A Behavioral Model of Non-Linear Peer Effects in Cognitive Achievement

Pilar Alcalde*

February 21, 2013

Abstract

Most work on achievement peer effects focuses on their shape and variables that transmit them. But these effects are generated by different mechanisms, with potentially different policy implications. Also, the endogeneity of unobserved effort may bias the estimation of peer effects. This paper proposes and estimates a model that explicitly considers the student effort decision and achievement production in a class. Three mechanisms for peer effects are considered: a production externality of peer effort, the effect of lagged achievement heterogeneity on production, and the effect of peer characteristics in the cost of effort. The model is estimated using data for Chile. Results show a large and positive production externality, and mixed coefficients for the other two mechanisms. Counterfactual changes on class composition or the level of peer characteristics have larger effects on achievement than on effort. Students compensate to some degree the changes in the quality of their peer group through their effort, and thus policies are predicted to have a smaller effect than when the student effort decision is not considered.

I Introduction

In recent years, economists have been increasingly interested in studying peer effects, i.e., interdependencies between individual decisions that are taken jointly by members of a group. Peer effects have been studied for various socio-economic situations: crime, obesity, smoking behavior, sexual initiation, employment search and several others (see Blume, Brock, Durlauf and Ioannides, 2010; Durlauf, 2004, for some surveys).

*I thank my thesis advisors for their dedication, support, and hours of their time: Petra E. Todd, Kenneth I. Wolpin, and Flavio Cunha. I am also grateful to Áureo de Paula, Leonardo Epstein and Claudio Lucarelli, and to participants at the seminars in P. Universidad Católica de Chile, Universidad Adolfo Ibáñez and Universidad Diego Portales for helpful comments.

One area of particular interest is the study of peer effects in cognitive achievement. The growing literature has provided strong evidence that classmates have an important effect on the determination of achievement. These studies also recognize the existence of different mechanisms that can generate these effects (Duflo, Dupas and Kremer, 2011; Hoxby and Weingarth, 2006). But in practice, little attention has been given to the identification and estimation of said mechanisms (Blume et al., 2010; Richards, 2012).[1]

These peer effects estimates are thus a *composite* of different mechanisms. Still, their study has given many valuable contributions. Among them, some consensus has been reached regarding the shape of peer effects, and what variables are most important in the transmission of social interactions. But the explicit consideration of the underlying mechanisms is important for two reasons.

On one hand, some context features require attention when studying peer effects. First, the outcome variable (e.g. test scores, graduation rates or some other measure of achievement) is not a behavioral variable per se, as in other topics where peer effects have been studied. It depends both on behavior and predetermined characteristics like ability or socio-economic background. Second, the behavioral variable, effort, is typically not observed. It may also depend on observable characteristics of the student and her peers, and this endogeneity may bias the estimation of peer effects, even providing estimates with the opposite sign of the true relationship (Cooley, 2009).

On the other hand, different mechanisms may potentially have different policy implications. The composite parameters of peer mechanisms deliver predictions that are only valid under the current policy regime, and may not be useful to inform the effect of out-of-sample predictions or alternative policies on achievement. This point is stressed in the experiment by Carrell, Sacerdote and West (2011). They used randomly allocated groups to estimate composite nonlinear peer effects, and then used their estimates to create "optimally" sorted peer groups. The design sorted

[1] Blume et al. (2010) write: "A final area that warrants far more research is the micro foundations of social interactions. In the econometrics literature, contextual and endogenous social interactions are defined in terms of types of variables rather than via particular mechanisms. This can delimit the utility of the models we have, for example, if the particular mechanisms have different policy implications." (p. 941)

incoming student into groups in a specific way to increase the grades of the lowest ability students. Contrarily, they observe a negative and significant treatment effect for those students, a positive and significant effect for middle students, and no effect for high ability students.

This highlights the danger in using composite estimates to perform out-of-sample predictions. The main problem is not the choice of an appropriate specification (Hurder, 2012), but the need of a deeper understanding of the endogenous response of the students and the underlying mechanisms that drive social interactions.

This paper proposes a behavioral model of student effort and achievement, with different mechanisms that generate peer effects, and estimates the magnitude and relevance of each peer mechanism. In this sense, the paper contributes to bridge the gap between the theoretical and empirical literatures, with a fully specified model of student effort and peer interactions that is identified and estimated.

The model proposed considers explicitly the role of own and peer effort. Students selected into a class choose their effort to maximize utility, which depends on their current achievement and the cost of effort. This effort is needed to produce current achievement, which occurs using a nonlinear value-added function. Both effort and achievement are observed to all members of the class, but are only observed with error by the econometrician. These features account for four common problems in the estimation of peer effects: self-selection, identification, measurement error in regressors, and omitted variables bias.

The model separately identifies three mechanisms for peer effects. First, the production of achievement depends directly on the average effort made by peers, as a production externality. Second, observable peer characteristics affect the individual cost of making effort. Third, the level of heterogeneity of lagged achievement affects the productivity parameters, in particular the productivity of own effort (Ding and Lehrer, 2007; Hoxby and Weingarth, 2006; Kang, 2007).

The model does not have an analytic solution; however it is possible to prove that the equilibrium exists and is unique under mild assumptions on the parameters. Uniqueness thus comes from the parametric assumptions of the model, and is not general to models of social interactions (Bisin, Moro and Topa, 2011). Due to the possibility of the Reflection Problem (Manski, 1993), identification is a key issue for the estimation of peer effects. Although this problem does not occur in nonlinear models (Brock and Durlauf, 2001*b*; Durlauf and Ioannides, 2010), other identification problems may be present, and thus the issue is given detailed attention. The parameters of the model are identified using information from the measurement equations for effort and achievement, from the moment conditions given by the production function of achievement, and from the first order condition of the student problem.

As previously stated, this paper contributes to bridge the gap between theoretical and empirical studies of peer effects. Several authors have provided models to study general contexts of social interactions (Bisin et al., 2011; Blume et al., 2010; Brock and Durlauf, 2001*a*,*b*; Durlauf and Ioannides, 2010). However, the two main differences between these papers and the context studied here (i.e. the unobservability of the behavioral variable and the mixed nature of the outcome variable) imply that the usual models and estimation techniques for peer effects models need to be adapted to the context of cognitive achievement.

Other theoretical models of peer effects are specific to cognitive achievement. Lazear (2001) presents a model of class size determination with peer externalities, but these are given and students do not choose how they behave. Cooley (2009) shows how the linear-in-means model of peer effects is not identified, beyond the Reflection Problem, when several unobserved mechanisms are included in the model and only one observed variable is used as a proxy. A related literature is the family of identity models, where students choose their social category and then school effort to match the ideal of the category (Akerlof and Kranton, 2002; Austen-Smith and Fryer, 2005). Cicala, Fryer and Spenkuch (2011) develop a Roy model of social interactions in which the production of two activities (study or mischief) occur within peer groups, and individuals choose their sector to maximize "social income".

A few empirical papers give more emphasis to peer effects mechanisms, but do not estimate directly the size of the different mechanisms. Cooley (2008) presents a nonparametric model of effort choice similar to the one presented here, but the peer mechanisms are not separately identified. Hoxby and Weingarth (2006) and Duflo et al. (2011) estimate peer effects, and discuss in detail the theoretical models and mechanisms that could give rise to the patterns they observe. Unfortunately, they do not estimate the models discussed directly. Lavy and Schlosser (2011) examine the existence of gender peer effects in different teacher and student variables that may affect achievement, but they do not estimate explicitly how these variables act as mechanisms on achievement.

The model proposed can be estimated using a Simulated Maximum Likelihood algorithm (Laroque and Salanie, 1989). This algorithm is applied to data from Chile's SIMCE, a standardized test that students take in 4th, 8th or 10th grade, depending on the survey year. The dataset contains students that gave the test twice: 4th grade in 2005 and 8th grade in 2009. It has several features that make it attractive to use: it is a large national database, including all students in the country, from both public and private schools, with detailed information about the students and their classroom, parents and teachers.

The parameters are precisely estimated, and provide evidence of significant peer effects. The effect of peer effort on production is positive and relatively large. The peer parameters on the cost of effort are not always positive: some variables have a positive effect on own effort by decreasing its cost, but others increase it or have no significant effect. Finally, class heterogeneity has a mixed impact on productivity: it increases some parameters of the production function but decreases others. Because the model is nonlinear, the parameter values are not very informative about the effect of the different variables on student outcomes, and thus marginal effects are calculated. The model predicts a rich pattern of marginal effects.

Finally, the estimated model is used in counterfactual analysis, this is, out-of-sample predictions of effort and achievement under different policies that affect peer group composition or characteristics. This is particularly useful to analyze policies where isolating the causal effect or obtaining an appropriate control group is hard. These counterfactuals help understand in more detail the strength and patterns of social interactions, as they exogenously change different dimensions of the model, one at a time. This sheds some light on the way in which peer effects operate in the classroom.

The model is used to analyze six different policies. The first counterfactual, providing tutors to the students, tries to isolate the endogenous effect, by exogenously lowering the cost function of the students. Counterfactuals 2 to 4 explore the effect of changes in peer characteristics, by exogenously increasing the number of inputs at home, or analyzing same-gender schools and tracking programs, this is, changing class composition according to gender or lagged achievement. Counterfactual 5 analyzes the effect of class size on the magnitude of peer effects. Finally, Counterfactual 6 shows how the model can be used to analyze the cream-skimming effect of school choice programs, and to separate the effect of student characteristics and class composition from the productivity of the school.

In general, the results are similar to studies that identify the causal effect of policies using quasi-experimental or experimental methods. The model proposed is able to separate the mechanisms that generate those effects.

The results show interesting peer effects. The endogenous effect is positive and significant, very small on effort but of moderate size on achievement. The contextual peer effect may be positive or negative depending on the variable considered, and this may generate counterintuitive effects on achievement. In general, observed contextual variables have small or negative effects; instead, the peer previous achievement has a moderate to large effect. Some policies have larger effects on effort than achievement, but others work the opposite way, especially those related to class composition. However, peer effects on achievement seem to be larger than on effort. Finally, students compensate to some degree the changes in the characteristics of their peer group through their effort. This implies that policies are predicted to have a smaller effect than when the student effort decision is not considered.

The paper proceeds as follows. Section II describes the evidence of nonlinear peer effects in cognitive achievement. Sections III and IV outline the model of student behavior and show the existence and uniqueness of the equilibrium of the student game. Section V describes the dataset and the variables used in the estimation. Section VI proves the identification of the parameters of the model. Section VII discusses the results from the estimation. Finally, Section VIII implements the counterfactual analyses, and Section IX presents the main conclusions and future lines of work.

II Empirical Evidence of Nonlinear Peer Effects

The estimation of peer effects has provided valuable insight into the shape of achievement peer effects, although these are composite estimates of different mechanisms. The first contribution has been the specification on the functional form. The baseline model is the linear-in-means specification (for example, Carrell, Fullerton and West, 2009; Hanushek, Kain and Rivkin, 2009; Hoxby, 2000*b*, among many others), but there has been growing evidence that peer effects are nonlinear, this is, that the size of peer effects depends on own and peer characteristics (Burke and Sass, 2008; Carrell et al., 2011; Cooley, 2008; Ding and Lehrer, 2007; Gibbons and Telhaj, 2008; Pinto, 2010; Sacerdote, 2001). This specification also may be theoretically preferrable, both because it has more interesting consequences for policy interventions, and because the linear model imposes very strict restrictions on the behavior of students and the production of achievement (Alcalde, 2012).[2]

The second contribution has been determining what variables are most important in the transmission of social interactions. The relevant variable seems to be peer ability or peer lagged achievement. Once appropriately controlling for this variable, some studies find that there are no other significant contextual variables (Gibbons and Telhaj, 2008; Kang, 2007; Vigdor and Nechyba, 2004; Zabel, 2008), or that only the gender composition of the class is relevant, with a positive effect of the proportion of females (Hoxby and Weingarth, 2006; Lavy and Schlosser, 2011).

The largest current discussion refers to the specific shape of these peer effects. There is a very broad set of specifications used to estimate nonlinear peer effects. This may be related to the

[2] Because in a linear model the peer effect is homogeneous, total achievement is invariant to the specific distribution of students across classes, and thus policy interventions only have distributional consequences and cannot result in global welfare gains.

limited attention that has been given to the underlying mechanisms, which may suggest criteria regarding the functional form to be estimated.

Hurder (2012) notes that the nonlinear models most commonly used are usually extensions of a linear-in-means model. The simplest nonlinear models are "linear-in-moments", and include the mean and standard deviation of peer baseline ability (Burke and Sass, 2008; Vigdor and Nechyba, 2004; Zabel, 2008). Other studies use a "linear-in-shares" specification, where current achievement depends on the share of peers in a specific segment of the baseline ability distribution (Lavy, Silva and Weinhardt, 2009; McEwan and Soderberg, 2006), or on interactions between the relative position of the student and these shares of peers (Burke and Sass, 2008; Carrell et al., 2011; Gibbons and Telhaj, 2008; Hoxby and Weingarth, 2006; Sacerdote, 2001; Zimmerman, 1999). The more sophisticated models are semiparametric: partial linear models, where the specification is linear but the mean peer variable enters nonparametrically (Ding and Lehrer, 2007; Rodriguez, 2010), single index models (Pinto, 2010) and quantile regressions (Cooley, 2008; Ding and Lehrer, 2007; Kang, 2007). Froelich and Michaelowa (2011) estimate a nonparametric model for textbooks. Finally, there is some evidence of peer effects on tracking experiments (Carrell et al., 2011; Duflo et al., 2011).

These studies find positive peer effects, i.e., having a better peer in terms of lagged achievement improves own achievement. But Ding and Lehrer (2007) and Pinto (2010) find increasing peer effects, this is, larger peer effects for students with high lagged achievement. Instead, Burke and Sass (2008) and Rodriguez (2010) find larger peer effects for students with low lagged achievement. Lavy et al. (2009) find negative peer effects arising only from bad peers at the very bottom of the ability distribution. Sacerdote (2001) and Carrell et al. (2011) find U-shaped peer effects: having a top student benefits low and top students more versus middle students. Finally, Burke and Sass (2008); Ding and Lehrer (2007); Kang (2007) and Zabel (2008) find a positive effect of mean peer lagged achievement and a negative effect of the standard deviation of peer lagged achievement.

As noted above, a few empirical papers give more emphasis to peer effects mechanisms.Hoxby and Weingarth (2006) rationalize their evidence on shares specifications with models where there is a positive effect of mean peer ability but also a positive effect from homogeneity, but they do not estimate these models directly. The same models rationalize the tracking evidence from Carrell et al. (2011) and Duflo et al. (2011). Lavy and Schlosser (2011) suggest that the positive peer effects of girls they find are due mostly to compositional change, namely due to having more girls in the classroom and not due to improved behavior of peers.

Therefore, despite the success in the contributions made by the literature so far, little attention has been given to the separate identification and estimation of the possible mechanisms that generate those estimates. As explained previously, these mechanisms may have potentially different policy implications. Also, the endogeneity of unobserved effort may bias the estimation of peer effects. The next section will make explicit these ideas with a model of student effort choice and achievement production.

III Model Description

The current section presents a model of student effort choice in which peer effects arise. The education process occurs in two stages, a first stage in which students are assigned to schools, and a second stage in which effort is chosen and achievement is produced. The model includes three separate mechanisms of peer effects. Both effort and achievement are observed to all members of the class, but are only observed with error by the econometrician. These features account for four common problems in the estimation of peer effects: self-selection, identification, measurement error in regressors, and omitted variable bias.

The model is as follows. Consider a population of students, indexed by i. In the first stage, each of them is assigned to a group or class, indexed by g. Each student i in g is characterized by a vector of observable characteristics X_{ig} of size K_X, and a measure of own achievement in the previous period A_{ig0}, or lagged achievement, unobserved to the econometrician but measured with error in the data. All of these characteristics are predetermined. Each group g of size m_g is characterized by an unobserved productivity parameter α_g that captures teacher and school quality, which is fixed across classmates.

Group assignment is not at random. The selection of students to each class follows a common rule based on unobservables that generates *sorting* of students, conditional on observable variables. These students also may have been classmates in the past and shared common factors (Gibbons and Telhaj, 2008; Hanushek, Kain, Markman and Rivkin, 2003). These two features imply that lagged achievement is correlated among classmates. In the same way, teachers are *matched* into student groups based on unobservable characteristics, conditional on observable variables, generating correlation between class productivity and lagged achievement of the pupils. Therefore, the joint probability distribution of lagged achievement and unobserved productivity in a class g is assumed to follow a normal multivariate distribution of size $m_g + 1$:

$$\begin{pmatrix} \alpha_g \\ A_{1g0} \\ \vdots \\ A_{mg0} \end{pmatrix} \sim N\left(\begin{pmatrix} Y_g\pi_t \\ Z_{1g}\pi_s \\ \vdots \\ Z_{mg}\pi_s \end{pmatrix}, \begin{pmatrix} \sigma_t^2 & \sigma_{st} & \dots & \sigma_{st} \\ \sigma_{st} & \sigma_s^2 & \dots & \sigma_p \\ \vdots & & \ddots & \\ \sigma_{st} & \sigma_p & \dots & \sigma_s^2 \end{pmatrix}\right)$$

where Y_g s a vector of observable school and teacher variables of size K_Y, and Z_{ig} is a partition of X_{ig} of size K_Z,[3] and the variance of lagged achievement depends on the gender of the student, this is, $\sigma_s^2 \in \{\sigma_{sm}^2, \sigma_{sf}^2\}$.[4]

In the second stage, once groups are assigned, each student i in group g chooses effort in the current period e_{ig}, which is needed to produce achievement. The utility function for the student is given by

$$U_{ig} = A_{ig1} - \left(c_0 + X_{ig}\beta + \bar{X}_{-ig}\delta\right)\frac{e_{ig}^2}{2}$$

[3] Z_{ig} is not necessarily a subset of X_{ig}; this condition is not required for equilibrium characterization or for identification. Nonetheless, because X_{ig} will determine achievement in the current period, consistency of the model requires that similar characteristics determine lagged achievement in the previous period. But because individual characteristics are not observed for the previous period, Z_{ig} includes the variables in X_{ig} that are time invariant.

[4] A standard and related selection correction in the literature is to use individual and school by year fixed effects (Cooley, 2008; Hanushek et al., 2003; Hoxby, 2000*b*). The lack of longtudinal data prevents its use here.

and depends on own achievement in the current period A_{ig1}, and the cost function $c_{ig} = c_0 + X_{ig}\beta + \bar{X}_{-ig}\delta$ that measures the cost of effort. This cost function depends on individual variables X_{ig}, and the average of these background characteristics for the peers of i, $\bar{X}_{-ig}$ where the subscript $-i$ indicates that student i is excluded from the peer group and from this average.

