Heterogeneous Peer Effects and Rank Concerns: Theory and Evidence

Michela M. Tincani

November 13, 2014
Job Market Paper
Please click here for most recent version

1Tincani: University College London, 30 Gordon Street, London, WC1H0BE, UK, m.tincani@ucl.ac.uk. I would like to thank Orazio Attanasio, Antonio Cabrales, Denis Chetverikov, Martin Cripps, Flavio Cunha, Mariacristina De Nardi, Christian Dustmann, Jan Eeckhout, Hannie Fang, David Figlio, Ed Hopkins, Tatiana Kornienko, Dennis Kristensen, Stephen Machin, Costas Meghir, Lars Nesheim, Imran Rasul, Nikita Roketskiy, Uta Schoenberg, Aureo de Paula, Petra Todd, Martin Weidner, Daniel Wilhelm, Ken Wolpin and participants at the Empirical Micro Seminar at Penn and at various conferences for insightful comments and discussions. I thank Julia Schmieder for excellent research assistance, John McGlynn and Kyle Small for assisting me with the computer cluster, and Rodrigo Astroza for helping me understand how earthquake damage propagates. Research funding from the Centre for Microdata Methods and Practice 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

While several studies have found that classroom composition affects student test scores, what gives rise to these peer effects is still poorly understood. This paper makes two contributions. First, it proposes a new method to identify peer effects and it implements it to estimate the impact of changing the variance of peer ability on the test scores of Chilean eighth graders. The method can detect heterogeneity of the treatment effect across students of different ability, while imposing minimal restrictions on the data. It combines semi-parametric techniques with a difference-in-differences approach that exploits the fact that the 2010 Chilean earthquake affected households differently depending on their distance from the rupture. I find that the effects are heterogeneous across students as well as subjects. These patterns are known to be hard to rationalize with standard models of peer effects. Therefore, as a second contribution this paper proposes a theoretical model of social interactions that explains those patterns as arising from rank concerns, that is, the concerns of students for their achievement relative to the achievement of their peers. The model could also explain similar heterogeneous effects found by Carrell, Sacerdote and West (Econometrica, 2013). A novel implication of the analysis is that policies that create homogeneous classrooms in terms of ability, like tracking, may improve the test scores of low-ability students rather than worsen them.
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. While a very large number of papers have examined the identification and quantification of peer effects, the mechanisms underlying them are still poorly understood. As a result, we are, for example, unable to predict the outcomes of regrouping students across peer groups using current empirical models (Carrell, Sacerdote, and West 2013).

This paper makes two contributions to the literature. First, it proposes a new method to identify peer effects that operate through the variance of peer ability, which are found to have heterogeneous impacts on the test scores of Chilean students. Second, to rationalize these findings and generalize them to a broader context, it proposes a theoretical model of social interactions where peer effects arise because of a preference for rank in terms of achievement, thus isolating a possible mechanism. The analysis has important implications for policy and for the estimation of peer effects.

The key empirical problem in the identification of peer effects with observational data is that classrooms with different distributions of student characteristics, for example, parental income, are different also in other unobserved ways that make it impossible to separate peer effects from other confounding effects, i.e., from correlated effects (Manski 1993). For example, certain types of teachers could be assigned to classrooms with a large proportion of disadvantaged students. Experiments with random assignment to classrooms are rare, especially beyond first grade, as noted in the survey by Epple and Romano (2011). I adopt a different approach.

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 and on the geological features of their location. To the extent that damage to a student’s home affected a student’s ability to study, changes to the distribution of damages within a classroom implied changes to the distribution of the students’ ability to study. I show that the classroom variance of earthquake damages is not systematically related to classroom characteristics nor to student or teacher reshuffling across classrooms, and I find strong evidence that teachers did not react to changes in this variance by, for example, adapting their focus of instruction. Therefore, the earthquake’s damage propagation provides useful variation to identify peer effects that operate through the variance of students’ ability to study.
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 allows me to exploit this variation to identify peer effects. For this purpose, I combine a semiparametric single-index model with a kernel-weighted difference-in-differences estimator. The model has several desirable features. First, the single-index model allows me to estimate the impact of seismic intensity at a student’s home on the student’s ability to study, modeled as a function of student characteristics, like lagged test scores and parental income. I refer to this scalar as a student’s type.

Second, the semiparametric econometric model imposes minimal assumptions on the technology of test score production and on how peer effects operate. Moreover, I make no assumptions on the shape of peer effects, which are broadly defined. There are peer effects if changing the distribution of student types in the classroom changes the mapping between an individual’s type and her test scores, i.e., two identical students may obtain different test scores in two classrooms that are identical except for the composition of peers. As a result of this model’s feature, the empirical findings are not driven by functional restrictions. In addition, the model can detect any pattern of heterogeneity in the peer effects across students, because they are obtained as a non-parametric function of student type.

Third, the difference-in-differences approach accounts for any spurious effects that the variance of seismic intensity may have on test scores. Classrooms that attract students who reside further from the school must have some desirable characteristics unobserved to the econometrician, but they also have larger measured seismic intensity variances, because this measure is based on students’ home locations. This could confound the peer effect estimates. For this reason, I jointly estimate the effect of changing seismic intensity variance on the post-earthquake cohort of students (2007-2011), affected by the earthquake, and on the pre-earthquake cohort (2005-2009), not affected by the earthquake. Any effect detected on the pre-earthquake cohort is spurious, and
it is differenced out from the post-earthquake effect. In estimation, the effects of increasing the variance of seismic intensity are obtained by performing a large number of pairwise comparisons among classrooms that have different seismic intensity variances, but that are otherwise very similar along many classroom characteristics, including classroom composition. This is achieved by kernel-weighting, in the spirit of Ahn and Powell (1993).

As a preliminary data analysis, I provide the first evaluation of the impact of the Chilean earthquake on student test scores. By exploiting the rich longitudinal data structure, and the fact that some Chilean regions were not affected by the earthquake, I estimate that being exposed to the earthquake reduced test scores by 0.05 sd, and this estimate is statistically significant at the 0.001 level. Moreover, every USD 100 in earthquake damages is accompanied by a reduction of 0.016 sd in test scores, also significant at the 0.001 level.

The main empirical finding is that the variance of peer types has a negative impact on some students, while benefitting others, and that these patterns are different for Mathematics and for Spanish test scores. While middle-ability students are always harmed by an increase in variance, and low-ability students always benefit, high-ability students in Spanish classes are harmed, whereas high-ability students in Mathematics classes benefit from the change. These patterns are hard to rationalize with standard models of peer effects.

As a second contribution I present a theoretical model of social interactions that can explain the empirical findings as arising from rank concerns. The model allows for standard technological spill-overs operating through the mean of peer ability, that have appeared in several papers in the literature. In addition, it allows for rank concerns among students. If students compete for grades, how much effort they exert depends on who their peers are. The model explains the observed outcomes as arising from the interplay of two incentives. The first one is an incentive to compete, and it predicts that students have a stronger incentive to exert effort in classrooms with more similarly able students, because improving their rank requires less effort. The second is an incentive to “rest on one’s laurels”, and it depends on the competition that a student faces from the students who are less able than her. Intuitively, if lower-ability students are exerting less effort, or if they are more numerous, then more able students

---

1This is obtained from a difference-in-differences model with continuous treatment intensity, in the spirit of Card (1992).

2See, for example, Arnott and Rowse (1987), Epple and Romano (1998), Epple and Romano (2008).
can reduce effort without risking a lower rank in achievement. While the net effect of these two incentives for low- and middle-ability students is the same across subjects, the incentive to compete prevails over the incentive to rest on one’s laurels in Mathematics classes, and the opposite is true in Spanish classes. The model can generate this finding whenever Mathematics and Spanish skills are not perfectly correlated.

This parsimonious theoretical model can explain the rich patterns of heterogeneity of the effects in a simple and intuitive way. Statistical tests of the model’s comparative statics result, as well as of another testable feature of the model, do not reject any of the model’s implications.

This is the first paper to analyze the interaction between rank concerns and changes to classroom composition. My findings have important implications for policy and for the estimation of peer effects. The assumption typically made in linear social interaction models, the workhorse of empirical research (Blume, Brock, Durlauf, and Jayaraman 2014), is that technological spill-overs do not work through the variance of peer ability. Under this restriction, if policies like tracking students by ability are found to be beneficial to the low-ability students, who are tracked with other low-ability students, then their impact must be ascribed to other factors, for example, teachers. However, if students have rank concerns, then tracking can be beneficial to low-ability students even in the absence of a teacher reaction. In fact, Tincani (2014) shows that this mechanism could explain why Carrell, Sacerdote, and West (2013) found that middle-ability students improved their test scores when tracked into homogeneous peer groups. A reduced-form model that includes ability variance as a regressor is able to capture the incentives that arise in a competitive environment remarkably well (Tincani 2014).

It would be very helpful to collect data that allow researchers to credibly measure rank concerns, either directly through the elicitation of preferences, or indirectly through randomized control trials that manipulate the incentives to compete. This knowledge could then be used to study the optimal organization of classrooms.

2 Literature Review

My paper is related to various strands in the literature. Various studies solve the problem of correlated effects (Manski 1993) by using data with random allocation of students to dorms. See, for example, Sacerdote (2001), Zimmerman (2003), Stinebrick-
ner and Stinebrickner (2006), Kremer and Levy (2008), and Garlick (2014). In contrast, very few randomized experiments with allocation to classrooms exist. Duflo, Dupas, and Kremer (2011) conduct an experiment with random assignment of Kenyan first graders. Whitmore (2005) uses data from the project STAR experiment in Tennessee, that randomly allocated kindergartners to classrooms. Finally, Carrell, Sacerdote, and West (2013) conduct an experiment that changes the allocation of students to squadrons in the U.S. Air Force Academy.

Previous studies in the Economics of Education literature have used data from earthquakes and natural disasters. They have typically viewed a natural disaster as a binary event that affects some individuals and does not affect others. For example, Cipollone and Rosolia (2007) identify the effect of the schooling achievement of young men on those of young women by exploiting the fact that boys living in areas affected by an earthquake in 1980 were more likely to graduate from high school than boys unaffected by the earthquake, because they were exempted from compulsory military service. In a recent paper, Imberman, Kugler, and Sacerdote (2012) exploit the forced relocations of students affected by hurricanes Katrina and Rita in 2005 to examine the impact of evacuees newly enrolled in a school on incumbent students. In a previous paper, Sacerdote (2008) examines the heterogeneous impact of such forced school relocations on different students, concluding that for students originally in the lowest performing schools there are long term gains from switching school that offset the costs of the disruptions.

Rank concerns have received much attention in contexts outside of education. Studies in the labor and personnel literatures have analyzed the importance of relative pay for job satisfaction (Clark and Oswald 1996, Card, Mas, Moretti, and Saez 2012, Hamermesh 2001, Brown, Gardner, Oswald, and Qian 2008), turnover decisions (Kwon and Meyerson Milgrom 2009), effort (Clark, Masclet, and Villeval 2010), happiness (Easterlin 1974, Frey and Stutzer 2002, Luttmer 2005, J Solnick and Hemenway 1998), and health (Marmot 2004). Relative productivity has been found to affect behavior (Blanes i Vidal and Nossol 2011), and research has shown that concerns for equity explain evidence from various laboratory experiments (Fehr and Schmidt 1999). Recent experimental evidence demonstrated an aversion to having a low rank (Kuziemko,

---

3See also Kang (2007), who studies peer effects among 7th and 8th graders in Korea using a quasi-randomization.

