimated.

[10] See also Bénabou (1996) for a theoretical analysis of the role of the variance of the peer ability distribution, and Lavy, Silva, and Weinhardt (2012) for a recent analysis of nonlinear peer effects.

domized experiments find that they are beneficial to students. While Duflo, Dupas, and Kremer (2011) attribute this positive impact to teachers in Kenyan schools, beneficial impacts of tracking due to peer-to-peer interactions have been found among low- and middle-ability students at the University of Amsterdam (Booij, Leuven, and Oosterbeek 2014), and among middle-ability students at the U.S. Air Force Academy, who were the only ones to be effectively tracked by the authors' intervention (Carrell, Sacerdote, and West 2013).[11]

3 A Theoretical Model of Social Interactions

I propose a simple theory of social interactions in the classroom and derive implications that can be tested empirically. The main implication is a comparative statics result on the effect of changing the ability variance in the classroom while keeping the ability range constant, which is the type of data variation generated by the natural experiment. Tracking changes classroom ability variance by reducing the classroom ability range instead. An advantage of my research approach is that I can investigate tracking even though I do not observe data variation akin to tracking. To do so, I first identify what type of rank preferences my data are compatible with. I then use the theoretical model under those preferences to infer how student incentives would be affected by tracking.

The model is an application of the theory of conspicuous consumption in Hopkins and Kornienko (2004), where individuals choose how much of their income to spend on a consumption good and how much on a positional good. Here, achievement is at the same time a consumption good and a positional good, and it can be produced at a cost. Specifically, students in a classroom choose how much costly effort e to exert, and effort increases achievement/test score y. Students are heterogeneous in terms of ability, i.e., how costly it is for them to exert effort.[12] The main model assumptions are the following:

A.1 Students' utility is increasing in own achievement.

[11] Using non-experimental approaches, Betts and Shkolnik (2000) and Lefgren (2004) find little evidence of benefits from tacking, while Lavy, Paserman, and Schlosser (2012) find that high-ability students benefit from other high-ability students and do not help average students.

[12] This corresponds to income heterogeneity in Hopkins and Kornienko (2004). Alternatively, students can be assumed to be heterogeneous in terms of how productive their effort is, and, under minor modifications to the assumptions on the utility function, the model would have the same implications.

A.2 There are technological spill-overs in the production of achievement, i.e., mean peer ability directly affects own achievement.

A.3 Students' utility is increasing in rank in terms of achievement.

Assumptions A.1 and A.2 are standard.[13] Assumption A.2 gives rise to exogenous peer effects (Manski 1993). Assumption A.3 is novel in the theoretical literature on educational peer effects. It introduces a competitive motive and it gives rise to endogenous peer effects (Manski 1993), because how much effort each student exerts is determined endogenously by the equilibrium of a game of status between students.

Students differ in terms of a type c: those with a higher c incur a larger cost of effort. Type c captures any student characteristics, physical or psychological, that affect her ability to study, such as cognitive skills, access to a computer or books, availability of an appropriate space for studying, parental help, etc. Notice that ability *reduces* c. Type c is distributed in the classroom according to c.d.f. $G(\cdot)$ on $[\underline{c}, \bar{c}]$. Each student's type c is private information, but the distribution of c in the classroom is common knowledge.

An appeal of the model is that it does not make distributional assumptions and, whenever feasible, functional form assumptions. However, some plausible shape restrictions are imposed to prove the results. The cost of effort is determined by an increasing and strictly quasi-convex function in effort $q(e; c)$. Higher types c incur higher costs for every level of effort e, i.e. $\frac{\partial q(e;c)}{\partial c} > 0$ for all e. Moreover, at higher types the marginal cost of effort is (weakly) higher: $\frac{\partial^2 q(e;c)}{\partial c \partial e} \geq 0$.

Effort determines achievement according to the production function $y(e) = a(\mu)e + u(\mu)$, where μ is the classroom mean of c. Parameters $a(\mu)$ and $u(\mu)$ capture technological spill-overs working through the mean of peer ability (assumption A.2). These can be indirect, i.e., working through the productivity of classroom specific factors, or direct, i.e. due to peer-to-peer contacts. An example of an indirect spill-over is teacher productivity depending on students' average abilities. An example of a direct spill-over is more able peers (lower μ) asking relevant questions in class and, in so doing, facilitating their classmates' learning. Notice that the model is flexible in that it allows these technological spill-overs to affect both the level of achievement (u) and the productivity of effort (a).

[13]For example, Blume, Brock, Durlauf, and Jayaraman (2014), Fruehwirth (2012), and De Giorgi and Pellizzari (2013) assume that student's utility is increasing in own achievement. Several papers model technological spill-overs as operating through mean peer characteristics, e.g. Arnott and Rowse (1987), Epple and Romano (1998), Epple and Romano (2008).

The utility function can be decomposed into two elements: a utility that depends only on own test score y and effort cost q, $V(y,q)$, embedding assumption A.1; and a utility that depends on rank in terms of achievement, embedding assumption A.3. The utility from achievement is non-negative, increasing and linear in achievement, decreasing and linear in q, and it admits an interaction between utility from achievement and cost of effort such that at higher costs, the marginal utility from achievement is (weakly) lower ($V_{12} \leq 0$).[14] No specific functional form assumptions are made on $q(\cdot)$ and on the interaction between y and q, therefore, results from the model are valid under a broad class of preferences. For example, more able peers (lower c) may (or may not) have higher marginal utilities from achievement.

The student's classroom rank in terms of achievement is given by the c.d.f. of achievement computed at a student's own achievement level, $F_Y(y)$. This is the fraction of students with achievement lower than one's own, and it is a standard way to model rank in theoretical models of status seeking (Frank 1985). Because achievement is an increasing and deterministic function of effort, rank in achievement is equal to rank in effort: $F_Y(y) = F_E(e)$. The utility from rank, $S(F_Y(y))$, is equal to $F_E(e) + \phi$, where ϕ is a non-negative constant.

Overall utility $U(y,q;c)$ is the product of utility from own achievement $V(y,q;c)$ and utility from rank $S(F_Y(y))$: $V(y,q;c)\,(F_E(e)+\phi)$. The parameter ϕ determines the type of rank concerns that students have. When $\phi > 0$, students have a minimum guaranteed level of utility even if they rank last ($F_E(e) = 0$). On the other hand, when $\phi = 0$ ranking low has dire consequences, and this will generate that students close to the bottom of the ability distribution will be desperate to avoid a low-rank. Because the two cases ($\phi = 0$ and $\phi > 0$) yield different testable implications, to the extent that I do not reject the model, I can identify which type of rank concerns is compatible with the data.

Each student chooses effort to maximize overall utility. Focusing on symmetric Nash equilibria in pure strategies, and assuming that the equilibrium strategy $e(c)$ is strictly decreasing and differentiable with inverse function $c(e)$, rank in equilibrium can be rewritten as $1 - G(c(e_i))$, and i's utility as $V(y(e_i), q(e_i, c_i))(1 - G(c(e_i)))$.[15] The

[14] All results are valid under an alternative set of assumptions for the utility from achievement. These are: strictly quasi-concave utility of achievement, decreasing and linear utility from cost of effort ($V_2 < 0, V_{22} = 0$) with a linear cost function ($\frac{d^2q}{d^2e} = 0$) and additive separability between utility from achievement and cost of effort ($V_{12} = 0$).

[15] The probability that a student i of type c_i with effort choice $e_i = e(c_i)$ chooses a higher effort

first-order condition then is:

$$\underbrace{V_1 \overbrace{a(\mu)}^{\text{Mg. increase in achiev.}}}_{\text{mg. ut. from increased achiev.}} + \underbrace{\frac{V(y,q)}{1-G(c(e_i))+\phi}\overbrace{g(c(e_i))(-c^{'}(e_i))}^{\text{Mg. increase in rank}}}_{\text{mg. ut. from increased rank}} = \underbrace{-V_2\frac{\partial q}{\partial e}}_{\text{mg. cost}} \tag{1}$$

and it implies the first-order differential equation reported in equation 7 in Online Appendix A.1. The solution to this differential equation is a function $e(c)$ that is a symmetric equilibrium of the game. The assumptions on the utility function, on the cost of effort function and on the achievement production function guarantee that the results in Hopkins and Kornienko (2004) apply under appropriate proof adaptations.[16] In particular, while the differential equation does not have an explicit solution, existence and uniqueness of its solution and comparative statics results concerning the equilibrium strategies can be proved for any distribution function $G(c)$ twice continuously differentiable and with a strictly positive density on some interval $[\underline{c}, \bar{c}]$, with $\underline{c} \geq 0$. The first theoretical result is summarized in the following Proposition:

Proposition 3.1 (Adapted from Proposition 1 in Hopkins and Kornienko (2004)). *The unique solution to the differential equation* (7) *with the boundary conditions* $e(\bar{c}) = \frac{1}{\bar{c}}$ *for* $\phi = 0$ *and* $e(\bar{c}) = e_{nr}(\bar{c})$ *for* $\phi > 0$*, where* e_{nr} *solves the first order condition in the absence of rank concerns (*$V_1 a(\mu)|_{e=e_{nr}} = -V_2\frac{\partial q}{\partial e}|_{e=e_{nr}}$*), is an (essentially) unique symmetric Nash equilibrium of the game of status. Equilibrium effort* $e(c)$ *is continuous and strictly decreasing in type* c*.*[17]

than another arbitrarily chosen individual j is $F(e_i) = Pr(e_i > e(c_j)) = Pr(e^{-1}(e_i) < c_j) = Pr(c(e_i) < c_j) = 1 - G(c(e_i))$, where $c(\cdot) = e^{-1}(\cdot)$. The function c maps e_i into the type c_i that chooses effort e_i under the equilibrium strategy. Strict monotonicity and differentiability of equilibrium $e(c)$ are initially assumed, and subsequently it is shown that equilibrium strategies must have these characteristics.

[16] To guarantee existence of an equilibrium when $\phi = 0$, one additional assumption must be made: each student has an upper bound on achievable test score, and students with higher ability (lower c) have a higher upper bound. For example, a student of type c can never achieve more than $\bar{y} = a(\mu)\frac{1}{c} + u(\mu)$. Under this technological constraint, no student of type c will exert more effort than $\frac{1}{c}$, because any unit of effort above this level does not increase achievement, but it is costly. Footnote 17 explains how this guarantees existence. One of the main differences with the model in Hopkins and Kornienko (2004) is that here equilibrium strategies $e(c)$ are decreasing in c, whereas there they are increasing. See the procurement auctions model in Hopkins and Kornienko (2007) for another example of decreasing strategies.

[17] The equilibrium is essentially unique, in the sense that the only source of multiplicity is at the point $\bar{c}$ when $\phi = 0$. In a symmetric equilibrium, the student with the highest cost, $\bar{c}$, has rank 0 when $\phi = 0$. Her equilibrium utility is 0, and the only way she can increase it is by increasing her

Proof: see Online Appendix A.1.

Notice that Proposition (3.1) rules out the case in which for large enough values of c students exert more effort. This would be akin to a backward-bending labor supply curve.[18] Because this result may appear restrictive, I empirically test it. The implication can be rephrased in terms of achievement, given that achievement is an increasing function of effort, to obtain the first testable implication:

Testable Implication 1: Achievement is decreasing in type c.

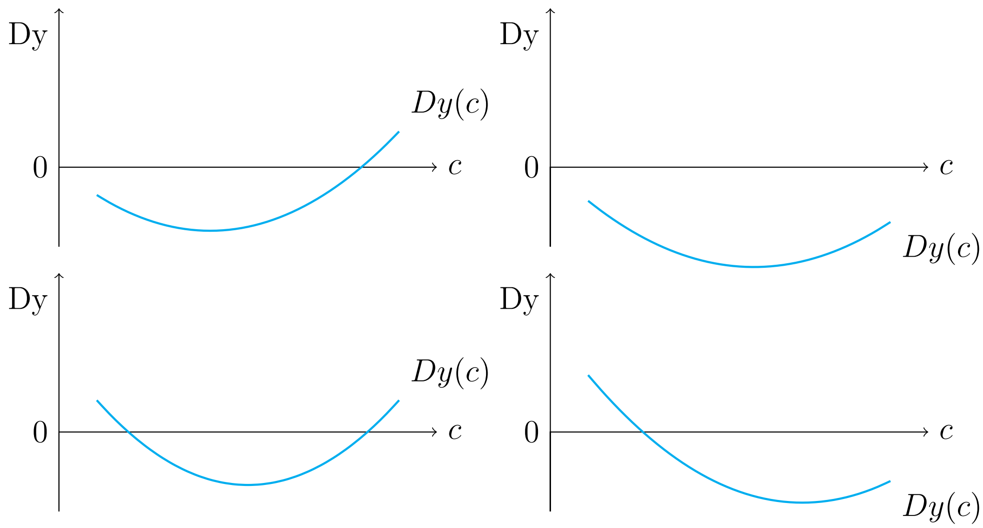
Now consider two distributions, $G_A(c)$ and $G_B(c)$, that are such that they have the same mean, and G_B has larger dispersion than G_A in the Unimodal Likelihood Ratio sense ($G_A \succ_{ULR} G_B$), defined in Online Appendix A.1. This happens when, for example, G_B is a mean-preserving spread of G_A. In informal terms, one can show that the effect of moving from G_A to G_B is heterogenous across individuals, depending on a student's rank in terms of c, and it depends on ϕ, i.e., on the type of rank concerns.[19] This result provides the second testable implication of the model:

Testable Implication 2: If students are averse to a low rank ($\phi = 0$), then when the dispersion of c increases, either all students perform more poorly, or all students except low-c (high-ability) students do more poorly. If students are not averse to a low rank ($\phi > 0$), then when the variance of c increases, middle-c students perform more poorly and high-c (low-ability) students perform better, while low-c (high-ability) students may perform better or worse. These patterns are represented graphically in Figure 1.

rank above 0. Therefore, for a strategy profile to be an equilibrium, it must be that the student with a slightly lower cost c exerts an amount of effort that is such that the least able student, $\bar{c}$, is unable to increase her rank by exerting more effort. Therefore, in equilibrium $\lim_{c \to \bar{c}^-} e(c) = \frac{1}{\bar{c}}$, where $\frac{1}{\bar{c}}$ is the maximum effort that student $\bar{c}$ can exert (see footnote 16), and the least able student has rank 0 and is indifferent between any level of effort between 0 and $\frac{1}{\bar{c}}$.

[18] For example, if the marginal utility from achievement tends to infinity as achievement approaches its lower bound, then it is not necessarily the case that students with a larger cost of effort exert less effort than those with a lower cost of effort. This is because for students with achievement close to the lower bound, decreasing effort would have a large cost in terms of utility. This is akin to income and substitution effects in labor supply: as the cost of effort increases, the substitution effect would induce individuals to exert less effort, but the income (in this case, achievement) effect would induce them to exert more effort to distance themselves from the achievement lower bound.

[19] The formal statement of the comparative statics result can be found in Proposition (A.1) in Online Appendix A.1.



(a) $\phi > 0$, no aversion to a low rank

(b) $\phi = 0$, aversion to a low rank

Figure 1: The function $Dy(c)$ traces the effect on achievement of increasing the variance of c, as a function of student type c. In the $\phi > 0$ case, it can cross the x-axis once or twice. If it crosses it once (upper panel a), the sequence of its signs, from low c to large c, is $-$, $+$. If it crosses it twice (lower panel a), the sequence of its signs, from low c to large c, is $+$, $-$, $+$. In the $\phi = 0$ case, it can cross the x-axis at most once. If it does not cross it (upper panel b), it must lie below it. If it crosses it (lower panel b), the sequence of its signs, from low c to large c, is $+$, $-$.

3.1 Model Intuition and Discussion of Assumptions

Intuition. Rank preferences imply that each student's incentives are affected by the fraction of peers below her ability, at her ability, and above her ability. First, as can be seen from the first order condition in 1, the marginal utility from increasing one's own rank depends positively on the density at one's own type c, $g(c)$. Intuitively, the more students there are of a similar type c to one's own (i.e., the larger this density), the more students can be surpassed in rank by exerting effort. Conversely, when fewer students can be surpassed for the same amount of effort, students have an incentive to "give up."

Second, the first order condition in 1 shows also that the marginal utility from increasing one's own rank increases when the fraction of less able students (i.e., $1 - G(c)$) decreases, and the extent of this increase depends on ϕ. This derives from the multiplicative specification of preferences, which implies that not only the utility from rank, but also the enjoyment of one's own absolute level of achievement is lower at lower ranks. The closer the students are to the bottom of the ability distribution (i.e., to $\bar{c}$), the fewer classmates there are with whom they can make favorable comparisons, and the stronger their incentive to exert effort to avoid a low achievement rank is. The comparative statics result indicates that this incentive is strongest when $\phi = 0$. For this reason, I call $\phi = 0$ the case of aversion to a low rank.