Same as lagged achievement, current achievement and effort are perfectly observed to the members of the group, but unobserved to the econometrician. Current achievement A_{ig1} is produced with a nonlinear value-added function (Todd and Wolpin, 2003):

$$A_{ig1} = C_{ig}(A_{ig0})^{\gamma}\alpha_g(1 + \bar{e}^{\phi}_{-ig})e^{\omega_g}_{ig}$$

that depends on lagged achievement A_{ig0}, productivity of the class α_g, own effort e_{ig} and average peer effort $\bar{e}_{-ig}$. The coefficients C_{ig} and ω_g vary with individual and class characteristics respectively:

$$\begin{aligned} C_{ig} &= W_{Cig}\kappa_C \\ \omega_g &= W_{\omega g}\kappa_\omega \end{aligned}$$

where W_{Cig} is a vector of individual and class characteristics of size K_C, and $W_{\omega g}$ is a vector of school variables of size K_ω. The coefficients γ and ϕ are fixed in the population.[5] A variable of particular interest that is assumed to enter W_{Cig} and $W_{\omega g}$ is the level of heterogeneity of the class, measured as the standard deviation of lagged achievement in the group. The model assumes that the heterogenity of lagged achievement in the class affects the learning productivity of the students, both in the level of production and in the marginal productivity of own effort (Ding and Lehrer, 2007; Duflo et al., 2011; Hoxby and Weingarth, 2006; Kang, 2007).

Effort e_{ig}, and therefore achievement, is determined by the equilibrium of the student game in the class. Under a few parameter constraints, this equilibrium exists and is unique, as it will be

[5]The model assumes that education technology is common knowledge, and that the only difference between schools is the level of inputs.

shown in the next section.

The model proposed allows for three different mechanisms by which peer effects can affect student effort. First, the average effort made by peers operates as an externality in production with parameter ϕ. Second, observable peer characteristics affect the cost of making effort for the student, measured through δ. Third, the level of class heterogeneity affects the productivity parameters C_{ig} and ω_g. These three mechanisms will affect the effort decision of the student and thus the production of achievement. Lagged achievement is predetermined in the model and thus is it another channel for contextual effects. The model assumes that peer lagged achievement has a direct effect on own productivity through class heterogeneity, and an indirect effect through the effort decision of the peers and the endogenous effect. Note there is no direct effect of mean peer lagged achievement on own effort; these different mechanisms will play a role on the counterfactual analysis.

IV Existence and Uniqueness of Equilibrium

This section proves that, although the model does not have an analytic solution, under mild sufficient parameter conditions the equilibrium of the student game exists and is unique. This implies that effort e_{ig}, and therefore achievement A_{ig1}, are uniquely determined by the equilibrium of the student game in the class.

The assumptions on the parameters are that the marginal productivities of own effort and of peer effort are restricted to the $[0, 1]$ interval, and the cost of making effort is always positive. Thus, uniqueness comes from the parametric assumptions of the model, and is not general to models of social interactions (Bisin et al., 2011). [6]

For simplicity, define the cost function c_{ig} and the predetermined variables that enter individual

[6] It is extremely difficult to find necessary conditions in this model. The equilibrium depends on the characteristics of the school and of every student in the class, and on the realizations of the unobserved variables, and thus it is possible that an equilibrium exists even if the constraints on the parameters do not hold.

achievement K_{ig}:

$$\begin{aligned} c_{ig} &= c_0 + X_{ig}\beta + \bar{X}_{-ig}\delta \\ K_{ig} &= C_{ig}(A_{ig0})^{\gamma}\alpha_g \end{aligned}$$

and thus achievement can be written as:

$$A_{ig1} = K_{ig}(1 + \bar{e}^{\phi}_{-ig})e^{\omega}_{ig}$$

IV.1 Existence of the Effort Best Response Function

The student problem is solved by a unique effort best response function if $0 < \omega_g \leq 1$, $0 \leq \phi$ and $c_{ig} > 0$. If these restrictions hold, then there is a unique optimal level of own effort for each level of peer effort.

The first order condition of the student problem is:

$$\omega_g K_{ig}(1 + \bar{e}^{\phi}_{-ig})e^{\omega_g - 1}_{ig} - c_{ig}e_{ig} = 0 \tag{1}$$

The solution to this equation satisfies the second order condition, and thus it is a maximum, if $\omega_g \leq 1$ and $c_{ig} > 0$.[7]

To analyze the existence of an equilibrium of this game, it is possible to obtain a more useful first order condition by re-writing (1) as a function of the class average effort instead of peer effort. The relation between peer effort $\bar{e}_{-ig}$ and average effort $\bar{e}_g$ is given by:

$$\bar{e}_{-ig} = \frac{1}{m_g - 1}\sum_{j \neq i} e_{jg} + \frac{1}{m_g - 1}e_{ig} - \frac{1}{m_g - 1}e_{ig} = \frac{m_g}{m_g - 1}\bar{e}_g - \frac{1}{m_g - 1}e_{ig}$$

[7] The second order condition of the student problem is given by

$$\omega_g(\omega_g - 1)K_{ig}(1 + \bar{e}^{\phi}_{-ig})e^{\omega_g - 2}_{ig} - c_{ig} < 0$$

which holds for any value of peer effort $\bar{e}_{-ig}$ if $\omega_g \leq 1$ and $c_{ig} > 0$.

And therefore the first order condition (1) is equivalent to:

$$\omega_g K_{ig}\left(1+\left(\frac{m_g}{m_g-1}\bar{e}_g-\frac{1}{m_g-1}e_{ig}\right)^{\phi}\right)e_{ig}^{\omega_g-1}-c_{ig}e_{ig}=0 \tag{2}$$

Unfortunately equation (2) has no analytic solution. The above equation $F_i(e_{ig},\bar{e}_g)=0$ defines implicitly the effort best response that depends on class average effort. It is possible that this effort best response is a correspondence, this is, that for the same level of class average effort $\bar{e}_g$, student i finds optimal to choose two or more different levels of own effort. If $0<\omega_g\leq 1$ and $0\leq\phi$, this effort best response is unique for each level of class average effort, and thus defines the implicit function $e_{ig}=\tilde{g}_i(\bar{e}_g)$. This can be argued by observing that:

$$\begin{aligned}
F_i(e_{ig},\bar{e}_g) &= A_i(e_{ig},\bar{e}_g)-B_i(e_{ig},\bar{e}_g)=0\\
&\text{where}\\
A_i(e_{ig},\bar{e}_g) &= \omega_g K_{ig}\left(1+\left(\frac{m_g}{m_g-1}\bar{e}_g-\frac{1}{m_g-1}e_{ig}\right)^{\phi}\right)e_{ig}^{\omega_g-1}\\
B_i(e_{ig},\bar{e}_g) &= c_{ig}e_{ig}
\end{aligned}$$

Because $B_i(e_{ig},\bar{e}_g)$ is a linear increasing function of e_{ig}, and $A_i(e_{ig},\bar{e}_g)$ is concave and decreasing in e_{ig} if $0<\omega_g\leq 1$ and $0\leq\phi$[8], then the solution to $F_i(e_{ig},\bar{e}_g)=0$ is unique for every value of $\bar{e}_g$.

Therefore, if $0<\omega_g\leq 1$, $0\leq\phi$ and $c_{ig}>0$, the student problem is solved by a unique effort best response as a function of class average effort, $e_{ig}=\tilde{g}_i(\bar{e}_g)$.

[8] The first derivative of A_i with respect to e_{ig} is given by:

$$\begin{aligned}
\frac{\partial A_i}{\partial e_{ig}}(e_{ig},\bar{e}_g) &= \omega_g(\omega_g-1)K_{ig}\left(1+\left(\frac{m_g}{m_g-1}\bar{e}_g-\frac{1}{m_g-1}e_{ig}\right)^{\phi}\right)e_{ig}^{\omega_g-2}\\
&\quad -\omega_g\frac{\phi}{m_g-1}K_{ig}\left(\frac{m_g}{m_g-1}\bar{e}_g-\frac{1}{m_g-1}e_{ig}\right)^{\phi-1}e_{ig}^{\omega_g-1}
\end{aligned}$$

which is negative for every value of e_{ig} and $\bar{e}_g$ if $0<\omega_g\leq 1$ and $\phi\geq 0$.

IV.2 Existence of a Game Equilibrium

This section argues that there is at least one equilibrium for the student game if $0 < \omega_g \leq 1$, $0 \leq \phi < 1$ and $0 < c_{ig}$. These are the same conditions that guaranteed the existence of a best response function for every student, plus the condition that $\phi < 1$.

As in Bisin et al. (2011), an equilibrium in class g is a vector of effort levels $\{e_{ig}\}_{i \in g}$ such that the first order condition (2) is satisfied jointly for all $i \in g$, and the equilibrium aggregate $\bar{e}_g$ satisfies the natural consistency condition $\bar{e}_g = \sum_{i \in g} e_{ig}$. First, it is possible to see that the effort best response $\tilde{g}_i(\bar{e}_g)$ is monotonically increasing. By the Implicit Function Theorem,

$$\tilde{g}_i'(\bar{e}_g) = -\frac{\partial F_i(\tilde{g}_i(\bar{e}_g), \bar{e}_g)/\partial \bar{e}_g}{\partial F_i(\tilde{g}_i(\bar{e}_g), \bar{e}_g)/\partial e_{ig}}$$

where

$$\begin{aligned}
\frac{\partial F_i}{\partial \bar{e}_g}(\tilde{g}_i(\bar{e}_g), \bar{e}_g) &= \omega_g K_{ig} \tilde{g}_i(\bar{e}_g)^{\omega_g - 1} \frac{\phi m_g}{m_g - 1} \left(\frac{m_g}{m_g - 1} \bar{e}_g - \frac{1}{m_g - 1} \tilde{g}_i(\bar{e}_g) \right)^{\phi - 1} \\
\frac{\partial F_i}{\partial e_{ig}}(\tilde{g}_i(\bar{e}_g), \bar{e}_g) &= \omega_g(\omega_g - 1) K_{ig} \tilde{g}_i(\bar{e}_g)^{\omega_g - 2} \left(1 + \left(\frac{m_g}{m_g - 1} \bar{e}_g - \frac{1}{m_g - 1} \tilde{g}_i(\bar{e}_g)\right)^{\phi} \right) \\
&\quad - \omega_g K_{ig} \tilde{g}_i(\bar{e}_g)^{\omega_g - 1} \frac{\phi}{m_g - 1} \left(\frac{m_g}{m_g - 1} \bar{e}_g - \frac{1}{m_g - 1} \tilde{g}_i(\bar{e}_g) \right)^{\phi - 1} - c_{ig}
\end{aligned}$$

If $0 < \omega_g \leq 1$ and $0 \leq \phi$, then

$$\frac{\partial \tilde{F}_i}{\partial \bar{e}_g}(\tilde{g}_i(\bar{e}_g), \bar{e}_g) \geq 0 \quad \text{and} \quad \frac{\partial \tilde{F}_i}{\partial e_{ig}}(\tilde{g}_i(\bar{e}_g), \bar{e}_g) < 0$$

and thus $\tilde{g}_i'(\bar{e}_g) \geq 0$ for every $\bar{e}_g$.

As the effort best responses are increasing, by adding over individuals in the class, it is possible to obtain the function of average class effort

$$\tilde{G}(\bar{e}_g) = \frac{1}{m_g} \sum_{i \in g} \tilde{g}_i(\bar{e}_g)$$

Therefore, an equilibrium of the student game is equivalent to finding the fixed points of the function $\tilde{G}$, this is, $\bar{e}_g = \tilde{G}(\bar{e}_g)$. By Brouwer's Fixed Point Theorem[9], a fixed point $\bar{e}_g$ always exists if $0 < \omega_g \leq 1$, $0 \leq \phi < 1$ and $0 < c_{ig}$.

This theorem applies if it is possible to bound the domain of the function $\tilde{G}$ to $D = [0, \hat{e}_g]$, such that $\tilde{G}(\bar{e}_g) < \bar{e}_g$ for all $\bar{e}_g \geq \hat{e}_g$, or in other words, such that $\tilde{G}'(\bar{e}_g) < 1$[10]. By definition, this is equivalent to finding a large $\hat{e}_g$ such that $\tilde{g}_i'(\bar{e}_g) < 1$ for all $\bar{e}_g \geq \hat{e}_g$ and all $i \in g$.

First, suppose that $\omega_g = 1$. Then, $\tilde{g}_i'(\bar{e}_g) < 1$ is equivalent to

$$\begin{aligned} \frac{\partial F_i}{\partial \bar{e}_g}(\tilde{g}_i(\bar{e}_g), \bar{e}_g) < -\frac{\partial F_i}{\partial e_{ig}}(\tilde{g}_i(\bar{e}_g), \bar{e}_g) & \qquad \Rightarrow \\ K_{ig}\phi\left(\frac{m_g}{m_g - 1}\bar{e}_g - \frac{1}{m_g - 1}\tilde{g}_i(\bar{e}_g)\right)^{\phi-1} \quad < \quad c_{ig} & \end{aligned}$$

For a sufficiently large $\bar{e}_g$, this inequality holds for every value of K_{ig} and c_{ig} if $0 \leq \phi < 1$ and $0 < c_{ig}$.

Second, suppose that $\omega_g \in (0, 1)$. Consider what happens to both derivatives as $\bar{e}_g$ grows. Because $\tilde{g}_i'(\bar{e}_g) \geq 0$, then $\tilde{g}_i(\bar{e}_g)$ grows with $\bar{e}_g$, and therefore if $0 \leq \phi < 1$:

$$\begin{aligned} \frac{\partial F_i}{\partial \bar{e}_g}(\tilde{g}_i(\bar{e}_g), \bar{e}_g) &\rightarrow 0 \\ \frac{\partial F_i}{\partial e_{ig}}(\tilde{g}_i(\bar{e}_g), \bar{e}_g) &\rightarrow -c_{ig} \end{aligned}$$

Therefore, $\tilde{g}_i'(\bar{e}_g)$ will eventually become less than 1 for some finite value as long as $0 < c_{ig}$.[11]

[9] Every continuous function from a closed ball of an Euclidean space to itself has a fixed point.

[10] And thus it is possible to discard all points $\bar{e}_g \geq \hat{e}_g$ as reasonable average class effort levels.

[11] For showing that $\partial F_i(\tilde{g}_i(\bar{e}_g), \bar{e}_g)/\partial e_{ig}$ converges to $-c_{ig}$, it is possible to simplify the first term of the derivative to

$$\frac{(x - g(x))^{\phi}}{g(x)^{2-\omega_g}}$$

Both the numerator and the denominator converge to infinity as x grows. Therefore, using L'Hopital's rule, it is possible to write

$$\frac{\phi(x - g(x))^{\phi-1}(1 - g'(x))}{(2 - \omega_g)g(x)^{1-\omega_g}g'(x)}$$

which converges to zero as long as $\frac{1-g'(x)}{g'(x)}$ doesn't grow too fast.

Therefore, if $0 < \omega_g \leq 1$, $0 \leq \phi < 1$ and $0 < c_{ig}$, by Brouwer's Fixed Point Theorem an equilibrium for the student game always exists.

IV.3 Uniqueness of the Game Equilibrium

Social interactions typically give rise to multiple equilibria, because of the externalities induced by peer behavior on own behavior (Bisin et al., 2011). This is not the case in this particular model of social interactions, which has a unique equilibrium if $0 < \omega_g \leq 1$, $0 \leq \phi < 1$ and $0 < c_{ig}$; these are the same conditions used to prove the existence of an equilibrium. Under them, $\tilde{G}''(\bar{e}_g) < 0$ for all $\bar{e}_g$, and because $\tilde{G}$ is monotonically increasing, it has at most one fixed point.

By definition,

$$\tilde{G}''(\bar{e}_g) = \frac{1}{m_g} \sum_{i \in g} \tilde{g}_i''(\bar{e}_g)$$

From the first order condition (1), it is possible to see that it is never optimal for the student to choose $e_{ig} = 0$, for any value of $\bar{e}_{-ig}$.[12] Then, the first derivative $\tilde{g}_i'(\bar{e}_g)$ can be re-written as:

$$\tilde{g}_i'(\bar{e}_g) = \frac{D(\bar{e}_g) m_g}{D(\bar{e}_g) + E(\bar{e}_g)}$$

where

$$\begin{aligned} D(\bar{e}_g) &= \frac{\omega_g \phi}{m_g - 1} K_{ig} \left(\frac{m_g}{m_g - 1} \bar{e}_g - \frac{1}{m_g - 1} \tilde{g}_i(\bar{e}_g) \right)^{\phi - 1} \\ E(\bar{e}_g) &= (2 - \omega_g) c_{ig} \tilde{g}_i(\bar{e}_g)^{1 - \omega_g} \end{aligned}$$

Note from the results above that $D(\bar{e}_g) \geq 0$ and $E(\bar{e}_g) > 0$ for all $\bar{e}_g$, and thus $\tilde{g}_i'(\bar{e}_g) \in [0, m_g)$.

[12] Note that $e_{ig}(\bar{e}_{-ig} = 0) = \left(\frac{\omega K_{ig}}{c_{ig}} \right)^{1/(2-\omega)} > 0$.

The second derivative of the individual effort best response is given by

$$\begin{aligned}\tilde{g}_i''(\bar{e}_g) &= \frac{D'(\bar{e}_g)m_g\left(D(\bar{e}_g)+E(\bar{e}_g)\right)-D(\bar{e}_g)m_g\left(D'(\bar{e}_g)+E'(\bar{e}_g)\right)}{\left(D(\bar{e}_g)+E(\bar{e}_g)\right)^2} \\ &= \frac{D'(\bar{e}_g)m_gE(\bar{e}_g)-D(\bar{e}_g)m_gE'(\bar{e}_g)}{\left(D(\bar{e}_g)+E(\bar{e}_g)\right)^2}\end{aligned}$$

and therefore $\tilde{g}_i''(\bar{e}_g) < 0$ if $D'(\bar{e}_g)E(\bar{e}_g) - D(\bar{e}_g)E'(\bar{e}_g) < 0$, where

$$\begin{aligned}D'(\bar{e}_g) &= (\phi-1)D(\bar{e}_g)\left(\frac{m_g}{m_g-1}\bar{e}_g-\frac{1}{m_g-1}\tilde{g}_i(\bar{e}_g)\right)^{-1}\left[\frac{m_g}{m_g-1}-\frac{1}{m_g-1}\tilde{g}_i'(\bar{e}_g)\right] \\ E'(\bar{e}_g) &= (2-\omega_g)(1-\omega_g)c_{ig}\tilde{g}_i(\bar{e}_g)^{-\omega_g}\tilde{g}_i'(\bar{e}_g)\end{aligned}$$

Note that $D'(\bar{e}_g) < 0$ for all $\bar{e}_g$, as $\phi < 1$ and $\tilde{g}_i'(\bar{e}_g) < m_g$. Additionally, $E'(\bar{e}_g) \geq 0$ as $\omega_g \leq 1$ and $\tilde{g}_i'(\bar{e}_g) \geq 0$. Therefore, $\tilde{g}_i''(\bar{e}_g) < 0$ for all $\bar{e}_g$, and thus $\tilde{G}'' < 0$, guaranteeing a unique equilibrium of the game.