4Tincani (2014) shows that a special case of the model of rank concerns presented in this paper can explain all their experimental results.
Buell, Reich, and Norton 2014). Rank concerns have been studied also in the literature on tournaments (Brown 2011, Genakos and Pagliero 2012). There exists a large theoretical literature on rank concerns dating back to at least the seminal work of Veblen (1899), and including models where individuals have preferences for status, defined by the ordinal rank in the distribution of consumption, income or wealth (Duesenberry 1949, Frank 1985, Robson 1992, Cooper, Garcia-Penalosa, and Funk 2001, Hopkins and Kornienko 2004). Finally, the neurological basis for this kind of preferences has been shown in studies with non-human primates, that find a positive correlation between rank and serotonin in the blood (Raleigh, McGuire, Brammer, and Yuwiler 1984, Zizzo 2002).

In contrast, only a limited number of papers have addressed competition and rank concerns in education. The idea that students might react to the composition of their peers because of rank concerns dates back to at least the concept of relative deprivation in education presented in Jencks and Mayer (1990). However, how peer group composition interacts with rank concerns to determine student outcomes has not been examined so far. To the best of my knowledge, the only other paper with a theoretical model of competition between students is Azmat and Iriberri (2010), where competition enters as a preference for having above average achievement. Unlike the model in this paper, their model is not used to predict the outcome of changing classroom composition. Using data on Spanish high-schoolers, they do not reject their model implications and conclude that students have rank concerns. Similarly, Tran and Zeckhauser (2012) test for the presence of rank concerns among Vietnamese university students and find support for them. Murphy and Weinhardt (2014) find that being highly ranked during primary school has large positive effects on secondary school achievement. Banerjee and Besley (1990) build a theoretical model with an informational externality, where observing own performance with respect to one’s peers is informative on own ability. While preferences are not directly increasing in one’s rank in their model, the externality introduces peer effects that operate through the entire distribution of outcomes in the peer group. They show that multiplicity of equilibria could explain why identical populations can be found exerting different levels of effort. Finally, research in psychology and education shows that students have competitive attitudes that increase with age (Madsen 1971, Johnson and Johnson 1974).

5 See also Cole, Mailath, and Postlewaite (1992).
6 See also Tao and Lee (2014), who develop an instrumental variable estimator for a social interaction model with an extreme order statistics. However, they do not analyze applications to education.
3 Data and Earthquake

This section describes the data sources and it explains how seismic intensity is calculated. In addition, it describes the impact that the exposure to the earthquake and the degree of exposure (seismic intensity) had on Chilean students’ test scores. This is the first evaluation of the impact of the Chilean 2010 earthquake on student outcomes.

3.1 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 8th grade students, for whom I have information on current and 4th grade Math and Spanish standardized test scores, father’s and mother’s education, household income, gender, and town of residence. 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 the school and at the student’s home. Classroom level information includes class size, and teaching experience, education, tenure at the school, gender, and type of contract for both Spanish and Math teachers.

One cohort is observed in the 8th grade in 2009, before the 2010 earthquake, while the other cohort is observed in the 8th grade in 2011, after the earthquake. Figure 1 illustrates the data time-line. Observing test scores twice for the same student allows me to use lagged test scores as a measure of initial student ability.

In addition, for both cohorts the dataset contains information on the amount of curriculum that Spanish teachers were able to cover. The amount of material covered in class is typically not observed in large, nationally representative datasets like the SIMCE. Additional information is available for students observed after the earthquake. This includes the number of books read by the students, and the student’s reported cost of exerting effort.

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

---

7This variable is not available for the Mathematics curriculum due to incompatibility of the survey questions across waves.
closed as a consequence of the earthquake. I use this list to identify evacuees (i.e., the students who attended them and had to change school), 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 the changes to classroom composition induced by the arrival of evacuees. 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.

When estimating the empirical model, I restrict attention to public school students, to account for housing quality as explained below. The resulting sample contains observations on 110,822 students, 2,579 schools and 3,712 classrooms.

3.2 Earthquake

Just a few days before the start of the new school year, on February 27th 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 lasted for about three minutes, it was felt strongly throughout 500 km (over 300 miles) along the country, covering six regions that together make up about 80 percent of the country’s population. In the specialized literature, it is referred to as a mega-earthquake.

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 USD 2.5 billion worth of subsidies. In the case of units that were destroyed, the Government committed to rebuilding on the same site to minimize disruptions.

By the time the SIMCE 2011 sample was collected, i.e., 20-22 months after the earthquake struck, only 24 percent of home reconstruction had been completed, consisting mostly of home repairs (Comerio 2013). Despite impressive efforts by the Government, not much new construction had been seen by the time the SIMCE data was collected, leading to frustration in the population, as shown in Figure 2.

8 These schools 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 to close because of low attendance.

9 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.
3.2.1 Seismic Intensity

I use results from the structural engineering literature to build a continuous measure of earthquake exposure, seismic intensity on the Medvedev-Sponheuer-Karnik (MSK) scale, that retains more information than a binary measure. The MSK scale uses observed damage to the structures to infer seismic intensity. Two main factors determine the magnitude of structural damage: type of housing and distance from the main asperity. An asperity is an area on a fault that is stuck or locked. The earthquake rupture usually begins at an asperity.

Conditional on a type of construction, seismic intensity on the MSK scale corresponds to a specific level of damage. House type is not observed in my dataset. However, Astroza, Ruiz, and Astroza (2012) report that the ∼ 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 ∼ 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. Table 1 shows the expected reconstruction costs for the buildings in my sample corresponding to various levels of MSK intensity, while Figure 10 shows pictures of four adobe houses that suffered three different grades of damage.

Distance from the main asperity matters both because the intensity of shaking decreases with distance from the earthquake’s rupture, and because it captures the type of soil a building rests on: shaking in alluvial and fluvial deposits tends to be stronger than on hard soil and rock. How seismic intensity propagates from the main asperity is captured by the intensity attenuation formula, which for the 2010 Chilean earthquake is:

\[ I(\Delta_A) = 19.781 - 5.927 \log_{10}(\Delta_A) + 0.00087\Delta_A \]  

where \( \Delta_A \) is the distance in kilometers from the earthquake’s centroid main asperity. The \( R^2 \) is 0.9894. I use this formula to measure MSK intensity in all 235 towns affected by the earthquake. Figure 5 reports the density of seismic intensity among...
the students in my sample who reside in the regions affected by the earthquake.

MSK intensity ranged from 1 (not perceptible) to 9 (destructive) in Constitución. Figure 10 shows the isoseismal map of the damaged area. As can be seen, even towns that are close to each other were subject to very different levels of seismic intensity. For example, Las Caboras (id 82) and Pichidegua (id 14) are only 7.83km (4.87m) apart. In Las Caboras, 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.12

Given these differences across neighboring towns, and because children in the same classroom come from different towns, it is not surprising that classmates suffered different levels of home damage. The modal range in damages within classrooms is approximately USD 730, corresponding to 1.62 times the average monthly income.13

3.3 Preliminary Data Analysis: Impact of Earthquake on Test Scores

This section presents the first evaluation of the effect of the 2010 Chilean earthquake on student test scores. Being exposed to the earthquake had a negative impact on test scores, and this negative impact was stronger at larger earthquake intensities.

To evaluate the earthquake’s impact on 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. This permits implementation of a difference-in-differences framework, which exploits also the longitudinal structure of the data through a value-added specification.14 Specifically, I estimate the within-school transformation of the following regression, for both Math and Spanish test scores:

\[ y_{ilsg} = \alpha + \lambda_s + y_{ils,(g-4)}\delta + x_i'\gamma + P_i\theta_P + E_i\theta_E + P_iE_i\theta + \epsilon_{ilsg} \]

where \( y_{ilsg} \) is the standardized test score of student \( i \) in classroom \( l \) in school \( s \) and in

---

12 See also Figure 1 in the online supplementary material, showing how dispersed damage was across neighboring towns.
13 Damage reconstruction costs were mostly covered by the government, therefore, this figure is intended to give an idea of how extensive the variation in damages was, rather than to represent the monetary cost that different households had to sustain.
14 This is similar to a difference-in-difference-in-differences approach, in that it accounts for individual, regional and cohort effects. However, technically the two approaches are different, because here own lagged test score enters as a regressor.
grade $g = 8$, $\lambda_s$ is an idiosyncratic school effect, $y_{ils(g-4)}$ is the student’s standardized test score in the fourth grade, and $x_i$ is a vector of student characteristics. The dummy variable $P_i$ indicates if a student belongs to the post-earthquake cohort, and the dummy variable $E_i$ indicates whether the student resides in a region affected by the earthquake. The parameter of interest is $\theta$, which captures the impact of being exposed to the earthquake on test score growth.\(^{15}\)

Table 2 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 (sd) (columns 4 and 8), with a larger negative impact in private schools, perhaps indicating that the Government devoted more reconstruction funds to public than to private schools.

To estimate the impact of seismic intensity at a student’s home, $I_i$, I calculate intensity for all students (in earthquake regions) in both cohorts. I then estimate the following regression: \(^{16}\)

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

The effect of earthquake intensity on test score growth is $\theta_{post} - \theta_{pre}$. 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).\(^{17}\) I find that increasing earthquake intensity by one category in the MSK-scale reduces test score growth by 0.008*** 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

---

\(^{15}\)Test score growth in this value added framework is $y_{ics8} - y_{ics(g-4)}\delta$. Denote it by $\Delta y$. Parameter $\theta$ is equal to $[\Delta y_{E=1,P=1} - \Delta y_{E=0,P=1}] - [\Delta y_{E=1,P=0} - \Delta y_{E=0,P=0}] = [\theta_E + \theta] - [\theta_E]$. By subtracting $\Delta y_{E=0,P=1}$ from $\Delta y_{E=1,P=1}$ in the first term, we eliminate the cohort effect from test score growth of post-earthquake students in an earthquake region. By subtracting $\Delta y_{E=0,P=0}$ from $\Delta y_{E=1,P=0}$ in the second term, we single out the region effect. Finally, by subtracting the region effect from the first term, we obtain the earthquake effect.

\(^{16}\)The estimation results can be found in Tables 6 and 7 in the online supplementary material.

\(^{17}\)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 validity of this assumption is verified in section 4.5. Moreover, the sample in this paper satisfies also the stronger assumption made in Card (1992), i.e. that treatment intensity is uncorrelated with student unobservables.
with unobservable student characteristics affecting outcomes. Using the unaffected 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 3). This indicates that seismic intensity is uncorrelated with student unobservables.

4 An Empirical Model of Social Interactions

In this section I build an econometric model that exploits information on seismic intensity at each student’s home to identify peer effects.

4.1 Achievement Production Function with Peer Spill-overs

I assume that a student’s achievement depends on her own characteristics, like innate ability and parental investments, and on classroom characteristics, like class size and teacher ability. There are peer effects because the composition of peers in a classroom can affect how a student’s characteristics map into achievement. For example, two students with identical characteristics may obtain different test scores in two classrooms that are identical except for the composition of peers.

I broadly define a student’s ability to study as a characteristic that captures how able a student is to learn. It is determined by all individual level inputs into the production of achievement like, for example, her innate ability and parental investments. Formally, 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 $c_i$. This dimensionality reduction assumption has two advantages. First, it makes a semiparametric approach feasible. Such an approach is desirable because it does not impose restrictions on the functional form of the achievement production technology. Second, as a result of this assumption, changes in classroom composition can be conveniently represented by changes in the classroom distribution of type $c$. The achievement production function of student $i$ in classroom $l$ is:

$$y_{il} = m_l(c_i) + \epsilon_{il} = e_l(c_i) + u_l + \epsilon_{il}$$

Not surprisingly, $\theta_{pre}$ in equation 2 is also not significantly different from zero. Moreover, I estimate a value-added regression similar to 2 without the regressor $(1-P_l)I_i$, and I obtain a coefficient on $P_lI_i$ that is very close to $\theta_{post} - \theta_{pre}$ in Table 2. See the online supplemental materials, where the results from these regressions are presented.
where \( y_{il} \) is test score, which measures achievement with error \( \epsilon_{il} \), with \( E(\epsilon_{il}|c_i, u_l, l) = 0 \). Type \( c_i \) is a linear function of student characteristics: \( c_i = \alpha_1 x_i \), where \( x_i \) contains initial ability (as measured by lagged test scores), father’s and mother’s education, household income, and gender. 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 \( l \). The term \( u_l \) contains all observable and unobservable characteristics of the classroom that affect achievement.