The comparative statics result is the net effect of these two incentives, which varies with a student's ability c and with the type of rank concerns ($\phi = 0$ or $\phi > 0$). Figure 2 shows two ability distributions $G(c)$ with different variances. As the variance increases, the density increases at the tails and decreases in the middle, therefore, irrespective of the value of ϕ, low- and high-c students have an increased incentive to exert effort and middle-c students have a lower incentive to exert effort. However, high-c students now face a larger fraction of less able students ($1 - G(c)$), which decreases their incentive to exert effort to avoid a low rank. Hence, high-c students face two opposite incentives. The model predicts that they reduce their effort when $\phi = 0$ and increase it when $\phi > 0$. Middle-c students have an incentive to lower their effort under both $\phi = 0$ and $\phi > 0$. This is because when $\phi = 0$ they face less competition from the less able students, while when $\phi > 0$ they face more competition from the less able students, but the model predicts that the incentive to reduce effort due to the lower density at their c level prevails. Under both $\phi = 0$ and $\phi > 0$, low-c students face two opposite incentives: the incentive to increase their effort due to the fatter tail at their end of the distribution, and the incentive to decrease their effort, because they face less competition from middle-c students and, therefore, they can reduce their effort cost without reducing their rank. Which effect prevails depends on how strong the preference for status is relative to the utility from achievement net of the effort cost. This is determined by $V(\cdot)$ and $q(\cdot)$, on which the model does not make specific functional form assumptions. Intuitively, a stronger preference for status leads to an increase in outcomes for low-c students, because the incentive to improve their rank prevails over the incentive to pay a lower effort cost.[20]

Discussion of assumptions. The assumption that overall utility is multiplicative in private utility V and social utility S may appear counterintuitive. However, this assumption makes the problem's structure similar to that of a first-price sealed-bid auction, where expected payoff is the product of the value of winning (V) and the probability of winning (F). As noted in Hopkins and Kornienko (2004), "it is this formal resemblance to an auction that permits clear comparative statics results." While it would be desirable to relax this assumption in favor of more general preference specifications, such a purely theoretical contribution would go beyond the scope of this paper. This paper is the first to apply existing theoretical tools in the rank concerns theoretical literature to the field of peer effects in education, and to test them in this context. Future extensions could specify less restrictive preferences and resort to numerical rather than analytical model solution tools.

Rank concerns can be modeled in many ways. As in Hopkins and Kornienko (2004), I consider only two ($\phi > 0$ and $\phi = 0$). While this is restrictive, these two cases alone can explain important and diverse evidence. Recent laboratory experiments show that individuals are averse to a low-rank, as predicted by the model when $\phi = 0$ (Kuziemko, Buell, Reich, and Norton 2014). Moreover, Tincani (2014) shows that the model with $\phi = 0$ can explain all of the (unexpected) results of the peer regrouping experiment in Carrell, Sacerdote, and West (2013), as well as more recent experimental evidence (Booij, Leuven, and Oosterbeek 2014). The case of $\phi > 0$, on the other hand, can explain this paper's empirical evidence.

[20]For example, it can be shown that if overall utility was $(y(e)-e)^{\alpha}\,(F(e)+\phi)$, then low-$c$ students would increase their effort when status has a larger weight than achievement net of effort cost ($\alpha < 1$) and decrease it otherwise ($\alpha \geq 1$).

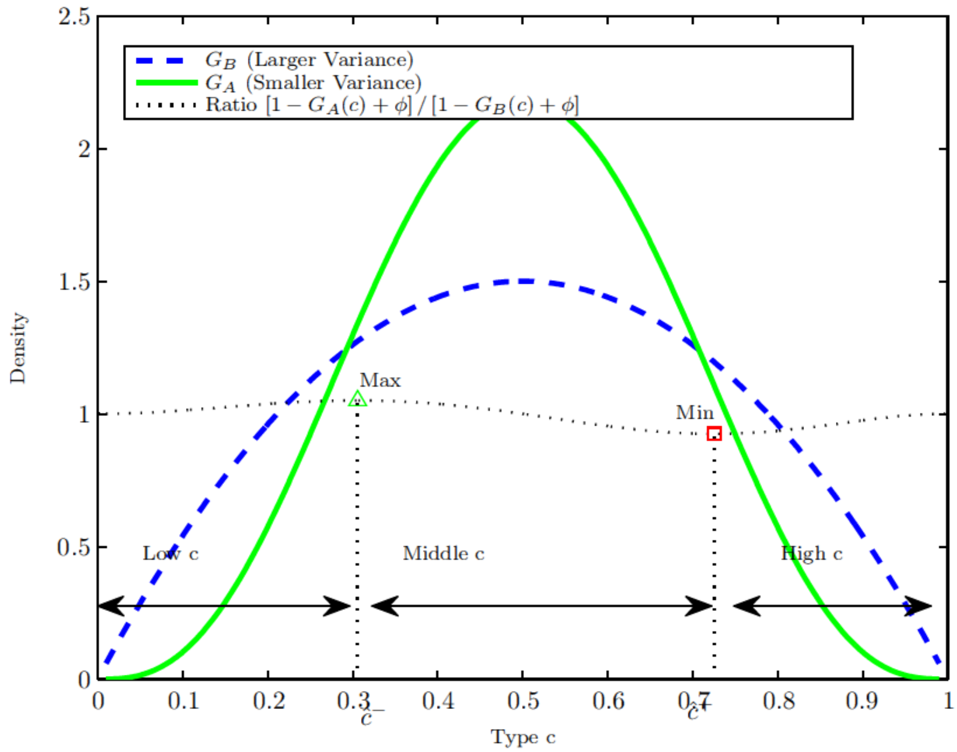


Figure 2: Type distributions in two classrooms, and cutoffs separating low-, middle- and high-c students.

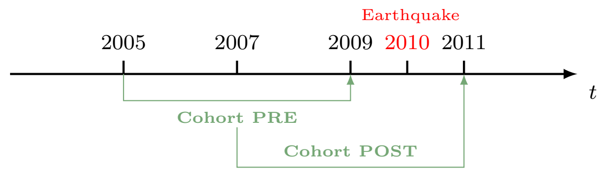


Figure 3: Data time-line.

4 Data and Earthquake

Data. I use two cohorts of students from the SIMCE dataset (*Sistema de Medición de la Calidad de la Educación*). For both cohorts I observe the universe of 8^{th} grade students, for whom I have information on current and 4^{th} grade Math and Spanish standardized test scores, father's and mother's education, household income, gender, and town of residence. Classroom level information includes class size, and the experience, education, tenure at the school, gender, and type of contract of both Spanish and Math teachers.

One cohort is observed in the 8^{th} grade in 2009, before the 2010 earthquake, while the other cohort is observed in the 8^{th} grade in 2011, after the earthquake, as shown in Figure 3. For both cohorts, identifiers are available at the student, teacher, classroom and school level. This makes this dataset ideal for the study of social interactions, because each student's classmates, as well as their school and Math and Spanish teachers, can be identified.

Finally, I obtained from the Chilean Ministry of Education a list of schools that closed as a consequence of the earthquake.[21] I use this list to identify evacuees (i.e., the students who attended the schools that closed), as well as the schools that they moved to. I drop from the sample the $5,988$ evacuees thus identified, and the 803 schools receiving at least one of them, to rid my estimates of any peer effect arising from changes to classroom composition induced by the arrival of evacuees.[22]

Earthquake. Just a few days before the start of the new school year, on February 27^{th}

[21] They closed either because their buildings became unsafe, or because most of the students' homes were so badly damaged, that students had to relocate and the schools had too low attendance.

[22] This distinguishes the identification strategy in this paper from that in Imberman, Kugler, and Sacerdote (2012), where the influx of Katrina evacuees is used as an exogenous source of change to classroom composition. Unlike the case of Katrina, evacuees in Chile were spread across a large number of schools. Therefore, each receiving school received only a small influx of evacuees, too small to detect any statistically significant impact.

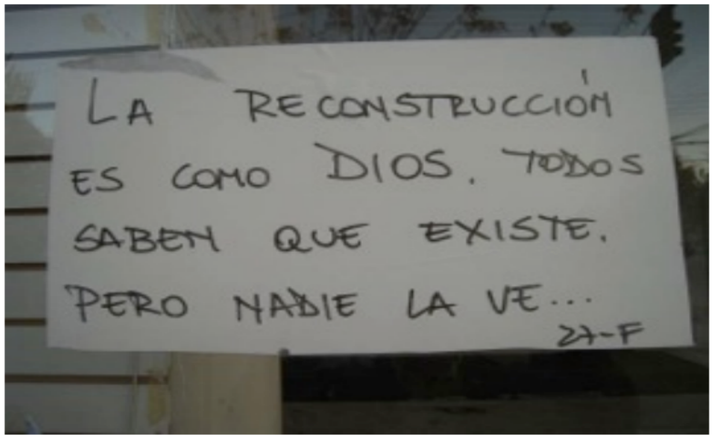


Figure 4: *Source*: Comerio (2013). Handmade sign found in Cauquenes, Chile, on February 2, 2012. Translation: "Reconstruction is like God. Everyone knows it exists, but nobody has seen it."

2010, at 3.34 am local time, Chile was struck by a magnitude 8.8 earthquake, the fifth-largest ever instrumentally recorded (Astroza, Ruiz, and Astroza 2012). Shaking was felt strongly throughout 500 km along the country, covering six regions that together make up about 80 percent of the country's population. The damage was widespread; $370,000$ housing units were damaged or destroyed. The government implemented a national reconstruction plan to rebuild or repair $220,000$ units of low- and middle-income housing. Estimated total costs are around \$2.5 billion. By the time the SIMCE 2011 sample was collected, i.e., 20-22 months after the earthquake struck, despite impressive efforts by the Government, only 24 percent of home reconstruction had been completed (Comerio 2013). This led to frustration in the population, as shown in Figure 4.

To explore how damages are distributed in my sample, I use the geographical coordinates of the town of the school and of the town of the home of each student to build a measure of local earthquake intensity at each school and student's home. To do so, I apply the intensity propagation formula (Astroza, Ruiz, and Astroza 2012), which yields seismic intensity on the Medvedev-Sponheuer-Karnik (MSK) scale as a function of a town's distance from the earthquake's main asperity, i.e., the point on the fault where the rupture starts.[23] MSK intensity indicates the extent of damage

[23]How earthquake damage propagates depends on the geological features of the affected area, therefore, structural engineers estimate this formula for each earthquake separately. First, they observe damages in a sample of towns, then, they estimate the parameters of a non-linear regression

observed in the buildings of a given town. The same MSK intensity corresponds to different levels of damage depending on the type of construction. However, conditional on a construction type, intensity can be mapped back into a specific level of damage. For example, in towns with low MSK intensity, 5, adobe constructions suffered on average damages for $20, while at the strongest intensity, 9, damages per adobe construction were on average $13,800.[24] House type is not observed in my dataset. However, Astroza, Ruiz, and Astroza (2012) report that the $\sim 60\%$ poorest Chileans live in one of two house types with very similar earthquake resistance: old traditional adobe constructions (6.1%) and unreinforced masonry houses (51.9%). Given the striking school stratification in Chile, with public school students belonging to the poorest $\sim 50\%$ of Chilean households, it is reasonable to expect that all public school students live in one of these two building types. To avoid measurement error due to house type unobservability, I restrict my sample to public school students. Finally, the main empirical analysis includes only students in the regions affected by the earthquake. The resulting sample contains observations on 110,822 students, 2,579 schools and 3,712 classrooms.

Astroza, Ruiz, and Astroza (2012) report that even towns that are close to each other were subject to very different levels of seismic intensity. For example, Las Cabras and Pichidegua are only 4.87 miles apart. In Las Cabras, the worst damage suffered by adobe houses has been moderate, ranging from fine cracks in the walls to the sliding of roof tiles, whereas in Pichidegua some adobe houses were destroyed. This suggests that in classrooms with students from different towns, we may observe that students' homes suffered different levels of damages, even though students in the same school tend to live in towns that are close to each other. In fact, in classrooms where not all student reside in the same town, the average standard deviation in damages is approximately USD 350, corresponding to 92 percent of the average monthly income. Of all classrooms in the regions affected by the earthquake, 48 percent are not composed of students who all reside in the same town. The map in Figure 4

of damage on distance from the asperity. For the 2010 Chilean earthquake, the estimated formula is $I(\Delta_A) = 19.781 - 5.927 \log_{10}(\Delta_A) + 0.00087\Delta_A$, where I is MSK intensity and Δ_A is distance in kilometers from the main asperity. The R^2 is 0.9894. Refer to the Online Supplementary Material for an explanation of how each sampled town is assigned an MSK intensity.

[24] See the Online Supplementary material on the author's website for details on how the reconstruction costs are calculated. Table 4 in Online Appendix A.2 shows reconstruction costs for adobe structures by MSK intensity, and Figure A.2 reports pictures of adobe houses that suffered different grades of damage.

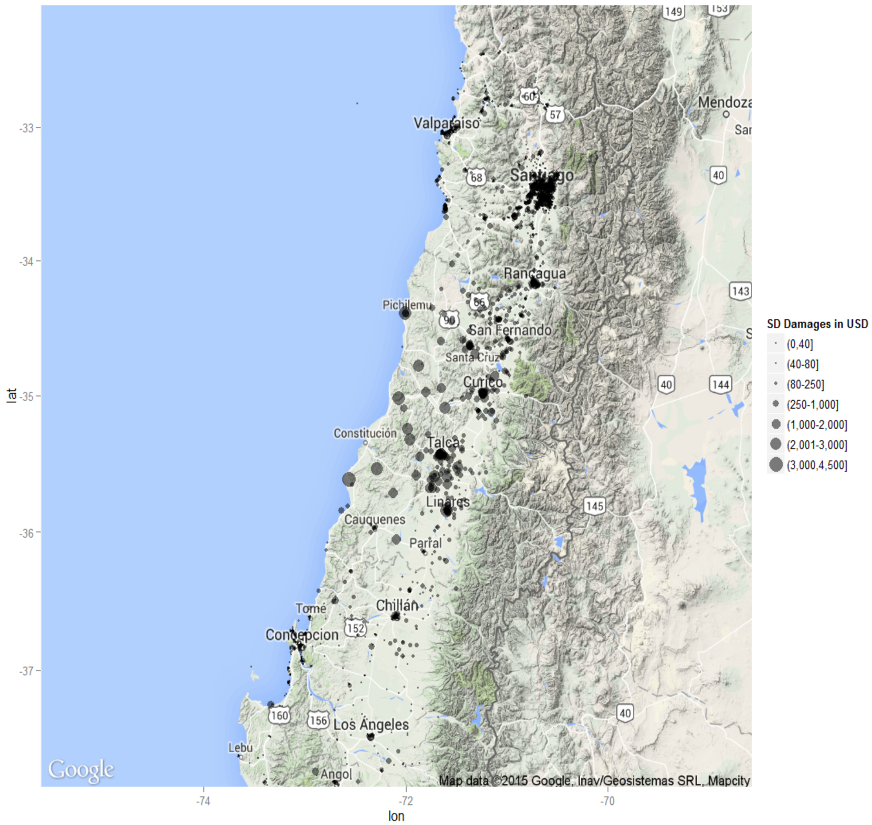


Figure 5: Schools that attract students from different towns, and within school standard deviation in damages.

plots these schools, and it indicates the standard deviation of damages within each school. As can be seen, there is considerable variation within and across schools. This valuable variation provides the basis for testing the model.

4.1 Descriptive Statistics and Preliminary Data Analysis

The modal damage in the sample is USD 365, with a mean of USD 950 and a standard deviation of USD $1,480$. Table 1 presents other sample descriptive statistics. The regions affected by the earthquake are poorer than those not affected. The main empirical analysis uses only earthquake regions. While one may worry that seismic intensity is correlated with student unobservable characteristics affecting outcomes, I show below that conditional on student observable characteristics, seismic intensity is uncorrelated with student outcomes.

I present the first evaluation of the effect of the 2010 Chilean earthquake on

Table 1: Sample descriptive statistics

	Earthquake Regions		**Non-Earthquake Regions**	
	Mean	**St Dev**	**Mean**	**St Dev**
Baseline Math Score	-0.095	0.949	-0.114	0.910
Baseline Spanish Score	-0.092	0.943	-0.087	0.810
Father's Education (years)	9.646	3.338	9.892	3.289
Mother's Education (years)	9.651	3.183	9.741	3.205
Monthly Household Income (USD)	363	341	430	408
Class size	26.7	10.8	25.7	12.2
% Classmates from same town	0.92	0.16	0.96	0.10
% Classmates from same town $\mid < 1$	0.83	0.20	0.88	0.13
% Classrooms with not all students residing in same town	0.48		0.38	
Math Teachers				
% Female	0.55		0.52	
% Postgraduate Degree	0.56		0.53	
Teaching Experience (years)	22.0	13.5	22.8	13.2
Tenure at school (years)	12.3	11.5	11.9	11.0
Spanish Teachers				
% Female	0.81		0.78	
% Postgraduate Degree	0.56		0.49	
Teaching Experience (years)	21.2	13.5	21.5	13.5
Tenure at school (years)	12.0	11.4	11.3	10.8

student test scores. I exploit the fact that the pre- and post-earthquake cohorts are observed both in regions affected and not affected by the earthquake, and that students are observed at two points in time, to estimate a difference-in-differences test score value-added regression. Table 5 in Online Appendix A.3 presents estimates for all schools and by school type (municipal, private subsidized, private unsubsidized). Being exposed to the earthquake reduced test score growth, on average, by 0.05 standard deviations (columns 4 and 8). This estimate is net of any individual, regional and/or cohort effects.

To estimate the impact of the continuous measure of seismic intensity at a student's home, I_i, I calculate MSK intensity for all students (in earthquake regions) in both cohorts. I then estimate the following regression of test scores y of student i in classroom l in school s and in grade $g = 8$ on past test score, individual characteristics $\mathbf{x_i}$, and dummies for belonging to the post-earthquake cohort, $P_i = 1$, and living in an earthquake region, $E_i = 1$ (estimation results can be found in Tables 6 and 7 in the Online Supplementary Material):

$$y_{ilsg} = \alpha + \lambda_s + y_{ils(g-4)}\delta + \mathbf{x_i}'\gamma + P_i\theta_P + E_i\theta_E + (1 - P_i)I_i\theta_{pre} + P_iI_i\theta_{post} + \epsilon_{ilsg}. \quad (2)$$

The effect of earthquake intensity on test score growth is $\theta_{post} - \theta_{pre}$.[25] I find that increasing earthquake intensity by one category in the MSK-scale reduces test score growth by 0.008^{***} standard deviations (sd) in all schools, and by 0.005^{***} and 0.006^{***} sd in municipal schools for Math and Spanish, respectively. This corresponds approximately to a reduction of 0.016 sd in test scores for every USD 100 in damages.