V Data

The model is estimated using data from Chile's SIMCE (Education Quality Measurement System). This is a national standardized test given every year to all students both in public and private schools, that are in 4th, 8th and/or 10th grade[13], and attempts to determine the degree of achievement related to the core curriculum. The SIMCE database has several features that make it attractive for estimation. It includes individual tests scores on several topics, at the national level, and detailed information of the students, their classes, and surveys to the parents and teachers.

The model is estimated using the information from the 2005 - 2009 panel. This is the same cohort of students who gave the SIMCE test twice, first in 2005 when they were in 4th grade, and then in 2009 when they were in 8th grade. A unique identifier allows to follow students even if

[13] 4th and 8th grade correspond to the 4th and last year of Elementary School (Educacion Basica), and 10th grade corresponds to the 2nd year of High School (Educacion Media).

they change school in the period. The sample contains full information for 130,853 students and their parents and teachers in 7,779 classes.

Section III assumed that achievement and effort were unobserved to the econometrician, but observed with error in the data. The SIMCE survey contains measures of four test scores: Math, Science, Language and Social Science. Of these, only the first two will be used in estimation, and thus the model will be estimated using a latent measure of mathematical achievement. There is a well documented link between mathematics achievement and labor market outcomes (Bishop, 1994; Murnane, Willett and Levy, 1995; NMAP, 2008). Furthermore, there is evidence that school effects are larger in mathematics than in language (Creemers, 2007; Teddlie and Stringfield, 2007), and thus it is expected that using math and science as the achievement measures would render clearer effects. The SIMCE survey also contains discrete measures of effort, self-reported by the student. Additionally, a continuous measure of effort is required for identification. The percentage of annual attendance to classes for each student is obtained from administrative data from the Ministry of Education (MINEDUC).

The variables used in estimation are described below:

- Measures of Individual Achievement: test scores for Math and Science.
- Measures of Individual Effort: 1 continuous and 4 categorical measures of effort.
    1. Percentage of annual attendance to classes.
    2. I do my homework even if I find it difficult (from 1, disagree, to 5, agree).
    3. My notebooks are generally complete (from 1, disagree, to 5, agree).
    4. During lectures I take notes of everything they teach us (from 1, disagree, to 5, agree).
    5. How many days (Mon-Fri) do you study or do homework after class? (from 1, never, to 4, every day)
- Vector X_{ig} of size $K_X = 8$: female, parent schooling defined as the maximum between mother or father schooling, parents have college education, both parents live at home, number of books at home is 10 to 50, number of books at home is 50 or more, computer at home, internet at home.
- Vector $\bar{X}_{-ig}$ of size $K_X = 8$: peer averages of the variables in vector X_{ig}.

- Vector W_{Cig} of size $K_C = 4$: constant, female, proportion of peer females, standard deviation of lagged achievement in the class.
- Vector $W_{\omega g}$ of size $K_\omega = 6$: constant, school is urban, school is public, school is subsidized, class size, standard deviation of lagged achievement in the class.[14]
- Vector Y_g of size $K_Y = 13$: constant, Math teacher gender, Science teacher gender, Math teacher experience, Science teacher experience, four dummies if any of the previous variables is missing for the class, school is urban, school is public, school is subsidized, class size.
- Vector Z_{ig} of size $K_Z = 7$: constant, female, parent schooling, parents have college education, school is urban, school is public, school is subsidized.

The distribution of the data is described in Table 1 for the test scores and individual and school variables, and in Table 2 for the discrete measures of effort.

VI Identification

The identification problem is key in the peer effects literature. In particular, it has been extensively discussed for linear models (Lee, 2007; Manski, 1993; Moffitt, 2001) and for binary choice models (Blume et al., 2010; Brock and Durlauf, 2001*a*,*b*; Durlauf and Ioannides, 2010). Although the Reflection Problem may not exist for nonlinear models (Brock and Durlauf, 2001*b*; Durlauf and Ioannides, 2010), these models may exhibit other forms of nonidentifiability (Blume et al., 2010). This section analyzes in detail the identification of the model proposed in Section III, using the measurement equations for the latent variables, and the moment restrictions given by the production function of achievement and the first order condition of the student problem.

As stated in the model description, the achievement and effort levels $\{A_{ig0}, A_{ig1}, e_{ig}\}$ are assumed to be unobserved and measured with error in the data. There are $L_0 = L_1 = 2$ measurements for lagged and current achievement available. Let T_{igtl} be the lth test score that measures achievement at time t ($t = 0$ for lagged achievement and $t = 1$ for current achievement) for student i

[14]Since 1982 Chile has a schooling voucher system, where any individual or institution can open a private school which is publicly funded according to the number of students it receives. I refer to these schools as "subsidized". In contrast, "public" schools are publicly funded and owned by the Municipality, and "private" schools are privately owned and do not receive any public funds.

in class g. Additionally, there are $L_2 = 5$ measurements for effort, one is a continuous variable and four are discrete. Therefore, let R_{igl} be the lth observed effort measurement for student i in class g, and let R^*_{igl} be the latent variable that generates the lth observed effort measurement; $l = 0$ represents the continuous variable, and $l = 1, \ldots, 4$ the discrete variables.

All test scores and discrete effort measures are assumed to have independent and normally distributed errors. Therefore test scores follow a continuous distribution, and the discrete effort measures follow an ordered probit. Instead, the continuous measure of effort (annual attendance) has a very particular distribution. It has a very long tail to the left with very few data, starts increasing around 83% of attendance, peaks around $95 - 98\%$ of attendance and then decreases sharply until it gets to 100%. This is, even when the variable could be theoretically censored at 1, very few individuals have complete attendance, only 7.7% of the sample, and there is no mass at 1. For this reason, the distribution of ε_{i20} will be approximated using a mixture of normals.

The measurement equations are independent across students and given by

$$
\begin{aligned}
T_{igtl} &= \eta_{tl} + \tau_{tl} A_{igt} + \varepsilon_{itl} \text{ , } t = 0, 1; l = 1, 2 \\
R^*_{igl} &= \tau_{2l} e_{ig} + \varepsilon_{i2l} \text{ , } l = 0, \ldots, 4 \\
\varepsilon_{i20} &\sim MN(\mu_{a1}, \mu_{a2}, \sigma^2_{a1}, \sigma^2_{a2}, w_{a1}, w_{a2}) \\
R_{igl} &= j \textit{ iff } c_{l,j-1} \leq R^*_{igl} \leq c_{l,j} \forall i \in g, l = 1, \ldots, 4
\end{aligned}
$$

where $\{\eta_0, \eta_1\}$ are vectors of constants, and $\{\tau_0, \tau_1, \tau_2\}$ are vectors of factor loadings, with the standard normalization that $\eta_{01} = \eta_{11} = 1$ and $\tau_{01} = \tau_{11} = \tau_{21} = 1$. The mixture of normals is defined by the means of the mixtures, μ_{aj} for $j = 1, 2$, the variances of the mixtures σ^2_{aj}, and the weights w_{aj}, with the standard normalizations that $w_{a1} + w_{a2} = 1$ and $E(\varepsilon_{i20}) = w_{a1}\mu_{a1} + w_{a2}\mu_{a2} = 0$. All other measurement errors have mean 0 and variances given by σ^2_{0l} for lagged test scores, σ^2_{1l} for current test scores, and σ^2_{2l} for the discrete effort measurements, which need to be normalized to $\sigma^2_{2l} = 1$ for $l = 1, \ldots, 4$.

The list of parameters to identify (84 parameters in total) is given by

- Utility function: $c_0, \{\beta_k\}_{k=1}^{K_X}, \{\delta_k\}_{k=1}^{K_X}$
- Production function: $\{\kappa_{Ck}\}_{k=1}^{K_C}$ that generate C_{ig},$\gamma, \phi, \{\kappa_{\omega k}\}_{k=1}^{K_\omega}$ that generate ω_g
- Type distribution: $\{\pi_{sk}\}_{k=1}^{K_Z}, \{\pi_{tk}\}_{k=1}^{K_Y}$ that generate the means of the types, with $\pi_{t1} = 1$, and $\sigma_{sm}^2, \sigma_{sf}^2, \sigma_t^2, \sigma_{st}, \sigma_p$.
- Measurement equations for test scores: η_0, η_1 with $\eta_{01} = \eta_{11} = 0$, τ_0, τ_1 with $\tau_{01} = \tau_{11} = 1$, and $\sigma_{0l}^2, \sigma_{1l}^2$.
- Measurement equations for effort measures: τ_2 with $\tau_{21} = 1$, $\{\mu_{aj}, \sigma_{aj}^2, w_{aj}\}_{j=1}^2$ for attendance, with $w_{a2} = 1 - w_{a1}$ and $\mu_{a2} = w_{a1}\mu_{a1}/w_{a2}$, and the cutoffs $\{c_{l,j}\}$ for the discrete measures of effort.

This parameters are identified, as shown below, given the following conditions in the data:

1. The columns in $W_{\omega g}, W_{Cig}, Y_g$ and Z_{ig} are linearly independent.
2. Two of the discrete effort measures have at least $J \geq 3$ categories.
3. The measurement errors are independent from the latent variables and other errors, and across classmates.
4. The vector Y_g includes a constant, which parameter π_{t1} is normalized to 1.
5. There is enough variation in observables across classmates, such that the columns in $[X_{ig}, \bar{X}_{-ig}]$ are linearly independent in the whole sample.

The conditions on the data are very mild, and thus identification comes mainly from the functional forms of the model.

VI.1 Test scores: Identification of Measurement Equations and Joint Distribution of Lagged Achievement

Using the measurement equations for test scores, it is possible to identify the factor loadings and error distributions. It is possible to use the following moments for test scores at $t = 0$:

$$
\begin{aligned}
E(T_{01}) &= E(A_0) \\
E(T_{02}) &= \eta_{02} + \tau_{02}E(A_0) \\
Cov(T_{01}, T_{02}) &= \tau_{02}V(A_0) \\
V(T_{01}) &= V(A_0) + \sigma_{01}^2 \\
V(T_{02}) &= \tau_{02}^2 V(A_0) + \sigma_{02}^2 \\
E(T_{02}|T_{01} = t) &= \eta_{02} + \tau_{02}E(A_0) + \frac{\tau_{02}V(A_0)}{V(A_0) + \sigma_{01}^2}(t - E(A_0)) \\
&= \eta_{02} + \frac{\tau_{02}V(A_0)}{V(A_0) + \sigma_{01}^2}t + \frac{\tau_{02}\sigma_{01}^2 E(A_0)}{V(A_0) + \sigma_{01}^2}
\end{aligned}
$$

These equations form a system that identifies the mean and variance of lagged achievement, $E(A_0), V(A_0)$, the vector of loadings τ_{02}, the vector of constants η_{02}, and the variance of the measurement errors $\sigma_{01}^2, \sigma_{02}^2$. The same equations for $t = 1$ identify $\eta_1, \tau_1, \sigma_{11}^2, \sigma_{12}^2$.

The measurement equations also provide the identification of the joint distribution of lagged achievement in the class. If the columns in Z_{ig} are linearly independent, then

$$
\begin{aligned}
E(T_{01}|Z) &= E(A_0|Z) = Z\pi_s \\
V(T_0|Z_2 = 0) &= \sigma_{sm}^2 + \sigma_{01}^2 \\
V(T_0|Z_2 = 1) &= \sigma_{sf}^2 + \sigma_{01}^2 \\
Cov(T_{ig01}, T_{jg01}) &= Cov(A_{ig0}, A_{jg0}) = \sigma_p
\end{aligned}
$$

VI.2 Effort measures: Identification of Measurement Equations

It is also possible to identify the measurement equations for effort measures, this is, the factor loadings, the parameters of the mixture of normals and the cutoffs. First, consider the continuous measure of effort, the percentage of annual attendance. The first raw moment, the mean $E(R_0^*)$, provides the identification of $E(e)$ given the normalization that the mixture has mean zero:

$$E(R_0^*) = E(e) + \sum_{j=1}^{2} w_{aj}\mu_{aj} = E(e)$$

Using the discrete measures for effort, it is possible to identify $V(e), \tau_2$ and the cutoffs $\{c_{l,j}\}$ for $j = 1, J, l = 1, 4$. These are $1 + 4 + 4(J-1)$ parameters, and therefore require the same number of equations. Because the measurement errors are normally distributed, the first $4(J-1)$ equations are given by:

$$\begin{aligned} \Pr(R_l \leq j) &= \Phi\left(\frac{c_{l,j} - \tau_{2l}E(e)}{\sqrt{\tau_{2l}^2 V(e) + 1}}\right) \\ &\Rightarrow \\ c_{l,j} - \tau_{2l}E(e) &= \Phi^{-1}\left(\Pr(R_l \leq j)\right)\sqrt{\tau_{2l}^2 V(e) + 1} \end{aligned} \tag{3}$$

This system of equations provides the first $4(J-1)$ equations. For the next 5 equations needed, consider the conditional probability $\Pr(R_l = 1|R_1 = 1)$ for $l > 1$:

$$\begin{aligned}
P_{l,1} &= Pr(R_l = 1|R_1 = 1) = \Phi\left(\frac{c_{l,1} - E(R_l^*|R_1 = 1)}{\sqrt{V(R_l^*|R_1 = 1)}}\right) \\
\text{where} & \\
E(R_l^*|R_1 = 1) &= \tau_{2l}E(e) - \frac{\tau_{2l}\tau_{21}V(e)}{\sqrt{\tau_{21}^2 V(e) + 1}}\lambda(x_1) \\
V(R_l^*|R_1 = 1) &= \frac{\tau_{2l}^2\tau_{21}^2 V(e)^2}{\tau_{21}^2 V(e) + 1}(1 - x_1\lambda(x_1) - \lambda^2(x_1)) + \tau_{2l}^2 V(e) + 1 \\
\lambda(x_1) &= \frac{\phi(x_1)}{1 - \Phi(x_1)} \\
x_1 &= \frac{c_{1,1} - \tau_{21}E(e)}{\sqrt{\tau_{21}^2 V(e) + 1}} \\
\text{and thus} & \\
c_{l,1} - \tau_{2l}E(e) &= \Phi^{-1}(P_{l,1})\sqrt{V(R_l^*|R_1 = 1)} - \frac{\tau_{2l}\tau_{21}V(e)}{\sqrt{\tau_{21}^2 V(e) + 1}}\lambda(x_1)
\end{aligned} \tag{4}$$

If $J \geq 3$ for at least 2 of the discrete effort measures, it is possible to form a system with 5 equations like (4). Therefore, the system of nonlinear equations formed by (3) and (4) has $5 + 4(J - 1)$ equations and unknowns, and thus it can be solved to identify $\{\tau_2, V(e)\}$ and the cutoffs $\{c_{l,j}\}$.

Finally, it is necessary to identify the parameters of the mixture of normals. For this, it is possible to use higher-order raw moments of the attendance variable:[15]

$$\begin{aligned}
E((R_0^*)^2) - E(e^2) &= \textstyle\sum_{j=1}^2 w_{aj}E(\varepsilon_{aj}^2) &&= \sum_{j=1}^2 w_{aj}(\mu_{aj}^2 + \sigma_{aj}^2) \\
E((R_0^*)^3) - E(e^3) &= \textstyle\sum_{j=1}^2 w_{aj}E(\varepsilon_{aj}^3) &&= \sum_{j=1}^2 w_{aj}(\mu_{aj}^3 + 3\mu_{aj}\sigma_{aj}^2) \\
E((R_0^*)^4) - E(e^4) &= \textstyle\sum_{j=1}^2 w_{aj}E(\varepsilon_{aj}^4) &&= \sum_{j=1}^2 w_{aj}(\mu_{aj}^4 + 6\mu_{aj}^2\sigma_{aj}^2 + 3\sigma_{aj}^4) \\
E((R_0^*)^5) - E(e^5) &= \textstyle\sum_{j=1}^2 w_{aj}E(\varepsilon_{aj}^5) &&= \sum_{j=1}^2 w_{aj}(\mu_{aj}^5 + 10\mu_{aj}^3\sigma_{aj}^2 + 15\mu_{aj}\sigma_{aj}^4)
\end{aligned}$$

[15] The flexibility of the mixture of normals implies that the variable may display non-trivial higher-order moments, even in the absence of such features within the components themselves.

The raw moments of R_0^* can be easily obtained from the data, and the raw moments for e come from similar equations using the other measures of effort. Because each element of the mixture ε_{aj} is a normal variable, it is straightforward to obtain the relation of these raw moments with $\{\mu_{aj}, \sigma_{aj}, w_{aj}\}$ shown in the third column above. These form a system of 4 equations, and together with the two normalizations in the weights w_{aj} and the mean μ_{a2}, provide the identification of the parameters of the mixture. This identifies all the measurement equations in the model.

VI.3 Identification of Joint Distribution of Latent Variables

Using the information from the previous section, it is possible to identify the joint distribution of achievement and effort in a class. The nonparametric identification of this joint distribution would require the availability of two continuous measurements for each latent variable (Cunha, Heckman and Schennach, 2010). Given the data, at most one continuous measurement is available for the set of latent variables, and thus the identification of the joint distribution relies on the parametric assumptions for the distribution of measurement errors.