Equation 3 is a semiparametric single-index model (Hall 1989, Ichimura 1993, Horowitz 2010). I jointly estimate the \( m_l(\cdot) \) function in each classroom, with kernel methods, and the \( \alpha \) parameters. The estimation algorithm is presented in Appendix A.1, where the details for the calculation of the standard errors are also presented. Notice that \( u_l \) is not separately identified from the constant in \( e_l(\alpha_1 x_i) \), therefore, peer effects cannot be separately identified from the effect of classroom characteristics at this stage.

### 4.2 Seismic Intensity as a Source of Identifying Variation

The ideal setting to evaluate the effect of peers on test scores is one where student allocation to classrooms is random or experimentally controlled. While college administrators sometimes adopt random assignment of peers to dorms, random assignment of peers to classrooms is rarely adopted in schools or colleges.\(^{19}\) Moreover, as noted in the survey by Epple and Romano (2011), experiments with random assignment to classrooms are rare, especially beyond primary school.\(^{20}\)

Given the limited availability of this kind of data, I adopt a different approach. To identify the effect that changing classroom composition has on test scores through its effect on \( e_l(c_i) \) in equation 3, one must be able to vary the classroom distribution of \( c \) independently of \( u_l \). The key identification problem in observational data is that classrooms with different distributions of student characteristics, like, for example, innate ability, are different also in other unobserved ways that make it impossible to separate peer effects from other confounding effects, i.e., correlated effects (Manski

---

\(^{19}\)Random assignment to dorms has been exploited to study peer effects among college roommates in, for example, Sacerdote (2001), Zimmerman (2003), Stinebrickner and Stinebrickner (2006), Kremer and Levy (2008), and Garlick (2014).

\(^{20}\)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.
1993). For this reason, I consider a shock to each student’s $c_i$ that is such that its classroom distribution, or at least a moment of this distribution, is not systematically related to unobserved classroom characteristics affecting outcomes.

This 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 it affects a student’s ability to learn, classrooms with different distributions of seismic intensity have different distributions of student types, even if the distributions of all other student characteristics are identical. This holds true independently of the channel through which seismic intensity affects a student’s ability to study, as long as there is an effect. One possible channel is structural damage to a student’s home, which determines medium- to long-term disruptions to a student’s environment. This may increase the opportunity cost of time, because students may be required to spend time helping their parents with home repairs. 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). Moreover, the severity of PTSD increases with proximity to the epicenter (Groome and Soureti 2004), indicating that the psychological impact should be measured by a continuous variable, like seismic intensity, rather than by a dummy variable for earthquake exposure.

In fact, survey evidence suggests that seismic intensity at a student’s home did 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 10. Second, I find that the negative impact of seismic intensity on test scores (see section 3.3) 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. Third, using a dataset...

---

21 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).

22 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, Dikaiaikou, Galanopoulou, and Yule 2006).

23 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 of this elicited measure of cost.

24 The additional effect is $-0.0173614$, p-value 0.049. The amount of homework assigned is observed...
collected only a few months after the earthquake, I find that students affected more badly by the earthquake report reading less books in the months immediately following the earthquake.\textsuperscript{25}

I allow seismic intensity to affect a student’s type $c_i$ for the 2011 cohort of students, and I estimate its impact. Formally, I assume that the ability to study for students affected by the earthquake is: $c_i = \alpha_1 x_i + \alpha_2 I_i + \alpha_3 I_i x_i$. This is the specification for $c_i$ that I estimate in the semiparametric single-index model in \textsuperscript{3} for this cohort of students. The interaction term $I_i x_i$ captures individual heterogeneity in the effects of seismic intensity on the student’s type $c_i$. For example, some parents may try to attenuate the impact of the earthquake by providing a new study environment for their child.\textsuperscript{26}

As I show in detail in sections \textsuperscript{4.5} and \textsuperscript{6} the variance of the distribution of seismic intensity in the classroom satisfies the exclusion restriction for the identification of peer effects; it changes the classroom distribution of $c_i$, but it does not affect other classroom characteristics. However, using the cohort of students that was not affected by the earthquake I find that the measure of seismic intensity variance has a spurious effect on test scores, because the measure of intensity is a function of a student’s location. Therefore, a larger measured seismic intensity variance signals that the classroom attracts students from farther away, indicating that it has some desirable characteristics unobserved to the econometrician. Table \textsuperscript{3} reports the parameter estimates from a value-added test score regression using the pre-earthquake student cohort, and it demonstrates the existence of spurious effects operating through the variance of seismic intensity. These effects must be accounted for.

\subsection*{4.3 Differencing out Spurious Effects}

I build an econometric model that performs comparisons between classrooms that suffered different variances of seismic intensity, but that are observationally identical otherwise. To account for the fact that the damage to school facilities following the earthquake could be another channel through which the earthquake affected test scores, only for Math classrooms in the cohort of students affected by the earthquake. See Table 9 in the online supplementary material.

\textsuperscript{25}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 accompanied by a reduction in effort.

\textsuperscript{26}The medical literature reports that the psychological impact is stronger on girls. My parameter estimates find support for this.
I only compare classrooms that suffered the same level of seismic intensity from the earthquake, as well as the same average seismic intensity among the students. This section shows how the method accounts for spurious effects of seismic intensity variance by measuring them in the pre-earthquake cohort and differencing them out.

I decompose the term $u_l$ into two terms:

$$u_l = \lambda(z_l) + \phi(z_l, F_l(x_i), \sigma^2_{Il}) = \lambda(z_l) + \phi_l$$

I do not make any functional form assumptions on the functions $\lambda(\cdot)$ and $\phi(\cdot)$. The first term is the direct impact of observed classroom characteristics $z_l$ on achievement. Vector $z_l$ contains damage to the school building, average seismic intensity at the students’ homes, teacher experience, teacher gender, whether the school is rural or urban, and class size.

The second term, $\phi(\cdot)$, captures the effect on test scores of unobserved classroom characteristics, like, for example, teacher ability. To capture non-random assignment of students to classrooms, I allow $\phi(\cdot)$ to depend on $z_l$ and on the distribution of $x_i$ in the classroom, $F_l(x_i)$. For example, a school might assign more able teachers to classrooms with more disadvantaged students. This term captures what Manski (1993) refers to as correlated effects. Moreover, this term depends also on $\sigma^2_{Il}$, the variance of seismic intensity in the classroom, to allow for spurious effects of this variance on test scores.

I calculate seismic intensity for all students, including those in the pre-earthquake cohort, using the formula in [1]. Any effect of the variance of seismic intensity on the achievement of pre-earthquake students is spurious, because no damages caused by the earthquake have occurred yet. 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, $z_l = z_{l'}$, and $F_l(x_i) = F_{l'}(x_i)$, but $\sigma^2_{Il} \neq \sigma^2_{Il'}$. In particular, $l$ has a smaller variance: $\sigma^2_{Il} - \sigma^2_{Il'} = \Delta \sigma^2_{Il'l'} < 0$. Because classroom composition, as measured by the distribution of $c$, is the same in the two classrooms, the mapping from type $c_i$ to test score is the same, i.e., $e_l(c) = e_{l'}(c)$ for every $c$. As shown in section [4.1] the function $e_l(c)$ is not identifiable, but the function $m_l(c)$ is. Taking the difference between the $m$ functions in these two classrooms at a given point
$c = \alpha_1 x_i$ gives:

$$m_l(c) - m_{l'}(c) = \Delta m_{ll'}^{pre}(c) = e_l(c) + \lambda(z_l) + \phi_l^{pre} + \epsilon_{il} - e_{l'}(c) - \lambda(z_{l'}) - \phi_{l'}^{pre} - \epsilon_{il'}$$

$$= \phi_l^{pre} - \phi_{l'}^{pre} + \epsilon_{il} - \epsilon_{il'}$$

$$= \Delta \phi_{ll'}^{pre} + \xi_{ill'}$$

(5)

where $pre$ indicates that the sample is the pre-earthquake cohort. $\Delta \phi_{ll'}^{pre}$ captures the spurious effect of the variance of seismic intensity on achievement. Consider now the post-earthquake cohort. The type $c_i$ for these students is affected also by the earthquake intensity:

$$y_{il} = e_l(\alpha_1 x_i + \alpha_2 I_i + \alpha_3 I_i x_i) + u_l + \epsilon_{il}$$

$$= m_l(\alpha_1 x_i + \alpha_2 I_i + \alpha_3 I_i x_i) + \epsilon_{il}.$$  

Like before, consider two classrooms, $s$ and $s'$, that share the same observed characteristics ($z_s = z_{s'}$, and $F_s(x_i) = F_{s'}(x_i)$), but where the variances of seismic intensity differ, i.e. $\sigma^2_{Is} \neq \sigma^2_{Is'}$. In particular, $\sigma^2_{Is} - \sigma^2_{Is'} = \Delta \sigma^2_{iss'} < 0$. 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 $c = \alpha_1 x_i + \alpha_2 I_i + \alpha_3 I_i x_i$, gives:

$$m_s(c_i) - m_{s'}(c) = \Delta m_{ss'}^{post}(c) = e_s(c) + \lambda(z_s) + \phi_{s}^{post} + \epsilon_{is} - e_{s'}(c) - \lambda(z_{s'}) - \phi_{s'}^{post} - \epsilon_{is'}$$

$$= e_s(c) - e_{s'}(c) + \phi_{s}^{post} - \phi_{s'}^{post} + \epsilon_{is} - \epsilon_{is'}$$

$$= \Delta e_{ss'}(c) + \Delta \phi_{ss'}^{post} + \xi_{iss'}$$

(6)

Consider now the four classrooms $l, l', s$ and $s'$ simultaneously. Suppose that $\Delta \sigma^2_{ll'} = \Delta \sigma^2_{iss'} < 0$, 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 $ss'$. That is, the treatment intensity is the same in the two pairs of classrooms. If these four classrooms share all other observed classroom characteristics (i.e. $z_i$ and the classroom
distribution of $x_i$, then the difference between the $\Delta m$ functions is:

$$
\Delta m_{ss'}^{\text{post}}(c) - \Delta m_{ss'}^{\text{pre}}(c) = \Delta e_{ss'}(c) + \Delta \phi_{ss'}^{\text{post}} + \xi_{ss'} - \Delta \phi_{ss'}^{\text{pre}} - \xi_{ss'}
$$

where $E(\xi_{ill'ss'}|x_i, z_i, I_i) = 0$. If $\Delta \phi_{ss'}^{\text{post}} = \Delta \phi_{ss'}^{\text{pre}}$, then the spurious effects cancel out,

$$
\Delta m_{ss'}^{\text{post}}(c) - \Delta m_{ss'}^{\text{pre}}(c) = \Delta e_{ss'}(c) + \Delta \xi_{ill'ss'},
$$

and the difference between the $\Delta m$ functions in the post- and pre-earthquake samples identifies the peer effect of the variance of student types, i.e., the change in achievement that is due to a change in the variance of $c$ in the classroom, $\Delta e_{ss'}(c)$.

A condition under which $\Delta \phi_{ss'}^{\text{post}} = \Delta \phi_{ss'}^{\text{pre}}$ is that the relationship between the intensity variance $\sigma_{ll}^2$ and the unobserved classroom characteristics remains unchanged after the earthquake strikes. This condition is discussed in detail in section 4.5.