Finally, I find that seismic intensity at the individual student level is uncorrelated with unobservable student characteristics affecting outcomes. Using the unaffected pre-earthquake cohort, I estimate a value-added regression of test scores on student characteristics, classroom characteristics and on own seismic intensity, and I find that the coefficient on the latter is not statistically different from zero (see Table 6 in Online Appendix A.3).[26] This indicates that, conditional on student observables,

[25] This technique is similar in spirit to Card (1992), with the difference that, in this context, I am able to construct treatment intensity also for the untreated pre-earthquake cohort. This allows me to make a weaker identifying assumption than in Card (1992). In fact, the estimated treatment effect here is consistent even if treatment intensity is correlated with unobserved student characteristics affecting outcomes, as long as this correlation is the same in the pre- and post-earthquake cohorts. The sample in this paper satisfies also the stronger assumption made in Card (1992), i.e. that treatment intensity is uncorrelated with student unobservables, as explained below.

[26] Not surprisingly, θ_{pre} in equation 2 is also not significantly different from zero. Moreover, I

seismic intensity is uncorrelated with student unobservables.

5 An Empirical Model of Social Interactions

Achievement Production Function with Peer Spill-overs. I assume that a student's achievement depends on her own characteristics and on classroom characteristics. There are peer effects because two students with identical characteristics may obtain different test scores in two classrooms that are identical except for the ability composition of peers.

I broadly define a student's ability to study as being determined by all individual level inputs into the production of achievement. I assume that it is a scalar obtained as a single index of a vector of student characteristics. I refer to this scalar as the student's type, and denote it by c_i. Type c_i, therefore, is the linear function $c_i = \alpha_1 \mathbf{x_i}$, where $\mathbf{x_i}$ contains student initial ability (as measured by lagged test scores), father's and mother's education, household income, and gender. Changes in classroom composition can be represented as changes in the classroom distribution of type c, $G_l(c)$. The achievement production function of student i in classroom l is:

$$y_{il} = m_l(c_i) + \epsilon_{il} = e(c_i; G_l(c)) + u_l(c_i) + \nu_{il} = e_l(c_i) + u_l(c_i) + \nu_{il} \tag{3}$$

where y_{il} is test score of student i in classroom l. The function $e_l(\cdot)$ maps individual type c_i into achievement, and it is indexed by l because of peer effects: it depends on the distribution of c in classroom l, $G_l(c)$. The function $u_l(c_i)$ captures the (heterogeneous) impact on test scores of all observable classroom characteristics. These are characteristics that are shared by all students in the classroom, and their effect may vary by student's type c_i. Specifically, $u_l(c_i) = u(c_i, \mathbf{z_l}, F_l(\mathbf{x_i}))$ where $\mathbf{z_l}$ contains teacher experience, teacher gender, whether the school is urban or rural, and class size; and $F_l(\mathbf{x_i})$ is the classroom distribution of student characteristics. The dependence on $F_l(\mathbf{x_i})$ captures the fact that teachers may teach differently depending on the characteristics of the students in the classroom, an indirect peer effect. Finally, there may be correlated effects (Manski 1993), as $E[\nu_{il}|l, c_i] \neq 0$. In particular, this conditional expectation is described by a function of type that depends on classroom

estimate a value-added regression similar to 2, without the regressor $(1 - P_i)I_i$, and I obtain a coefficient on $P_i I_i$ that is very close to $\theta_{post} - \theta_{pre}$ in equation 2. See the Online Supplementary Material, where the results from these regressions are presented.

characteristics, $\psi_l(c_i) = \psi(c_i, \mathbf{z_l}, F_l(\mathbf{x_i}))$, so that $\nu_{il} = \psi_l(c_i) + \epsilon_{il}$ with $E(\epsilon_{il}|l, c_i) = 0$. The shock ϵ_{il} is a measurement error. Correlated effects may arise because, for example, more motivated students are found in classrooms with better characteristics $\mathbf{z_l}$, or because certain types of teachers are assigned to classrooms with certain student compositions $F_l(\mathbf{x_i})$. In a similar spirit to $u_l(c_i)$, the function $\psi_l(c_i)$ captures the heterogeneous impact on students of unobserved classroom characteristics. I do not make any functional form assumptions on the functions $e_l(\cdot)$, $u_l(\cdot)$ and $\psi_l(\cdot)$, and any distributional assumptions on ϵ_{il}.

Equation 3 is a semiparametric single-index model (Hall 1989, Ichimura 1993, Horowitz 2010). I jointly estimate the $m_l(c_i)$ function in each classroom l, using kernel methods, and the α parameters.[27] The estimation algorithm is presented in Online Appendix A.4, where the details for the calculation of the standard errors are also presented. Notice that $e_l(c_i)$ is not separately identified from $u_l(c_i) + \psi_l(c_i)$. Therefore, at this stage peer effects cannot be separately identified from the effect of observed and unobserved classroom characteristics, or, in other words, from indirect peer effects (u_l) and correlated effects (ψ_l).

Seismic Intensity as a Source of Identifying Variation. The ideal setting to evaluate the direct effect of peers on test scores is one where student allocation to classrooms is random or experimentally controlled, so that $e_l(c_i)$ can be independently varied from $u_l(c_i) + \psi_l(c_i)$. While college administrators sometimes adopt random assignment of peers to dorms, random assignment of peers to classrooms is rarely adopted in schools or colleges.[28] Moreover, as noted in the survey by Epple and Romano (2011), experiments with random assignment to classrooms are rare, especially beyond primary school.[29]

Given the limited availability of this kind of data, I adopt a different approach. The goal is to vary moments of the distribution of student types in the classroom,

[27] I impose the restriction that c_i does not depend on unobservable student characteristics, because allowing for an unobserved shock to affect c_i would require to assume that the $m(c_i)$ function is monotonic. I do not impose shape restrictions on $m(\cdot)$ because I want to test for its monotonicity, to test the first theoretical model's implication.

[28] Random assignment to dorms has been used by, for example, Sacerdote (2001), Zimmerman (2003), Stinebrickner and Stinebrickner (2006), Kremer and Levy (2008), and Garlick (2014).

[29] One such experiment was conducted among Kenyan first graders and studied by Duflo, Dupas, and Kremer (2011). Whitmore (2005) studies peer effects among kindergartners using data from the project STAR experiment in Tennessee. See also Kang (2007), who studies peer effects among 7^{th} and 8^{th} graders in Korea using a quasi-randomization.

$G_l(c)$, and quantify how $e_l(c_i)$ is affected, net of any change in $u_l(c_i)+\psi_l(c_i)$. Varying the distribution of student types could be achieved by comparing classrooms with different distributions of student characteristics $\mathbf{x_i}$. However, as $u_l(\cdot)$ and $\psi_l(\cdot)$ depend on $F_l(\mathbf{x_i})$, the effect on $e_l(c_i)$, i.e., the peer effect of interest, would not be separately identified from correlated effects and from indirect peer effects. To overcome this obstacle, I consider a shock to each student's type c_i that is such that its distribution in the classroom, or at least a moment of this distribution, is not systematically related to unobserved classroom characteristics or to the productivity of teachers.

The shock is seismic intensity at a student's home for students who were affected by the 2010 Chilean earthquake, I_i. To the extent that I_i affects a student's ability to learn, c_i, classrooms with different distributions of seismic intensity have different distributions of student types, even if the distributions of all other student characteristics, $F_l(\mathbf{x_i})$, are identical. This holds true independently of the channel through which seismic intensity affects a student's ability to study. One possible channel is the disruption to the home environment. It may increase the opportunity cost of time, because students may be required to spend time helping their parents with home repairs.[30] Additionally, students may not have access anymore to the areas of the home that they used for doing their homework. Another possible channel involves psychological well-being. The medical literature finds that earthquake exposure affects brain function and that it can cause Post Traumatic Stress Disorder (PTSD).[31] Moreover, the severity of PTSD increases with seismic intensity (Groome and Soureti 2004).

Survey evidence suggests that seismic intensity at a student's home did in fact affect a student's ability to study. First, conditional on student initial ability and on parental education and income, students more affected by the earthquake report that it is more costly for them to study, as shown in Table A.3 in Online Appendix A.3.[32]

[30] This is particularly likely to have occurred among the low-income Chilean families that my sample focuses on, because most of the government subsidies were in the form of vouchers for purchasing the materials needed for the repairs, and families were expected to perform the repairs themselves (Comerio 2013).

[31] This may last for several months after the earthquake. See, for example, Altindag, Ozen, et al. (2005), Lui, Huang, Chen, Tang, Zhang, Li, Li, Kuang, Chan, Mechelli, et al. (2009), Giannopoulou, Strouthos, Smith, Dikaiakou, Galanopoulou, and Yule (2006).

[32] As shown earlier, conditional on these student observables, seismic intensity is not correlated with unobservables affecting test scores. Therefore, I am confident that this effect can be interpreted as causal. Students were asked to rate how much they agree with sentences such as "It costs me to concentrate and pay attention in class" and "Studying Mathematics costs me more than it costs my classmates". I combine the answers to these questions into a single factor using factor modeling, and I estimate the impact of seismic intensity at a student's home on this elicited measure of cost.

Second, I find that the negative impact of seismic intensity on test scores (see section 4.1) is larger in classrooms in which the teacher assigns homework more frequently. This suggests that home damage affected the productivity of study time at home.[33] Third, using a dataset collected only a few months after the earthquake, I find that students affected more badly by the earthquake report reading less books.[34]

In the empirical model, I assume that the ability to study for the 2011 cohort of students affected by the earthquake is: $c_i = \alpha_1 \mathbf{x_i} + \alpha_2 I_i + \alpha_3 I_i \mathbf{x_i}$. The interaction term $I_i \mathbf{x_i}$ captures individual heterogeneity in how seismic intensity affects a student's type c_i. For example, wealthier parents may try to attenuate the impact of the earthquake by providing a new study environment for their child.[35] As I show in detail below, I find that the variance of seismic intensity in the classroom satisfies an exclusion restriction that allows me to use it in a similar fashion to an Instrumental Variable (IV) for variance of student types. However, unlike in the case of IV, the observability of c_i (and of its variance) is not required in this framework; the single-index model estimates c_i. Given that a student's ability to study is not directly observed, a standard IV approach would not be feasible here because the instrumented variable would not be observed.

Differencing out the confounding effect of geographic dispersion. Seismic intensity is based on a student's home location. Therefore, seismic intensity variance is positively related to the geographic dispersion of the students in a classroom. Using the cohort of students that was not affected by the earthquake, I construct future seismic intensity for each student in the sample, and I find that the coefficient on seismic intensity variance in a test score value-added regression is large and positive, as shown in Table 6 in Online Appendix A.3. This indicates that geographic dispersion of the students in the classroom is (positively) correlated with student outcomes. This could be because classrooms that attract students from further away have some desir-

[33] The additional effect is $-.0173614$, p-value 0.049, see Table 9 in the Online Supplementary Material. The amount of homework assigned is observed only for Math classrooms in the cohort of students affected by the earthquake, therefore, a difference-in-differences strategy that accounts for cohort effects cannot be adopted here.

[34] See Figure 4 in the Online Supplementary Material. This is compatible with an increase in cost of effort/decrease in the ability to study, that was followed by a reduction in effort. This evidence is only suggestive as it comes from a cross-section.

[35] The medical literature reports that the psychological impact of earthquake exposure is stronger on girls, and my parameter estimates find support for this.

able characteristics unobserved to the econometrician, and/or because the students in those classrooms are different in unobserved ways such as their motivation. Regardless of the nature of this correlation, if not accounted for it could confound the estimate of the peer effect of interest, i.e., of the effect of the variance of peer ability ("instrumented" by variance of seismic intensity) on own test scores.

To account for this correlation, I let the functions $u_l(\cdot)$ and $\psi_l(\cdot)$ depend on the variance of seismic intensity in the classroom, σ^2_{Il}. As a result, when we compare classrooms in the post-earthquake cohort that, *ceteris paribus*, suffered different levels of variance of seismic intensity, student outcomes are different for two reasons: direct peer effects, i.e., the variance of c_i is different in these classrooms and, as a result, $e_l(c_i)$ is different; and geographic dispersion effects due to the fact that those classrooms are different in unobserved ways (different $\psi_l(\cdot)$) and/or their observed classroom characteristics have different productivities (different $u_l(\cdot)$). However, when we compare classrooms in the pre-earthquake cohort that, *ceteris paribus*, have different levels of variance of seismic intensity, student outcomes are different only for one reason: geographic dispersion. This is because in the pre-earthquake cohort, student ability to study, c_i, has not been affected by seismic intensity yet, therefore, variance of seismic intensity in the classroom is not related to variance of student ability, but only to geographic dispersion. This suggests an identification strategy: the geographic dispersion effects can be computed in both cohorts of students, and differenced out from the post-earthquake cohort.

Specifically, I model classroom effects in both cohorts as $u(c_i, \mathbf{z_l}, F_l(\mathbf{x_i}), I_l, \bar{I}_l, \sigma^2_{Il})$ and $\psi(c_i, \mathbf{z_l}, F_l(\mathbf{x_i}), I_l, \bar{I}_l, \sigma^2_{Il})$, where I_l is seismic intensity in the school's town, $\bar{I}_l$ is the average and σ^2_{Il} is the variance of seismic intensity suffered by the students in classroom l. I let $u_l(\cdot)$ and $\psi_l(\cdot)$ depend on I_l and $\bar{I}_l$, and not just on σ^2_{Il}, for two reasons: first, to allow for the fact that actual seismic intensity at the school and average intensity of the students in the classroom can directly impact the outcomes of the student in the post-earthquake cohort (e.g. through the damage to school facilities); second, to allow for the fact that, in both cohorts of students, geographic location as captured by these variables may be spuriously related to student test scores if, for example, schools affected more strongly are in poorer/wealthier areas. In estimation, I compare only classrooms with the same values of I_l and $\bar{I}_l$. Therefore, my approach is robust to the earthquake affecting student outcomes through multiple channels, and it is robust to spurious effects related to the geographic location of the school.

Suppose that there are two classrooms in the pre-earthquake cohort, l and $l^{'}$, that are identical in everything, except in the variance of seismic intensity. Specifically, $\mathbf{z_l} = \mathbf{z_{l'}}$, $I_l = I_{l'}$, $\bar{I}_l = \bar{I}_{l'}$ and $F_l(\mathbf{x_i}) = F_{l'}(\mathbf{x_i})$, but $\sigma^2_{Il} \neq \sigma^2_{Il'}$. Assume w.l.o.g. that l has a smaller variance: $\sigma^2_{Il} - \sigma^2_{Il'} = \Delta\sigma^2_{Ill'} < 0$. Because the distribution of c is the same in the two classrooms, $e_l(c) = e_{l'}(c)$ for every c. The function $e_l(c)$ is not identifiable, but the function $m_l(c)$ is. Letting $\phi_l(c) = u_l(c) + \psi_l(c)$ and taking the difference between the m functions in these two classrooms at a given point $c = \alpha_1\mathbf{x_i}$ gives:

$$
\begin{aligned}
m_l(c) - m_{l'}(c) = \Delta m^{pre}_{ll'}(c) &= e_l(c) + \phi^{pre}_l(c) + \epsilon_{il} \\
&\quad - e_{l'}(c) - \phi^{pre}_{l'}(c) - \epsilon_{il'} \\
&= \phi^{pre}_l(c) - \phi^{pre}_{l'}(c) + \epsilon_{il} - \epsilon_{il'} \\
&= \Delta\phi^{pre}_{ll'}(c) + \xi_{ill'} \qquad (4)
\end{aligned}
$$

where *pre* indicates that the sample is the pre-earthquake cohort. $\Delta\phi^{pre}_{ll'}(c)$ is the geographic dispersion effect. Consider now the post-earthquake cohort. Type c_i for these students is affected also by the earthquake intensity. Consider two classrooms, s and $s^{'}$, that, as before, share the same characteristics, but where the variances of seismic intensity differ, i.e. $\sigma^2_{Is} \neq \sigma^2_{Is'}$. W.l.o.g., $\sigma^2_{Is} - \sigma^2_{Is'} = \Delta\sigma^2_{Iss'} < 0$. Because in the post-earthquake cohort seismic intensity affects c_i, the difference in the intensity variances in the two classrooms causes a difference in the variance of c. As a result, if there are peer effects, they will cause a difference between the $e(\cdot)$ functions in the two classrooms, i.e., $e_s(c) \neq e_{s'}(c)$ for at least some c. Taking the difference of the m functions in these two classrooms, at a given point $c = \alpha_1\mathbf{x_i} + \alpha_2 I_i + \alpha_3 I_i\mathbf{x_i}$, gives:

$$
\begin{aligned}
m_s(c_i) - m_{s'}(c) = \Delta m^{post}_{ss'}(c) &= e_s(c) + \phi^{post}_s(c) + \epsilon_{is} \\
&\quad - e_{s'}(c) - \phi^{post}_{s'}(c) - \epsilon_{is'} \\
&= e_s(c) - e_{s'}(c) + \phi^{post}_s(c) - \phi^{post}_{s'}(c) + \epsilon_{is} - \epsilon_{is'} \\
&= \Delta e_{ss'}(c) + \Delta\phi^{post}_{ss'}(c) + \xi_{iss'} \qquad (5)
\end{aligned}
$$

Consider now the four classrooms l, l', s and s' simultaneously. Suppose that $\Delta\sigma^2_{Ill'} = \Delta\sigma^2_{Iss'} < 0$, i.e. the difference in intensity variances within the pre-earthquake pair ll' is identical to the difference in intensity variances within the post-earthquake pair

ss'. If these four classrooms share all other observed classroom characteristics (i.e. $\mathbf{z_l}$, I_l, $\bar{I}_l$ and the classroom distribution of $\mathbf{x_i}$), then the difference between the Δm functions is:

$$\begin{aligned}\Delta m^{post}_{ss'}(c) - \Delta m^{pre}_{ll'}(c) &= \Delta e_{ss'}(c) + \Delta\phi^{post}_{ss'}(c) + \xi_{iss'} - \Delta\phi^{pre}_{ll'}(c) - \xi_{ill'} \\ &= \Delta e_{ss'}(c) + \Delta\phi^{post}_{ss'}(c) - \Delta\phi^{pre}_{ll'}(c) + \Delta\xi_{ill'ss'}\end{aligned}$$

where $E(\Delta\xi_{ill'ss'}|\mathbf{x_i}, \mathbf{z_l}, I_l, \bar{I}_l) = 0$. If $\Delta\phi^{post}_{ss'}(c) = \Delta\phi^{pre}_{ll'}(c)$, then the geographic dispersion effects cancel out,

$$\Delta m^{post}_{ss'}(c) - \Delta m^{pre}_{ll'}(c) = \Delta e_{ss'}(c) + \Delta\xi_{ill'ss'}, \tag{6}$$

and the difference between the Δm functions in the post- and pre-earthquake samples identifies the effect on a student of type c of increasing the variance of student types, i.e., the peer effect of interest, $\Delta e_{ss'}(c)$. See Figures 10 and 11 in Online Appendix A.5 for a graphical representation of this differencing technique.