Once the factor loadings are identified, take one continuous measure for each of the latent variables (lagged achievement, current achievement and effort) such that the loadings are different than zero. Then, it is possible to rewrite the measurement equations as

$$\begin{aligned} \frac{T_{igt1}}{\tau_{t1}} &= \frac{\eta_{t1}}{\tau_{t1}} + A_{igt} + \frac{\varepsilon_{it1}}{\tau_{t1}} \\ \frac{R^*_{ig0}}{\tau_{20}} &= e_{ig} + \frac{\varepsilon_{i20}}{\tau_{20}} \end{aligned}$$

For every class g define the vector of latent variables of size $3m_g$

$$\theta_g = (\{A_{ig0}\}_{i\in g}, \{A_{ig1}\}_{i\in g}, \{e_{ig}\}_{i\in g})$$

and the vectors for the measurement, also of size $3m_g$

$$W_g = \left(\{\frac{T_{ig01} - \eta_{01}}{\tau_{01}}\}_{i \in g}, \{\frac{T_{ig11} - \eta_{11}}{\tau_{11}}\}_{i \in g}, \{\frac{R^*_{ig0}}{\tau_{20}}\}_{i \in g}\right) \tag{5}$$

$$\omega_g = \left(\{\frac{\varepsilon_{i01}}{\tau_{01}}\}_{i \in g}, \{\frac{\varepsilon_{i11}}{\tau_{11}}\}_{i \in g}, \{\frac{\varepsilon_{i20}}{\tau_{20}}\}_{i \in g}\right) \tag{6}$$

Then, the measurement equation can be written as $W_g = \theta_g + \omega_g$. If the measurement errors are independent from the latent variables, then the probability distribution $f_{Wg}(W)$ is given by the convolution of the two independent distributions $f_{\theta g}(\theta)$ and $f_{\omega g}(\omega)$:

$$f_{Wg}(W) = (f_{\theta g} * f_{\omega g})(W) = \int_{-\infty}^{\infty} f_{\theta g}(s) f_{\omega g}(W - s) \partial s \tag{7}$$

The convolution f_{Wg} is a weighted average of the function $f_{\theta g}$ where the weighting is given by $f_{\omega g}$. Using the characteristic functions of f_{Wg} from the data and of $f_{\omega g}$ from the parametric assumptions on the distribution of errors, it is possible to to obtain the joint distribution of latent variables in the class, $f_{\theta g} = f(\{A_{ig0}\}_{i \in g}, \{A_{ig1}\}_{i \in g}, \{e_{ig}\}_{i \in g})$. This can always be done, as the required characteristic functions always exist, and the characteristic function of the normal (and thus, of the mixture of normals) is always different than zero.

VI.4 Identification of Production Function Parameters

With knowledge of the joint distribution of $(\{A_{ig0}\}_{i \in g}, \{A_{ig1}\}_{i \in g}, \{e_{ig}\}_{i \in g})$ for the whole class, it is possible to identify the parameters of the production function $\{\kappa_C, \gamma, \phi, \kappa_\omega\}$ and of the unobserved teacher quality distribution $\{\pi_t, \sigma_t^2, \sigma_{st}\}$. These are $K_C + 2 + K_\omega + K_Y + 2$ parameters.

Define $\Omega_{ig} = (A_{ig0}, \{e_{ig}\}_{i \in g}, W_{Cig}, W_{\omega g}, Y_g, Z_{ig})$. The information from the measurement equations at $t = 0$ identified the distribution of $A_{ig0} \sim N(Z_{ig}\pi_s, \sigma_s^2)$. From the joint distribution of latent variables, it is possible to compute the conditional distribution $p(A_{ig1}|\Omega_{ig})$ and the conditional mean function given by:[16]

$$\begin{aligned} E(A_{ig1}|\Omega_{ig}) &= C_{ig}(A_{ig0})^{\gamma}E(\alpha_g|A_{ig0},Y_g)(1+\bar{e}^{\phi}_{-ig})e^{\omega_g}_{ig} \\ &= (W_{Cig}\kappa_C)(A_{ig0})^{\gamma}\left(Y_g\pi_t+\frac{\sigma_{st}}{\sigma^2_s}(A_{ig0}-Z_{ig}\pi_s)\right)(1+\bar{e}^{\phi}_{-ig})e^{(W_{\omega g}\kappa_{\omega})}_{ig} \end{aligned}$$

This function can be evaluated at different values of Ω_{ig} at discretion. First, take pairs of values for $\{A^0_{ig0}, A^1_{ig0}\}$ and $\{\bar{e}^0_{-ig}, \bar{e}^1_{-ig}=1\}$, for Z_{ig} such that $Z_{ig}\pi_s = A_{ig0}$ for both pairs of values, and fixed values Ω^0_{ig} for all other variables. Then, it is possible to identify the parameters of the production function that are fixed across classes:

$$\begin{aligned} \gamma &= \frac{\ln(E(A_{ig1}|\Omega^0_{ig},A^0_{ig0},Z^0_{ig}))-\ln(E(A_{ig1}|\Omega^0_{ig},A^1_{ig0},Z^1_{ig}))}{\ln(A^0_{ig0})-\ln(A^1_{ig0})} \\ \phi &= \frac{1}{\ln(\bar{e}^0_{-ig})}\left(2\frac{E(A_{ig1}|\Omega^0_{ig},\bar{e}^0_{-ig})}{E(A_{ig1}|\Omega^0_{ig},\bar{e}^1_{-ig}=1)}-1\right) \end{aligned}$$

Second, to identify π_t and σ_{st}, note that it is possible to redefine $Y_{ig}=[Y_g, \frac{1}{\sigma^2_s}(A_{ig0}-Z_{ig}\pi_s)]$ and $\tilde{\pi}_t=[\pi_t,\sigma_{st}]$, such that $Y_{ig}\tilde{\pi}_t = Y_g\pi_t+\sigma_{st}/\sigma^2_s(A_{ig0}-Z_{ig}\pi_s)$. Then, take K_Y+2 values of Y_{ig}, value of $e_{ig}=1$, and fixed values Ω^0_{ig} for all other variables, and build a system of equations composed of equations of the form:

$$\begin{aligned} \frac{E(A_{ig1}|\Omega^0_{ig},e_{ig}=1,Y^0_{ig}=y)}{E(A_{ig1}|\Omega^0_{ig},e_{ig}=1,Y^1_{ig}=\tilde{y})} &= \frac{y\tilde{\pi}_t}{\tilde{y}\tilde{\pi}_t} \\ \left(\frac{E(A_{ig1}|\Omega^0_{ig},e_{ig}=1,Y^0_{ig}=y)}{E(A_{ig1}|\Omega^0_{ig},e_{ig}=1,Y^1_{ig}=\tilde{y})}\tilde{y}-y\right)\tilde{\pi}_t &= 1-\frac{E(A_{ig1}|\Omega^0_{ig},e_{ig}=1,Y^0_{ig}=y)}{E(A_{ig1}|\Omega^0_{ig},e_{ig}=1,Y^1_{ig}=\tilde{y})} \end{aligned}$$

It is possible to build in this way a linear system of K_Y+1 equations and unknowns. A normalization is required: the vector Y_g contains a constant with parameter $\pi_{t1}=1$.[17] Then, the system has a unique solution if the vector Y_g is linearly independent. This identifies the values of π_t and σ_{st}.

[16]For a bivariate normal distribution, the conditional distribution is given by $X_2|X_1 \sim N(\mu_2+\frac{\sigma_{12}}{\sigma^2_1}(X_1-\mu_1), \sigma^2_2-\frac{\sigma^2_{12}}{\sigma^2_1})$.

[17]The assumption is required for identification, otherwise in every equation it would be possible to divide both $y\tilde{\pi}_t$ and $\tilde{y}\tilde{\pi}_t$ by any $a\in\Re$ and maintain the equality with the ratio of moments from the data.

Third, to identify the parameters of the production function that depend on class and individual characteristics, take $K_C + 1$ values for W_{Cig}, $e_{ig} = 1$, values for Z_{ig} consistent with W_{Cig} for those variables that both vectors have in common, $A_{ig0} = Z_{ig}\pi_s$, and fixed values Ω^0_{ig} for all other variables. Then it is possible to build a system of K_C equations of the form:

$$\frac{E(A_{ig1}|\Omega^0_{ig}, A^0_{ig0}, e_{ig} = 1, W^0_{Cig}, Z^0_{ig})}{E(A_{ig1}|\Omega^0_{ig}, A^1_{ig0}, e_{ig} = 1, W^1_{Cig}, Z^1_{ig})} = \frac{(W^0_{Cig}\kappa_C)(Z^0_{ig}\pi_s)^\gamma}{(W^1_{Cig}\kappa_C)(Z^1_{ig}\pi_s)^\gamma}$$

Because W_{Cig} includes a constant, this system of equations can be solved with knowledge of π_s and γ, and has a unique solution if the columns in W_{Cig} are linearly independent.

Fourth, to identify the parameters of the marginal productivity of effort, take $K_\omega + 1$ values for $W_{\omega g}$, values for Y_g consistent with $W_{\omega g}$ for those variables that both vectors have in common, and fixed values Ω^0_{ig} for all other variables. It is possible to build a system of equations of the form:

$$\frac{E(A_{ig1}|\Omega^0_{ig}, W^0_{\omega g}, Y^0_g)}{E(A_{ig1}|\Omega^0_{ig}, W^1_{\omega g}, Y^1_g)} = \frac{Y^0_g\pi_t + \frac{\sigma_{st}}{\sigma^2_s}(A^0_{ig0} - Z^0_{ig}\pi_s)}{Y^1_g\pi_t + \frac{\sigma_{st}}{\sigma^2_s}(A^0_{ig0} - Z^0_{ig}\pi_s)}(e^0_{ig})^{(W^0_{\omega g} - W^1_{\omega g})\kappa_\omega}$$

Because $W_{\omega g}$ includes a constant, this system of equations can be solved with knowledge of π_t, and has a unique solution if the columns in $W_{\omega g}$ are linearly independent.

Finally, it is possible to identify the variance of the unobserved teacher productivity σ^2_t using the conditional variance:

$$\begin{aligned} V(A_{ig1}|\Omega_{ig}) &= C^2_{ig}(A_{ig0})^{2\gamma}V(\alpha_g|A_{ig0}, Y_g)(1 + \bar{e}^\phi_{-ig})^2 e^{2\omega_g}_{ig} \\ &= C^2_{ig}(A_{ig0})^{2\gamma}\left(\sigma^2_t - \frac{\sigma^2_{st}}{\sigma^2_s}\right)(1 + \bar{e}^\phi_{-ig})^2 e^{2\omega_g}_{ig} \end{aligned}$$

VI.5 Identification of Utility Parameters

In the last step, the identification of the utility parameters $\{c_0, \{\beta_k\}_{k=1}^K, \{\delta_k\}_{k=1}^K\}$ comes from the effort decision, which is given by the first order condition of the student problem:

$$\begin{aligned} 0 &= \omega_g C_{ig}(A_{ig0})^\gamma \alpha_g (1+\bar{e}^\phi_{-ig}) e_{ig}^{\omega_g - 1} - c_{ig} e_{ig} \\ \Rightarrow & \\ e_{ig} &= \left(\frac{\omega_g C_{ig}(A_{ig0})^\gamma \alpha_g (1+\bar{e}^\phi_{-ig})}{c_0 + X_{ig}\beta + \bar{X}_{-ig}\delta} \right)^{1/(2-\omega_g)} \end{aligned}$$

The utility parameters are identified from the moment conditions:

$$E\left(\left(\frac{\omega_g C_{ig}(A_{ig0})^\gamma \alpha_g (1+\bar{e}^\phi_{-ig})}{c_0 + X_{ig}\beta + \bar{X}_{-ig}\delta} \right)^{1/(2-\omega_g)} - e_{ig} | X_{ig}, \bar{X}_{-ig} \right) = 0$$

given global identification, this is, the above expectation is different from zero for any value of the parameters different from the true. This holds if the columns in $[X_{ig}, \bar{X}_{-ig}]$ are linearly independent. This completes the identification of the model.

VII Estimation and Results

The model is estimated using a Simulated Maximum Likelihood method (Laroque and Salanie, 1989); see Appendix A for details on the specific algorithm used. Table 4 in the Appendix shows the fit of the model with respect to the true data. The proposed model does a good job at fitting the distribution of each of the outcome variables, both for the measures of achievement and effort. The difference between true and simulated moments is negligible.

Table 3 shows the parameter values, the variance and standard deviation of the estimates, and the confidence intervals. All parameters are precisely estimated and significant at the 95% confidence level, except for four parameters,[18] which may be related to the parametric assumptions and

[18]The parameters that are statistically zero are δ_1, the effect of the percentage of females in the class in the cost

the large sample size.[19] In particular, the productivity externality of peer effort ϕ is relatively large: the marginal productivity of peer effort is 0.7025, versus the marginal productivity of own effort ω_g, which is around 0.1128 depending on the class characteristics. The effects of peer characteristics on the cost of effort δ are not always positive: variables like peer parent schooling or if peer parents live at home have a positive effect on own effort by decreasing the cost of effort, but peer inputs at home increase it and percentage of peer females has no significant effect on the cost of effort. It is necessary to note this is not the total effect of these variables in own effort: peer characteristics affect directly the cost of own effort, but they also affect peer effort and thus interact with the endogenous effect. Finally, the SD of lagged achievement in the class has a mixed impact on productivity: it increases the constant (κ_{C4}=0.0029) but it decreases the marginal productivity of effort ($\kappa_{\omega 6}$=-0.0014); the final effect depends on the level of heterogeneity and other characteristics that affect C_{ig} and ω_g.

Because the model is nonlinear, the parameter values are not very informative about the magnitude of the effect of the different variables on student outcomes. For this, it is better to look at the marginal effects of the variables, in Tables 5 and 6. The marginal effect is defined as the percentual change in equilibrium achievement and effort due to a unitary increase in the value of a variable, for every student, keeping all other variables constant. Thus, the marginal effect does not consider consistency between the variables. For example, the marginal effect of being female considers changing the value of "female" from 0 to 1 in the whole sample, but the variable "% of females in the class" is kept at the true values. Table 5 shows this marginal effect for the dummy variables in the model, and for the variables in the $[0, 1]$ interval, and Table 6 for the continuous variables in the model (or more precisely, discrete variables with a support larger than [0,1]).

The model predicts a rich pattern of marginal effects. All of them are significant at the 95%

function, $\kappa_{\omega 4}$, the parameter of whether the school is subsidized in the marginal productivity of effort, π_{t5}, the effect of Science teacher experience on the mean of class productivity, and mu_{a1}, the mean of the first component in the mixture of normals for attendance.

[19] Ideally the significance of the model could be stated using a measure of goodness of fit. Unfortunately, these measures do not control for sample size, and thus it is very easy to reject any model with a sample of the size available here.

confidence level, except for the marginal effect of being in a urban school on achievement, and the marginal effect of class size on effort. The variables that measure the effect of peer characteristics have marginal effects with the same sign for effort and achievement as the parameter value. The percentage of peer females and the average peer inputs at home (books, computer, internet) have negative and small effects on both outcomes (-10 to -20% of a SD of moving from 0 to 100%). As a reference, an effect of more than 10% of a standard deviation is considered large in the education literature. The average parent schooling, peers with parents with college and peers with parents at home have instead positive and moderate effects on both outcomes. Interestingly, peer parent schooling has a convex effect on outcomes; the marginal effect of increasing peer parent schooling in one year increases with the number of years considered. The effect on achievement is small, but the effect on effort is considerably large, and even larger than own parent schooling.

Other variables have the expected effects. Individual variables have positive effects on both outcomes, except for being female (negative effect on achievement but positive effect on effort), having parents with college (positive effect on achievement but negative on effort) and having internet at home (negative effect on both outcomes). The effect of parent schooling is very large and concave: each year of parent schooling increases achievement by around 5% of a SD. School variables have a very large effect on achievement and effort: being in a public school decreases achievement by -84% of a SD and effort by -10.8%, and being in a subsidized school decreases achievement by -58% of a SD and effort by -5.1%. Class size has a positive, concave and moderate effect: each extra student in the class increases achievement around 0.7% of a SD. In constrast, teacher variables have almost negligible but significant effects on achievement and effort, and this suggests they may not be good measures of teacher quality.

VIII Counterfactuals

One of the advantages of having a behavioral model is that it is possible to do out-of-sample predictions and counterfactual analysis. This section analyzes the changes in effort and achievement predicted for the students under different policies that affect their peer group composition or characteristics. This is particularly useful to analyze policies where isolating the causal effect or obtaining an appropriate control group is hard.

These counterfactuals help understand in more detail the strength and pattern of social interactions, as they exogenously change different dimensions of the model, one at a time. This sheds some light to the manner in which peer effects operate in the classroom. The model is used to analyze six different policies. The first counterfactual, providing tutors to the students, tries to isolate the endogenous effect, by exogenously lowering the cost function of the students. Counterfactuals 2 to 4 explore the effect of changes in peer characteristics, by exogenously increasing the number of inputs at home, or analyzing same-gender schools and tracking programs, this is, changing class composition according to gender or lagged achievement. Counterfactual 5 analyzes the effect of class size on the magnitude of peer effects. Finally, Counterfactual 6 shows how the model can be used to analyze the cream-skimming effect of school choice programs, and to separate the effect of student characteristics and class composition from the productivity of the school.

In general, the results are similar to studies that identify the causal effect of policies using quasi-experimental or experimental methods. The model proposed is able to separate the mechanisms that generate those effects.

The results show interesting peer effects. The endogenous effect is positive and significant, very small on effort but of moderate size on achievement. The contextual peer effect may be positive or negative depending on the variable considered, and this may generate counterintuitive effects on achievement. In general, observed contextual variables have small or negative effects; instead, the peer previous achievement has a moderate to large effect. Some policies have larger effects on effort than achievement, but others work the opposite way, especially those related to class composition. However, peer effects on achievement seem to be larger than on effort. Finally, students compensate to some degree the changes in the characteristics of their peer group through their effort. This implies that policies are predicted to have a smaller effect than when the student effort decision is not considered.

VIII.1 A Policy that Decreases the Cost of Effort

Suppose the social planner wants to increase the achievement of students in the lower third of the lagged achievement distribution ("worst" students), by providing each student with a tutor, who is able to lower the cost of making effort for that student, c_{ig}, in 10%. Similar policies are evaluated by Banerjee, Cole, Duflo and Linden (2007); Cabezas, Cuesta and Gallego (2011) and Rothman and Henderson (2011), where tutors are usually assigned to poor students or students lagging behind, but they do not analyze if any peer effects appear in those policies.

This counterfactual stresses the role of the endogenous effect in the determination of effort and achievement. Peer characteristics are kept fixed, but the exogenous shift in the cost of effort has a direct effect on own effort and an indirect effect through the effort decision of the peers.

The tutors are assigned directly to the worst students; the results are shown in column (1) of Table 7. Math test scores increase significantly by 10.4% of a SD and effort increases by 6.55% on average for treated students. This corresponds to 5.3 score points; as a comparison, it is close to giving their parents 2 extra years of education.[20] Additionally, the model predicts interesting spillovers to the peers due to a positive endogenous effect: the other students in the class increase effort by 0.35% and test scores by 3.75% of a SD.

It is possible to explore the strength of the endogenous effect using two modifications of the policy. The first alternative is to pick the classes where these worst students are present, and provide the tutors to everybody in that class. This involves more than doubling the number of tutors. The results are shown in column (2) of Table 7. Mean math test scores increase significantly by 15.3% of a SD on average, and mean effort by 7% on average. For the worst students, this increase corresponds to 7.49 score points, which is larger than giving their parents 2 extra years of education.