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 $i = \text{pre}$ and $i = \text{post}$ from the two separate cohorts are of the same type $c_i$, then they have different characteristics $x_i$, specifically: $x_{\text{pre}} = \frac{\alpha_1 + \alpha_3 I_{\text{post}}}{\alpha_1} x_{\text{post}} + \frac{\alpha_2}{\alpha_1} I_{\text{post}}$. 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 $\alpha$. It is the fact that I jointly estimate the $m(\cdot)$ functions and the $\alpha$ parameters that allows me to make the appropriate comparisons between pre- and post-earthquake students. Simpler comparisons based on observable student characteristics cannot capture the learning disruptions induced by the earthquake.

### 4.4 Implementation of the Differencing Technique

The method for differencing out spurious effects has two key aspects. First, it requires doing pair-wise comparisons between classrooms. This is because the treatment is defined only in relative terms: if classroom A has a larger seismic variance than classroom B, but a lower seismic variance than classroom C, then it is the treated classroom when
compared to B, while it is the untreated classroom when compared to C. To account for the intrinsically relative nature of the treatment, I perform all possible pair-wise comparisons, and consider the classroom with a larger variance in each pair as the treated classroom. In a sample with \( n \) classrooms, there are \( \frac{1}{2} \binom{n}{2} = \frac{1}{2} n(n-1) \) such comparisons. The treatment effect is then estimated by averaging over these pairwise comparisons.

The second key aspect of the method is that these pairwise comparisons must be made among pairs of classrooms that, except for the variance of seismic intensity, are identical in terms of all classroom observable characteristics, such as teacher characteristics, and in terms of the distribution of student characteristics within the classroom, such as the distribution of lagged test scores. Because only a limited number of classroom pairs are going to be exactly identical, I use kernel weighting, where pairs of classrooms that are very similar obtain higher weights than pairs of classrooms that are less similar. This is similar to the analysis in Powell (1987) and Ahn and Powell (1993).\(^27\) Moreover, when “conditioning” on the distribution of student characteristics, I reduce the dimensionality of the problem by considering a finite number of moments of this distribution: the mean, the variance, the skewness and the kurtosis. Intuitively, for this kernel weighting method to deliver the desired result, two sets of conditions must be met. First, the classroom effects, the correlated effects, the spurious effects and the function \( e_l(c) \) must vary smoothly with classroom characteristics and with the distribution of student characteristics.\(^28\) 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** \( \lambda(\cdot) \) in equation (4) is continuous.

**DA.2** \( \phi(z_l, F_l(x_i), \sigma^2_{I_l}) \) in equation (4) is similar to \( \tilde{\phi}(z_l, W_1, \sigma^2_{I_l}) \) \( \forall (z_l, W_1, \sigma^2_{I_l}) \), where \( W_1 \) is a vector containing the mean, variance, skewness and kurtosis of the elements of \( x_i \).\(^29\)

**DA.3** \( \tilde{\phi}(\cdot, \cdot, \cdot) \) is continuous in its first two arguments.

**DA.4** If two classrooms in the pre-earthquake cohort are such that \( W_1 \cong W_P \), then \( e_l(c) \cong e_P(c) \) \( \forall c \). If two classrooms in the post-earthquake cohort are such that \( [W_s, \mu_s, \sigma^2_{I_s}] \cong [W_s', \mu_s', \sigma^2_{I_s'}] \), where \( \mu_l \) and \( \sigma^2_l \) are the mean and variance of seismic

\(^{27}\)In Ahn and Powell (1993), pairs of observations obtain larger kernel weights when they are similar in terms of variables determining sample selection. In their context, this ensures that sample selection bias is differenced out.

\(^{28}\)In Ahn and Powell (1993), this assumption corresponds to the assumption of continuity of the selection function (see page 9 in their paper).

\(^{29}\)I ignore the cross-moments of \( x_i \) to make the numerical implementation tractable.
intensity in the class, then \( e_s(c) \equiv e_s'(c) \ \forall c. \)

These assumptions allow me to build an approximated counterpart to equation 7. Assumption DA.1 implies that if two classrooms are similar, then the classroom effects captured by \( \lambda(\cdot) \) are also similar. Assumptions DA.2 and DA.3 together mean that if two classrooms are similar, then \( \tilde{\phi}(W_l, z_l, \sigma^2 I_l) - \tilde{\phi}(W_v, z_v, \sigma^2 I_v) \approx \Delta \phi_{l'v} \). That is, the difference between the \( \tilde{\phi} \) functions computed in classrooms with similar vectors \( W \) and \( z \) is a good approximation to the spurious effect in equations 5 and 6. 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 \( W_1 \equiv W_v, z_1 \equiv z_v \) and \( \sigma_{l1}^2 - \sigma_{l'v}^2 = \Delta \sigma_{l'v}^2 < 0 \), I can build the approximated counterpart to equation 5:

\[
m_l(c) - m_{l'}(c) \approx \Delta m_{ll'}^{pre}(c) = \Delta \phi_{ll'}^{pre} + \xi_{ll'}. \tag{8}
\]

Similarly, for any two classrooms \( s, s' \) in the post-earthquake cohort with \( W_s \equiv W_{s'}, z_s \equiv z_{s'} \) and \( \sigma_{s1}^2 - \sigma_{s'v}^2 = \Delta \sigma_{s's'}^2 < 0 \), I can build the approximated counterpart to equation 6:

\[
m_s(c) - m_{s'}(c) \approx \Delta m_{ss'}^{post}(c) = \Delta e_{ss'}(c) + \Delta \phi_{ss'}^{post} + \xi_{ss'}. \tag{9}
\]

If four classrooms, \( l, l' \) from the pre-earthquake cohort and \( s, s' \) from the post-earthquake cohort, are such that \( W_1 \equiv W_v \equiv W_s \equiv W_{s'}, z_1 \equiv z_v \equiv z_s \equiv z_{s'}, and 0 > \Delta \sigma_{l'v}^2 \equiv \Delta \sigma_{s's'}^2 < 0 \), then subtracting equation 8 from equation 9 yields an approximation to \( \Delta m_{ss'}^{post} - \Delta m_{ll'}^{pre} \), the quantity of interest.

To ensure that the classrooms are similar, I assign declining weights to quadruples that are more dissimilar in terms of \( W \) and \( z \) and \( \Delta \sigma_{j}^2 \). 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 \( Z_t = [W_t, z_t] \) for \( t = l, l', s, s' \), I assign the weight \( \frac{1}{h} k \left( \frac{Z_t - Z_{tt'}}{h} \right) \) 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 \( W \) and \( z \) (\( tt' = ll', ss' \)), and also that across cohorts the two pairs of classrooms are similar (\( tt' = sl \)).

\[30\] 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.
I build a weight that is declining in $|\Delta \sigma^2_{Il'} - \Delta \sigma^2_{Iss'}|$, to guarantee that the pre- and post-earthquake pairs differ in terms of $\sigma^2_I$ in a similar way, i.e., that the intensities of treatment in the two pairs of classrooms are very similar: $\frac{1}{h_{\Delta \sigma}} k \left( \frac{\Delta \sigma^2_{Il'} - \Delta \sigma^2_{Iss'}}{h} \right)$. The weight for the quadruple, $\omega_{ll's's'}$, is the product of these four kernel weights:

$$\omega_{ll's's'} = d_{ll's's'} \prod_{tt' \in \{ll's's', sl's'\}} \frac{1}{h_{\Delta \sigma}} k \left( \frac{Z_t - Z_{t'}}{h} \right)$$

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

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

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

$$\hat{\gamma} = \hat{\Delta e}(c) = \frac{\sum_{l=1}^{N_{\text{pre}-1}} \sum_{l'=1}^{N_{\text{pre}}} \sum_{s=1}^{N_{\text{post}-1}} \sum_{s'=1}^{N_{\text{post}}} \omega_{ll's's'} (\hat{m}_{ss'}^{\text{post}} - \hat{m}_{Il'}^{\text{pre}})}{\sum_{l=1}^{N_{\text{pre}-1}} \sum_{l'=1}^{N_{\text{pre}}} \sum_{s=1}^{N_{\text{post}-1}} \sum_{s'=1}^{N_{\text{post}}} \omega_{ll's's'}}$$

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

### 4.5 Identifying assumption

For the variance of seismic intensity to satisfy the required exclusion restriction for identifying peer effects, the following assumption must be satisfied:

**IA.** Constancy of spurious effects, i.e. $\Delta \phi_{Il'}^{\text{pre}} = \Delta \phi_{ss'}^{\text{post}} \forall l, l', s, s'$ t.s. $\sigma^2_{Il} - \sigma^2_{Il'} = \sigma^2_{Is} - \sigma^2_{Is'}$.

For this identifying assumption to be met, classroom quality and parental sorting across schools must not endogenously change as a reaction to changes in seismic intensity variance. If, for example, schools decided to assign more experienced teachers to classrooms that suffered a larger variance in damages, this assumption would not be
satisfied. As another example, the assumptions would be violated if motivated parents systematically avoided classrooms that suffered a large variance in damages. Intuitively, the earthquake must not have changed the pre-existing (spurious) relationship between intensity variance and unobserved classroom characteristics.

Table 5 examines endogenous 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. As can be seen in the row named “Effect of Classroom Intensity Variance”, the estimated effect is never statistically different from zero. This suggests that parents have not reallocated across schools as a reaction to intensity variance.

Tables 6 and 6 show that also observed classroom quality did not change as an effect of changes in seismic intensity variance. None of the estimates in the row named “Effect of Classroom Intensity Variance” is significantly different from zero. Although it could still be the case that unobserved classroom characteristics reacted to the earthquake, the fact that observed characteristics did not gives me confidence in the validity of the identifying assumption.

5 Estimation Results

Table 8 presents the parameter estimates. The coefficient on lagged test score is normalized to $-1$, because only the ratios among the $\alpha$ parameters are identified. Under this normalization, student type $c_i$ can be interpreted as a cost of exerting effort, and a larger lagged test score causes a lower cost of effort. As expected, earthquake intensity is estimated to increase student type. Figure 6 shows an example of the function $m_i(c_i)$ estimated in two classrooms. The higher a student’s type $c_i$ is, the lower achievement is. The model fit is very good, as can be seen in Table 9.

The effect of changing the variance 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 $\gamma$ in equation 11 estimated for Mathematics and Spanish test scores, over a grid of values for $c$. These are the subject-specific estimated functions $\hat{\Delta}e(c)$. Going from low to high $c$, this function is negative and then positive for Spanish test scores, while it is positive, then negative.

---

31 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.
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.

Carrell, Sacerdote, and West (2013) also find heterogeneous impacts on test scores of changes to the variance of student types (ability), and they show that they are not captured by a peer-effect specification commonly used in empirical work. These patterns are hard to explain with the standard models of peer effects that do not include the variance of peer variables as a regressor. These linear social interaction models are the workhorse of empirical research (Blume, Brock, Durlauf, and Jayaraman 2014). Moreover, my finding that the effects vary by subject is novel. It is important to understand what gives rise to these patterns, and whether they generalize to a broader context.