An important feature of this technique is that the difference in the differences is computed for every point c, i.e., for every student type, and student types are determined differently in the two cohorts of data. Therefore, if two students $i = pre$ and $i = post$ from the two separate cohorts are of the same type c_i, then they must have different characteristics $\mathbf{x_i}$, specifically: $\mathbf{x_{pre}} = \frac{\alpha_1+\alpha_3 I_{post}}{\alpha_1}\mathbf{x_{post}} + \frac{\alpha_2}{\alpha_1}I_{post}$. For example, to be of the same type, the student affected by the earthquake must have a larger initial ability (lagged test score) than the student unaffected by the earthquake, to compensate for the fact that seismic intensity reduced her ability to study. By how much it must be larger depends on the value of α. It is the fact that I jointly estimate the $m(\cdot)$ functions and the α parameters that allows me to make the appropriate comparisons between pre- and post-earthquake students.[36]

To implement the differencing technique, a large number of pair-wise comparisons between classrooms must be performed. This is because the treatment (increasing

[36] Because student types are a function of different covariates in the pre- and post-earthquake cohorts, the standard quantile difference-in-differences (QDID) framework cannot be adopted here (Athey and Imbens 2006). The QDID approach can potentially be extended to allow for different covariates in the treated and untreated sub-populations. However, if feasible, such an extension would require a stronger assumption on the error term than the conditional mean independence assumed in equation 3, a full independence assumption would be needed.

variance of student types) is defined only in relative terms: if classroom A has a larger variance than classroom B, but a lower variance than classroom C, then it is the treated classroom when compared to B, while it is the untreated classroom when compared to C. I perform all possible pair-wise comparisons, and consider the classroom with the larger variance in each pair as the treated classroom. The treatment effect is then estimated by averaging over these pair-wise comparisons. Notice that I average over various treatment intensities, so that the definition of treatment is an increase - of no specific value - in the variance of students' ability to study. This is all that is needed to test the second implication of the theoretical model.

A second key aspect of the method is that these pair-wise comparisons must be made among pairs of classrooms that are identical in terms of all characteristics, except for the variance of seismic intensity. As one of the characteristics is the distribution of students' $\mathbf{x_i}$ variables, a high-dimensional object, only a limited number of classrooms are exactly identical. Therefore, I use kernel weighting, where pairs of classrooms that are very similar obtain higher weights than pairs of classrooms that are less similar, in the spirit of Powell (1987) and Ahn and Powell (1993). Moreover, when conditioning on the distribution of student characteristics, I consider only the mean, the variance, the skewness and the kurtosis of this distribution, to reduce the dimensionality of the matching. Online Appendix A.5 presents the details of the method's implementation, as well as the technical assumptions that must be made to introduce kernel weighting and dimensionality reduction.

I estimate $\Delta e_{ss'}(c)$ on a fine grid of values for c. To test the second theoretical model's prediction, it is sufficient to trace the sign of this function over its domain. This reveals how the effect of increasing the variance of student types varies across students.

Identifying assumption. In the model, the geographic dispersion effects ($\Delta\phi(c)$) are additively separable from the peer effects ($\Delta e(c)$). This is in line with most of the peer effect literature, where confounding effects like correlated effects are typically additive. Given additivity, these potentially confounding effects cancel out if they are identical in the pre- and post-earthquake cohorts. Therefore, the main identifying assumption is:

IA. Constancy of geographic dispersion effects, i.e. $\forall\, c$, $\Delta\phi^{pre}_{ll'}(c) = \Delta\phi^{post}_{ss'}(c)\; \forall l, l', s, s'$

s.t. $\sigma^2_{Il} - \sigma^2_{Il'} = \sigma^2_{Is} - \sigma^2_{Is'}$.

This assumption would not be met if the relationship between student geographic dispersion and unobserved classroom characteristics (i.e., $\psi_l(\cdot, \sigma^2_{Il})$) changed between the two cohorts. This would happen if, for example, schools assigned more experienced teachers or more resources to classrooms that suffered larger variance in damages, or if motivated parents systematically avoided classrooms that suffered larger variance in damages. Table 8 in Online Appendix A.6 examines parental sorting. It presents results from difference-in-differences regressions similar to 2, where the unit of observation is the classroom rather than the student, and where the dependent variables are mean student characteristics in the classroom. As can be seen, the estimated effect of classroom intensity variance is never statistically different from zero. This means that the relationship between intensity variance (i.e., geographic dispersion) and student characteristics is the same in the pre- and post-earthquake cohorts, suggesting that parents did not reallocate across schools as a reaction to intensity variance.[37]

Using the same difference-in-differences framework, Tables 9 and 10 show that also the relationship between observed classroom quality and geographic dispersion did not change after the earthquake; none of the estimates of the effect of classroom intensity variance is significantly different from zero.

The identifying assumption would not be met also if the relationship between student geographic dispersion and the productivity of observed classroom characteristics (i.e., $u_l(\cdot, \sigma^2_{Il})$) changed before and after the earthquake. The main threat to identification in this case would be a reaction of teachers to the variance of damages in the classroom. For example, teachers in classrooms where some students were badly affected while others were not could have changed their focus on instruction. This is an indirect (peer-to-teacher) peer effect that, without data on teacher's practices, usually challenges the identification of direct (peer-to-peer) peer effects. Using survey data on the amount of curriculum covered by Spanish teachers in the pre- and in the post-earthquake cohort, and using the same difference-in-differences framework described above, I evaluate whether the relationship between geographic dispersion (as measured by seismic intensity variance) and amount of curriculum covered changed before and after the earthquake. As can be seen in Table 11, this relationship did

[37]This is not surprising, considering that the sample does not include the students who were forced to relocate because their school closed as an effect of the earthquake, nor does it include the schools that received these evacuees.

Table 2: Parameter Estimates (bootstrapped standard errors in parentheses)

Parameter	Coefficient on	Math	Spanish
α_{12}	Parental Education	-0.01162^{***}	-0.02116^{***}
		(0.00516)	(0.00446)
α_{13}	High Income Dummy	-0.05596^{***}	-0.03560^{**}
		(0.01620)	(0.01749)
α_{14}	Female	0.129037^{***}	-0.23034^{***}
		(0.01953)	(0.03504)
α_2	Seismic Intensity	0.032588	0.09463
		(0.05962)	(0.14377)
α_3	Seismic Intensity*High Income	-0.00037^{***}	-0.00037
		(0.0000)	(0.00271)
α_4	Seismic Intensity*Female	-0.00313	0.05500^{*}
		(0.028773)	(0.03341)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

not change.[38]

Together, these pieces of evidence give me confidence in the plausibility of the identifying assumption.

6 Estimation Results and Statistical Tests of the Theoretical Model's Predictions

Table 2 presents the parameter estimates. The coefficient on lagged test score is normalized to -1, because only the ratios among the α parameters are identified. Under this normalization, student type c_i can be interpreted as a cost of exerting effort, because a lower lagged test score is expected to increase the cost of effort. As expected, earthquake intensity is estimated to increase student type. The model fit is very good, as can be seen in Table 3.

[38]Teachers were given a list of topics, and had to indicate in how much detail they covered each topic. I aggregated the answer into a percentage. While the amount of curriculum covered is not a fine measure of teachers' practice, focus of instruction, and/or effort, it is the only one available. Most studies of peer effects do not have measures of teacher effort. There is strong reason to believe that a change in teachers' focus of instruction in the classroom would be reflected in the speed at which they teach and, therefore, in the percentage of curriculum covered.

Table 3: Model Fit, Test Scores

	Mathematics		Spanish	
	Actual	Model	Actual	Model
Pre-Earthquake Cohort				
Overall	-.185	-.189	-.121	-.123
Female	-.304	-.283	-.050	-.063
Male	-.058	-.089	-.196	-.186
Female				
Urban	-.300	-.279	-.052	-.064
Rural	-.322	-.302	-.043	-.056
Male				
Urban	-.035	-.066	-.180	-.172
Rural	-.159	-.188	-.262	-.249
Female				
Lower Income	-.414	-.387	-.148	-.155
Higher Income	-.130	-.120	.104	.083
Male				
Lower Income	-.222	-.246	-.348	-.328
Higher Income	.155	.116	.001	-.003
Post-Earthquake Cohort				
Overall	-.222	-.228	-.153	-.156
Female	-.307	-.292	-.058	-.078
Male	-.132	-.159	-.254	-.239
Female				
Urban	-.302	-.287	-.071	-.086
Rural	-.329	-.315	.001	-.039
Male				
Urban	-.120	-.148	-.257	-.246
Rural	-.180	-.205	-.242	-.209
Female				
Lower Income	-.414	-.388	-.146	-.160
Higher Income	-.151	-.151	.071	.042
Male				
Lower Income	-.237	-.262	-.351	-.322
Higher Income	-.0004	-.0304	-.133	-.136

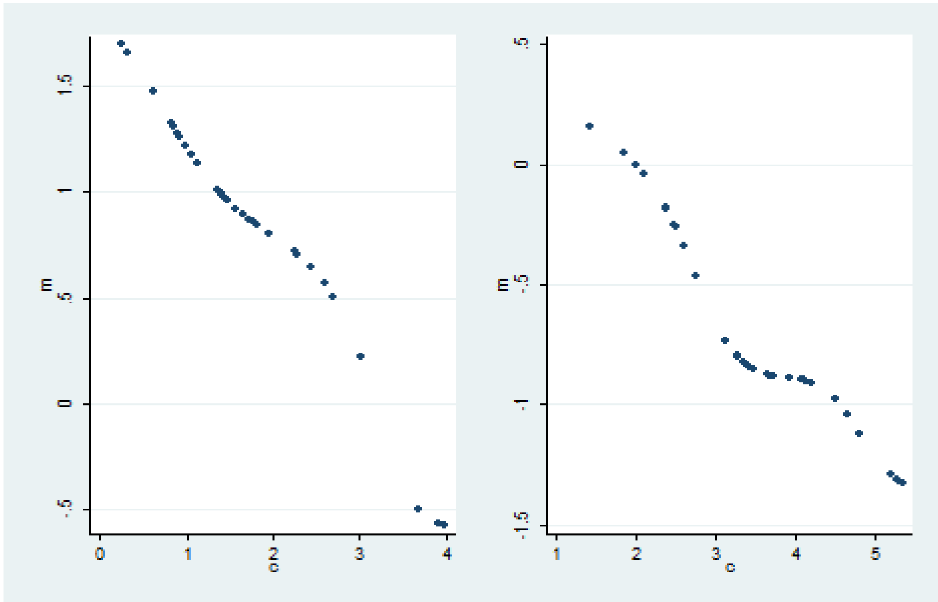


Figure 6: Examples of estimated $m(c)$ functions in two classrooms.

Testing the first theoretical model's implication. The identifiable function $m_l(c)$ is the sum of two functions: $e_l(c) + \phi_l(c)$. The function $e_l(c)$ is the empirical counterpart of the theoretical's model equilibrium effort function. The first theoretical model's prediction is that this function is decreasing. I cannot directly test the monotonicity of $e_l(c)$, however, I can test the monotonicity of $m_l(c)$. If $\phi_l(c)$ is constant, then a decreasing $m_l(c)$ implies that $e_l(c)$ is decreasing. If $\phi_l(c)$ is not constant, then a decreasing $m_l(c)$ implies that it is not true that $e_l(c)$ and $\phi_l(c)$ are both increasing. In either case, an increasing $m_l(c)$ would be reason for concern, while a decreasing $m_l(c)$ would be compatible with the theoretical model's prediction.

Figure 6 shows an example of the function $m_l(c_i)$ estimated in two classrooms. The higher a student's type c_i is, the lower achievement is. I formally test monotonicity of $\hat{m}(\hat{c}_i)$ using the method developed in Chetverikov (2013). For all classrooms, the null hypothesis that the m function is decreasing is not rejected at the $\alpha = 0.10$ significance level. Details of the method can be found in Online Appendix A.7. The values of the test statistics and critical values can be found in the online supplementary material, where plots of the $\hat{m}$ functions in a large number of classrooms are also presented.

Testing the second theoretical model's implication. The effect of increasing the vari-

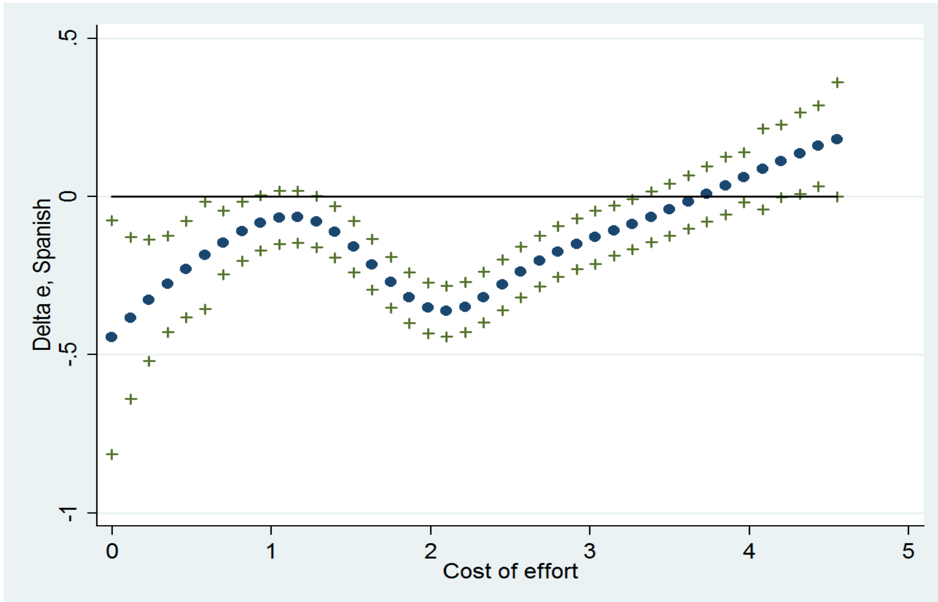


Figure 7: Estimated $\widehat{\Delta e}(c)$ for Spanish test scores. One-sided 90 percent confidence interval reported.

ance of c on student test scores is heterogeneous depending on a student's type. Figures 7 and 8 report the estimates and one-sided point-by-point 90 percent confidence intervals for $\widehat{\Delta e}(c)$ for Mathematics and Spanish test scores, over a grid of values for c. Going from low to high c, this function is negative and then positive for Spanish test scores, while it is positive, then negative and then positive for Math test scores. This means that increasing the variance of c has a negative impact on the test scores of middle-cost (middle-ability) students, and a positive impact on the test scores of high-cost (low-ability) students, while it has a negative impact on low-cost (high-ability) Spanish students, and a positive impact on low-cost (high-ability) Mathematics students.