From the perspective of the worst students, the variation between the two scenarios is that

[20]With 1 extra year of parental education, the scores of these students increase 2.7 points; with 2 extra years, 6.14 points, and with 3 extra years, 9.8 points.

their peers are also benefited from the policy now. The difference in outcomes then is due to the endogenous effect. Effort is 0.4% larger and achievement 4.3% of a SD larger when their peers also get tutors versus when they do not.

A second alternative is to keep the number of tutors fixed, but instead of allocating the tutor to the worst student directly, assign him to the smartest student in the class, in terms of mean lagged achievement. From the perspective of the worst students, the policy compares receiving no tutor at all, and only the smartest students in the class receiving one. The results are shown in column (3) of Table 7. The benefited students increase their effort and achievement similarly to what the worst students did in column (1). But now the endogenous effect to the worst students is larger: they increase effort significantly by 0.67% and achievement by 6.44% of a SD. Math test scores for the worst students increase 3.27 score points in average, which corresponds to giving their parents 1 extra year of education.

The results from this exercise are smaller to what is found in true data, which means that in practice tutors may decrease the cost of effort by more than a 10%. Banerjee et al. (2007) find an effect for all children in treatment schools of 14%SD in the first year and 28%SD in the second year. Cabezas et al. (2011) find a treatment effect in a reading test of between 15%SD and 20%SD.

In conclusion, this counterfactual shows a positive and significant endogenous effect, which is very small on effort but of moderate size in achievement. This indicates that the endogenous peer effect has an externality effect that is not being internalized by the students.

VIII.2 A Policy that Provides Home Inputs

Suppose the social planner wants to increase now the achievement of students in the lower third of the education distribution, called "underprivileged", and measured by those students whose parents have 11 years of schooling or less. The social planner will provide home inputs to these students: books at home, computer at home or computer with internet at home. This policy stresses the role of contextual effects; the shift on individual inputs has a direct effect on own cost of effort and an indirect effect through peer cost of effort. This will further influence own effort and achievement through the endogenous effect. Therefore, this policy will add one additional channel to the previous policy.

The results show that these policies exhibit negative contextual peer effects, conditional on lagged achievement and other observable variables. This negative indirect effect counteracts the direct effect of the inputs, thus not changing or decreasing (sometimes largely) the achievement of the students intended to benefit. The results are similar to studies that identify the causal effect of policies using quasi-experimental or experimental methods. Here the model is able to separate what are the mechanisms that generate those effects.

VIII.2.1 Providing Books at Home

The first policy is to increase the number of books at home. Underprivileged students who do not have books at home receive 10, and those who have 10 receive 50 or more. The first column of Table 8 shows the results, which are all significant at the 95% level. Math test scores decrease -0.76% of a SD for the underprivileged students, a significant but negligible effect: it corresponds to -0.39 points, much less than taking 1 year of parent schooling which decreases test scores in -1.1 points. This counterintuitive result can be explained by a positive direct effect, a negative contextual effect and a positive endogenous effect. In fact, the underprivileged students decrease their cost of effort by -0.88% and increase their effort by 0.33% (positive direct effect), but their peers increase their cost of effort by 1.1% (negative contextual effect) and decrease effort by -0.6%. The direct effect of own effort on achievement is smaller than the endogenous effect.

As before, it is possible to explore the strength of the contextual effect using two modifications of the policy. First, one may give books to everybody in the class. The results are shown in column (2) of Table 8. Test scores decrease -1.59% of a SD for the underprivileged students, and mean effort -0.65%, a small but significant effect. It corresponds to -0.79 score points, less than taking 1 year of parent scholing from those students.

The difference in outcomes between the two scenarios is due to the contextual effect of peers having more books. The cost of effort is 1.8% larger when their peers also get books versus when they do not. This makes effort -1% lower and achievement -0.8% of a SD lower, similar to the contextual effect peers faced in column (1).

The second alternative is to give the books instead to the smartest student in the class, in terms of mean lagged achievement. The results are shown in column (3) of Table 8. Math test score decreases -0.73% of a SD for the underprivileged students. Their cost of effort increases 1.6% (negative contextual effect), and their effort decreases -0.9% for the underprivileged students. Thus the size of the contextual effect is similar when underprivileged students receive the books or not.

In conclusion, the effects of this policy are significant but negligible. This is consistent with the evidence by Glewwe, Kremer and Moulin (2009), who find no effect of providing textbooks on test scores. On the contrary, Froelich and Michaelowa (2011) find a large effect of textbooks when the number of books is small. But both papers note that the results depend on the production of achievement that is being considered. In the model here, even if the marginal effects of having more books at home are large (as Evans, Kelley, Sikora and Treiman, 2010, also find), the positive effects of own books cancels out with the negative effect of having peers with more books, and thus the policy does not have a relevant effect on achievement, conditional on lagged achievement.

VIII.2.2 Providing Computer at Home

The second policy is to increase the number of computers at home. The literature of the effects of computer access and use on achievement presents mixed results (Bussière and Gluszynski, 2004). Papers that are more successful at isolating the causal effect of computers find negative or zero effects on test scores (Angrist and Lavy, 2002; Fairlie and Robinson, 2012; Malamud and Pop-Eleches, 2011; Vigdor and Ladd, 2010).

Underprivileged students who do not have a computer at home receive one for free; the results are shown in column (1) of Table 9. Math test scores decrease -3.75% of a SD for the underprivileged students, which corresponds to -1.9 score points, larger than taking 1 year of parent schooling from those students. The decomposition shows that underprivileged students increase their cost of effort by 2.4% and decrease their effort by -1.6%. The results show a negative contextual effect: their classmates increase their cost of effort by 1.2% and decrease effort by -0.7% and achievement -0.3% of a SD.

If the computers are assigned to everybody in the class (column (2) of Table 9), Math test scores decrease -5.6% of a SD for the underprivileged students, or -2.8 score points, which is equal to taking more than 2 years of education from their parents. Effort decreases -2.6% for the underprivileged students. The contextual effect of peers having more computers is small but not negligible. The cost of effort is 1.5% larger when their peers also get computers versus when they do not. This makes effort -1% lower and achievement -1.88% of a SD lower, larger to the contextual effect peers faced in column (1).

Again, another alternative is to give the computers instead to the smart students in the class. The results are shown in column (3) of Table 9. Math test scores decrease -3.28% for the smart students, which corresponds to -1.09 points, similar to taking 1 year of schooling from the parents. The contextual effect is slightly larger when the underprivileged students do not receive the computers. Their cost of effort increases 1.8% and their effort decreases -1.14%. The proportion of computers in the class has a slightly larger effect than the proportion of books on both effort and achievement.

VIII.2.3 Providing Computer and Internet at Home

Finally, the third policy is to increase the number of computers with internet at home. Vigdor and Ladd (2010) find that the introduction of home computer and high-speed internet access is associated with modest but statistically significant negative impacts on student math and reading test scores.

The model forces the students with internet at home to also have a computer, so the effect of the policy will be a composition of the two effects. Underprivileged students who do not have a computer at home receive a computer with internet connection for free, and those who do have a computer but do not have internet connection receive one. Now, the negative effects are significant and very large.

Column (1) of Table 10 shows that Math test scores decrease -7.96% of a SD for the underprivileged students, which corresponds to -4 score points, and is rather large; it is equivalent to taking 3 years of education from the parents. Their cost of effort increases 6.3% and their effort decreases -3.88%. The contextual effect on peers is the largest of all the input policies: effort decreases -1.4% and achievement decreases -3.48% of a SD for their classmates.

When the computers with internet are ssigned to everybody in the class (column (2) of Table 10), Math test scores decrease -12.41% of a SD for the underprivileged students, or -6.3 score points, and equivalent to taking more than 3 years of education from the parents. Their cost of effort increases 9.66% and effort decreases -5.79%. This represents a substantial negative contextual effect. The cost of effort is 3.1% larger when their peers also get computers with internet versus when they do not. This makes effort -2% lower and achievement -4.45% of a SD lower, larger than the contextual effect peers had in column (1).

Again, another alternative is to give the computers instead to the smart students in the class. In this case, (column (3) of Table 10), Math test scores decrease -4.84% of a SD for the underprivileged, which corresponds to -2.46 points, similar to taking 2 year of education from their parents.The contextual effect is similar when the underprivileged students do not receive the computers. Their cost of effort increases 3.4% for the underprivileged, and effort decreases -2.2%, again larger than the contextual effect peers had in column (1). This is consistent with Vigdor and Ladd (2010), who find larger negative effects for students with lower socioeconomic background.

VIII.3 Same-Gender Schools

This counterfactual analyzes the effect of the gender composition of the class. This peer variable has received a lot of attention in the literature, and some authors argue it is the only contextual variable to have significant effects on achievement conditional on previous peer achievement (Hoxby, 2000*b*; Hoxby and Weingarth, 2006; Lavy and Schlosser, 2011).

There is a relatively large discussion on the effect of same-gender schools versus coeducational schools on outcomes (Park, Behrman and Choi, 2012; Sullivan, 2009). Even controlling adequately for selection of students into each type of school, some studies find single-sex schooling is associated with higher average scores for both genders (Park et al., 2012), and others find no significant effect (Jackson, 2012). Graham, Imbens and Ridder (2010) find that making classes slightly more segregated by sex would have a positive effect on girls and have a negative effect on boys of approximately the same size.

To abstract from variation in school and teacher characteristics, it is possible to evaluate the policy using simulated data: 120,000 students in 6,000 classes of size $m_g = 20$ in 3,000 schools. Half of the students in each school belong to each gender. Individual characteristics are simulated according to the mean and standard deviation of characteristics in the true data, and school and teacher characteristics are set to the mean values of the true data. Students are first randomly asigned to each class, and then separated according to their gender, and compared.

Table 11 shows the results. The model predicts that moving to a same-gender class significantly increases Math test score 3.25% of a SD for females and decreases -2.13% of a SD for males. This corresponds to 1.65 and -1.08 score points respectively. On average, mean test scores increase significantly 0.55% of a SD. This is a modest effect: it is equivalent to half of the effect of keeping the classes mixed but giving 1 year of parent schooling to females and taking 1 year of schooling from males. The results are almost the same with smaller class sizes (size 10).

The results of this counterfactual show the mechanisms that affect effort and achievement. Table 3 reports that having more females in the class significantly productivity (the variable also increases the cost of effort but the parameter is not significant at the 95% level), so having more peer females should decrease achievement for females. In fact, the cost of effort decreases -0.139% for males and increases 0.157% for females, and both are significant. Nonetheless, mean effort increases nonsignificantly for males and 1.051% for females.

Having more peer females increases effort for females in the same-gender class because the new classmates make more effort than boys, and thus increase achievement through the endogenous effect. The results are opposite to those of Lavy and Schlosser (2011), who find that a higher proportion of females increases test scores, but not because of a change in individual behavior, but due to compositional effect on the level of disruption and violence, relationships and teacher's fatigue.

VIII.4 Tracking Inside the School

This section analyzes the most important of the contextual effects in the literature: the previous achievement or ability of the classmates, by focusing on the tracking of students inside a school.

As other counterfactuals, the estimation of the effect of tracking on student outcomes is complicated by the difficulty of setting an appropriate control group. The usual finding that tracking has a negative effect on low ability classes, and a positive effect on high-ability classes may be due to a selectivity bias problem (Betts and Shkolnik, 2000*a*,*b*; Figlio and Page, 2002). Other studies find insignificant effects(Betts and Shkolnik, 2000*a*; Figlio and Page, 2002). Instead, the scarce experimental evidence indicates that it may benefit all students. Duflo et al. (2011) find that all students benefited from tracking, and those near the median of the baseline test distribution benefited similarly whether assigned to the lower or upper section. In contrast, Garlick (2012) finds for college roommates that tracking reduces grades, driven by very large negative effects on the lower tail of the distribution and insignificant effects on the upper tail.

There is a related literature on input allocations (Bhattacharya, 2009; Graham, Imbens and Ridder, 2009; Graham et al., 2010), where a large number of inputs have to be allocated simultaneously. They nonparametrically estimate the expected outcomes for different allocation rules, using discrete types (Bhattacharya, 2009; Graham et al., 2010), or continuous types (Graham et al., 2009). Two dimensions of the problem are different to the one studied here. First, they assume that the average response to group composition remains unchanged across reallocations, opposite to the findings by Carrell et al. (2011). Second, the inputs allocated differ only in one dimension, where the students in the model differ in multiple dimensions that affect achievement of the group.

Because tracking may generate optimal changes in the behavior of teachers, schools and parents, this counterfactual should be considered only a partial evaluation of a true tracking policy, dealing only with changes in effort made by the students.

As in the previous counterfactual, 120,000 students in 6,000 classes of size $m_g = 20$ in 3,000 schools are simulated. Individual characteristics are simulated according to the mean and standard deviation of characteristics in the true data, and school and teacher characteristics are set to the mean values. Students are first randomly assigned to each class, and then tracked according to their mean lagged achievement, and compared.

The results are reported in Table 12 in the Appendix. Mean math test scores increase significantly 8.53% of a SD for the top students, and decreases non-significantly -3.7% of a SD for the worst students. These correspond to 4.34 and -1.88 score points respectively. These are moderate to large changes; they are equivalent to keeping the class composition constant, but substracting less than 1 years of parental schooling to the students with lower lagged achievement, and giving 2 years of parental schooling to the students with higher lagged achievement. The mean math score for the whole school increases significantly 2.41% of a SD, or 1.22 points.

This change in math test score can be decomposed in the following way. The cost of effort increases 4.48% for the worst students and decreases -4.22% for the top students, due to the contextual effect. This generates a non-significant decrease in mean effort for the worst students (-0.646%) and a significant increase in mean effort for the top students (4.93%). It is interesting to note the differences of the effort response. For the top students, the effect of the decrease in the cost of effort is amplified by the endogenous effect of having peers that make more effort. The worst students compensate for the negative contextual and endogenous effect, and effort does not change statistically. The results are similar when the class size is fixed to 10, and when tracking is done with the true data, with all students in schools with two or more classes.

This analysis is useful even as a partial evaluation. There is a moderate and positive peer effect for students in the higher track, and no significant peer effects for students in the lower track. In the same direction, the results in Zimmer (2003) suggest that, controlling for the peer level, the school and teacher response after tracking has no effect on high-ability students, and has a positive effect on low and average ability students. These results taken together can rationalize the effects in Duflo et al. (2011).

With this counterfactual, it is also possible to simulate the effects found in the experiment by Carrell et al. (2011). Instead of tracking their students, they put the middle students together in a class, and the top and worst students together in a second class. They find that middle students increase their score in 0.067, and worst and top students decrease their score in -0.054 and -0.004 grade points respectively. It is possible to perform the same analysis with the simulated data; see Table 13 for the results. The direction of the result is the same as in Carrell et al. (2011): the middle students improve their score, and worst and top students decrease. The magnitude of the effects is very small, and not significant. This provides an external validation for the model.

VIII.5 Changes in Class Size

The model also has interesting predictions regarding changes in class size. A larger class size decreases the marginal productivity of effort ω_g,which corresponds to the congestion effect outlined by Lazear (2001). Also, a larger class size increases teacher and school productivity α_g. This occurs because the number of teacher hired per student is one of the main determinants of the cost of education (Krueger, 2003). Additionally, in Chile school funding is closely related to its size,[21] thus a larger class size may further increase the productivity of the school.

The final result will depend on which effect is larger. Konstantopoulos (2008) and Krueger and Whitmore (2001) find negative effects of class size on achievement, consistent with the first effect

[21] In the public and subsidized sector, the value of the voucher is per student per day attended, and in the private sector, a tuition is charged by the school to every student.

being larger than the second one. Hoxby (2000*a*) and Duflo, Dupas and Kremer (2009) find no significant effect of class size, consistent with both effects being of equal size.

This section analyses the effect on effort and achievement of changes in class size. Again, a sample of 120,000 students in 3,000 schools is simulated, using the mean and standard deviation of student characteristics in the original sample. All teacher and school characteristics are the same and equal to the sample mean. In each school, the 40 students are randomly assigned to classes of equal size, of 5, 10, 20 or 40 students.

The analysis is performed from a "short-run" and a "long-run" perspective. In the "short-run" perspective, teacher productivity is not allowed to vary with class size, and thus class size only affects the productivity of effort. In the "long-run" perspective, class size also affects the productivity of the teacher. This separates between the two effects outlined above.

The results for the "short-run" perspective are shown in Table 14. Decreases in class size have a positive and significant effect on both mean achievement and effort. Take the 5 worst students of each school in terms of lagged achievement. Their math test score increases 4.1% of a SD when decreasing class size from 40 to 20 students, 6% of a SD when decreasing class size to 10, and 6.7% of a SD when decreasing class size to 5. The results are larger for the 5 top students of each school: 5, 7.7 and 10.1% of a SD respectively. These results are consistent with by Konstantopoulos (2008) who finds that class size reductions have a positive effect in all students but increases the achievement gap. Effort also increases as class size decreases, and the size of the effect is almost the same for top and worst students.

The results for the "long-run" perspective are shown in Table 15. Decreases in class size have a negative and significant effect on mean achievement and a very small, significant and positive effect on mean effort. Take the 5 worst students of each school in terms of lagged achievement. Their math test score decreases -12.5% of a SD when decreasing class size from 40 to 20 students, -18.9% of a SD when decreasing class size to 10, and -22.408% of a SD when increasing class size to 5. For the 5 top students of each school, the results are again larger.

All these analyses assumed that the different classes in the school have the same number of students. Nonetheless, Lazear (2001) suggests in his model that bad students should go in small classes and good students in large classes to maximize school output. This rule can be tested; the results depend on the horizon considered.

Table 16 shows the predicted test score when there are only two classes in a school of 40, and the worst students of the school in terms of lagged achievement are tracked together in a class of 5 up to 35 students. Panel (a) shows the "short-run" effect, when teacher productivity is fixed across classes of different size. School output has a significant inverted-U shape: average test score is maximized when classes are of the same size. The effect is a result of the interaction of changes in the productivity of effort and a tracking component. If we look at the 5 top and 5 worst students in each school, the effect of class size is monotone. For every group, Math test score is larger when they have less peers - each group is in a smaller class - but moving from one class size to the next not always has significant results.

Panel (b) of Table 16 shows the "long-run" effect, when teacher productivity varies with class size. The result is completely opposite, and in line to the idea suggested by Lazear (2001). School output has a significant U shape: average test score is maximized when classes are of very different size. If we look at the 5 top and 5 worst students in each school, the effect of class size is again monotone, but now for each group Math test score is larger when they have more peers.