6 Robustness

As a first robustness check 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. I then compare the estimates with those obtained from the semi-parametric technique. As can be seen from Table 10, 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 importance of adopting an approach that can detect the heterogeneity of the impact across students, because the mean treatment effects mask considerable heterogeneity.\footnote{32\ding{102}Ding and Lehrer (2007) find that a larger variance of peer ability in Chinese secondary schools has a negative average effect on college entrance examination scores. However, the authors do not explore the heterogeneity of these effects.}

Second, I ask whether the observed patterns of heterogeneity can be explained by unobserved changes in teachers’ effort, focus of instruction or productivity, which could have endogenously reacted to changes in the variance of seismic intensity. For example, teachers in classrooms where some students were badly affected while others were not could have changed their focus of instruction. This could have generated the observed heterogeneous impacts.
This alternative mechanism can be excluded for two reasons. First, I observe how much of the Spanish curriculum teachers were able to cover in the pre- and in the post-earthquake cohorts. I estimate a difference-in-differences regression like the one in equation 2 to evaluate the effect of seismic intensity variance on this variable. If teachers changed their behavior by, for example, focusing only on certain students, or by exerting more or less effort, or if their productivity changed as an effect of seismic intensity variance, then this would be reflected in the amount of curriculum that they covered. However, the effect of seismic intensity variance on this variable is not statistically different from zero, as can be seen in Table 11. Second, for the teacher channel to explain the 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. Therefore, *ad hoc* theories are needed. No previous study provides support for such theories.

7 A Theoretical Model of Social Interactions

In this section I propose a simple theory that can readily explain all the empirical findings. The model is a game of social interactions that allows also for standard technological spill-overs working through the mean of peer ability. I present the two main model results as testable implications, that I formally test in the next section.

The model is based on 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. In the application of their theory that I propose here, achievement is at the same time a consumption good and a positional good. Therefore, there is no trade-off between two distinct goods. However, producing achievement is costly. Specifically, students in a classroom choose how much effort $e$ to exert, and effort increases achievement/test score $y$. Students are heterogeneous in terms of how costly it is for them to exert effort. The main model assumptions are the following:

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

A.2 There are technological spill-overs in the production of achievement.

---

$^{33}$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.3 Students’ utility is increasing in rank in terms of achievement.

Assumptions A.1 and A.2 are standard. 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).

Students differ in terms of a type $c$: those with a higher $c$ incur a larger cost of effort and can potentially achieve less, i.e. if they exert maximal effort, they achieve less than a student with lower $c$ exerting maximal effort. Type $c$ captures any student characteristics, physical or psychological, that affect learning ability, such as cognitive skills, access to a computer or books, availability of an appropriate space for studying, parental help, etc. 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.

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 $y(e) = \frac{1}{\mu} e + u$, where $\mu$ is the classroom mean of $c$ and $u > 0$ incorporates classroom specific factors like, for example, teacher ability. Technological spill-overs (assumption A.2) are represented by $\frac{1}{\mu}$, which affects the marginal product of effort: the larger mean $c$ in the classroom, the lower the productivity of own effort. For example, more able peers (lower mean $c$) are beneficial to own learning because they ask relevant questions in class. In addition, $\mu$ interacts with a student’s type $c$ to determine the maximum achievable level of achievement: $\bar{y} = \frac{1}{\mu c} + u$. As a result, no student chooses effort above $\frac{1}{\bar{c}}$, because any unit of effort above $\frac{1}{\bar{c}}$ does not increase achievement, but it is

34 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).

35 The assumption that there is an upper bound on achievement that depends on type $c$ is needed to guarantee the existence of an equilibrium, because it yields a boundary condition for the differential equation characterizing equilibrium strategies. Although it is not standard in models of peer effects, it is reasonable and intuitive.

36 This assumption means that, for example, a student with low cognitive skills (low $\frac{1}{\bar{c}}$) who is exerting maximal effort achieves less than a classmate with high cognitive skills (low $\frac{1}{\bar{c}}$) who is also exerting maximal effort. At the same time, if placed in a better classroom (in terms of $\mu$ and $u$), the low cognitive skills student can potentially achieve more than he did in the worse classroom.
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 ($V_{12} \leq 0$).

37 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 ($\frac{dq}{de} = 0$) and additive separability between utility from achievement and cost of effort ($V_{12} = 0$).

38 The probability that a student $i$ of type $c_i$ with effort choice $e_i = c(e_i)$ chooses a higher effort 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 $e_i$.

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 a standard approach to modeling 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 is equal to rank plus a parameter $\phi$: $S(F_Y(y)) = F_E(e) + \phi$, with $\phi \geq 0$.

Overall utility is the product of utility from own achievement and utility from rank. Therefore, the parameter $\phi$ measures the level of competitiveness in the classroom. When $\phi > 0$, students have a minimum guaranteed level of utility even if they rank last ($F_E(e) = 0$). As $\phi$ becomes larger, the importance of rank in determining behavior decreases, while the utility from the absolute level of achievement becomes more important. On the other hand, when $\phi = 0$ ranking low has dire consequences, and students try to avoid a low-rank. Recent laboratory evidence provides support for the existence of aversion to a low-rank (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).

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 $V(y(e_i), q(e_i, c_i))(1 - G(c(e_i)))$.

Overall utility is the product of utility from own achievement and utility from rank.
order condition then is:

\[
\frac{V_1}{mg. \text{ ut. from increased achiev.}} \frac{1}{\mu} + \frac{V(y, q) g(c(e_i))(-c'(e_i))}{1 - G(c(e_i)) + \phi} \frac{\partial q}{\partial e} = -V_2 \frac{\partial q}{\partial e}
\]

and it implies the first-order differential equation reported in equation 13 in Appendix A.2. The solution to this differential equation is a symmetric equilibrium of the game. The assumptions I make on the utility function, one 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. 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 \([c, \bar{c}]\), with \(c \geq 0\). The first theoretical result is summarized in the following Proposition:

**Proposition 7.1** (Adapted from Proposition 1 in Hopkins and Kornienko (2004)).

The unique solution to the differential equation (13) 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 \(V_1 \frac{1}{\mu} |_{e=e_{nr}} = -V_2 \frac{\partial q}{\partial e} |_{e=e_{nr}}\), is an (essentially) unique symmetric Nash equilibrium of the game of status. Equilibrium effort \(e(c)\) is continuous and strictly decreasing in type \(c\).

Proof: see Appendix A.2.

Notice that Proposition 7.1 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. The implication can be rephrased in terms of achievement, given that achievement is an increasing function of effort, to obtain the first testable implication:

under the equilibrium strategy.

One of the main differences with the model in Hopkins and Kornienko (2004) is that here equilibrium strategies are decreasing, whereas there they are increasing. Hopkins and Kornienko (2007) apply the model in Hopkins and Kornienko (2004) to the case of procurement auctions, where strategies are decreasing like in the current application.

The equilibrium is essentially unique, in the sense that the only source of multiplicity is at the point \(\bar{c}\). This is because the boundary condition for \(\phi = 0\) does not predict a unique behavior for the student of the highest type, \(\bar{c}\).

For example, if the marginal utility from achievement tends to infinity as achievement approaches its lower bound, then when the cost of effort increases, students who are close to the lower bound of achievement will not necessarily reduce effort, because the cost of doing so would be large.
**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 Appendix A.2. 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 $\phi$, i.e., on the strength of the competitive motive.\(^{42}\)

If students care a lot about their 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. Therefore, the difference between equilibrium achievement in a more dispersed, $y(e_B(c))$, and less dispersed classroom, $y(e_A(c))$, is either always negative, or positive for low-$c$ students, and negative for all other students (lower panels of Figure 9).

If students care less about their 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. The difference $y(e_B(c)) - y(e_A(c))$ is either positive for low-$c$ students, negative for middle-$c$ students and positive for high-$c$ students, or it is negative for low- and middle-$c$ students, and positive for high-$c$ students (lower panels of Figure 10).

In both the $\phi = 0$ and the $\phi > 0$ cases, the model is agnostic on the reaction of low-$c$ students. This is because it depends on the shape of the distribution functions, on which the model makes no assumptions. This result provides the second testable implication of the model:

**Testable Implication 2:** When the dispersion in $c$ increases, the effect on student’s achievement is positive for some students and negative for others, and the only admissible patterns of heterogeneity are those listed in proposition (A.1) in Appendix A.2 which are represented graphically in Figures 9 and 10. In particular, the function $y(e_B(c)) - y(e_A(c))$ (in the lower panels of the Figures), that traces the effect on achievement as a function of student type, can cross the x-axis at most twice. If it does not cross the x-axis, it must be negative. If it crosses it once, it can cross it from above or from below. If it crosses it twice, the sequence for its sign, moving from low to high $c$, is $+, -, +$.

\(^{42}\)The formal statement of the comparative statics result can be found in Proposition (A.1) in Appendix A.2.
7.1 Model Intuition: the Competition and the Resting on Your Laurels Incentives

The intuition behind the comparative statics result can be explained by the interplay of two different incentives. The net effect of these two incentives for students of different type $c$ levels depends on two factors: the strength of the preferences for rank, and the shape of the distribution of student types, on which the model makes no assumptions.

The first incentive is what I call an incentive to compete. As can be seen from the first order condition in (12), the marginal utility from increased rank depends positively on the density at one’s own type $c$, $g(c)$. Intuitively, the more people there are of a similar type to one’s own (i.e., the larger this density), the more people can be surpassed in rank by exerting effort. Therefore, as the density around one’s type $c$ increases, the incentive to exert effort increases. Conversely, as the density around one’s type decreases, students have an incentive to “give up”, because it becomes more costly to surpass the higher-ability students, who are now significantly more able than them. The incentive to compete works through changes to the density at one’s own type.

The second incentive is what I call an incentive to rest on one’s laurels. As can be seen in the first order condition in (12), the marginal utility from increased rank is lower at higher ranks (i.e., higher $1 - G(c(e_i))$). To see why, recall that preferences are such that at lower ranks the enjoyment of one’s own private utility is lower. Therefore, high-$c$ students, who are most at risk of having a low rank, compete fiercely to avoid ranking low if there are not many other high-$c$ types with whom they can make favorable comparisons. As the number of students with similar or worse $c$ levels increases, high-$c$ students are not as desperate to avoid a low rank. Hence, they have a weaker incentive to exert effort. In equilibrium, this incentive trickles down to lower values of $c$.

\[43\text{In the discrete case, the key feature that corresponds to a change in density is the change in distance between successive students. For example, if a student is type } c_1 = 10 \text{ and the next student is type } c_2 = 2, \text{ the type 10 student might “give up” because surpassing a type 2 would be too costly in terms of effort cost. However, if the next student becomes a type 9, the type 10 student has an increased incentive to compete, because surpassing the next student becomes less costly. The comparative statics results in this paper are proved with a continuum of students. Recent theoretical work shows that, in this class of models, the equilibrium strategy with a discrete number of players tends to the equilibrium strategy with a continuum of agents as the number of agents goes to infinity (Bhaskar and Hopkins 2013).}\]

\[44\text{For example, if } c \text{ can take on values between 1 and 10, a student of type } c = 9 \text{ exerts substantial effort to avoid ranking low if there are only few students with a type } c \text{ as high or higher than hers. On the other hand, if the classroom composition changes, and many students are a type 9 or higher, then the same student can make more favorable comparisons, and she does not fear a low rank. Therefore,}\]
because changes to the behavior of high-$c$ students affect the equilibrium behavior of all other students. The incentive to rest on one’s laurels works through changes to the cumulative distribution function at one’s own type.

To understand how these two incentives produce the comparative statics result, consider the effect on different portions of $G(\cdot)$ of increasing its dispersion, as illustrated in Figure 11. The density at the tails increases, therefore, low- and high-$c$ students have an increased incentive to compete. Conversely, middle-$c$ students have a lower incentive to compete, because the density around their type $c$ level decreases.