The patterns observed in the data are consistent with the model's comparative statics result when $\phi > 0$, i.e., when students have a minimum guaranteed level of utility. The one-sided point-by-point 90 percent confidence intervals indicate that the function $\widehat{\Delta e}(c)$ for Spanish test scores is statistically negative and then statistically positive, moving from low to high c.[39] This corresponds to the upper-left panel of

[39]While $\widehat{\Delta e}(c)$ is not statistically different from zero in the neighborhood of $c = 1$, the p-value for the alternative hypothesis that it is greater than zero is 0.84, indicating that this alternative would be rejected at all reasonable significance levels. Therefore, the behavior of the function near $c = 1$ does not reject the model's implication, because the model allows this function to become arbitrarily

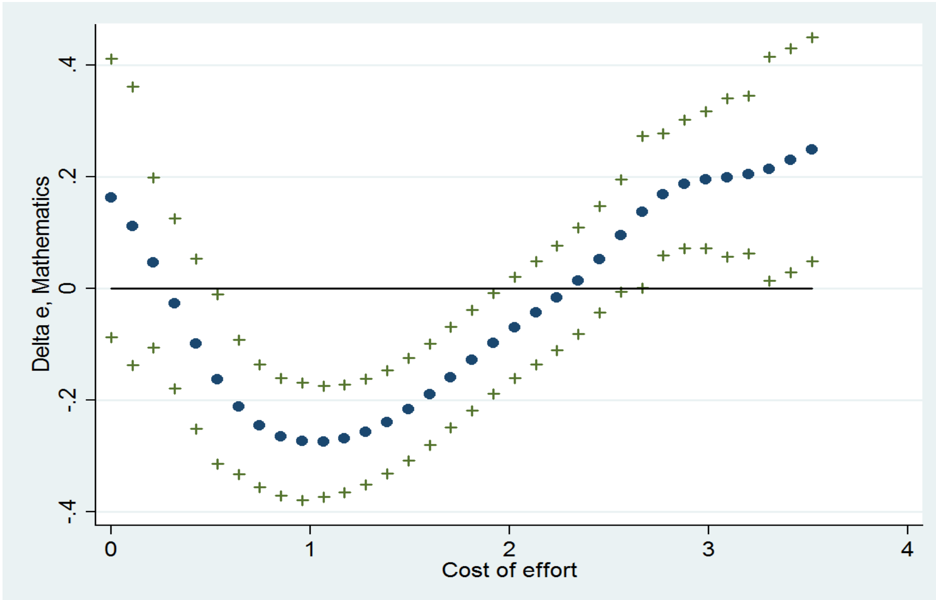


Figure 8: Estimated $\widehat{\Delta e}(c)$ for Math test scores. One-sided 90 percent confidence interval reported.

Figure 3 in the theory section. As typically occurs with non-parametric estimators, the variance is larger near the boundaries, where the data density is smaller. This affects inference for the case of Mathematics. As the confidence intervals indicate, for very small values of c, the null that $\widehat{\Delta e}(c) = 0$ cannot be rejected at the 10 percent significance level. This does not reject the model's implication under $\phi > 0$, which allows this function to be either positive or negative for low values of c.[40] While the point estimate of $\widehat{\Delta e}(c)$ for Math corresponds to the lower-left panel of Figure 3, a pattern like the one in the upper-left panel cannot be excluded.

The model can easily explain the difference in the patterns of $\widehat{\Delta e}(c)$ for Spanish and Math test scores, as discussed in the model intuition section. Whether low-cost (high-ability) students increase or decrease their effort as the variance of student types increases depends on the relative importance that they give to rank versus achievement in the utility. The point estimates suggest higher levels of competition for grades in Math, but the case in which Math and Spanish patterns are similar cannot be rejected statistically. It is possible, however, to reject the model with

close to zero.

[40]This could potentially be problematic if the true function crossed the x-axis multiple times near the boundary. To rule out this occurrence, in ongoing work I am bootstrapping the estimated function.

$\phi = 0$, indicating that aversion to a low rank is not a feature of the data.

While I do find evidence of rank concerns in my data, I am not able to identify whether they are intrinsic or extrinsic. There is anecdotal evidence indicating that Chilean public school 8^{th} graders face extrinsic incentives to care about their achievement rank.[41] However, intrinsic rather than extrinsic rank concerns may be generating the observed patterns.

Implications for Ability Tracking. The two cases $\phi > 0$ and $\phi = 0$ have very different implications for tracking. Under $\phi = 0$, the case of aversion to a low rank, students who have a low ability rank but who become the best in their class once they are tracked may have an incentive to lower their effort, because, being a "big fish in a small pond", they do not face anymore the risk of ranking low in achievement. When $\phi > 0$, this incentive is not present. All students in tracked classrooms have an incentive to exert more effort. This is because tracking, by partitioning the domain of the ability distribution, increases for all students the fraction of peers of a similar ability. Therefore, the marginal utility of effort increases for all under $\phi > 0$, regardless of a student's position in the ability distribution. The type of rank concerns for which I find evidence in the data ($\phi > 0$) imply that tracking students by ability would give an incentive to all students to increase their effort, and compete more fiercely among equals.

7 Robustness and Alternative Mechanisms

First, I estimate the impact of seismic intensity variance on test scores using the parametric difference-in-differences approach with continuous treatment outlined in equation 2. As can be seen in Table 12 in Online Appendix A.8, the estimated impact is not significantly different from zero for Mathematics, and it is -0.25 sd for Spanish test scores. These results are both qualitatively and quantitatively compatible with the results from the semi-parametric approach. Moreover, they underline the

[41] Chilean parents are more likely to choose a private (subsidized) school in high-school rather than primary school (which runs from 1^{st} to 8^{th} grade). Therefore, at the end of the 8^{th} grade many public school students apply for admission to private (subsidized) schools. At the same time, they are required to take a national exam, and many private schools select their students also on the basis of past grades in primary school and on the national exam. This gives an extrinsic incentive to students in the 8^{th} grade to compete for grades.

importance of adopting an approach that can detect the heterogeneity of the impact across students, because the mean treatment effects mask considerable heterogeneity.

Second, I ask whether the observed patterns of heterogeneity can be explained by alternative mechanisms. As a first mechanism I consider the teacher channel. This channel can be excluded for two reasons. First, as outlined in section 5, I do not find evidence of teachers changing their teaching practices as a reaction to the classroom variance in damages. Second, for this channel to explain the empirical results, one must assume that Math and Spanish teachers reacted differently to a change in intensity variance, with Spanish teachers focusing only on the high-cost individuals, and Mathematics teachers focusing on the high- and on the low-cost individuals. No previous study provides support for such *ad hoc* theories.

As a second alternative mechanism I consider the theory of social cognitive learning, which posits that students learn from similar classmates (Bandura 1986, Schunk 1996). This theory is rejected by my data, because it would require that low-cost students in both Mathematics and Spanish classes increase their test scores when the variance increases, because of the larger density at their ability level. However, I find that low-cost students in Spanish classes obtain lower test scores.

Related to this mechanism is the mechanism of self-selection into peer subgroups formed mainly of peers with a similar ability, a mechanism proposed by Carrell, Sacerdote, and West (2013) to explain their experimental findings. If only mean peer ability matters and if students choose more often to become friends with similarly able peers when their availability increases, then we should see a worsening of the outcomes of low-ability students when there are more low-ability students in the classroom. In this context, this mechanism can be excluded because low-ability students in both Math and Spanish classes increase their test scores when the number of low-ability students in the classroom increases following a variance increase.

A final mechanism I consider is cooperative behavior between students affected by the earthquake. I call this mechanism the “good samaritan”: if students who were less affected by the earthquake helped the more affected ones, for example by offering them to study with them at their less affected homes, then the estimation of the impact of damage at a student’s home on a student’s type would be biased. As a result, the variance of seismic intensity would provide a biased measure of the variance of student types. A key implication of the good samaritan mechanism is that if such a mechanism were present, then the econometric model would systematically underestimate the variance in student types, because this mechanism would work as an insurance between students that lowers the heterogeneity of the earthquake impacts on students' ability to study. Identification relies on contrasting classrooms with different variances in seismic intensity, and on considering the classroom with the larger variance as the treated classroom. As long as the systematic under-estimation of student type variance preserves the relative ordering of classrooms in terms of true student type variance, the paper's conclusions are not affected. For example, if the bias is the same in all classrooms, then the relative ordering is preserved and the study's conclusions do not change. It would be worrisome if classrooms with a larger variance in seismic intensity variance had a larger degrees of insurance. A symptom of this would be that in classrooms with larger variance in seismic intensity, seismic intensity at a student's home is a bad predictor of a student's test scores. However, I find that seismic intensity always decreases in a statistically significant way students' test scores. Therefore, I am confident that the good samaritan mechanism, if present, does not pose a threat to the validity of this paper's empirical findings.

8 Conclusions

This paper is the first to study rank concerns as a mechanism underlying peer effects in education. There is mounting evidence that the standard linear-in-means model of peer effects is not an appropriate model to describe peer effects in test scores (Sacerdote 2014). Yet, there is still no consensus on what constitutes a plausible model.

The theoretical framework that I propose generates non-linear peer effects that can help us understand some of the existing evidence, and it does so by introducing reasonable assumptions on the primitives, i.e., on student preferences. There is evidence that in many contexts students have intrinsic (Tran and Zeckhauser 2012, Azmat and Iriberri 2010) and extrinsic (Cotton, Hickman, and Price 2014) reasons to care about classroom rank. However, the existing literature has so far ignored the implications that these preferences have for peer effects. I find that they generate an interaction between academic competition and classroom ability composition that can be exploited by policymakers in the design of classroom allocation rules.

In light of the encouraging new results of this paper, it appears worthwhile to further explore the role of rank concerns in determining peer effects. There are at least two possible directions for future research. First, the theoretical model in this paper shows that the impact of changing peer group composition depends on the strength of rank concerns. Therefore, it would be very useful to collect data that allow researchers to credibly measure student rank concerns, either through the elicitation of preferences with surveys, or through laboratory experiments. This knowledge could then be used to study optimal peer allocation.

Second, there currently exist two separate strands in the experimental literature in Education: studies that randomly vary the incentives for students to compete, for example, through merit fellowships, affirmative action programs, and financial awards, and studies that randomly vary student composition. This paper's findings indicate that it would be useful to consider both variations simultaneously, because of potential complementarities between the two types of interventions. Results from such experiments would inform policy makers on the optimal combination of competition incentives and peer group composition.

References

AHN, H., AND J. L. POWELL (1993): "Semiparametric estimation of censored selection models with a nonparametric selection mechanism," *Journal of Econometrics*, 58(1), 3–29.

ALTINDAG, A., S. OZEN, ET AL. (2005): "One-year follow-up study of posttraumatic stress disorder among earthquake survivors in Turkey," *Comprehensive psychiatry*, 46(5), 328–333.

ANGRIST, J., AND V. LAVY (2009): "The effects of high stakes high school achievement awards: Evidence from a randomized trial," *The American Economic Review*, pp. 1384–1414.

ANGRIST, J. D. (2014): "The perils of peer effects," *Labour Economics*.

ARNOTT, R., AND J. ROWSE (1987): "Peer group effects and educational attainment," *Journal of Public Economics*, 32(3), 287–305.

ASTROZA, M., S. RUIZ, AND R. ASTROZA (2012): "Damage Assessment and Seismic Intensity Analysis of the 2010 (Mw 8.8) Maule Earthquake," *Earthquake Spectra*, 28(S1), S145–S164.

Athey, S., and G. W. Imbens (2006): “Identification and inference in nonlinear difference-in-differences models,” *Econometrica*, 74(2), 431–497.

Azmat, G., and N. Iriberri (2010): “The importance of relative performance feedback information: Evidence from a natural experiment using high school students,” *Journal of Public Economics*, 94(7), 435–452.

Bandura, A. (1986): *Social foundations of thought and action.* Englewood Cliffs, NJ Prentice Hall.

Bénabou, R. (1996): “Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance,” *The American Economic Review*, pp. 584–609.

Betts, J. R., and J. L. Shkolnik (2000): “The effects of ability grouping on student achievement and resource allocation in secondary schools,” *Economics of Education Review*, 19(1), 1–15.

Blume, L. E., W. A. Brock, S. N. Durlauf, and R. Jayaraman (2014): “Linear social interactions models,” *forthcoming, Journal of Political Economy.*

Booij, A., E. Leuven, and H. Oosterbeek (2014): “Ability Peer Effects in University: Evidence from a Randomized Experiment,” *Unpublished Working Paper.*

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009): “Peer effects and social networks in education,” *The Review of Economic Studies*, 76(4), 1239–1267.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013): “From natural variation to optimal policy? The importance of endogenous peer group formation,” *Econometrica*, 81(3), 855–882.

Chetverikov, D. (2013): “Testing regression monotonicity in econometric models,” *mimeo, UCLA.*

Cipollone, P., and A. Rosolia (2007): “Social interactions in high school: Lessons from an earthquake,” *The American economic review*, pp. 948–965.

Coleman (1966): “Equality of educational opportunity,” *Washington, dc*, pp. 1066–5684.

Comerio, M. C. (2013): *Housing recovery in Chile: A qualitative mid-program review*. Pacific Earthquake Engineering Research Center Headquarters at the University of California.

Cotton, C., B. R. Hickman, and J. P. Price (2014): "Affirmative Action and Human Capital Investment: Evidence from a Randomized Field Experiment," Discussion paper, National Bureau of Economic Research.

De Giorgi, G., and M. Pellizzari (2013): "Understanding social interactions: Evidence from the classroom," *The Economic Journal*.

Ding, W., and S. F. Lehrer (2007): "Do peers affect student achievement in China's secondary schools?," *The Review of Economics and Statistics*, 89(2), 300–312.

Dubey, P., and J. Geanakoplos (2010): "Grading exams: 100, 99, 98,Ě or A, B, C?," *Games and Economic Behavior*, 69(1), 72–94.

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

Elsner, B., and I. Isphording (2015): "A big fish in a small pond: Ability rank and human capital investment," *mimeo, Intitute for the Study of Labor*.

Epple, D., and R. Romano (2008): "Educational Vouchers and Cream Skimming," *International Economic Review*, 49(4), 1395–1435.

——— (2011): "Peer effects in education: A survey of the theory and evidence," *Handbook of social economics*, 1(11), 1053–1163.

Epple, D., and R. E. Romano (1998): "Competition between private and public schools, vouchers, and peer-group effects," *American Economic Review*, pp. 33–62.

Frank, R. H. (1985): *Choosing the right pond: Human behavior and the quest for status*. Oxford University Press.

Fruehwirth, J. C. (2012): "Can achievement peer effect estimates inform policy? A view from inside the black box," *Review of Economics and Statistics*, (00).

——— (2013): “Identifying peer achievement spillovers: Implications for desegregation and the achievement gap,” *Quantitative Economics*, 4(1), 85–124.

Fryer, R. G. (2010): “Financial incentives and student achievement: Evidence from randomized trials,” Discussion paper, National Bureau of Economic Research.

Fu, C., and N. Mehta (2015): “Ability Tracking, School and Parental Effort, and Student Achievement: A Structural Model and Estimation,” Discussion paper, Working Paper, University of Western Ontario.

Garlick, R. (2014): “Academic Peer Effects with Different Group Assignment Rules: Residential Tracking versus Random Assignment,” Discussion paper, Working paper, World Bank Development Research Group, Washington, DC.

Ghosal, S., A. Sen, and A. W. Van Der Vaart (2000): “Testing monotonicity of regression,” *Annals of statistics*, pp. 1054–1082.

Giannopoulou, I., M. Strouthos, P. Smith, A. Dikaiakou, V. Galanopoulou, and W. Yule (2006): “Post-traumatic stress reactions of children and adolescents exposed to the Athens 1999 earthquake,” *European Psychiatry*, 21(3), 160–166.

Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (1996): “Crime and Social Interactions,” *Quarterly Journal of Economics*, 111(2).

Graham, B. S. (2008): “Identifying social interactions through conditional variance restrictions,” *Econometrica*, 76(3), 643–660.

Groome, D., and A. Soureti (2004): “Post-traumatic stress disorder and anxiety symptoms in children exposed to the 1999 Greek earthquake,” *British Journal of Psychology*, 95(3), 387–397.

Hall, P. (1989): “On projection pursuit regression,” *The Annals of Statistics*, pp. 573–588.

Hopkins, E., and T. Kornienko (2004): “Running to keep in the same place: consumer choice as a game of status,” *American Economic Review*, pp. 1085–1107.

——— (2007): “Cross and double cross: comparative statics in first price and all pay auctions,” *The BE Journal of Theoretical Economics*, 7(1).

Horowitz, J. L. (2010): *Semiparametric and nonparametric methods in econometrics*, vol. 692. Springer.

Hoxby, C. M., and G. Weingarth (2005): "Taking race out of the equation: School reassignment and the structure of peer effects," Discussion paper, Working paper.

Ichimura, H. (1993): "Semiparametric least squares (SLS) and weighted SLS estimation of single-index models," *Journal of Econometrics*, 58(1), 71–120.

Imberman, S. A., A. D. Kugler, and B. I. Sacerdote (2012): "Katrina's children: Evidence on the structure of peer effects from hurricane evacuees," *The American Economic Review*, 102(5), 2048–2082.

Jencks, C., and S. E. Mayer (1990): "The social consequences of growing up in a poor neighborhood," *Inner-city poverty in the United States*, 111, 186.

Kang, C. (2007): "Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea," *Journal of Urban Economics*, 61(3), 458–495.

Kremer, M., and D. Levy (2008): "Peer effects and alcohol use among college students," *The Journal of Economic Perspectives*, 22(3), 189–189.

Kremer, M., E. Miguel, and R. Thornton (2009): "Incentives to learn," *The Review of Economics and Statistics*, 91(3), 437–456.

Kuziemko, I., R. Buell, T. Reich, and M. Norton (2014): "Last-place Aversion: Evidence and Redistributive Implications," *Quarterly Journal of Economics*, 129(1), 105–149.

Lavy, V., M. D. Paserman, and A. Schlosser (2012): "Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom*," *The Economic Journal*, 122(559), 208–237.

Lavy, V., and A. Schlosser (2011): "Mechanisms and impacts of gender peer effects at school," *American Economic Journal: Applied Economics*, pp. 1–33.

LAVY, V., O. SILVA, AND F. WEINHARDT (2012): "The good, the bad, and the average: evidence on ability peer effects in schools," *Journal of Labor Economics*, 30(2), 367–414.

LEFGREN, L. (2004): "Educational peer effects and the Chicago public schools," *Journal of urban Economics*, 56(2), 169–191.