In conclusion, the model predicts strong effects of class size, and the effects are positive of negative depending if changes in teacher productivity are allowed or not. The estimated effects of class size on teacher productivity are so strong, that counteract the negative effect of class size on the productivity of effort, and also the effect of tracking students.

VIII.6 Magnet Schools

The proposed model can also be applied to the analysis of Magnet, Selective or Exam schools. Several papers argue that their positive effects observed in the data are due to composition effects:

the selection of good students and the interaction of peer effects explains most of the increase in test scores (Valenzuela and Allende, 2012). The main problem of these studies is the determination of an appropriate control group. Pischke and Manning (2006) show that controlling for baseline test scores or even using instrumental variables is not enough to eliminate selection effects. Most quasi-experimental and experimental studies find small effects of exam school offers (Abdulkadiroglu, Angrist and Pathak, 2011; Clark, 2007; Cullen and Jacob, 2007). A problem with this method is that students that do not get into a specific exam school tend to get into other good school, thus the method is only comparing between specific selective schools. Ballou, Goldring and Liu (2006) use acceptance in a lottery as an instrument to predict school attendance, and find a positive magnet school effect.

To illustrate the use of the model, this section focuses on a program implemented in Chile during 2011, the "Bicentenario" Schools (MINEDUC, 2011). The program selected 60 publicly funded schools in the country, and gave them resources to help improve their results.[22] The schools commit to be in the top 10% of the SIMCE distribution[23] and in the top 5% of the PSU distribution[24] for each cohort.

There are three types of schools in the policy: new schools (18, or 30% of the total), restructured schools (40, or 67%, must make changes and implement a new educational project), and expanded schools (2, or 3%, have good results but must increase number of students). All schools must provide positions for at least 80 new students.

The model proposed can be used to analyze the policy from two perspectives. The first part focuses on the predicted effect of cream-skimming the best students in a district, without improving the efficiency of the school. This effect is cited in most studies that analyze Selective Schools, but there has been no attempt to estimate it, besides the work of Altonji, Huang and Taber (2010). The

[22] The program gives financial resources to upgrade educational projects, infrastructure and equipment, provides pedagogical and methodological support, and monitors the learning of the students with frequent evaluations.

[23]The SIMCE (Education Quality Measurement System) is a national test taken by all schools either in 4th grade, 8th grade or 10th grade, and measures the degree of achievement related to the core curriculum.

[24]The PSU (University Selection Test) is the national college entrance test and, together with the school grades, is the sole determinant of college acceptance in most universities in the country.

second analysis deals with the implementation of the policy, and shows how the proposed model can be used to separate the effect of observable class composition, including selection and peer effects, to those of school productivity. Given the implementation date, a complete evaluation of the policy is premature, and will be left for future work.

VIII.6.1 Effect of Skimming in a District

It is usually assumed that Magnet Schools attract some or all of the best students in an area, generating a positive effect for these students, because of the positive peer interactions that they experiment, but also a negative effect for the students left behind in the schools of origin, because they are deprived of the influence of their strongest peers.

This effect is largely cited, but there has been virtually no attempt to quantify it. Even if higher achieving schools cream-skim advantaged students (Bifulco, Ladd and Ross, 2009), this is not enough to compute the effect on those left behind. Altonji et al. (2010) use a model with contextual effects to generate a formula that calculates the cream-skimming effect. This effect is increasing in the degree of heterogeneity within schools, the school choice takeup rate of strong students relative to weak students, and the importance of peers. The effect is small of any of the channels is small, and thus they find a negative but small cream-skimming effect of a voucher program.

This section uses a counterfactual policy to ilustrate the size of the effects on the students that move and those left behind when the behavior of the student is accounted for. For a given district or geographic area, a school will be selected at random from the top of the distribution of predicted math test score in that district, and named a “Magnet School”. [25] Then, new students will be selected at random from the top of the distribution of lagged achievement in the district, conditional on observed test scores, and moved to this Magnet School. Math test scores will be predicted before and after the policy using observable variables for three types of students: new

[25]The true selection mechanism was based on the quality of the educational project proposed, the quality of the school administration, the strategy to attract vulnerable students, geographical location, the availability of some technical or financial support network in order to improve their processes, and the funding solicited by the school.

students that move to the Magnet School, old students that were in the Magnet School before the policy and remain there, and students left behind in the schools of origin.

This analysis only focuses on the peer effect of cream-skimming students, and there is no change in the productivity or the resources of the school (which is not the case with the true policy).

The effect of skimming will depend largely on the size of the district. Two districts are selected: Lebu, a small district with 6 schools and data for 62 students in 8th grade 2009, and Puente Alto, a very large district with 104 schools and data for 4,063 students in 8th grade 2009. The Magnet School and the elegible students in each district are selected at random from the top 30% of the distribution in the small district, and from the top 10% of the distribution in the large district (over 248.72 and 279.25 SIMCE points respectively). The selected Magnet Schools have data on 10 students and 66 students each; assume they will open the same number of slots for new students. Two situations are considered: restructured schools, where new students mix with the old ones, and new schools, where they do not mix and the reference group is composed only of new students.

There are differences regarding the peer environment that the new students originally had. In the small district, only 2 of the 10 students came from a top school,[26] and so 80% of students are moving from an environment with worse achieving peers to better achieving peers. In the large district, 21 of the 66 students came from a top school, and so only 68% of the students are moving to an environment with better peers. Also, 38 come from top 20% schools.

Also, given dispersion in school quality, the origin of these new students is not random. There is some clustering in the school of origin, which may create negative peer effects for the old classmates. But the clustering changes with the size of the district. In the small district, half of the new students come from a single school (5 out of a school of 17, in other words 29% of the 8th grade students of that school move out). In the large district, the larger concentration are 7 new students out of 154.

[26] This is, a school that also was in the top 30% of the distribution of the district

One should then expect both the positive and negative effects of the skimming policy to be different in districts of different size. This is consistent with the factors outlined by Altonji et al. (2010) that affect the formula for the cream-skimming effect.

Table 17 shows the results. They are not significant, due to the small sample size, but the point estimates are still interesting. As expected, they depend dramatically on the size of the district. In the small district, the new students in the Magnet School have moderate sized effects, and there is a very small positive effect for the students left behind in the old school. When the Magnet School is New, this is, the new students do not mix with the old ones, the new students gain 12.84% of a SD; when the school is Restructured, the new students gain 15.36% of a SD, and the old students lose -5.3% of a SD. The students left behind have a positive effect of 1.62% of a SD. This counterintuitive effect occurs because there is a substantial decrease in the cost of effort for the students left behind, and this contextual effect increases effort for those students.

These results are equivalent to keeping the class composition constant, and give between 1 and 2 years of parental schooling to the new students, take between 0 and 1 year from the old students in the magnet schools, and give 0 years to the students left behind[27].

In the large district, results are completely different. The new students have a negligible increase in math scores, and the students left behind have a small and negative effect. When the Magnet School is New, the new students gain 1.31% of a SD; when the school is Restructured, the new students gain 0.047% of a SD, and the old students in the magnet school gain 0.959% of a SD. The students left behind have a negative effect of -1.67% of a SD. The total effects therefore are negligible.

This analysis of a skimming policy suggests that it may be better to assign Magnet Schools to small rather than large districts. This is not the design that has been followed in the case of Chile: of the 60 Magnet Schools, only 18 of them are located in districts with less than 1,000 students in 8th grade, and 10 are located in districts with more than 4,000 students.

[27] They get a small positive effect from the increased peer parental schooling.

VIII.6.2 Effect of the Implementation of the Magnet School Policy

The evaluation of selective schools policies is difficulted by the need of an appropriate control group. This can be overcame using the model to separate the effects of class composition from those of school productivity. This is illustrated using the Magnet School policy implemented in Chile.[28] Of the 60 magnet schools, 14 have students in the 8th grade 2009 SIMCE sample, and 20 in the 2011 SIMCE sample.[29]

The effect of the class composition (including selection based on observables and peer effects) and school productivity can be separated in the following way. For each of these schools, it is possible to compute the observed test score ("In Sample Score"), and predict using the model the test score that the school should obtain given the class composition, the peer effects, and other observable characteristics ("Predicted Score"). Also, it is possible to compute the school productivity parameter α_g given the observable school and teacher characteristics ("In Sample Productivity"). Finally, for every school it is possible to estimate what should be the value of α_g so that the model, given the observed student composition, predicts the same average test score observed in the sample ("Predicted Productivity").

The Predicted Score considers the whole effect that class composition, possible selection and peer effects have on the equilibrium effort and achievement of the students, conditional on observable variables. Therefore, if a school has Predicted Score similar to the In Sample Score, it means that the school is as productive as the non-magnet schools, given its observable characteristics, and that all the change in test score after the implementation of the policy can be explained due to class composition. On the other hand, if a school has a Predicted Score lower than the In Sample Score, it means that the school is doing better than what is predicted using the class composition and the characteristics of the students, and thus the school is more productive than the non-magnet schools.

[28] A complete evaluation of the policy is premature, and will be left for future work. The model proposed and estimation methodology require some adjustments for a full evaluation. For example, the model only specifies the mean of the distribution of lagged achievement as a function of observable variables. It may be posible to condition the distribution of lagged achievement on observed lagged test scores to predict A_{ig0}, to gain more predictive power at the top of the distribution, where these students may belong.

[29] Of the 60 schools, 21 of them only have students in high school, and 15 of them are new.

The results can be found on Table 18. In 2009, previous to the implementation of the policy, 8 of the 14 schools in the data are more productive that what is apparent from their observable characteristics (57% of the schools selected later), 4 of them slightly more productive and 4 of them substantially more productive.[30] In 2011, after the implementation of the policy, 13 of the 20 schools are more productive than apparent from their observable characteristics. This is, a 65% of the observable magnet schools have test scores that are larger that what can be predicted given their class composition and observable characteristics. Of these 20 schools, 11 are old (with data observed for 2009) and 9 are new (no data for 2009). Of the old schools, 7 (63%) saw an increase in their Predicted Productivity from 2009 to 2011. Between them, two schools stand out due to their large productivity increase: the Purén school increased its Predicted Productivity from 0.9548 to 1.1776 (24.4% increase) and the Aisén school increased its Predicted Productivity from 1.0497 to 1.1094 (5.6% increase). Of the new schools, 7 (77.9%) are largely more productive that what is apparent. For example, the Villarrica school obtained a math score of 332.2 in the sample, when its class composition predicts a score of 228.9 (below the national average), this is the score is 45% larger than predicted. This implies that the Predicted Productivity is 1.277 instead of 1.03, or 24% larger. Similar cases can be seen in the Puente Alto school and San Antonio school, with test scores 32.2% and 29.3% larger that what is predicted given their class composition, and Predicted Productivity 19% and 20% larger respectively.

The previous analysis assumed, as in the model, that the students in the magnet schools have the same distribution of lagged achievement than the population. In other words, they are drawn at random from the population, conditional on observed variables. This assumption may be subject to special critique in the case of Magnet Schools; it is plausible to think that the skimming of students implies that these students are better than average, or that their distribution of lagged achievement conditional on observables is not the same as the rest of the students. It is possible to repeat the previous analysis, but assuming that for these students lagged achievement conditional

[30]For example, the school that will become Magnet School of Puerto Montt has Predicted Productivity of 1.118 instead of In Sample Productivity of 1.042, or a 7.2% more productive

on observables is drawn from the top 10% of the distribution.[31] This also is an extreme assumption, particularly in the case of the Restructured schools that have kept their previous students, but it is useful to set a lower bound on the productivity of the Magnet Schools. Most naturally, the true situation is in between the two analyses.

The results are shown in Table 19. Now for 2009, previous to the implementation of the policy, in this extreme case only 1 of the 14 schools in the data is more productive that what is apparent from their observable characteristics (7%). After the implementation of the policy, in 2011, only 1 of the 20 schools is more productive than apparent (5%). This is, this extremely successful school has test scores that are larger that what can be predicted even if their students are drawn from the top 10% of the distribution of lagged test score given the class composition and observable characteristics. This is a new school, the Villarrica school, who obtained a math score of 332.2 in the sample, and if their students belonged to the top 10% conditional on observables they would get 316.6 points. The San Antonio school, also new, obtained a math score of 327.4 in the sample, and if their students belonged to the top 10% conditional on observables they would get 335.9 points. Even if this school is not more productive than apparent in this extreme case, it gets fairly close. For the 11 old schools in the data, even if none of the old schools is more productive than apparent in this extreme case, 5 (45.5%) saw an increase in their Predicted Productivity from 2009 to 2011 even if their students are highly selective. Between them, the Purén school stands out; the model predicts an increase in its productivity of 17% even if its students are highly selective in the two years.

To the best of my knowledge, there are no other educational policies in the period that may be affecting the results. To make sure that the results are not spurious due to a trend in the educational system, it is possible to perform two simple analysis. First, I run two regressions, using average Math test scores for the schools in 2009 and 2011 against a dummy that takes value 1 if the school is a Magnet School. The coefficient for the 2009 regression is 4.09 but non significant, indicating that

[31] To operationalize this, the mean lagged achievement is calculated from observables as in the model, but the shock to lagged achievement in each simulation is sampled from a normal truncated from below, so that only values from the top 10% of the distribution are sampled.

the Magnet Schools are not selected at random from the population[32] The coefficient for the 2011 regression is 19.87 and statistically significant; this is, the relationship between the schools that would become Magnet Schools and their test scores changed substantially after the implementation of the policy. The second analysis performs an analysis of variance analysis (ANOVA) for the difference between 2011 and 2009 scores for the two groups (Magnet and non-Magnet Schools). This test rejects the hypothesis of equality of means between the groups. In case there was a trend in the system, the increase in test scores is larger for these selected schools.

These results indicate that the implementation of the policy may have been sucessfull until now. The magnet schools show very good results in terms of productivity between 2009 and 2011, given the observable characteristics of their students. A more comprehensive evaluation will be left for future work.

IX Conclusion

The growing literature on achievement peer effects has provided strong evidence that classmates have an important effect on the determination of achievement. These studies also have recognized the existence of various different mechanism that can generate these effects, but in practice little attention has been given to the identification and estimation of said mechanisms.

The explicit consideration of the underlying mechanisms of peer effects is important because they may have potentially different policy implications, and because the endogeneity of unobserved effort may bias the estimation of peer effects. For this reason, this paper proposes and estimates a behavioral model that explicitly considers the student effort decision and achievement production in a class. The model proposed considers explicitly the role of own and peer effort, and identifies three mechanisms for peer effects: a production externality of peer effort, the effect of lagged achievement heterogeneity on production, and the effect of peer characteristics in the cost of effort. In this sense, this paper represents a contribution to bridge the gap between the theoretical and the

[32] In 2009 the policy has not been announced yet.

empirical literatures.

The estimation of the model using data for Chile results in precisely estimated parameters that provide evidence of significant peer effects. In particular, the effect of peer effort on production is positive and relatively large. The peer parameters on the cost of effort are not always positive: some variables have a positive effect on own effort by decreasing the cost of effort, but others increase it or have no significant effect on the cost of effort. Finally, class heterogeneity has a mixed impact on productivity: it increases some parameters of the production function but decreases others.

The counterfactual analysis performed on six different policies help understand in more detail the strength and patterns of social interactions, as they exogenously change different dimensions of the model, one at a time. the results are similar to studies that identify the causal effect of policies using quasi-experimental or experimental methods, and show interesting peer effects. The endogenous effect is positive and significant, very small on effort but of moderate size on achievement. The contextual peer effect may be positive or negative depending on the variable considered, and this may generate counterintuitive effects on achievement. In general, observed contextual variables have small or negative effects; instead, the peer previous achievement has a moderate to large effect. Some policies have larger effects on effort than achievement, but others work the opposite way, especially those related to class composition. However, peer effects on achievement seem to be larger than on effort. Finally, students compensate to some degree the changes in the characteristics of their peer group through their effort. This implies that policies are predicted to have a smaller effect than when the student effort decision is not considered.

Some aspects of the model necessary for counterfactual analysis require more attention. First, it would be desirable to include the effort decision of teachers or the input selection decision of parents in the model. Both agents may respond to changes in peer composition and this may change the predictions of the model. Second, it is necessary to evaluate the possibility of making the production function of achievement more flexible, but keeping identification. This would allow to evaluate which predictions of the model are robust to its specification, and which depend on the complementarity of inputs implicit in the production function. Finally, it is possible to extend the model to evaluate the policy of Bicentenario or magnet schools. All of these aspects are beyond the extend of this paper, and are left for future work.

References

Abdulkadiroglu, Atila, Joshua Angrist and Parag Pathak (2011), The Elite Illusion: Achievement effects at Boston and New York exam schools, NBER working paper no. 17264, National Bureau of Economic Reseach.

Akerlof, George and Rachel Kranton (2002), 'Identity and schooling: Some lessons for the economics of education', *Journal of Economic Literature* **40**, 1167–1201.

Alcalde, Pilar (2012), On the specification and estimation of the Linear-in-Means achievement peer effects model, Working paper, University of Pennsylvania.

Altonji, Joseph, Ching-I Huang and Christopher Taber (2010), Estimating the Cream Skimming Effect of School Choice, NBER working paper no. 16579, National Bureau of Economic Reseach.

Angrist, Joshua and Victor Lavy (2002), 'New evidence on classroom computers and pupil learning', *The Economic Journal* **112**, 735–765.

Austen-Smith, David and Ronald Fryer (2005), 'An economic analysis of 'acting white'', *Journal of Economic Literature* **120**, 551–583.

Ballou, Dale, Elen Goldring and Keke Liu (2006), Magnet schools and student achievement, Technical report, National Center for the Study of Privatization in Education, Teachers College, Columbia University.

Banerjee, Abhijit, Shawn Cole, Esther Duflo and Leigh Linden (2007), 'Remedying education: Evidence from two randomized experiments in India', *The Quarterly Journal of Economics* **122**, 1235–1264.

Betts, Julian and Jamie Shkolnik (2000*a*), 'The effects of ability grouping on student achievement and resource allocation in secondary schools', *Economics of Education Review* **19**, 1–15.

Betts, Julian and Jamie Shkolnik (2000*b*), 'Key difficulties in identifying the effects of ability grouping on student achievement', *Economics of Education Review* **19**, 21–26.

Bhattacharya, Debopam (2009), 'Inferring optimal peer assignment from experimental data', *Journal of the American Statistical Association* **104**, 486–500.

Bifulco, Robert, Helen Ladd and Stephen Ross (2009), 'The effects of Public School Choice on those left behind: Evidence from Durham, NC', *Peabody Journal of Education* **84**, 130–149.

Bishop, John (1994), 'Is the test score decline responsible for the productivity growth decline?', *American Economic Review* **79**, 178–197.