When the dispersion of $c$ increases, high-$c$ students have also an incentive to rest on their laurels, because they can make more favorable comparisons. Hence, high-$c$ students face two opposite incentives. The model predicts that, when $\phi = 0$, the resting on your laurels incentive prevails, and high-$c$ students reduce their effort; when $\phi > 0$, the incentive to compete prevails, and high-$c$ students increase their effort. For middle-$c$ students, when $\phi = 0$ the incentive to (not) compete and to rest on their laurels (because they face less competition from higher-$c$ students) go in the same direction, resulting in a reduction of effort. When $\phi > 0$, the model predicts that the lower incentive to compete due to the lower density at their $c$ level prevails over the fact that they face more competition from higher-$c$ students. Therefore, middle-$c$ students reduce their effort under all values of $\phi$. For low-$c$ students, the incentive to compete and to rest on their laurels go in different directions. On one hand, they have a stronger incentive to compete in a classroom with fatter tails. On the other, they have an incentive to exert less effort because they face less competition from middle-$c$ students, who are exerting less effort. Without further distributional assumptions, the model does not predict unambiguously which incentive prevails for low-$c$ students.

8 Formal Tests of the Model Implications

The semiparametric approach has two features that make it ideal for formally testing the model implications. First, it imposes minimal restrictions on the data and no distributional assumptions. Importantly, it makes no functional form assumptions.

Intuitively, the strength that the resting on your laurels effect has by the time it trickles down from the high-$c$ to the low-$c$ students depends on the shape of the distribution of $c$, on which the model makes no assumptions.

A restriction imposed by the model is that it does not allow cost $c_{i}$ to depend on unobservable student characteristics. This is because estimation of such a model requires to assume that the $m$
on $e(\cdot)$, which is the counterpart of the equilibrium effort function in the theoretical model, on which the model does not impose any functional form. Second, it recovers the distribution of the student types, i.e., the cost of effort in the model. This is essential to test both predictions. As shown in the robustness section 6, a simpler regression approach is unable to test for the heterogeneity of the peer effects across students, because it cannot recover student types.

8.1 Testing Model Implication 1

The function $e_l(c)$ is decreasing if and only if $m_l(c)$ is decreasing. I formally test monotonicity of $\hat{m}(\cdot)$ using the method developed in 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. 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 method can be found in Appendix A.3. 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.

8.2 Testing Model Implication 2

The patterns observed in the data are consistent with the model’s comparative statics result when $\phi > 0$, i.e., when rank concerns are weaker because students have a minimum guaranteed level of utility. Formally, to test the second model implication, it suffices to perform one-sided hypothesis tests on $\hat{\gamma}$ over a grid of values for $c$. The one-sided point-by-point 90 percent confidence intervals reported in Figures 7 and 8 can be used to perform these one-sided tests at the 10 percent significance level. As can be seen, the function for Spanish test scores is statistically negative and then statistically positive, moving from low to high $c$. As typically occurs with non-parametric estimation, the function is monotonic. Under such an assumption, the theoretical model’s implication of monotonicity could not be tested. I am working on an extension of the econometric model that allows cost $c_i$ to depend on unobservables without assuming monotonicity.

47 The constant term $u_l$ may be different within categories because of different values of the $z_l$ variables, however, this does not affect the monotonicity $m$.

48 While $\hat{\Delta}e(c)$ 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 close
tors, the variance is larger near the boundaries, where the data density is smaller. This affects inference for the case of Mathematics. For very small values of $c$, the null that $\hat{\Delta}_e(c) = 0$ cannot be rejected at the 10 percent significance level. This does not reject the model’s implication, which allows this function to be either positive or negative for low values of $c$.\footnote{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.}

The fact that the estimated $\Delta_e(c)$ function behaves differently for Spanish and Math test scores can be easily explained without the need to make \textit{ad hoc} assumptions on different technologies or preferences. The theoretical model predicts that the two observed behaviors, corresponding to Figure \footnote{For example, in my sample the correlation between Mathematics and Spanish lagged test scores is 0.76.} can occur if the distribution of cost of effort is different in Math and in Spanish classes. In this context, where Math and Spanish students are the same students, this occurs if Math and Spanish skills (captured by the cost of effort in each subject) do not have a correlation equal to 1. This fact is uncontroversial.\footnote{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.}

The intuition is simple. Low-$c$ (high-ability) students face two opposing incentives when the variance of student types increases: an incentive to exert more effort to surpass in ranking students of a similar type, who are now more numerous, and an incentive to exert less effort, because they face less competition from below, i.e., from the middle-$c$ (middle-ability) students. If the middle-ability students are significantly less able than the high-ability students, then the fact that the high ability students face less competition from them does not affect their behavior. Therefore, the incentive to compete prevails. If the middle-ability students are similar to the high-ability students in terms of ability, then their behavior matters for the high-ability students, and the incentive to exert less effort because of the lower competition from below prevails.

\section{Implications and Conclusions}

My findings have important implications for policy and for the estimation of peer effects. Tracking students by ability is proposed as a way to improve student outcomes. Typically, to the extent that students benefit from more able peers, tracking is expected to negatively impact the achievement of low-ability students, who are tracked
with other low-ability students, unless teachers’ productivity increases because they can better target their instruction in more homogeneous classrooms. However, this is true only if direct spill-overs between peers do not work through the variance of peer ability, which is the assumption typically made in linear social interaction models, the workhorse of empirical research (Blume, Brock, Durlauf, and Jayaraman 2014). Under this restriction, if tracking is found to be beneficial to low-ability students, its impact must be ascribed to other factors, like, for example, teachers. While this is a realistic assumption in a setting like the one in Duflo, Dupas, and Kremer (2011), with first-grade students who are unlikely to have rank concerns, this assumption could be more problematic in higher grades. In fact, research in the psychology of education shows that competitive attitudes increase as students get older (Madsen 1971, Johnson and Johnson 1974).

If students benefit from having higher-achieving peers, but they also have rank concerns, then tracking can be beneficial to the test scores of low-ability students even in the absence of a teacher reaction. This is because tracking introduces a trade-off between a lower mean peer ability, and a lower variance. In a homogeneous classroom, some students may have a higher incentive to exert effort in order to compete with their similarly-able classmates. In fact, Tincani (2014) shows that this mechanism could explain why Carrell, Sacerdote, and West (2013) found that middle-ability students at the U.S. Air Force Academy improved their test scores on average by 0.082 sd when tracked into homogeneous squadrons. The theoretical model 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 rank concerns, either directly through the elicitation of preferences, or indirectly through randomized control trials that manipulate the incentives to compete. This knowledge could then be used to study optimal peer allocation.

This paper has implications also for the estimation of peer effects. When students

\footnote{Tincani (2014) shows that all the results of the experiment, including the negative impact on low-ability students who are placed in bimodal classrooms, can be rationalized by the special case with $\phi = 0$ of the model of rank concerns presented here. Preliminary results from an experiment with University of Amsterdam students tracked to classrooms of different ability compositions also indicate that reducing the variance of ability increases test scores, without a teacher reaction. Garlick (2014) finds that reducing the variance of peer ability has, on average, a negative effect compared to random allocation to peer groups. However, he considers tracking of university students to dormitories, rather than to classrooms. Like the author notices, peer effects in this context are likely to operate through time use or the transfer of soft skills, as suggested also in Stinebrickner and Stinebrickner (2006). Students in the same dormitories are not generally direct competitors for grades, therefore, it is less likely that peer effects arise from rank concerns in this context.}
have rank concerns, the most appropriate specification for the estimation of peer effects should include the variance of peer ability as a regressor. As shown in the Monte Carlo experiment in Tincani (2014), this guarantees good out-of-sample policy predictions. Blume, Brock, Durlauf, and Jayaraman (2014) show that linear social interaction models like, for example, the linear-in-means model, which do not include variance as a regressor, are indirectly based on the assumption that peer effects generate conformity within peer groups. This idea is exploited also by the excess-variance methods (Glaeser, Sacerdote, and Scheinkman 1996, Graham 2008). However, rank concerns give rise to peer effects without necessarily implying within group conformity. This could explain why, as surveyed in Sacerdote (2014), linear social interaction models estimate larger peer effects in crime and drinking behavior, where conformity is a realistic assumption, than in test scores.

To summarize, this paper makes two contributions. First, it presents a new method to identify and estimate peer effects among Chilean eight graders. Second, it analyzes, for the first time, the interaction between rank concerns and policies that change classroom composition. While rank concerns have received much attention in the Labor, Experimental, and Theory literatures, their role in the development of human capital had not been explored before.
10 Tables and Figures
Figure 1: Data time-line.

Figure 2: This Figure is reported from Comerio (2013). It depicts a handmade sign found in Cauquenes, Chile, on February 2, 2012. Translation: “Reconstruction is like God. Everyone knows it exists, but nobody has seen it.”
Figure 3: 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).
Figure 4: Seismic intensity distribution and isoseismal map in the area damaged by the 27 February 2010 earthquake. Roman numbers refer to the MSK scale and Arabic numbers are the identifiers of the sampled towns. Source: The picture and its description are reported from Astroza, Rui and Astroza (2012). The names of the sampled towns can be found in Figure 2 in the Online Supplementary Material.
Table 1: Estimated reconstruction costs by MSK-intensity category (Adobe constructions)

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

Figure 5: Density of Seismic Intensity in the Sample of Students in Earthquake Regions
Table 2: Difference-in-differences evaluation of earthquake impact on test scores, dependent variables Spanish and Mathematics test scores in eighth grade

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spanish</td>
<td>Mathematics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish test score in fourth grade</td>
<td>0.656*** (0.00249)</td>
<td>0.642*** (0.00240)</td>
<td>0.595*** (0.00651)</td>
<td>0.645*** (0.00167)</td>
<td>0.635*** (0.00234)</td>
<td>0.655*** (0.00231)</td>
<td>0.646*** (0.00623)</td>
<td>0.645*** (0.00159)</td>
</tr>
<tr>
<td>Math test score in fourth grade</td>
<td>0.635*** (0.00234)</td>
<td>0.642*** (0.00240)</td>
<td>0.595*** (0.00651)</td>
<td>0.645*** (0.00167)</td>
<td>0.635*** (0.00234)</td>
<td>0.655*** (0.00231)</td>
<td>0.646*** (0.00623)</td>
<td>0.645*** (0.00159)</td>
</tr>
<tr>
<td>Household income (CLP)</td>
<td>3.61e-08** (1.10e-08)</td>
<td>2.86e-08*** (6.38e-09)</td>
<td>1.51e-08 (9.60e-09)</td>
<td>2.52e-08*** (4.70e-09)</td>
<td>7.68e-08*** (1.03e-08)</td>
<td>4.98e-08*** (5.91e-09)</td>
<td>6.44e-08*** (8.46e-09)</td>
<td>6.03e-08*** (4.35e-09)</td>
</tr>
<tr>
<td>Father's education (yrs)</td>
<td>0.00940*** (0.000795)</td>
<td>0.00756*** (0.000782)</td>
<td>0.0137*** (0.00208)</td>
<td>0.00885*** (0.000748)</td>
<td>0.00634*** (0.000748)</td>
<td>0.00756*** (0.000748)</td>
<td>0.0115*** (0.00183)</td>
<td>0.00715*** (0.000497)</td>
</tr>
<tr>
<td>Mother's education (yrs)</td>
<td>0.0102*** (0.000819)</td>
<td>0.00870*** (0.000830)</td>
<td>0.0133*** (0.00191)</td>
<td>0.00982*** (0.000765)</td>
<td>0.00806*** (0.000765)</td>
<td>0.00785*** (0.000765)</td>
<td>0.0162*** (0.00169)</td>
<td>0.00859*** (0.000516)</td>
</tr>
<tr>
<td>Female</td>
<td>0.134*** (0.00439)</td>
<td>0.121*** (0.00415)</td>
<td>0.101*** (0.0104)</td>
<td>0.124*** (0.00920)</td>
<td>-0.0999** (0.00410)</td>
<td>-0.110*** (0.00385)</td>
<td>-0.0880*** (0.00914)</td>
<td>-0.104*** (0.00268)</td>
</tr>
<tr>
<td>Household lives in earthquake region (E)</td>
<td>0.0725+ (0.0417)</td>
<td>0.0431 (0.0337)</td>
<td>0.0653 (0.0766)</td>
<td>0.0566* (0.0247)</td>
<td>0.110** (0.0390)</td>
<td>0.00636 (0.0313)</td>
<td>-0.0186 (0.0677)</td>
<td>0.0395+ (0.0229)</td>
</tr>
<tr>
<td>Cohort 2007-2011, affected by earthquake (P)</td>
<td>0.0529*** (0.00815)</td>
<td>0.057*** (0.00864)</td>
<td>0.0469+ (0.0271)</td>
<td>0.0548*** (0.00580)</td>
<td>0.0497*** (0.00755)</td>
<td>0.0380*** (0.00799)</td>
<td>0.0209 (0.0238)</td>
<td>0.0429*** (0.00535)</td>
</tr>
<tr>
<td>P*E</td>
<td>-0.0377*** (0.00969)</td>
<td>-0.0440*** (0.00981)</td>
<td>-0.115*** (0.0291)</td>
<td>-0.0492*** (0.00667)</td>
<td>-0.0284*** (0.00900)</td>
<td>-0.0633*** (0.00908)</td>
<td>-0.0519*** (0.0256)</td>
<td>-0.059*** (0.00616)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.429*** (0.0300)</td>
<td>-0.274*** (0.0283)</td>
<td>-0.180** (0.0769)</td>
<td>-0.333*** (0.0200)</td>
<td>-0.373*** (0.0288)</td>
<td>-0.0999*** (0.0263)</td>
<td>0.0526 (0.0679)</td>
<td>-0.185*** (0.0185)</td>
</tr>
<tr>
<td>School Fixed Effects</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>97057</td>
<td>116446</td>
<td>20389</td>
<td>233892</td>
<td>97658</td>
<td>117011</td>
<td>20501</td>
<td>235170</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001
Table 3: Pre-Earthquake Cohort, dependent variables Spanish and Math test scores in eighth grade