LEVITT, S., J. LIST, AND S. SADOFF (2011): "The effect of performance-based incentives on educational achievement: Evidence from a randomized experiment," *Unpublished Manuscript, University of Chicago.*

LUI, S., X. HUANG, L. CHEN, H. TANG, T. ZHANG, X. LI, D. LI, W. KUANG, R. C. CHAN, A. MECHELLI, ET AL. (2009): "High-field MRI reveals an acute impact on brain function in survivors of the magnitude 8.0 earthquake in China," *Proceedings of the National Academy of Sciences*, 106(36), 15412–15417.

LYLE, D. S. (2009): "The effects of peer group heterogeneity on the production of human capital at West Point," *American Economic Journal: Applied Economics*, pp. 69–84.

MANSKI, C. F. (1993): "Identification of endogenous social effects: The reflection problem," *The review of economic studies*, 60(3), 531–542.

MURPHY, R., AND F. WEINHARDT (2014): "Top of the Class: The Importance of Ordinal Rank," Discussion paper, CESifo Group Munich.

PAGAN, A., AND A. ULLAH (1999): *Nonparametric econometrics*. Cambridge university press.

POWELL, J. L. (1987): *Semiparametric estimation of bivariate latent variable models*. University of Wisconsin–Madison, Social Systems Research Institute.

RIVKIN, S. G., E. A. HANUSHEK, AND J. F. KAIN (2005): "Teachers, schools, and academic achievement," *Econometrica*, 73(2), 417–458.

SACERDOTE, B. (2008): "When the saints come marching in: Effects of Hurricanes Katrina and Rita on student evacuees," Discussion paper, National Bureau of Economic Research.

——— (2011): "Peer effects in education: How might they work, how big are they and how much do we know thus far?," *Handbook of the Economics of Education*, 3, 249–277.

——— (2014): "Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?," *Annual Review of Economics*, (0).

Sacerdote, B. I. (2001): "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, 116(2), 681–704.

Schunk, D. H. (1996): "Learning theories," *Printice Hall Inc., New Jersey.*

Stinebrickner, R., and T. R. Stinebrickner (2006): "What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds," *Journal of public Economics*, 90(8), 1435–1454.

Tincani, M. (2014): "On the Nature of Social Interactions in Education: An Explanation for Recent Evidence," Discussion paper, UCL.

Todd, P. E., and K. I. Wolpin (2014): "Estimating a Coordination Game in the Classroom," Discussion paper, UPenn.

Tran, A., and R. Zeckhauser (2012): "Rank as an inherent incentive: Evidence from a field experiment," *Journal of Public Economics*, 96(9), 645–650.

Vigdor, J., and T. Nechyba (2007): "Peer effects in North Carolina public schools," *Schools and the equal opportunity problem*, pp. 73–101.

Whitmore, D. (2005): "Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment," *American Economic Review, P&P*, pp. 199–203.

Zimmerman, D. J. (2003): "Peer effects in academic outcomes: Evidence from a natural experiment," *Review of Economics and Statistics*, 85(1), 9–23.

A Online Appendix - NOT FOR PUBLICATION

A.1 Model Details and Proofs

A.1.1 Differential Equation

The first-order differential equation characterizing equilibrium strategies is obtained by rearranging the first order condition in 1, and substituting $c^{'}(e) = \frac{1}{e^{'}(c)}$:

$$
\begin{aligned}
e^{'}(c_i) &= \left(\frac{g(c_i)}{1 - G(c_i) + \phi}\right)\left(\frac{V(y(e), q(e,c))}{a(\mu)V_1 + V_2\frac{\partial q}{\partial e}}\right). \qquad (7)\\
&= \frac{g(c_i)}{1 - G(c_i) + \phi}\psi(e_i, c_i).
\end{aligned}
$$

A.1.2 Proof of Proposition 3.1

The proof is an adaptation of the proof in Hopkins and Kornienko (2004), where equilibrium strategies are strictly increasing and where the consumption and positional goods are two separate goods. Here I report the proof that if the strategy $e^*(c)$ is a best response to other students' effort choices, then it is decreasing.[42] I also provide an intuition for why the strategy is continuous and unique.

It is easy to show that the boundary conditions in the statement of the Proposition are optimal for the student with the highest cost, $\bar{c}$. For the case $\phi > 0$, the student with the highest type, $\bar{c}$, chooses the effort function that maxims utility V in the absence of rank concerns, as specified by the boundary condition in the statement of the Proposition. To see why, notice that in equilibrium her utility from rank is zero, therefore, she maximizes V because $V \times F + \phi \times V = V \times 0 + \phi \times V = \phi \times V$. For the case $\phi = 0$, footnote 17 explains why the student at $\bar{c}$ is indifferent between exerting any level of effort in $[0, \frac{1}{\bar{c}}]$. I assume the boundary condition $e(\bar{c}) = \frac{1}{\bar{c}}$ because it preserves continuity of the equilibrium effort function (because $\lim_{c \to \bar{c}^-} = \frac{1}{\bar{c}}$).

Proof If a student i of type c_i exerts effort $e_i = e^*(c_i)$ and this is a best response to the efforts of the other students as summarized by the effort distribution $F_E(\cdot)$, then it must be that $e_i \geq e_{nr}(c_i)$, where $e_{nr}(c_i)$ solves the first-order condition in the

[42]The remaining part of the proof, showing that the equilibrium strategy is strictly decreasing, continuous and differentiable is a lengthy adaptation of the proofs in Hopkins and Kornienko (2004), and it is available from the author upon request.

absence of rank concerns, i.e., $V_1 a(\mu)|_{e=e_{nr}} = -V_2 \frac{\partial q}{\partial e}|_{e_{nr}}$. This is because if $e < e_{nr}(c_i)$, then $F_E(e) + \phi < F_E(e_n) + \phi$ and $V\left(y(e(c)), q(e(c), c)\right) < V\left(y(e_{nr}(c)), q(e_{nr}(c), c)\right)$. Therefore, $V\left(y(e), q(e, c)\right)\left(F_E(e) + \phi\right) < V\left(y(e_{nr}), q(e_{nr}(c), c)\right)\left(F_E(e_{nr}) + \phi\right)$, i.e., any level of effort below the no-rank-concerns level is strictly dominated by the no-rank-concerns level. Suppose that equality holds, so $e_i = e_{nr}(c_i)$. Then $e^*(\cdot)$ is decreasing because $e_n(c_i)$ is decreasing. This follows from the assumptions on utility that $V_{11} = 0$, $V_{22} = 0$, $V_{ij} \leq 0$ for $i \neq j$, and from the assumptions on the cost of effort function that $\frac{\partial q}{\partial c} > 0$, $\frac{\partial q}{\partial e} > 0$, $\frac{\partial^2 q}{\partial^2 e} > 0$ and $\frac{\partial^2 q}{\partial e \partial c} \geq 0$. To see why, let $FOC(e, c) = V_1 a(\mu) + V_2 q_1$ and notice that by the Implicit Function Theorem:

$$\frac{de_{nr}}{dc} = -\frac{\partial FOC/\partial c}{\partial FOC/\partial e}.$$

The numerator is:

$$\frac{\partial FOC}{\partial c} = a(\mu) V_{12} \frac{\partial q}{\partial c} + V_{22} \frac{\partial q}{\partial e} \frac{\partial q}{\partial c} + V_2 \frac{\partial^2 q}{\partial e \partial c} \leq 0.$$

The denominator is:

$$\frac{\partial FOC}{\partial e} = a(\mu)^2 V_{11} + a(\mu) V_{12} \frac{\partial q}{\partial e} + \left(a(\mu) V_{21} + V_{22} \frac{\partial q}{\partial e}\right) \frac{\partial q}{\partial e} + V_2 \frac{\partial^2 q}{\partial^2 e} \leq 0.$$

As a result, $e^*(\cdot)$ is decreasing in c when it is equal to optimally chosen effort in the absence of rank concerns, because $\frac{de_{nr}}{dc} \leq 0$.

If equality does not hold, we want to show that if e_i is a best-response and $e_i > e_{nr}(c_i)$, then it is still the case that e_i is decreasing in c_i. First, I show that for any other choice $\tilde{e}_i \in (e_{nr}(c_i), e_i)$,

$$\frac{\partial V}{\partial c_i}\left(y(e_i), q(e_i, c_i)\right)\left(F_E(e_i) + \phi\right) < \frac{\partial V}{\partial c_i}\left(y(\tilde{e}_i), q(\tilde{e}_i, c_i)\right)\left(F_E(\tilde{e}_i) + \phi\right). \tag{8}$$

Rewrite the left-hand side as:

$$\frac{\partial V}{\partial c_i}\left(y(e_i), q(e_i, c_i)\right)\left(F_E(\tilde{e}_i) + \phi\right) + \frac{\partial V}{\partial c_i}\left(y(e_i), q(e_i, c_i)\right)\left(F_E(e_i) - F_E(\tilde{e}_i)\right).$$

The first term is smaller or equal to the right-hand side of equation 8, because $\frac{\partial V}{\partial c}$ is decreasing in e by the assumptions that $V_{21} \leq 0, V_{22} = 0, \frac{\partial q}{\partial c} > 0, V_2 < 0$, and $\frac{\partial^2 q}{\partial c \partial e} \geq 0$. To see why, notice that $\frac{\partial^2 V}{\partial c \partial e} = \left(V_{21} a(\mu) + V_{22} \frac{\partial q}{\partial e}\right) \frac{\partial q}{\partial c} + V_2 \frac{\partial q}{\partial c \partial e} \leq 0$. The second term is strictly negative, because first, $\frac{\partial V}{\partial c_i}$ is strictly negative by virtue of the assumptions that $V_2 < 0$ and $\frac{\partial q}{\partial c} > 0$, and second, $(F_E(e_i) - F_E(\tilde{e}_i)) > 0$. To see why the latter is true, notice that for $e > e_{nr}$, $V(y(e), q(e, c))$ is decreasing in e. Therefore, if e is a best-response, it must be the case that $F_E(e_i) > F_E(\tilde{e}_i)$, otherwise a student could lower effort and obtain a higher utility, while not lowering her status. This establishes the inequality in 8, so that at e_i, the overall marginal utility with respect to c ($\frac{\partial}{\partial c}(V(y,q)(F_E(e) + \phi))$) is strictly decreasing in e. This implies that an increase in type c leads to a decrease in the marginal return to e, therefore, the optimal choice of effort e must decrease. Q.E.D.

To show that if an effort function is an equilibrium strategy, then it must be continuous, suppose not. That is, suppose that that there was a jump downwards in the equilibrium effort function $e^*(c)$ at $\tilde{c}$, so that $\lim_{c \to \tilde{c}^+} e^*(c) = \tilde{e} < e^*(\tilde{c})$. Then, there would exist an $\epsilon > 0$ small enough, such that the student of type $\tilde{c} - \epsilon$ can reduce her effort to $\tilde{e}$, which is below $e^*(\tilde{c} - \epsilon)$, and obtain a discrete increase in utility because of the lower effort, while her rank would decrease by less, by continuity of the rank function $S(\cdot)$ at $\tilde{c}$. Therefore, there exists a student with an incentive to deviate, and such discontinuous $e^*(c)$ function cannot be an equilibrium strategy.

Uniqueness of the solution to the differential equation in 7, and therefore uniqueness of the equilibrium, follows from the Lipschitz continuity of the equilibrium effort function. This could potentially not be satisfied in the case $\phi = 0$ at the boundary condition $e(\bar{c}) = \frac{1}{\bar{c}}$, because the denominator on the right hand side of the differential equation in 7 is zero in this case. However, it can be shown that the assumptions on the second and cross derivatives of V guarantee uniqueness of the solution also in this case, refer to Hopkins and Kornienko (2004) for more details.

Intuition for equilibrium uniqueness. Intuitively, uniqueness of the equilibrium in the case $\phi > 0$ follows from two key assumptions: achievement gives utility per se, i.e., irrespective of the status it provides, and individuals have different costs of producing achievement. A common type of multiplicity in this class of games is when all individuals exert the same amount of effort. If this were an equilibrium, there would be an infinite number of equilibria. However, all students playing the same level of effort e^* is not an equilibrium, because students with a high enough cost (i.e., with a cost above a certain cutoff that depends on e^*, $c > cutoff(e^*)$), have an incentive to reduce effort, obtain zero rank, and enjoy their private utility from achievement. Therefore, the classical problem of multiplicity of equilibria in coordination games does not arise. In the case $\phi = 0$, this problem could in principle arise, as noted in Hopkins and Kornienko (2004). Suppose that all students play effort e^*, then, all students have the highest status, because $F(e^*) = 1$. In this scenario, no student has an incentive to decrease effort, because she would obtain a rank of 0 and, therefore, a utility of 0. However, no student has an incentive to increase effort either, because each student carries a weight equal to zero, therefore, she would not increase her status but her cost of effort would increase. As a result, all students exerting effort e^* could be an equilibrium, and this is true for multiple values of e, yielding multiple equilibria. Following Hopkins and Kornienko (2004), I rule out this case by assuming that individuals have a dislike for ties, so that rank can in fact be written as $\gamma F(y) + (1 - \gamma)F^-(x)$, where $\gamma \in [0, 1)$ and $F^-(y) = \lim_{y \to y^-} F(y)$ is the mass of individuals with consumption strictly less than x. As a result, when all students exert the same level of effort e^*, each student has an individual incentive to deviate and exert more effort, because this would increase her utility from rank by eliminating her tie with the other students. Therefore, I assume that $F(y)$ in the body of the paper in fact corresponds to $\gamma F(y) + (1 - \gamma)F^-(x)$, and uniqueness follows.

A.1.3 Comparative Statics

Definition Two distributions G_A, G_B with support on $[\underline{c}, \bar{c}]$ satisfy the Unimodal Likelihood Ratio (ULR) order, $G_A \succ_{ULR} G_B$, if the ratio of their densities $L(c) = g_A(c)/g_B(c)$ is strictly increasing for $c < \tilde{c}$ and strictly decreasing for $c > \tilde{c}$ for some $\tilde{c} \in [\underline{c}, \bar{c})$ and if $\mu_A \geq \mu_B$.

In particular, if B has the same mean but higher variance than A, then $G_A \succ_{ULR} G_B$. Define the cutoffs $\hat{c}^-$ and $\hat{c}^+$ as the extremal points of the ratio $(1 - G_A(c) + \phi)/(1 - G_B(c) + \phi)$ when $G_A \succ_{ULR} G_B$. It can be shown that these cutoffs are such that $\underline{c} < \hat{c}^- < \hat{c}^+ \leq \bar{c}$, and they can be conveniently interpreted as cutoffs that separate type categories.[43] Low c students are those with $c \in [\underline{c}, \hat{c}^-)$, middle c students as those with $c \in (\hat{c}^-, \hat{c}^+)$, and high c students as those with $c \in (\hat{c}^+, \bar{c}]$. The model has the following prediction:

[43]The proof is available upon request from the author. It is a modification of the proof in Hopkins and Kornienko (2004), there the c.d.f. functions and not their complement appear in the ratio.

Proposition A.1 (Adapted from Proposition 4 in Hopkins and Kornienko (2004)). *Suppose $e_A(c)$ and $e_B(c)$ are the equilibrium choices of effort for distributions G_A and G_B. If $G_A \succ_{ULR} G_B$ and $\mu_A = \mu_B$, then:*

- *If $\phi = 0$: $y(e_A(c))$ crosses $y(e_B(c))$ at most once. Moreover, $y(e_A(c)) > y(e_B(c))$ for all $c \in [\hat{c}^-, \bar{c})$ with a possible crossing on $[\underline{c}, \hat{c}^-)$.*
- *If $\phi > 0$: $y(e_A(c))$ crosses $y(e_B(c))$ at most twice. Moreover, $y(e_A(c)) < y(e_B(c))$ for all $c \in [\hat{c}^+, \bar{c})$ with a crossing in $(\tilde{c}, \hat{c}^+)$ so that $y(e_A(c)) > y(e_B(c))$ for all $c \in [\hat{c}^-, \tilde{c}]$, with a possible crossing on $[\underline{c}, \hat{c}^-)$.*

Proof The proof is a lengthy adaptation of the proof in Hopkins and Kornienko (2004). It is available from the author upon request.