Bisin, Alberto, Andrea Moro and Giorgio Topa (2011), The empirical content of models with multiple equilibria in economies with social interactions, NBER working paper no. 17196, National Bureau of Economic Reseach.

Blume, Lawrence, William Brock, Steven Durlauf and Yannis Ioannides (2010), Identification of social interactions, *in* J.Benhabib, A.Bisin and M.Jackson, eds, 'Handbook of Social Economics', Vol. 1, Elsevier, chapter 18.

Brock, William and Steven Durlauf (2001*a*), 'Discrete choice with social interactions', *Review of Economic Studies* **68**, 235–260.

Brock, William and Steven Durlauf (2001*b*), Interactions-based models, *in* J.Heckman and E.Leamer, eds, 'Handbook of Econometrics', Vol. 5, Elsevier, chapter 54, pp. 3297–3380.

Burke, Mary and Tim Sass (2008), Classroom peer effects and student achievement, Working Paper 18, National Center for Analysis of Longitudinal Data in Education Research.

Bussière, Patrick and Tomasz Gluszynski (2004), The impact of computer use on reading achievement of 15-year-olds, Technical report, Learning Policy Directorate, Human Resources and Skills Development Canada.

Cabezas, Veronica, Jose Ignacio Cuesta and Francisco Gallego (2011), Effects of short-term tutoring on cognitive and non-cognitive skills: Evidence from a randomized evaluation in Chile, Working paper, P. Universidad Catolica de Chile.

Carrell, Scott, Bruce Sacerdote and James West (2011), From natural variation to optimal policy? The Lucas Critique meets peer effects, Working paper.

Carrell, Scott, Richard Fullerton and James West (2009), 'Does your cohort matter? Measuring peer effects in college achievement', *Journal of Labor Economics* **27**, 439–464.

Cicala, Steve, Roland Fryer and Joerg Spenkuch (2011), A roy model of social interactions, NBER working paper no. 16880, National Bureau of Economic Reseach.

Clark, Damon (2007), Selective schools and academic achievement, IZA discussion paper no. 3182, Institute for the Study of Labor.

Cooley, Jane (2008), Desegregation and the Achievement Gap: Do diverse peers help?, Working paper, University of Wisconsin, Madison.

Cooley, Jane (2009), Can achievement peer effect estimates inform policy? A view from inside the black box, Working paper, University of Wisconsin, Madison.

Creemers, Bert (2007), Educational effectiveness and improvement: the development of the field in mainland Europe, *in* T.Townsend, ed., 'International Handbook of School Effectiveness and Improvement', Vol. 17, Springer, pp. 223–241.

Cullen, Julie and Brian Jacob (2007), Is gaining access to selective elementary schools gaining ground? Evidence from randomized lotteries, NBER working paper no. 13443, National Bureau of Economic Reseach.

Cunha, Flavio, James Heckman and Susanne Schennach (2010), 'Estimating the technology of cognitive and noncognitive skill formation', *Econometrica* **78**, 883–931.

Ding, Weili and Steven Lehrer (2007), 'Do peers affect student achievement in china's secondary schools?', *The Review of Economics and Statistics* **89**, 300–312.

Duflo, Esther, Pascaline Dupas and Michael Kremer (2009), Additional resources versus organizational changes in education: Experimental evidence from Kenya, Working paper, Massachussetts Institute of Technology.

Duflo, Esther, Pascaline Dupas and Michael Kremer (2011), 'Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya', *American Economic Review* **101**, 1739–1774.

Durlauf, Steven (2004), Neighborhood effects, *in* J. V.Henderson and J.-F.Thisse, eds, 'Handbook of Regional and Urban Economics', Vol. 4, Elsevier, chapter 50.

Durlauf, Steven and Yannis Ioannides (2010), 'Social interactions', *Annual Review of Economics* **2**, 451–478.

Evans, Mariah, Jonathan Kelley, Joanna Sikora and Donals Treiman (2010), 'Family scholarly culture and educational success: Books and schooling in 27 nations', *Research in Social Stratification and Mobility* **28**, 171–197.

Fairlie, Robert and Jonathan Robinson (2012), Experimental evidence on the effects of home computers on academic achievement among schoolchildren, Technical report, University of California, Santa Cruz.

Figlio, David and Marianne Page (2002), 'School choice and the distributional effects of ability tracking: Does separation increase inequality?', *Journal of Urban Economics* **51**, 497–514.

Froelich, Markus and Katharina Michaelowa (2011), 'Peer effects and textbooks in African primary education', *Labour Economics* **18**, 474–486.

Garlick, Robert (2012), Academic peer effects when residential groups are tracked or randomly assigned, working paper, University of Michigan.

Gibbons, Stephen and Shqiponja Telhaj (2008), Peers and achievement in England's Secondary Schools, Discussion paper no.1, SERC.

Glewwe, Paul, Michael Kremer and Sylvie Moulin (2009), 'Many children Left Behind? Textboks and test scores in Kenya', *Americal Economic Journal: Applied Economics* **1**, 112–135.

Graham, Bryan, Guido Imbens and Geert Ridder (2009), Complementarity and aggregate implications of Assortative Matching: A nonparametric analysis, NBER working paper no. 14860, National Bureau of Economic Reseach.

Graham, Bryan, Guido Imbens and Geert Ridder (2010), Measuring the effects of segregation in the presence of social spillovers: A nonparametric approach, NBER working paper no. 16499, National Bureau of Economic Reseach.

Hanushek, Eric, John Kain, Jacob Markman and Steven Rivkin (2003), 'Does peer ability affect student achievement', *Journal of Applied Econometrics* **18**, 527–544.

Hanushek, Eric, John Kain and Steven Rivkin (2009), 'New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement', *Journal of Labor Economics* **27**, 349–383.

Hoxby, Caroline (2000*a*), 'The effects of class size on student achievement: New evidence from population variation', *The Quarterly Journal of Economics* **115**, 1239–1285.

Hoxby, Caroline (2000*b*), Peer effects in the classroom: Learning from gender and race variation, Working Paper 7867, National Bureau of Economic Research.

Hoxby, Caroline and Gretchen Weingarth (2006), Taking race out of the equation: School reassignment and the structure of peer effects, Working paper, Harvard University.

Hurder, Stephanie (2012), Evaluating econometric models of peer effects with experimental data, Working paper, Harvard Business School.

Jackson, Kirabo (2012), 'Single-sex schools, student achievement, and course selection: Evidence from rule-based student assignments in Trinidad and Tobago', *Journal of Public Economics* **96**, 173–187.

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

Konstantopoulos, Spyros (2008), 'Do small classes reduce the achievement gap between low and high achievers? Evidence from Project STAR', *Elementary School Journal* **108**, 1–17.

Krueger, Alan (2003), 'Economic considerations and class size', *The Economic Journal* **113**, F34–F63.

Krueger, Alan and Diane Whitmore (2001), Would smaller classes help close the black-white achievement gap?, *in* J.Chubb and T.Loveless, eds, 'Bridging the Achievement Gap', Brookings Institution Press.

Laroque, Guy and Bernard Salanie (1989), 'Estimation of multi-market fix-price models: An application of pseudo maximum likelihood methods', *Econometrica* **57**, 831–860.

Lavy, Victor and Analia Schlosser (2011), 'Mechanisms and impacts of gender peer effects at school', *American Economic Journal: Applied Economics* **3**, 1–33.

Lavy, Victor, Olmo Silva and Felix Weinhardt (2009), The Good, the Bad and the Average: Evidence on the scale and nature of ability peer effects in schools, NBER working paper no. 15600, National Bureau of Economic Reseach.

Lazear, Edward (2001), 'Educational production', *The Quarterly Journal of Economics* **116**, 777–803.

Lee, Lung-fei (2007), 'Identification and estimation of econometric models with group interactions, contextual factors and fixed effects', *Journal of Econometrics* **140**, 333–374.

Malamud, Ofer and Cristian Pop-Eleches (2011), 'Home computer use and the development of human capital', *The Quarterly Journal of Economics* **126**, 987–1027.

Manski, Charles (1993), 'Identification of endogenous social effects: The Reflection Problem', *Review of Economic Studies* **60**, 531–542.

McEwan, Patrick and Kristen Soderberg (2006), 'Roommate effects on grades: Evidence from first-year housing assignments', *Research in Higher Education* **47**, 347–370.

MINEDUC (2011), Resumen informe de implementación liceos bicentenario de excelencia, Technical report, Subsecretaría de Educación, Ministerio de Educación.

Moffitt, Robert (2001), Policy interventions, low-level equilibria and social interactions, *in* S.Durlauf and P.Young, eds, 'Social Dynamics', MIT Press.

Murnane, Richard, John Willett and Frank Levy (1995), 'The growing importance of cognitive skills in wage determination', *Review of Economics and Statistics* **77**, 251–266.

NMAP, National Mathematics Advisory Panel (2008), Fundations for success: The final report of the National Mathematics Advisory Panel, Technical report, Washington DC, U.S. Department of Education.

Park, Hyunjoon, Jere Behrman and Jaesung Choi (2012), 'Causal effects of single-sex schools on college entrance exams and college attendance: random assignment in Seoul High Schools', *Demography* .

Pinto, Cristina (2010), Semiparametric estimation of peer effects in classrooms: Evidence from Brazilian schools in 2003, Working paper, University of California, Berkeley.

Pischke, Joern-Steffen and Alan Manning (2006), Comprehensive versus Selective schooling in England and Wales: What do we know?, NBER working paper no. 12176, National Bureau of Economic Reseach.

Richards, Seth (2012), Peer effects in sexual initiation: Separating supply and demand mechanisms, working paper, Carnegie Mellon University.

Rodriguez, Jorge (2010), La no linealidad del efecto par educacional: Evidencia para Chile, Working paper.

Rothman, Terri and Mary Henderson (2011), 'Do school-based tutoring programs significantly improve student performance on standardized tests?', *Research in Middle Level Education* **34**(6).

Sacerdote, Bruce (2001), 'Peer effects with random assignment: Results for Dartmouth roommates', *The Quarterly Journal of Economics* **116**, 681–704.

Sullivan, Alice (2009), 'Academic self-concept, gender and single-sex schooling', *British Educational Research Journal* **35**, 259–288.

Teddlie, Charles and Sam Stringfield (2007), A history of school effectiveness and improvement research in the USA, *in* T.Townsend, ed., 'International Handbook of School Effectiveness and Improvement', Vol. 17, Springer, pp. 131–166.

Todd, Petra and Kenneth Wolpin (2003), 'On the specification and estimation of the Production Function for cognitive achievement', *The Economic Journal* **113**, F3–F33.

Valenzuela, Juan Pablo and Claudio Allende (2012), Logros en Liceos Publicos de Excelencia en Chile: valor agregado o solo descreme de la elite?, Technical report, Universidad de Chile.

Vigdor, Jacob and Helen Ladd (2010), Scaling the digital divide: Home computer technology and student achievement, NBER working paper no. 16078, National Bureau of Economic Reseach.

Vigdor, Jacob and Thomas Nechyba (2004), Peer effects in North Carolina public schools, Working paper, Duke University.

Zabel, Jeffrey (2008), 'The impact of peer effects on student outcomes in New York City public schools', *Education Finance and Policy* **3**, 197–249.

Zimmer, Ron (2003), 'A new twist in the educational tracking debate', *Economics of Education Review* **22**, 307–315.

Zimmerman, David (1999), Peer effects in academic outcomes: Evidence from a natural experiment, Technical Report 52, Williams College.

Appendix

A Estimation Algorithm

The model is estimated using a Simulated Maximum Likelihood algorithm. The achievement and effort of each student is influenced by the characteristics and behavior of his whole group or class, therefore this is the relevant unit of observation. The algorithm used for estimation is:

1. For each class g, perform a Monte Carlo simulation. For each simulation draw d, sample shocks for $(\{A^d_{ig0}\}_{i\in g}, \alpha^d_g)$ for the whole class.

2. Compute optimal effort e^d_{ig} for every $i \in g$. Use a Newton algorithm to solve the nonlinear system of equations given by the students FOC, and find the game equilibrium given $(\{A^d_{ig0}\}_{i\in g}, \alpha^d_g)$. Then, compute optimal achievement A^d_{ig1} using the production function.

3. Using the measurement equations, compute the likelihood of test scores and effort conditional on each draw d for each student $i \in g$:

   - For lagged test scores, the likelihood is given by:
   $$p^d(T_{ig0}) = \phi(T_{ig0}; \eta_0 + \tau_0 A^d_{ig0}, \Sigma_{\varepsilon 0})$$
   - For current test scores, the likelihood is given by:
   $$p^d(T_{ig1}) = \phi(T_{ig1}; \eta_1 + \tau_1 A^d_{ig1}, \Sigma_{\varepsilon 1})$$
   - The likelihood of the effort measures is given by a mixture of normals for the continuous measure, and an ordered probit for the discrete measures:
   $$\begin{aligned} p^d(R_{ig}) &= p^d(R_{ig0}) \prod_{l=1}^{4} \prod_{j=1}^{L} p^d(R_{igl} = j)^{1\{R_{igl}=j\}} \\ p^d(R_{ig0}) &= \sum_{j=1}^{2} w_{aj} \phi(R_{ig0}; e^d_{ig} + \mu_{aj}, \sigma^2_{aj}) \\ p^d(R_{igl} = j) &= \Phi\left(c_{l,j} - \tau_{2l} e^d_{ig}\right) - \Phi\left(c_{l,j-1} - \tau_{2l} e^d_{ig}\right) \forall l = 1, \ldots, 4 \end{aligned}$$
   - The likelihood of observing the measures for each student is thus given by:
   $$p^d_{ig} = p^d(T_{ig0}) p^d(T_{ig1}) p^d(R_{ig})$$

4. Finally, the objective function is given by:
$$\Lambda = \prod_{g=1}^{G} \int \prod_{i\in g} (p^d_{ig} | \{A^d_{ig0}\}_{i\in g}, \alpha^d_g) dF(\{A^d_{ig0}\}_{i\in g}, \alpha^d_g)$$

   This function is maximized using numerical methods, finishing the estimation algorithm.

Table 1:
Distribution of Test Scores and Individual and Class Characteristics

	Individual Characteristics N=130853			
Variable	Mean	Std. Dev.	Min	Max
current math	267.15	51.36	134.35	402.42
current cience	266.16	50.58	133.59	402.74
lagged math	260.06	51.48	90.83	363.62
lagged science	269.16	49.42	108.58	363.69
attendance	0.9398	0.0512	0.6	1
female	0.5262	0.4993	0	1
parents scholing	12.021	2.94	1	20
parents have college	0.3613	0.4804	0	1
parents live at home	0.5852	0.4927	0	1
books at home 10-50	0.4806	0.4996	0	1
books at home 50-more	0.2671	0.4424	0	1
computer at home	0.6932	0.4611	0	1
internet at home	0.4842	0.4998	0	1

	Teacher and Class Characteristics N=7779			
Variable	Mean	Std. Dev.	Min	Max
math teacher female	0.4899	0.4999	0	1
science teacher female	0.6116	0.4874	0	1
math teacher experience	16.0062	13.5998	0	62
science teacher experience	15.0878	13.3415	0	66
missing math teacher female	0.1426	0.3496	0	1
missing science teacher female	0.1498	0.3569	0	1
missing math teacher experience	0.1545	0.3615	0	1
missing science teacher experience	0.1625	0.3689	0	1
school is urban	0.8275	0.3779	0	1
school is public	0.501	0.5	0	1
school is subsidized	0.416	0.4929	0	1
class size	16.8213	8.5812	2	41

Table 2:
Distribution of Discrete Effort Measures

		Category				
		Complet. Disagree	Disagree	Indiff	Agree	Complet. Agree
	Measure	1	2	3	4	5
1	Does homework	2.66%	7.15%	22.94%	38.09%	29.16%
2	Notebooks complete	5.16%	12.8%	17.41%	31.53%	33.1%
3	Takes notes	7.85%	13.97%	31.25%	30.44%	16.48%
		Category				
		Never	1 or 2 days	3 or 4 days	Every day	
	Measure	1	2	3	4	
4	How many days studies	14.31%	57.37%	20.67%	7.65%	

Table 3:
Parameter Values and Variances

Parameter Vector		Value	Variance	St. Dev.	Confidence Interval		P-value
Cost Function							
c_0	constant	15.2371	6.38E-04	2.53E-02	15.1876	15.2867	0.00
β_1	female	-1.7941	9.60E-05	9.80E-03	-1.8133	-1.7749	0.00
β_2	parents scholing	0.0976	4.75E-06	2.18E-03	0.0934	0.1019	0.00
β_3	parents have college	0.2319	2.16E-04	1.47E-02	0.2031	0.2608	0.00
β_4	parents live at home	-0.9178	8.91E-05	9.44E-03	-0.9363	-0.8993	0.00
β_5	books at home 10-50	-0.5299	1.14E-04	1.07E-02	-0.5509	-0.509	0.00
β_6	books at home 50-more	-1.0766	1.80E-04	1.34E-02	-1.1029	-1.0502	0.00
β_7	computer at home	-0.0921	1.77E-04	1.33E-02	-0.1181	-0.066	0.00
β_8	internet at home	0.3206	1.78E-04	1.33E-02	0.2944	0.3467	0.00
δ_1	% of females	0.0092	3.24E-04	1.80E-02	-0.0261	0.0445	0.61
δ_2	peer parents scholing	-0.0417	1.39E-05	3.73E-03	-0.049	-0.0344	0.00
δ_3	% of parents have college	-2.5015	6.05E-04	2.46E-02	-2.5497	-2.4532	0.00
δ_4	% of parents live at home	-1.6644	2.90E-04	1.70E-02	-1.6978	-1.631	0.00
δ_5	% of books at home 10-50	1.542	4.29E-04	2.07E-02	1.5014	1.5826	0.00
δ_6	% of books at home 50-more	0.6905	5.96E-04	2.44E-02	0.6427	0.7383	0.00
δ_7	% of computer at home	1.9675	5.74E-04	2.40E-02	1.9206	2.0145	0.00
δ_8	% of internet at home	1.0934	4.74E-04	2.18E-02	1.0507	1.1361	0.00
Production Function							
κ_{C1}	constant	1.0455	1.39E-07	3.73E-04	1.0448	1.0462	0.00
κ_{C2}	female	-0.0121	3.55E-08	1.88E-04	-0.0125	-0.0118	0.00
κ_{C3}	% of females	-0.0078	1.36E-07	3.68E-04	-0.0085	-0.0071	0.00
κ_{C4}	SD of lagged achievement	0.0029	5.58E-10	2.36E-05	0.0029	0.003	0.00
γ	constant	0.8471	2.90E-10	1.70E-05	0.8471	0.8471	0.00
ϕ	constant	0.7025	2.82E-06	1.68E-03	0.6992	0.7058	0.00
$\kappa_{\omega 1}$	constant	0.1128	1.73E-08	1.31E-04	0.1125	0.1131	0.00
$\kappa_{\omega 2}$	school is urban	-0.0095	4.28E-09	6.55E-05	-0.0096	-0.0094	0.00
$\kappa_{\omega 3}$	school is public	-0.0054	6.24E-09	7.90E-05	-0.0056	-0.0053	0.00
$\kappa_{\omega 4}$	school is subsidized	-7.18E-06	5.93E-09	7.70E-05	-1.58E-04	1.44E-04	0.93
$\kappa_{\omega 5}$	class size	-1.02E-04	5.56E-12	2.36E-06	-1.07E-04	-9.78E-05	0.00

(table continues...)