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

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>
<thead>
<tr>
<th></th>
<th>top 33 percent</th>
<th>bottom 33 percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother’s education</td>
<td>.0007012</td>
<td>.0020148*</td>
</tr>
<tr>
<td></td>
<td>(.0007415)</td>
<td>(.0008893)</td>
</tr>
<tr>
<td>Father’s education</td>
<td>-.0007132</td>
<td>.0013843</td>
</tr>
<tr>
<td></td>
<td>(.0007134)</td>
<td>(.0008544)</td>
</tr>
<tr>
<td>Household income</td>
<td>1.76e−08+</td>
<td>−2.32e−08+</td>
</tr>
<tr>
<td></td>
<td>(1.02e-08)</td>
<td>(1.23e-08)</td>
</tr>
<tr>
<td>Math test score t-1</td>
<td>−.0549353***</td>
<td>.1004088***</td>
</tr>
<tr>
<td></td>
<td>(.0019868)</td>
<td>(.0023838)</td>
</tr>
<tr>
<td>Seismic intensity at student’s home</td>
<td>.0128674***</td>
<td>−.0105234***</td>
</tr>
<tr>
<td></td>
<td>(.0022842)</td>
<td>(.002741)</td>
</tr>
<tr>
<td>Observations</td>
<td>46059</td>
<td>46059</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
+ $p < 0.10$,  * $p < 0.05$,  ** $p < 0.01$,  *** $p < 0.001$
Figure 6: Examples of estimated $m(c)$ functions in two classrooms.

Figure 7: Estimated $\hat{\Delta}e(c)$ for Spanish test scores. One-sided 90 percent confidence interval reported.
Table 5: No Evidence that Intensity Variance Affected Classroom Composition

<table>
<thead>
<tr>
<th></th>
<th>(1) Mean Spanish Lagged Test Score</th>
<th>(2) Mean Math Lagged Test Score</th>
<th>(3) Mean hh income (CLP)</th>
<th>(4) Mean father’s education (yrs)</th>
<th>(5) Mean mother’s education (yrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>School is in earthquake region (E)</td>
<td>-0.282*** (0.0364)</td>
<td>-0.358*** (0.0377)</td>
<td>-48529.1*** (9599.6)</td>
<td>1.050*** (0.150)</td>
<td>0.538*** (0.140)</td>
</tr>
<tr>
<td>Cohort 2007-2011, affected by earthquake (P)</td>
<td>0.0241 (0.0152)</td>
<td>-0.0370* (0.0157)</td>
<td>22318.3*** (3987.5)</td>
<td>0.0541 (0.0622)</td>
<td>0.0393 (0.0583)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(1-P)*Classroom Intensity Mean, $\theta_{\mu}^{pre}$</th>
<th>(1-P)*Classroom Intensity Mean, $\theta_{\mu}^{post}$</th>
<th>(2-P)*Classroom Intensity Mean, $\theta_{\mu}^{pre}$</th>
<th>(2-P)*Classroom Intensity Mean, $\theta_{\mu}^{post}$</th>
<th>(1-P)*Classroom Intensity Variance, $\theta_{\sigma^2}^{pre}$</th>
<th>(1-P)*Classroom Intensity Variance, $\theta_{\sigma^2}^{post}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1-P)*Classroom Intensity Mean, $\theta_{\mu}^{pre}$</td>
<td>0.0388*** (0.00598)</td>
<td>0.0522*** (0.00618)</td>
<td>2086.0 (1581.7)</td>
<td>-0.216*** (0.0247)</td>
<td>-0.102*** (0.0231)</td>
<td></td>
</tr>
<tr>
<td>(1-P)*Classroom Intensity Mean, $\theta_{\mu}^{post}$</td>
<td>0.0334*** (0.00599)</td>
<td>0.0532*** (0.00619)</td>
<td>-306.2 (1572.6)</td>
<td>-0.218*** (0.0245)</td>
<td>-0.101*** (0.0230)</td>
<td></td>
</tr>
<tr>
<td>Effect of Classroom Intensity Mean, $\theta_{\mu}^{post} - \theta_{\mu}^{pre}$</td>
<td>-0.0054516+ (0.0030045)</td>
<td>0.0009108 (0.0031042)</td>
<td>-2392.183** (787.414)</td>
<td>-0.0020608 (0.0122813)</td>
<td>0.0014217 (0.015087)</td>
<td></td>
</tr>
</tbody>
</table>

|                                   | (1-P)*Classroom Intensity Variance, $\theta_{\sigma^2}^{pre}$ | (1-P)*Classroom Intensity Variance, $\theta_{\sigma^2}^{post}$ | (2-P)*Classroom Intensity Variance, $\theta_{\sigma^2}^{pre}$ | (2-P)*Classroom Intensity Variance, $\theta_{\sigma^2}^{post}$ | Effect of Classroom Intensity Variance, $\theta_{\sigma^2}^{post} - \theta_{\sigma^2}^{pre}$ |
|-------------------------------|-----------------------------------|--------------------------------|--------------------------|-----------------------------------|----------------------------------|----------------------------------|
| (1-P)*Classroom Intensity Variance, $\theta_{\sigma^2}^{pre}$ | -0.0373 (0.119) | -0.0306 (0.123) | 8081.1 (31829.6) | -1.526** (0.496) | -1.228** (0.465) |
| (1-P)*Classroom Intensity Variance, $\theta_{\sigma^2}^{post}$ | 0.174* (0.0842) | 0.0601 (0.0871) | 2292.5 (22021.0) | -0.827* (0.343) | -0.326 (0.322) |
| Effect of Classroom Intensity Variance, $\theta_{\sigma^2}^{post} - \theta_{\sigma^2}^{pre}$ | 0.210873 (0.1456232) | 0.0907525 (0.1505045) | -5788.624 (38715.2) | 0.6992537 (0.6033702) | 0.9024236 (0.5656727) |

| Constant | -0.122*** (0.0108) | -0.135*** (0.0111) | 234675.0*** (2861.5) | 9.342*** (0.0446) | 9.223*** (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 6: No Evidence that the Relationship Between Classroom Intensity Variance and Classroom Quality Changed After the Earthquake Struck (Spanish teachers)

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

Standard errors in parentheses

+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001
Table 7: No Evidence that the Relationship Between Classroom Intensity Variance and Classroom Quality Changed After the Earthquake Struck (Math Teachers)

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

Standard errors in parentheses

$+ p < 0.10$, $* p < 0.05$, $** p < 0.01$, $*** p < 0.001$
Table 8: Parameter Estimates (bootstrapped standard errors in parentheses)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient on</th>
<th>Math</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{12}$</td>
<td>Parental Education</td>
<td>$-0.01162^{***}$</td>
<td>$-0.02116^{***}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00516)</td>
<td>(0.00446)</td>
</tr>
<tr>
<td>$\alpha_{13}$</td>
<td>High Income Dummy</td>
<td>$-0.05596^{***}$</td>
<td>$-0.03560^{**}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01620)</td>
<td>(0.01749)</td>
</tr>
<tr>
<td>$\alpha_{14}$</td>
<td>Female</td>
<td>$0.129037^{***}$</td>
<td>$-0.23034^{***}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01953)</td>
<td>(0.03504)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>Seismic Intensity</td>
<td>0.032588</td>
<td>0.09463</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.05962)</td>
<td>(0.14377)</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>Seismic Intensity*High Income</td>
<td>$-0.00037^{***}$</td>
<td>$-0.00037$ $(0.0000)$ $(0.00271)$</td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>Seismic Intensity*Female</td>
<td>-0.00313</td>
<td>0.05500*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.028773)</td>
<td>(0.03341)</td>
</tr>
</tbody>
</table>

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

Figure 8: Estimated $\widehat{\Delta e}(c)$ for Math test scores. One-sided 90 percent confidence interval reported.
<table>
<thead>
<tr>
<th></th>
<th>Pre-Earthquake Cohort</th>
<th>Post-Earthquake Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mathematics Actual</td>
<td>Mathematics Model</td>
</tr>
<tr>
<td>Overall</td>
<td>-0.185</td>
<td>-0.189</td>
</tr>
<tr>
<td>Female</td>
<td>-0.304</td>
<td>-0.283</td>
</tr>
<tr>
<td>Male</td>
<td>-0.058</td>
<td>-0.089</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>-0.300</td>
<td>-0.279</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.322</td>
<td>-0.302</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>-0.035</td>
<td>-0.066</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.159</td>
<td>-0.188</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower Income</td>
<td>-0.414</td>
<td>-0.387</td>
</tr>
<tr>
<td>Higher Income</td>
<td>-0.130</td>
<td>-0.120</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower Income</td>
<td>-0.222</td>
<td>-0.246</td>
</tr>
<tr>
<td>Higher Income</td>
<td>0.155</td>
<td>0.116</td>
</tr>
</tbody>
</table>
Equilibrium Achievement Functions Do Not Cross