Table 4: Estimated reconstruction costs by MSK-intensity category (Adobe constructions)

MSK Intensity	**Expected cost (USD)**	**Expected cost (USD) over average household monthly income**
V	20	0.04
V$\frac{1}{2}$	120	0.26
VI	220	0.49
VI $\frac{1}{2}$	950	2.10
VII	1,680	3.72
VII $\frac{1}{2}$	4,210	9.32
VIII	6,740	14.92
VIII $\frac{1}{2}$	10,270	22.73
IX	13,800	30.54

A.2 Earthquake Damages

A.3 Effect of Earthquake on Test Scores

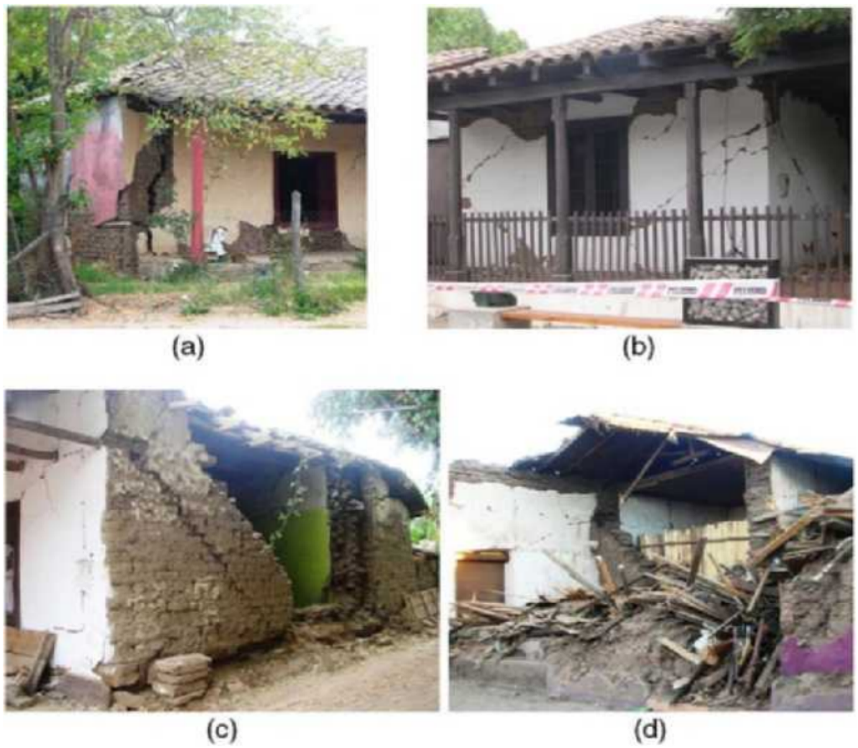


Figure 9: Typical damages to adobe structures and their corresponding grades of damage. (a) Vertical cracks at wall corner, G3; (b) diagonal crack in wall, G3; (c) wall collapse through out-of-plane, G4; (d) collapse of the roof, G5. *Source:* The picture and damage descriptions are reported from Astroza, Rui and Astroza (2012).

Table 5: Difference-in-differences evaluation of earthquake impact on test scores, dependent variables Spanish and Mathematics test scores in eighth grade

	Spanish				Mathematics			
	(1) Municipal	(2) Voucher	(3) Unsubsidized	(4) All Schools	(5) Municipal	(6) Voucher	(7) Unsubsidized	(8) All Schools
Spanish test score in fourth grade	0.656***	0.642***	0.595***	0.645***				
	(0.00249)	(0.00240)	(0.00651)	(0.00167)				
Math test score in fourth grade					0.635***	0.655***	0.646***	0.645***
					(0.00234)	(0.00231)	(0.00623)	(0.00159)
Household income (CLP)	3.61e-08**	2.86e-08***	1.51e-08	2.52e-08***	7.68e-08***	4.98e-08***	6.44e-08***	6.03e-08***
	(1.10e-08)	(6.38e-09)	(9.60e-09)	(4.70e-09)	(1.03e-08)	(5.91e-09)	(8.46e-09)	(4.35e-09)
Father's education (yrs)	0.00940***	0.00756***	0.0137***	0.00885***	0.00634***	0.00756***	0.0115***	0.00715***
	(0.000795)	(0.000782)	(0.00208)	(0.000538)	(0.000741)	(0.000724)	(0.00183)	(0.000497)
Mother's education (yrs)	0.0102***	0.00870***	0.0133***	0.00982***	0.00806***	0.00785***	0.0162***	0.00859***
	(0.000819)	(0.000830)	(0.00191)	(0.000557)	(0.000765)	(0.000768)	(0.00169)	(0.000516)
Female	0.134***	0.121***	0.101***	0.124***	-0.0999***	-0.110***	-0.0880***	-0.104***
	(0.00439)	(0.00415)	(0.0104)	(0.00290)	(0.00410)	(0.00385)	(0.00914)	(0.00268)
Household lives in earthquake region (E)	0.0725+	0.0431	0.0653	0.0566*	0.110**	0.00636	-0.0186	0.0395+
	(0.0417)	(0.0337)	(0.0766)	(0.0247)	(0.0390)	(0.0313)	(0.0677)	(0.0229)
Cohort 2007-2011, affected by earthquake (P)	0.0529***	0.0579***	0.0469+	0.0548***	0.0497***	0.0380***	0.0209	0.0429***
	(0.00815)	(0.00864)	(0.0271)	(0.00580)	(0.00755)	(0.00799)	(0.0238)	(0.00535)
P*E	-0.0377***	-0.0440***	-0.115***	-0.0492***	-0.0284**	-0.0633***	-0.0519*	-0.0509***
	(0.00969)	(0.00981)	(0.0291)	(0.00667)	(0.00900)	(0.00908)	(0.0256)	(0.00616)
Constant	-0.429***	-0.274***	-0.180*	-0.333***	-0.373***	-0.0999***	0.0526	-0.185***
	(0.0309)	(0.0283)	(0.0769)	(0.0200)	(0.0288)	(0.0263)	(0.0679)	(0.0185)
School Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes
Observations	97057	116446	20389	233892	97658	117011	20501	235170

Standard errors in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Pre-Earthquake Cohort, dependent variables Spanish and Math test scores in eighth grade

	(1) Spanish	(2) Math
Spanish test score in fourth grade	0.676***	
	(0.00408)	
Math test score in fourth grade		0.638***
		(0.00383)
Household income (CLP)	2.99e-08	9.66e-08***
	(1.82e-08)	(1.73e-08)
Father's education (yrs)	0.00874***	0.00584***
	(0.00126)	(0.00118)
Mother's education (yrs)	0.00731***	0.00559***
	(0.00130)	(0.00122)
Female	0.138***	-0.109***
	(0.00715)	(0.00674)
Own Intensity	0.0149	0.0146
	(0.0129)	(0.0120)
Classroom Intensity Mean	0.00943	0.0580
	(0.0688)	(0.0648)
Classroom Intensity Variance	0.223+	0.319**
	(0.128)	(0.120)
Constant	-0.472	-0.681+
	(0.425)	(0.400)
School Fixed Effects	yes	yes
Observations	35175	35302

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Only municipal schools are in the estimation sample.

Table 7: Probit regression, marginal probability estimates reported. Dependent variables: being a at top or bottom third of the distribution of elicited cost of effort

	top 33 percent	bottom 33 percent
Mother's education	.0007012	.0020148*
	(.0007415)	(.0008893)
Father's education	-.0007132	.0013843
	(.0007134)	(.0008544)
Household income	$1.76e-08^{+}$	$-2.32e-08^{+}$
	(1.02e-08)	(1.23e-08)
Math test score t-1	$-.0549353^{***}$	$.1004088^{***}$
	(.0019868)	(.0023838)
Seismic intensity at student's home	$.0128674^{***}$	$-.0105234^{***}$
	(.0022842)	(.002741)
Observations	46059	46059

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.4 Algorithm for the Estimation of the Semiparametric Single-Index Model

1. Normalize to a constant one of the elements of α_1, because only the ratios among the components of α are identified. I normalize to -1 the coefficient on initial ability (lagged test score).

2. Make an initial guess for all the other elements of α.

3. Form c_i $\forall i$ according to $c_i = \alpha_1 \mathbf{x_i}$ if i belongs to the pre-earthquake cohort, and $c_i = \alpha_1 \mathbf{x_i} + \alpha_2 I_i + \alpha_3 I_i \mathbf{x_i}$ if i belongs to the post-earthquake cohort. I_i is interacted with household income and student gender.

4. Estimate $E(y_i|c, l; \alpha)$ $\forall l$ by Nadaraya-Watson kernel regression with weights w_i:

$$\hat{m}_l(c; \alpha) = \frac{\sum_{i \in l} w_i K\left(\frac{c_i - c}{h}\right) y_i}{\sum_{i \in l} w_i K\left(\frac{c_i - c}{h}\right)}$$

   with a standard normal Kernel: $K(\psi) = (2\pi)^{-\frac{1}{2}} exp(-0.5\psi^2)$ and optimal bandwidth $h = 1.06\hat{\sigma}_c n^{-1/5}$, minimizing the Approximated Mean Integrated Squared Error (AMISE).[44] The weights w_i are such that only observations i where the p.d.f. of c at c_i exceeds a small positive number are used (see Ichimura (1993) and Horowitz (2010)). Observation i is excluded from the calculation of $\hat{m}$ at c_i.

5. Compute the sum of squared residuals in each l at the current guess for α: $SSR_l(\alpha) = \sum_{i \in l} w_i (y_i - \hat{m}_l(c_i; \alpha))^2$. The weights are the same as those used in the kernel estimator of m.

6. Update guess for α using Generating Set Search algorithm (HOPSPACK).

7. Repeat steps 1-6 until convergence to the minimizer of $\sum_l SSR_l(\alpha)$.

Notice that unlike in the standard semiparametric single-index model, here the $SSR(\alpha)$ is computed in each classroom l, and its sum over classrooms is minimized. The

[44] The MISE is equal to $E\{\int [\hat{m}(c) - m(c)]^2 dx\} = \int \left[(Bias\hat{m})^2 + V(\hat{m})\right] dc$, and AMISE substitutes the expressions for the bias and variance of $\hat{m}$ with approximations. See Pagan and Ullah (1999), p. 24.

dataset is clustered at the classroom level. While the functions m are allowed to differ by classrooms, the parameter α is restricted to be identical in all classrooms. To account for the clustered sample design in the estimation of the standard errors of the α parameters, I bootstrap 100 samples stratified at the classroom level, and I estimate α in each bootstrapped sample to obtain the standard errors.

The standard errors of $\widehat{\Delta e(\cdot)}$, which are needed to test the comparative statics result, cannot be easily bootstrapped for computational reasons.[45] Instead, I use the result in Ichimura (1993), who proves that the asymptotic variance of $\hat{m}_l(c)$ in the appropriately weighted semiparametric single-index model above is identical to the asymptotic variance of a non-parametric conditional mean estimator. The variance of such estimator is $V(\hat{m}_l(c)) = \frac{\sigma_l^2}{n_l h_l f_l(c)} \int K^2(\psi)d\psi + o(n^{-1}h^{-1})$, where σ^2 is the variance of ϵ_{il}, h_l is the bandwidth, n_l is the size of classroom l (on average this is almost 30), and $f_l(c)$ is the density at c in classroom l. The kernel $K(\cdot)$ is the normal kernel, resulting in $\int K^2(\psi)d\psi = 0.2821$. I estimate the asymptotic variance of $\hat{m}_l(c)$ $\forall l$ on a fine grid for c. I substitute $f(c)$ with its kernel estimator, and σ_l^2 with its estimator obtained by averaging the squared residuals in each classroom: $\hat{\sigma}_l^2 = \frac{\sum_{i \in l}(y_i - \hat{y}_i)^2}{n_l - 1}$. I assume that the covariances between the $\hat{m}_l(c)$ belonging to different classrooms l are zero $\forall c$, and I obtain the following expression for the variance of $\widehat{\Delta e(c)}$:

$$V\left(\widehat{\Delta e(c)}\right) = \sum_{l=1}^{N^{pre}-1} \sum_{l'=l+1}^{N^{pre}} \sum_{s=1}^{N^{post}-1} \sum_{s'=s+1}^{N^{post}} \kappa_{ll'ss'}^2 \left(V\left(\hat{m}_s^{post}(c)\right) + V\left(\hat{m}_{s'}^{post}(c)\right) + V\left(\hat{m}_l^{pre}(c)\right) + V\left(\hat{m}_{l'}^{pre}(c)\right)\right).$$

The weights $\kappa_{ll'ss'}$ are given by:

$$\kappa_{ll'ss'} = \frac{\omega_{ll'ss'}}{\sum_{l=1}^{N^{pre}-1} \sum_{l'=l+1}^{N^{pre}} \sum_{s=1}^{N^{post}-1} \sum_{s'=s+1}^{N^{post}} \omega_{ll'ss'}}$$

where $\omega_{ll'ss'}$ is defined in equation 9.

[45]This would require submitting around 4,000 jobs of duration 72 hours each.

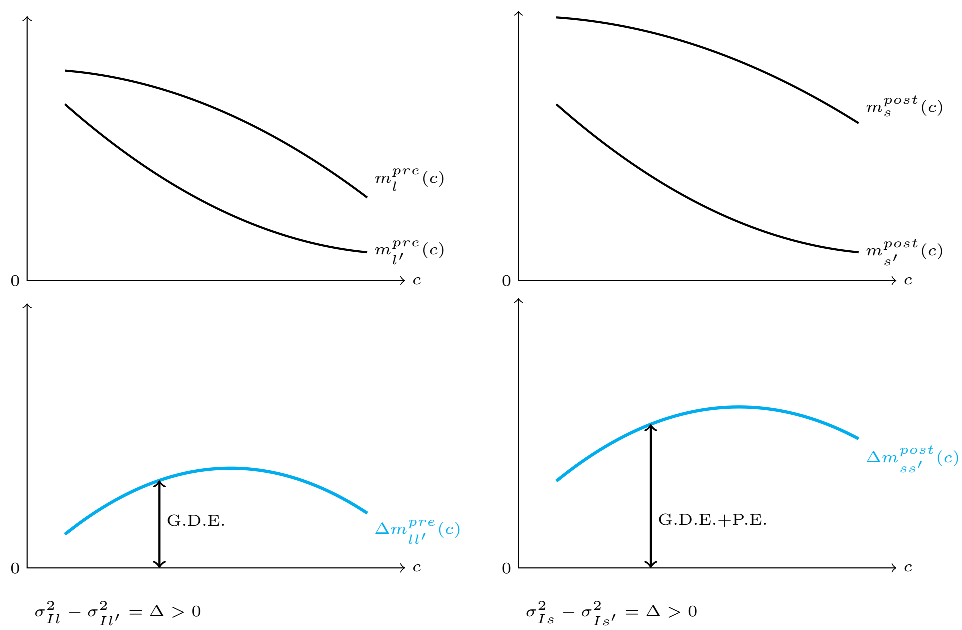


Figure 10: Classrooms l, l', s, and s' are identical in terms of classroom and student characteristics. G.D.E. is the effect of geographic dispersion, P.E. is the peer effect of interest.

A.5 Differencing technique

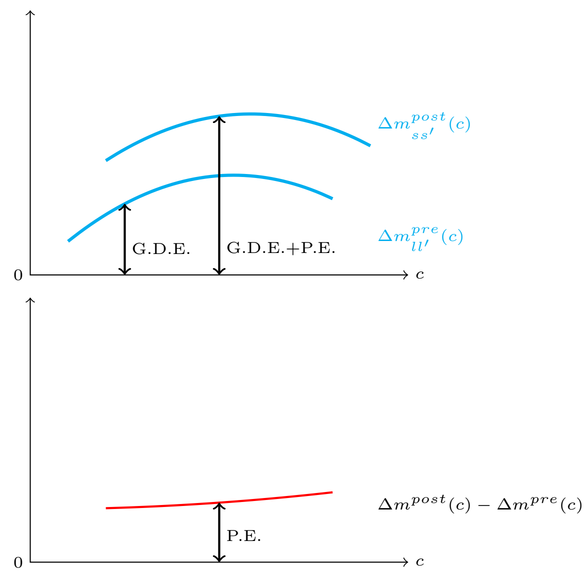


Figure 11: Differencing out the geographic dispersion effects. Notice that the difference between the Δm^{post} and Δm^{pre} functions can be taken only over the overlapping portion of the two domains. The domain of Δm^{post} is shifted to the right with respect to the domain of Δm^{pre}, because all students have been affected by the earthquake and, therefore, their cost of effort is larger than for pre-earthquake students.

Intuitively, for the kernel weighting method and dimensionality reduction to deliver the desired result, two sets of conditions must be met. First, $\psi_l(\cdot)$ and $u_l(\cdot)$ must vary smoothly with $\mathbf{z_l}, F_l(\mathbf{x_i}), I_l, \bar{I}_l$. This guarantees that there are no jumps in these functions when two classrooms that are similar, but not identical, in these variables are compared. Similarly, when two classrooms are similar but not identical in these variables, the distributions of c are similar but not identical. To avoid jumps in the $e(c; G_l(c))$ function, one needs to assume that this function varies smoothly with $G_l(c)$.[46] Second, it must be innocuous to consider only a finite number of moments rather than the entire distribution of student characteristics. Formally, I make the following additional assumptions:

DA.1 $u(c_i, \mathbf{z_l}, F_l(\mathbf{x_i}), I_l, \bar{I}_l, \sigma^2_{Il})$ and $\psi(c_i, \mathbf{z_l}, F_l(\mathbf{x_i}), I_l, \bar{I}_l, \sigma^2_{Il})$ are continuous in $\mathbf{z_l}, F_l(\mathbf{x_i}), I_l, \bar{I}_l$.

DA.2 $\phi(c_i, \mathbf{z_l}, F_l(\mathbf{x_i}), I_l, \bar{I}_l, \sigma^2_{Il})$ (where $\phi(\cdot) = u(\cdot) + \psi(\cdot)$) is similar to $\tilde{\phi}(c_i, \mathbf{z_l}, \mathbf{W_l}, I_l, \bar{I}_l, \sigma^2_{Il})$ $\forall (c, \mathbf{z_l}, \mathbf{W_l}, \sigma^2_{Il})$, where $\mathbf{W_l}$ is a vector containing the mean, variance, skewness and kurtosis of the elements of $\mathbf{x_i}$.[47].

DA.3 $\tilde{\phi}(\cdot)$ is continuous in $\mathbf{z_l}, I_l, \bar{I}_l$ and $\mathbf{W_l}$.

DA.4 If two classrooms in the pre-earthquake cohort are such that $\mathbf{W_l} \cong \mathbf{W_{l'}}$, then $e_l(c) \cong e_{l'}(c)$ $\forall c$. If two classrooms in the post-earthquake cohort are such that $[\mathbf{W_s}, \bar{I}_s, I_s, \sigma^2_{Is}] \cong [\mathbf{W_{s'}}, \bar{I}_{s'}, I_{s'}, \sigma^2_{Is'}]$, then $e_s(c) \cong e_{s'}(c)$ $\forall c$.

These assumptions allow me to build an approximated counterpart to equation 6. Assumption DA.1 implies that if two classrooms are similar, then the classroom effects are also similar for every student type c. Assumptions DA.2 and DA.3 together mean that if two classrooms are similar, then $\tilde{\phi}(c_i, \mathbf{z_l}, \mathbf{W_l}, I_l, \bar{I}_l, \sigma^2_{Il}) - \tilde{\phi}(c_i, \mathbf{z_{l'}}, \mathbf{W_{l'}}, I_{l'}, \bar{I}_{l'}, \sigma^2_{Il'}) \cong \Delta\phi_{ll'}(c)$. That is, the difference between the $\tilde{\phi}$ functions computed in classrooms with similar vectors $\mathbf{W}$ and $\mathbf{z}$ is a good approximation to the spurious effect in equations 4 and 5. Finally, assumption DA.4 means that when the distribution of c varies in a classroom, if this affects $e_l(c)$ because there are peer effects, then $e_l(c)$ varies smoothly with the change in the distribution of c. As a result of these four assumptions, for any two classrooms l, l' in the pre-earthquake cohort with $\mathbf{W_l} \cong \mathbf{W_{l'}}$, $\mathbf{z_l} \cong \mathbf{z_{l'}}$, $I_l \cong I_{l'}$, $\bar{I}_l \cong \bar{I}_{l'}$ and $\sigma^2_{Il} - \sigma^2_{Il'} = \Delta\sigma^2_{Ill'} < 0$, I can build the approximated counterpart to $\Delta m^{pre}_{ll'}(c)$, $\tilde{\Delta} m^{pre}_{ll'}(c)$. Similarly, for two similar classrooms in the post-earthquake cohort, I can build the approximated counterpart to $\Delta m^{post}_{ss'}(c)$, $\tilde{\Delta} m^{post}_{ss'}(c)$. If four

[46]In Ahn and Powell (1993), this assumption corresponds to continuity of the selection function (see page 9 in their paper).

[47]I ignore the cross-moments of $\mathbf{x_i}$ to make the numerical implementation tractable.

classrooms, l, l' from the pre-earthquake cohort and s, s' from the post-earthquake cohort, are such that $\mathbf{W_l} \cong \mathbf{W_{l'}} \cong \mathbf{W_s} \cong \mathbf{W_{s'}}$, $\mathbf{z_l} \cong \mathbf{z_{l'}} \cong \mathbf{z_s} \cong \mathbf{z_{s'}}$, $\mathbf{I_l} \cong \mathbf{I_{l'}} \cong \mathbf{I_s} \cong \mathbf{I_{s'}}$, $\mathbf{\bar{I}_l} \cong \mathbf{\bar{I}_{l'}} \cong \mathbf{\bar{I}_s} \cong \mathbf{\bar{I}_{s'}}$ and $0 > \Delta\sigma^2_{Ill'} \cong \Delta\sigma^2_{Iss'} < 0$, then subtracting $\tilde{\Delta} m^{pre}_{ll'}(c)$ from $\tilde{\Delta} m^{post}_{ll'}(c)$ yields an approximation to $\Delta e_{ss'}(c) + \Delta\xi_{ill'ss'}$, the quantity of interest.

To ensure that the classrooms are similar, I assign increasing weights to quadruples that are more similar in terms of $\mathbf{W}$, $\mathbf{z}$, I, $\bar{I}$ and $\Delta\sigma^2_I$. I construct weights using multivariate standard normal kernel functions. As before, let ll' index a pre-earthquake classroom pair, and ss' a post-earthquake classroom pair. Letting $Z_t = [\mathbf{W_t}, \mathbf{z_t}, I_t, \bar{I}_t]$ for $t = l, l', s, s'$, I assign the weight $\frac{1}{h} k\left(\frac{Z_t - Z_{t'}}{h}\right)$ to each of the pairs $tt' \in \{ll', ss', sl\}$. This ensures that the pairs within the pre- and post-earthquake cohorts are composed of classrooms that are similar to each other in terms of $\mathbf{W}$ and $\mathbf{z}$ ($tt' = ll', ss'$), and also that across cohorts the two pairs of classrooms are similar ($tt' = sl$).[48] Finally, I build a weight that is declining in $|\Delta\sigma^2_{Ill'} - \Delta\sigma^2_{Iss'}|$, to guarantee that the pre- and post-earthquake pairs differ in terms of σ^2_I in a similar way, i.e., that the intensities of treatment in the two pairs of classrooms are very similar: $\frac{1}{h_{\Delta\sigma}} k\left(\frac{\Delta\sigma^2_{Ill'} - \Delta\sigma^2_{Iss'}}{h}\right)$. The weight for the quadruple, $\omega_{ll'ss'}$, is the product of these four kernel weights:

$$\omega_{ll'ss'} = d_{ll'ss'} \frac{1}{h_{\Delta\sigma}} k\left(\frac{\Delta\sigma^2_{Ill'} - \Delta\sigma^2_{Iss'}}{h}\right) \prod_{tt' \in \{ll', ss', sl\}} \frac{1}{h} k\left(\frac{Z_t - Z_{t'}}{h}\right) \tag{9}$$

where $d_{ll'ss'}$ is a dummy variable equal to one if $\Delta\sigma^2_{Ill'} < 0$ and $\Delta\sigma^2_{Iss'} < 0$. For each value of c, the estimator of $\Delta e_{ss'}(c)$ is obtained by estimating γ in the following weighted regression on a constant:

$$(\Delta\widehat{m^{post}_{ss'}} - \Delta\widehat{m^{pre}_{ll'}}) = \omega_{ll'ss'}\gamma + \xi_{ll'ss'}$$

with $E[\xi_{ll'ss'}] = 0$. The OLS estimator of γ is the following weighted sample mean:

$$\hat{\gamma} = \widehat{\Delta e}(c) = \frac{\sum_{l=1}^{N^{pre}-1} \sum_{l'=l+1}^{N^{pre}} \sum_{s=1}^{N^{post}-1} \sum_{s'=s+1}^{N^{post}} \omega_{ll'ss'} (\Delta\widehat{m^{post}_{ss'}}(c) - \Delta\widehat{m^{pre}_{ll'}}(c))}{\sum_{l=1}^{N^{pre}-1} \sum_{l'=l+1}^{N^{pre}} \sum_{s=1}^{N^{post}-1} \sum_{s'=s+1}^{N^{post}} \omega_{ll'ss'}}. \tag{10}$$

The parameter γ computed at a value for c yields the value of the $\Delta e(\cdot)$ function at one point. Estimating γ over a grid of values for c allows me to trace the behavior of

[48] Notice that I use a unique bandwidth. Following Pagan and Ullah (1999), I normalize the variables Z_t so that they all have the same standard deviation, and using a unique bandwidth is admissible.

this function over its domain, and to detect how the effects of changing the variance of student types (peer effects) vary across student types. Computing $\hat{\gamma}$ at each grid point requires doing a number of calculations of the order of 10^{12}, therefore, parallel processing is required.

A.6 Identifying Assumption

Table 8: No Evidence that Intensity Variance Affected Classroom Composition

	(1) Mean Spanish Lagged Test Score	(2) Mean Math Lagged Test Score	(3) Mean hh income (CLP)	(4) Mean father's education (yrs)	(5) Mean mother's education (yrs)
School is in	-0.282***	-0.358***	-48529.1***	1.050***	0.538***
earthquake region (E)	(0.0364)	(0.0377)	(9599.6)	(0.150)	(0.140)
Cohort 2007-2011,	0.0241	-0.0370*	22318.3***	0.0541	0.0393
affected by earthquake (P)	(0.0152)	(0.0157)	(3987.5)	(0.0622)	(0.0583)
(1-P)*Classroom Intensity	0.0388***	0.0522***	2086.0	-0.216***	-0.102***
Mean, θ_μ^{pre}	(0.00598)	(0.00618)	(1581.7)	(0.0247)	(0.0231)
P*Classroom Intensity	0.0334***	0.0532***	-306.2	-0.218***	-0.101***
Mean, θ_μ^{post}	(0.00599)	(0.00619)	(1572.6)	(0.0245)	(0.0230)
Effect of Classroom	-0.0054516+	0.0009108	-2392.183**	-0.0020608	0.0014217
Intensity Mean, $\theta_\mu^{post} - \theta_\mu^{pre}$	(0.0030045)	(0.0031042)	(787.414)	(0.0122813)	(.0115087)
(1-P)*Classroom Intensity	-0.0373	-0.0306	8081.1	-1.526**	-1.228**
Variance, $\theta_{\sigma^2}^{pre}$	(0.119)	(0.123)	(31829.6)	(0.496)	(0.465)
P*Classroom Intensity	0.174*	0.0601	2292.5	-0.827*	-0.326
Variance, $\theta_{\sigma^2}^{post}$	(0.0842)	(0.0871)	(22021.0)	(0.343)	(0.322)
Effect of Classroom	0.210873	0.0907525	-5788.624	0.6992537	0.9024236
Intensity Variance, $\theta_{\sigma^2}^{post} - \theta_{\sigma^2}^{pre}$	(0.1456232)	(0.1505045)	(38715.2)	(0.6033702)	(0.5656727)
Constant	-0.122***	-0.135***	234675.0***	9.342***	9.223***
	(0.0108)	(0.0111)	(2861.5)	(0.0446)	(0.0419)
Observations	10477	10480	10086	10077	10083

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: No Evidence that the Relationship Between Classroom Intensity Variance and Classroom Quality Changed After the Earthquake Struck (Spanish teachers)

	(1)	(2)	(3)	(4)	(5)
			Spanish Teacher		
	Class size	Permanent contract	Experience	Postgrad degree	Female
School is in earthquake region (E)	1.784+	0.137	0.273	0.0761	-0.344**
	(0.913)	(0.114)	(1.247)	(0.113)	(0.129)
Cohort 2007-2011, affected by earthquake (P)	-2.269***	0.203***	-0.536	0.336***	0.0560
	(0.381)	(0.0482)	(0.547)	(0.0478)	(0.0565)
(1-P)*Classroom Intensity Mean, θ_{μ}^{pre}	-0.278+	-0.0058	-0.121	0.0297	0.0666**
	(0.150)	(0.0188)	(0.207)	(0.0187)	(0.0214)
P*Classroom Intensity Mean, θ_{μ}^{post}	-0.324*	-0.0309+	-0.105	0.0094	0.0697**
	(0.150)	(0.0186)	(0.204)	(0.0185)	(0.0212)
Effect of Classroom Intensity Mean, $\theta_{\mu}^{post} - \theta_{\mu}^{pre}$	-0.0457	-0.0252**	0.0162	-0.0203*	0.0031
	(0.0749)	(0.0095)	(0.1065)	(0.0094)	(0.0111)
(1-P)*Classroom Intensity Variance, $\theta_{\sigma^2}^{pre}$	-7.374*	-0.240	-5.653	0.585	-0.331
	(2.936)	(0.374)	(4.121)	(0.400)	(0.416)
P*Classroom Intensity Variance, $\theta_{\sigma^2}^{post}$	-9.218***	0.0429	1.255	-0.105	-0.379
	(2.095)	(0.275)	(2.867)	(0.266)	(0.283)
Effect of Classroom Intensity Variance, $\theta_{\sigma^2}^{post} - \theta_{\sigma^2}^{pre}$	-1.8437	0.2832	6.9077	-0.69014	-0.0476
	(3.6082)	(0.4645)	(5.0200)	(0.4807)	(0.5037)
Constant	25.60***	0.139**	22.38***	-0.146**	0.780***
	(0.381)	(0.0487)	(0.563)	(0.0485)	(0.0576)
Observations	10339	9128	8358	9128	8529

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: No Evidence that the Relationship Between Classroom Intensity Variance and Classroom Quality Changed After the Earthquake Struck (Math Teachers)

	(6)	(7)	(8)	(9)
	Math Teacher			
	Permanent Contract	Experience	Postgrad degree	Female
School is in earthquake region (E)	0.171	1.307	-0.319**	-0.0557
	(0.113)	(1.230)	(0.113)	(0.114)
Cohort 2007-2011, affected by earthquake (P)	0.265***	0.329	0.307***	0.0398
	(0.0482)	(0.543)	(0.0476)	(0.0502)
(1-P)*Classroom Intensity Mean, θ_{μ}^{pre}	-0.0045	-0.249	0.0847***	0.0267
	(0.0187)	(0.204)	(0.0187)	(0.0188)
P*Classroom Intensity Mean, θ_{μ}^{post}	-0.0420*	-0.415*	0.0735***	0.0146
	(0.0185)	(0.201)	(0.0186)	(0.0186)
Effect of Classroom Intensity Mean, $\theta_{\mu}^{post} - \theta_{\mu}^{pre}$	-0.0375***	-0.1655	-0.0112	-0.0121
	(0.0094)	(0.1053)	(0.0094)	(0.0097)
(1-P)*Classroom Intensity Variance, $\theta_{\sigma^2}^{pre}$	-0.425	-7.541+	0.743+	-0.574
	(0.370)	(4.100)	(0.419)	(0.378)
P*Classroom Intensity Variance, $\theta_{\sigma^2}^{post}$	-0.142	-0.170	0.220	-0.122
	(0.261)	(2.750)	(0.259)	(0.257)
Effect of Classroom Intensity Variance, $\theta_{\sigma^2}^{post} - \theta_{\sigma^2}^{pre}$	0.2826	7.3715	-0.5231	0.4523
	(0.4529)	(4.9383)	(0.4925)	(0.4573)
Constant	0.224***	23.25***	-0.0611	0.0621
	(0.0484)	(0.552)	(0.0480)	(0.0511)
Observations	9229	8508	9229	8623

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.7 Testing Monotonicity of $m(c)$

The procedure that I use is an application of Chetverikov (2013). It would be computationally unfeasible to perform the test in all classrooms. Therefore, I create 72 categories of classrooms that have similar distributions of c, and test monotonicity within each category. Consider classroom categories containing approximately 60 classrooms each. Classrooms in the same category share similar mean and variance of c. Therefore, the theoretical model predicts that the equilibrium e functions should be very similar across classrooms within each category. The monotonicity of m is tested within each one of these categories. Separating the sample in categories makes this procedure feasible from a computational point of view. In all categories, the null hypothesis that the m function is decreasing is not rejected at the $\alpha = 0.10$ significance level.

Details of the simulations can be found in the Online Supplementary Material on the author's webpage. I follow the choice of bandwidth recommended in Ghosal, Sen, and Van Der Vaart (2000), and I adopt the plug-in approach to simulate the critical values.

An important distinction with Chetverikov (2013) is that the c_i values in my sample are estimated (and not observed); $\hat{c}_i = \hat{\alpha}_1 \mathbf{x_i} + \hat{\alpha}_2 I_i P_i + \hat{\alpha}_3 I_i P_i \mathbf{x_i}$. However, this additional noise is asymptotically negligible because the bandwidth used in the kernel weighting functions goes to zero as the sample size increases, and because $\hat{\alpha}$ is root-n consistent (as shown in Ichimura (1993)), therefore, it is faster than the nonparametric rates appearing in the derivations in Chetverikov (2013).

A.8 Robustness

Table 11: Difference-in-differences evaluation of the effect of the seismic intensity mean and variance on coverage of the Spanish curriculum

	% of Spanish curriculum covered
Spanish teacher experience	-0.000137
	(0.000211)
School is in earthquake region (E)	-0.000161
	(0.0151)
Cohort 2007-2011, affected by earthquake (P)	0.0283***
	(0.00675)
(1-P)*Classroom Intensity Mean	-0.00149
	(0.00248)
P*Classroom Intensity Mean	-0.00110
	(0.00246)
Effect of Earthquake Intensity Mean	0.0003955
	(0.0013076)
(1-P)*Classroom Intensity Variance	0.0192
	(0.0574)
P*Classroom Intensity Variance	0.0415
	(0.0346)
Effect of Earthquake Intensity Variance	0.0222701
	(0.0670429)
Constant	0.661***
	(0.00874)
Observations	6438

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other included regressors: Spanish teacher characteristics (tenure at school, type of contract, possession of postgraduate degree, gender), classroom characteristics (class size, and mean and variance of lagged test scores in Math and Spanish).

Table 12: Difference-in-Differences Evaluation of the Effect of Seismic Intensity Variance on Test Scores. Dependent variables Spanish and Math test scores in eighth grade.

	(1) Math	(2) Spanish
Math test score in 4^{th} grade	0.638***	
	(0.00254)	
Spanish test score in 4^{th} grade		0.657***
		(0.00272)
[1em] Household lives in earthquake region (E)	0.338*	0.0874
	(0.133)	(0.143)
Cohort 2007-2011, affected by the earthquake (P)	0.0371***	0.0528***
	(0.00936)	(0.00997)
(1-P)*Classroom Intensity Mean	0.0753	-0.193+
	(0.102)	(0.111)
P*Classroom Intensity Mean	0.0676	-0.191+
	(0.102)	(0.111)
Effect of Intensity Mean	-0.0076943	0.0020984
	(0.0158669)	(0.0169411)
(1-P)*Classroom Intensity Variance	-0.113	0.238*
	(0.101)	(0.108)
P*Classroom Intensity Variance	-0.0243	-0.0171
	(0.0633)	(0.0719)
Effect of Intensity Variance	0.0885604	-0.2549768^{*}
	(0.1149763)	(0.124782)
Constant	-0.675	0.331
	(0.449)	(0.484)
Observations	83295	81307

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other included regressors: household income, seismic intensity at home interacted with cohort dummy, father's education, mother's education, student gender, teacher gender, teacher experience, class size.