Equilibrium Achievement Functions Cross Once

Figure 9: No minimum guaranteed level of utility ($\phi = 0$). The upper panels show two admissible pairs of equilibrium achievement functions $y_A$ and $y_B$ for distributions $G_A$ and $G_B$ respectively, with $G_A \succ_{ULR} G_B$ and $\phi = 0$. The two functions can cross at most once. The lower panels show the difference between the equilibrium functions: $Dy(c) = y_B(c) - y_A(c)$. The function tracing the difference can cross the x-axis at most once. If it does not cross it (panel a), it must lie below it. If it crosses it (panel b), the sequence of its signs, from low $c$ to large $c$, is $+, −$. 
Table 10: Difference-in-Differences Evaluation of the Effect of Seismic Intensity Variance on Test Scores. Dependent variables Spanish and Math test scores in eighth grade.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Math</td>
<td>Spanish</td>
</tr>
<tr>
<td>Math test score in 4th grade</td>
<td>0.638***</td>
<td>0.657***</td>
</tr>
<tr>
<td></td>
<td>(0.00254)</td>
<td>(0.00272)</td>
</tr>
<tr>
<td>Spanish test score in 4th grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Item] Household lives in earthquake region (E)</td>
<td>0.338*</td>
<td>0.0874</td>
</tr>
<tr>
<td></td>
<td>(0.133)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>Cohort 2007-2011, affected by the earthquake (P)</td>
<td>0.0371***</td>
<td>0.0528***</td>
</tr>
<tr>
<td></td>
<td>(0.00936)</td>
<td>(0.00997)</td>
</tr>
<tr>
<td>(1-P)*Classroom</td>
<td>0.0753</td>
<td>-0.193+</td>
</tr>
<tr>
<td>Intensity Mean</td>
<td>(0.102)</td>
<td>(0.111)</td>
</tr>
<tr>
<td>P*Classroom</td>
<td>0.0676</td>
<td>-0.191+</td>
</tr>
<tr>
<td>Intensity Mean</td>
<td>(0.102)</td>
<td>(0.111)</td>
</tr>
</tbody>
</table>

| Effect of Intensity Mean | -0.0076943 | 0.0020984 |
| (1-P)*Classroom          | -0.113     | 0.238*    |
| Intensity Variance       | (0.101)    | (0.108)   |
| P*Classroom              | -0.0243    | -0.0171   |
| Intensity Variance       | (0.0633)   | (0.0719)  |

| Effect of Intensity Variance | 0.0885604 | -0.2549768* |
| Constant                   | -0.675    | 0.331      |
| (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.
Table 11: Difference-in-differences evaluation of the effect of the seismic intensity mean and variance on coverage of the Spanish curriculum

<table>
<thead>
<tr>
<th></th>
<th>% of Spanish curriculum covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish teacher experience</td>
<td>-0.000137 (0.000211)</td>
</tr>
<tr>
<td>School is in earthquake region (E)</td>
<td>-0.000161 (0.0151)</td>
</tr>
<tr>
<td>Cohort 2007-2011, affected by earthquake (P)</td>
<td>0.0283*** (0.00675)</td>
</tr>
<tr>
<td>(1-P)*Classroom Intensity Mean</td>
<td>-0.00149 (0.00248)</td>
</tr>
<tr>
<td>P*Classroom Intensity Mean</td>
<td>-0.00110 (0.00246)</td>
</tr>
<tr>
<td>Effect of Earthquake Intensity Mean</td>
<td>0.0003955 (0.0013076)</td>
</tr>
<tr>
<td>(1-P)*Classroom Intensity Variance</td>
<td>-0.00192 (0.00574)</td>
</tr>
<tr>
<td>P*Classroom Intensity Variance</td>
<td>0.0415 (0.0346)</td>
</tr>
<tr>
<td>Effect of Earthquake Intensity Variance</td>
<td>0.0222701 (0.0670429)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.661*** (0.00874)</td>
</tr>
</tbody>
</table>

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).
Figure 10: Minimum guaranteed level of utility ($\phi > 0$). The upper panels show two admissible pairs of equilibrium achievement functions $y_A$ and $y_B$ for distributions $G_A$ and $G_B$ respectively, with $G_A \succ_{ULR} G_B$ and $\phi > 0$. The two functions cross once or twice. The lower panels show the difference between the equilibrium functions: $Dy(c) = y_B(c) - y_A(c)$. The function tracing the difference crosses the x-axis once or twice. If it crosses it once (panel a), the sequence of its signs, from low $c$ to large $c$, is $-,$ $+$. If it crosses it twice (panel b), the sequence of its signs, from low $c$ to large $c$, is $+,$ $-,$ $+$. 
Figure 11: Type distributions in two classrooms, and cutoffs separating low-, middle- and high-\(c\) students.
References


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


56


A Appendix

A.1 Algorithm for the Estimation of the Semiparametric Single-Index Model

• Normalize to a constant one of the elements of $\alpha_1$, because only the ratios among the components of $\alpha$ are identified. I normalize to -1 the coefficient on initial ability.

• Make an initial guess for all the other elements of $\alpha$.

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

• 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 $h = 1.06\hat{\sigma}_n^{-1/5}$, minimizing the Approximated Mean Integrated Squared Error (AMISE). 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 $c_i$.

• Compute the sum of squared residuals in each $l$ at the current guess for $\alpha$:

$$SSR_l(\alpha) = \sum_{i \in l} w_i (y_i - \hat{m}_l(c_i;\alpha))^2.$$ The weights are the same as those used in the kernel estimator of $m$.

• Loop over $\alpha$ to find the parameter value that minimizes $\sum_l SSR_l(\alpha)$

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 dataset is clustered at the classroom level. While the functions $m$ are allowed to differ

---

52 The MISE is equal to $E\{\int [\hat{m}(c) - m(c)]^2 dx\} = \int [(Bias\hat{m})^2 + V(\hat{m})] dc$, and AMISE substitutes the expressions for the bias and variance of $\hat{m}$ with approximations. See Pagan and Ullah (1999), p. 24.
by classrooms, the parameter $\alpha$ is restricted to be identical in all classrooms. To account for the clustered sample design in the estimation of the standard errors of the $\alpha$ parameters, I bootstrap 100 samples stratified at the classroom level, and I estimate $\alpha$ in each bootstrapped sample to obtain the standard errors.

The standard errors of $\hat{\mu} \Delta e(\cdot)$, which are needed to test the comparative statics result, cannot be easily bootstrapped for computational reasons. 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 $V(\hat{m}_l(c)) = \frac{\sigma_l^2}{n_l h_l(c)} \int K^2(\psi) d\psi + o(n^{-1}h^{-1})$, where $\sigma^2$ is the variance of $\epsilon$, $h_l$ is the bandwidth, $n_l$ is the size of classroom $l$ (on average this is 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 $\hat{m}_l(c)$ on a fine grid for $c$. I substitute $f(c)$ with its kernel estimator, and $\sigma^2_l$ 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 $\frac{1}{\mu} \Delta e(c)$:

$$V \left( \frac{1}{\mu} \Delta e(c) \right) = \sum_{l=1}^{N_{pre}-1} \sum_{l'=l+1}^{N_{pre}} \sum_{s=1}^{N_{post}-1} \sum_{s'=s+1}^{N_{post}} \kappa_{ll's's'} \left( V(\hat{m}_{s'}^{post}(c)) + V(\hat{m}_{s'}^{post}(c)) + V(\hat{m}_l^{pre}(c)) + V(\hat{m}_l^{pre}(c)) \right).$$

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

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

where $\omega_{ll's's'}$ is defined in equation [10].

A.2 Model Details and Proofs

A.2.1 Differential Equation

The first-order differential equation characterizing equilibrium strategies is obtained by rearranging the first order condition in [12] and substituting $c'(e) = \frac{1}{e(c)}$:

\[53\]This would require submitting around 4,000 jobs of duration 72 hours each.
\[ e'(c_i) = \left( \frac{g(c_i)}{1 - G(c_i) + \phi} \right) \left( \frac{V(y(e), q(e, c))}{\mu V_1 + V_2 \frac{\partial q}{\partial c}} \right) . \] (13)

\[ = \frac{g(c_i)}{1 - G(c_i) + \phi} \psi(e_i, c_i). \]

A.2.2 Proof of Proposition 7.1

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.\(^{54} \) The key differences with the proofs in Hopkins and Kornienko (2004) derive from the fact that here the consumption good is the same as the positional good, there is a direct cost for producing the positional good, and agents’ types, corresponding to income in Hopkins and Kornienko (2004), increase the cost.

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 absence of rank concerns, i.e., \( V_1^1 \Big|_{e=e_{nr}} = -V_2^2 \frac{\partial q}{\partial c} \Big|_{e_{nr}} \). This is because if \( e < e_{nr}(c_i) \), then \( F_E(e) + \phi < F_E(e_n) + \phi \) and \( V(y(e(c)), q(e(c), c)) < V(y(e_{nr}(c)), q(e_{nr}(c), c)) \).

Therefore, \( V(y(e), q(e, c)) (F_E(e) + \phi) < V(y(e_{nr}(c)), q(e_{nr}(c), c)) (F_E(e_{nr}) + \phi) \), 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 \( e_{nr}(c_i) \) 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 \( \frac{\partial q}{\partial c} > 0 \), \( \frac{\partial q}{\partial e} > 0 \), \( \frac{\partial^2 q}{\partial e^2} > 0 \) and \( \frac{\partial^2 q}{\partial e \partial c} \geq 0 \). To see why, let \( FOC(e, c) = V_1^1 \Big|_{\mu} + V_2 q_1 \) and notice that by the Implicit Function Theorem:

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

The numerator is:

\(^{54} \)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.
\[ \frac{\partial FOC}{\partial c} = \frac{1}{\mu} V_{12} \frac{\partial q}{\partial c} + V_{22} \frac{\partial q}{\partial e} \frac{\partial q}{\partial c} + V_2 \frac{\partial^2 q}{\partial e \partial c} \leq 0. \]

The denominator is:
\[ \frac{\partial FOC}{\partial e} = \frac{1}{\mu^2} V_{11} + \frac{1}{\mu} V_{12} \frac{\partial q}{\partial e} + \left( \frac{1}{\mu} V_{21} + V_{22} \frac{\partial q}{\partial e} \right) \frac{\partial q}{\partial e} + V_2 \frac{\partial^2 q}{\partial e^2} \leq 0. \]

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

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 \( \tilde{e}_i \in (e_{nr}(c_i), e_i) \),
\[ \frac{\partial V}{\partial c_i} (y(e_i), q(e_i, c_i)) (F_E(e_i) + \phi) < \frac{\partial V}{\partial c_i} (y(\tilde{e}_i), q(\tilde{e}_i, c_i)) (F_E(\tilde{e}_i) + \phi). \]

Rewrite the left-hand side as:
\[ \frac{\partial V}{\partial c_i} (y(e_i), q(e_i, c_i)) (F_E(\tilde{e}_i) + \phi) + \frac{\partial V}{\partial c_i} (y(e_i), q(e_i, c_i)) (F_E(e_i) - F_E(\tilde{e}_i)). \]

The first term is smaller or equal to the right-hand side of equation [14] 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 \( \frac{\partial^2 q}{\partial e \partial c} \geq 0 \). To see why, notice that \( \frac{\partial^2 V}{\partial c \partial e} = \left( V_{21} \frac{1}{\mu} + V_{22} \frac{\partial q}{\partial e} \right) \frac{\partial q}{\partial c} + V_2 \frac{\partial q}{\partial c} \leq 0 \). The second term is strictly negative, because first, \( \frac{\partial V}{\partial c} \) 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 [14] 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.

A.2.3 Comparative Statics

**Definition** Two distributions \( G_A, G_B \) with support on \( [c, \bar{c}] \) 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. 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:

**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 $c \in [\hat{c}^-, \hat{c})$ 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 $(\hat{c}, \hat{c}^+)$ so that $y(e_A(c)) > y(e_B(c))$ for all $c \in [\hat{c}^-, \hat{c}]$, 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.

### A.3 Testing Monotonicity of $m(c)$

The procedure that I use is an application of Chetverikov (2013). 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 while the constant in $m$ might differ across classrooms, the derivative of $m$ should be very similar across classrooms. 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.

---

55The 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.
Details of the simulations can be found in the Online Supplementary Material. 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, because $\hat{c}_i = \hat{\alpha}_1 x_i + \hat{\alpha}_2 I_i P_i + \hat{\alpha}_3 I_i P_i x_i$. 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).