Parameter Vector		Value	Variance	St. Dev.	Confidence Interval		P-value
$\kappa_{\omega 6}$	SD of lagged achievement	-0.0014	2.59E-11	5.09E-06	-0.0014	-0.0014	0.00
Mean of Teacher Productivity							
π_{t2}	Math teacher female	0.0012	6.61E-09	8.13E-05	0.001	0.0014	0.00
π_{t3}	Science teacher female	0.0013	8.76E-09	9.36E-05	0.0011	0.0015	0.00
π_{t4}	Math teacher experience	7.31E-07	1.11E-11	3.33E-06	-6.72E-05	-5.42E-05	0.00
π_{t5}	Science teacher experience	-3.63E-03	1.20E-11	3.47E-06	-6.07E-06	7.53E-06	0.83
π_{t6}	missing Math teacher female	-0.0036	9.21E-08	3.03E-04	-0.0042	-0.003	0.00
π_{t7}	missing Science teacher female	-0.0085	1.26E-07	3.55E-04	-0.0092	-0.0078	0.00
π_{t8}	missing Math teacher experience	-0.0027	8.01E-08	2.83E-04	-0.0032	-0.0021	0.00
π_{t9}	missing Science teacher experience	0.0112	1.12E-07	3.35E-04	0.0106	0.0119	0.00
π_{t10}	school is urban	0.0081	1.03E-07	3.21E-04	0.0075	0.0087	0.00
π_{t11}	school is public	0.0089	1.30E-07	3.60E-04	0.0082	0.0096	0.00
π_{t12}	school is subsidized	-0.0089	1.25E-07	3.54E-04	-0.0096	-0.0082	0.00
π_{t13}	class size	0.0014	1.31E-10	1.14E-05	0.0014	0.0015	0.00
Mean of Lagged Achievement							
π_{s1}	constant	250.5181	1.83E-03	4.27E-02	250.4343	250.6018	0.00
π_{s2}	female	-8.2883	7.64E-04	2.76E-02	-8.3424	-8.2341	0.00
π_{s3}	parents scholing	2.7106	1.45E-05	3.81E-03	2.7031	2.718	0.00
π_{s4}	parents have college	2.8783	9.41E-04	3.07E-02	2.8182	2.9384	0.00
π_{s5}	school is urban	6.8412	3.23E-03	5.68E-02	6.7298	6.9526	0.00
π_{s6}	school is public	-35.5727	3.27E-03	5.72E-02	-35.6847	-35.4607	0.00
π_{s7}	school is subsidized	-23.2813	3.38E-03	5.81E-02	-23.3953	-23.1674	0.00
Joint distribution of Lagged Achievement and Teacher Productivity							
σ_t	SD of teacher prod	0.0011	2.02E-09	4.50E-05	0.001	0.0012	0.00
σ_{sm}	SD of lagged achievement for males	25.423	1.48E-04	1.22E-02	25.3991	25.4469	0.00
σ_{sf}	SD of lagged achievement for females	25.0354	1.40E-04	1.18E-02	25.0122	25.0585	0.00
σ_{st}	Covariance between teachers and students	0.0174	5.05E-07	7.11E-04	0.0161	0.0188	0.00
σ_p	Covariance between peers	392.4848	8.82E-02	2.97E-01	391.9026	393.067	0.00
Measurement Equations for Achievement							
η_{02}	constant for lagged Science	9.422	2.62E-02	1.62E-01	9.1049	9.739	0.00
τ_{02}	loading for lagged Science	0.9986	3.83E-07	6.19E-04	0.9974	0.9998	0.00
η_{12}	constant for current Science	8.6853	3.29E-02	1.81E-01	8.3298	9.0407	0.00
τ_{12}	loading for current Science	0.9638	4.53E-07	6.73E-04	0.9625	0.9651	0.00

(table continues...)

Parameter Vector		Value	Variance	St. Dev.	Confidence Interval		P-value
σ_{01}	SD of lagged Math error	41.4842	9.56E-04	3.09E-02	41.4236	41.5448	0.00
σ_{02}	SD of lagged Science error	39.0929	7.86E-04	2.80E-02	39.0379	39.1478	0.00
σ_{11}	SD of current Math error	38.4141	7.77E-04	2.79E-02	38.3595	38.4687	0.00
σ_{12}	SD of current Science error	38.7109	6.61E-04	2.57E-02	38.6605	38.7613	0.00
Measurement Equations for Effort							
τ_{21}	loading for discrete effort 1	-0.5223	6.82E-07	8.26E-04	-0.5239	-0.5206	0.00
τ_{22}	loading for discrete effort 2	1.4107	7.63E-07	8.73E-04	1.409	1.4124	0.00
τ_{23}	loading for discrete effort 3	0.871	4.91E-07	7.01E-04	0.8696	0.8724	0.00
τ_{24}	loading for discrete effort 4	2.2631	8.05E-07	8.97E-04	2.2614	2.2649	0.00
μ_{a1}	mean error 1 in mixture	5.07E-06	4.45E-06	2.11E-03	-4.13E-03	4.14E-03	1.00
σ_{a1}	SD error 1 in mixture	0.1328	1.30E-07	3.61E-04	0.1321	0.1335	0.00
σ_{a2}	SD error 2 in mixture	0.0922	3.58E-08	1.89E-04	0.0918	0.0926	0.00
w_{a1}	weight error 1 in mixture	0.4593	9.42E-05	9.70E-03	0.4403	0.4783	0.00
$c_{1,1}$	cutoffs for discrete effort	-2.5438	5.15E-06	2.27E-03	-2.5482	-2.5393	0.00
$c_{1,2}$	cutoffs for discrete effort	-1.9032	1.98E-06	1.41E-03	-1.906	-1.9005	0.00
$c_{1,3}$	cutoffs for discrete effort	-1.0549	1.17E-06	1.08E-03	-1.0571	-1.0528	0.00
$c_{1,4}$	cutoffs for discrete effort	-0.0578	1.52E-06	1.23E-03	-0.0602	-0.0554	0.00
$c_{2,1}$	cutoffs for discrete effort	-0.0191	3.83E-06	1.96E-03	-0.0229	-0.0153	0.00
$c_{2,2}$	cutoffs for discrete effort	0.7024	1.33E-06	1.15E-03	0.7001	0.7046	0.00
$c_{2,3}$	cutoffs for discrete effort	1.2515	9.78E-07	9.89E-04	1.2495	1.2534	0.00
$c_{2,4}$	cutoffs for discrete effort	2.0757	1.39E-06	1.18E-03	2.0734	2.078	0.00
$c_{3,1}$	cutoffs for discrete effort	-0.4155	2.57E-06	1.60E-03	-0.4186	-0.4123	0.00
$c_{3,2}$	cutoffs for discrete effort	0.2254	1.20E-06	1.10E-03	0.2233	0.2276	0.00
$c_{3,3}$	cutoffs for discrete effort	1.0862	1.01E-06	1.01E-03	1.0842	1.0882	0.00
$c_{3,4}$	cutoffs for discrete effort	1.9861	1.61E-06	1.27E-03	1.9836	1.9886	0.00
$c_{4,1}$	cutoffs for discrete effort	1.5267	2.30E-06	1.52E-03	1.5237	1.5296	0.00
$c_{4,2}$	cutoffs for discrete effort	3.2147	1.40E-06	1.18E-03	3.2124	3.217	0.00
$c_{4,3}$	cutoffs for discrete effort	4.0885	2.59E-06	1.61E-03	4.0853	4.0916	0.00

Table 4:
Model Fit: Mean and Standard Deviation of Measures

	Mean		St. Dev.	
	True	Simulated	True	Simulated
Current Math	267.15	266.14	51.362	50.933
Current Science	266.16	265.22	50.583	50.411
Attendance	0.9398	0.9392	0.0512	0.0567
Effort1	3.8394	3.8434	1.0103	1.0093
Effort2	3.746	3.744	1.1905	1.1894
Effort3	3.3372	3.3353	1.1418	1.1415
Effort4	2.2167	2.2098	0.7804	0.7784

Table 5:
Marginal Effect for Discrete Variables

	Unit change in value of variable	
	On Achievement	On Effort
female	-7.774%SD**	5.372%**
	(0.9141)	(0.2582)
parents have college	3.986%SD**	-0.392%**
	(0.2225)	(0.0448)
parents live at home	9.12%SD**	4.018%**
	(0.6616)	(0.1223)
books at home 10-50	5.16%SD**	2.27%**
	(0.3736)	(0.068)
books at home 50-more	10.785%SD**	4.76%**
	(0.7931)	(0.1548)
computer at home	0.922%SD**	0.398%**
	(0.0676)	(0.0126)
internet at home	-3.216%SD**	-1.376%**
	(0.2361)	(0.043)
% of females (0 to 100%)	-4.911%SD**	-0.5%**
	(0.2476)	(0.0068)
% of parents have college (0 to 100%)	26.141%SD**	11.799%**
	(2.1859)	(0.578)
% of parents live at home (0 to 100%)	16.516%SD**	7.4%**
	(1.2197)	(0.2472)
% of books at home 10-50 (0 to 100%)	-15.717%SD**	-6.513%**
	(1.186)	(0.224)
% of books at home 50-more (0 to 100%)	-7.349%SD**	-3.061%**
	(0.5595)	(0.1137)
% of computer at home (0 to 100%)	-20.666%SD**	-8.393%**
	(1.5658)	(0.287)

(table continues...)

	Unit change in value of variable	
	On Achievement	On Effort
% of internet at home (0 to 100%)	-10.982%SD**	-4.615%**
	(0.8244)	(0.162)
Math teacher female	0.783%SD**	0.075%**
	(0.037)	(0.0004)
Science teacher female	0.836%SD**	0.08%**
	(0.0395)	(0.0005)
school is urban	1.104%SD	-5.723%**
	(1.4839)	(0.39)
school is public	-84.725%SD**	-10.851%**
	(4.9263)	(0.4228)
school is subsidized	-58.05%SD**	-5.107%**
	(3.2976)	(0.2755)

Standard errors are in parenthesis below.

** These results are significant at a 95% confidence level.

Table 6:
Marginal Effect for Continuous Variables

	Marginal Effect on Achievement Unitary change in the value of the variable				
Variable	0-1	5-6	10-11	15-16	20-21
Parents Schooling	5.083%SD**	5.038%SD**	4.995%SD**	4.954%SD**	4.915%SD**
	(0.0544)	(0.0539)	(0.0532)	(0.0526)	(0.052)
Peer Parent Schooling	0.398%SD**	0.407%SD**	0.415%SD**	0.424%SD**	0.434%SD**
	(0.028)	(0.0287)	(0.0294)	(0.0302)	(0.031)
Math teacher experience	-0.039%SD**	-0.039%SD**	-0.039%SD**	-0.039%SD**	-0.039%SD**
	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)
Science teacher experience	0.0005%SD**	0.0005%SD**	0.0005%SD**	0.0005%SD**	0.0005%SD**
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Class size		0.724%SD**	0.721%SD**	0.718%SD**	0.715%SD**
		(0.0332)	(0.0331)	(0.0331)	(0.033)

	Marginal Effect on Effort Unitary change in the value of the variable				
Variable	0-1	5-6	10-11	15-16	20-21
Parents Schooling	0.195%**	0.174%**	0.156%**	0.14%**	0.126%**
	(0.0318)	(0.0286)	(0.0258)	(0.0235)	(0.0214)
Peer Parent Schooling	0.174%**	0.177%**	0.18%**	0.182%**	0.185%**
	(0.0045)	(0.0046)	(0.0047)	(0.0049)	(0.005)
Math teacher experience	-0.004%**	-0.004%**	-0.004%**	-0.004%**	-0.004%**
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Science teacher experience	0.00005%**	0.00005%**	0.00005%**	0.00005%**	0.00005%**
	(0.0000003)	(0.0000003)	(0.0000003)	(0.0000003)	(0.0000003)
Class size		0.004%	0.003%	0.002%	0.001%
		(0.0048)	(0.0049)	(0.0049)	(0.005)

Standard errors are in parenthesis below.

** These results are significant at a 95% confidence level.

Table 7:
Counterfactual 1: Decrease the cost of effort c_{ig} for worst students

	Lower c_{ig} in 10%		
Tutor assigned to	Worst student	Every student	Smart student
Number of benefits	37000	82372	37000
SIMCE points for worst	5.298	7.493	3.281
Equivalence in Schooling	Close to 2 years	More than 2 years	More than 1 year
Change in Mean Math score, Percentage of Standard Deviation			
For the classes:	6.74%SD**	15.37%SD**	6.71%SD**
	(1.784)	(1.035)	(1.794)
For benefited students:	10.4%SD**	15.37%SD**	10.61%SD**
	(2.068)	(1.035)	(2.004)
For worst students:	10.4%SD**	14.71%SD**	6.44%SD**
	(2.068)	(0.934)	(0.457)
For the rest of their class:	3.75%SD**	15.91%SD**	1.59%SD**
	(0.313)	(1.048)	(0.149)
Percentual Change in Mean cost of effort			
For the classes:	-4.309%	-10%	-4.535%
	(-)	(-)	(-)
For benefited students:	-10%	-10%	-10%
	(-)	(-)	(-)
For worst students:	-10%	-10%	0%
	(-)	(-)	(-)
For the rest of their class:	0%	-10%	0%
	(-)	(-)	(-)
Percentual Change in Mean effort			
For the classes:	3.14%**	7.011%**	3.114%**
	(0.875)	(0.043)	(0.883)
For benefited students:	6.55%**	7.011%**	6.547%**
	(0.225)	(0.043)	(0.229)
For worst students:	6.55%**	7.01%**	0.671%**
	(0.225)	(0.042)	(0.026)
For the rest of their class:	0.357%**	7.012%**	0.151%**
	(0.023)	(0.045)	(0.01)

Standard errors are in parenthesis below.

** These results are significant at a 95% confidence level.

Table 8:
Counterfactual 2.1: Public Provision of Home Inputs, Books

	Increase Number of Books		
Tutor assigned to	Worst student	Every student	Smart student
Number of benefits	37006	108269	37006
SIMCE points for worst	-0.392	-0.799	-0.376
Equivalence in Schooling	Less than -1 year	Less than -1 year	Less than -1 year
Change in Mean Math score, Percentage of Standard Deviation			
For the classes:	-0.62%SD**	-1.47%SD**	-0.53%SD**
	(0.06)	(0.146)	(0.051)
For benefited students:	-0.77%SD**	-1.47%SD**	-0.71%SD**
	(0.083)	(0.146)	(0.079)
For worst students:	-0.77%SD**	-1.57%SD**	-0.74%SD**
	(0.083)	(0.156)	(0.065)
For the rest of their class:	-0.54%SD**	-1.41%SD**	-0.31%SD**
	(0.049)	(0.141)	(0.028)
Percentual Change in Mean cost of effort			
For the classes:	0.443%**	0.982%**	0.352%**
	(0.03)	(0.075)	(0.03)
For benefited students:	-0.879%**	0.982%**	-0.799%**
	(0.058)	(0.075)	(0.055)
For worst students:	-0.879%**	0.941%**	1.687%**
	(0.058)	(0.081)	(0.057)
For the rest of their class:	1.101%**	1.003%**	0.643%**
	(0.035)	(0.072)	(0.038)
Percentual Change in Mean effort			
For the classes:	-0.281%**	-0.653%**	-0.243%**
	(0.02)	(0.049)	(0.019)
For benefited students:	0.338%**	-0.653%**	0.298%**
	(0.035)	(0.049)	(0.033)
For worst students:	0.338%**	-0.655%**	-0.924%**
	(0.035)	(0.053)	(0.032)
For the rest of their class:	-0.6%**	-0.652%**	-0.35%**
	(0.02)	(0.047)	(0.021)

Standard errors are in parenthesis below.

** These results are significant at a 95% confidence level.

Table 9:
Counterfactual 2.2: Public Provision of Home Inputs, Computer

	Increase Number of Computers		
Tutor assigned to	Worst student	Every student	Smart student
Number of benefits	37006	108269	37006
SIMCE points for worst	-1.914	-2.873	-1.095
Equivalence in Schooling	More than	More than	
	-1 year	-2 years	-1 year
Change in Mean Math score, Percentage of Standard Deviation			
For the classes:	-2.34%SD**	-4.31%SD**	-1.87%SD**
	(0.177)	(0.325)	(0.137)
For benefited students:	-3.76%SD**	-4.31%SD**	-3.29%SD**
	(0.289)	(0.325)	(0.243)
For worst students:	-3.76%SD**	-5.64%SD**	-2.15%SD**
	(0.289)	(0.432)	(0.161)
For the rest of their class:	-1.6%SD**	-3.63%SD**	-0.68%SD**
	(0.12)	(0.27)	(0.047)
Percentual Change in Mean cost of effort			
For the classes:	1.645%**	3.05%**	1.307%**
	(0.055)	(0.101)	(0.045)
For benefited students:	2.468%**	3.05%**	2.067%**
	(0.086)	(0.101)	(0.075)
For worst students:	2.468%**	4.095%**	1.829%**
	(0.086)	(0.139)	(0.061)
For the rest of their class:	1.235%**	2.529%**	0.525%**
	(0.04)	(0.083)	(0.027)
Percentual Change in Mean effort			
For the classes:	-1.063%**	-1.937%**	-0.854%**
	(0.037)	(0.067)	(0.03)
For benefited students:	-1.636%**	-1.937%**	-1.398%**
	(0.059)	(0.067)	(0.05)
For worst students:	-1.636%**	-2.605%**	-1.146%**
	(0.059)	(0.091)	(0.04)
For the rest of their class:	-0.768%**	-1.594%**	-0.329%**
	(0.026)	(0.054)	(0.016)

Standard errors are in parenthesis below.

** These results are significant at a 95% confidence level.

Table 10:
Counterfactual 2.3: Public Provision of Home Inputs, Internet

	Increase Number of Computer+Internet		
Tutor assigned to	Worst student	Every student	Smart student
Number of benefits	37006	108269	37006
SIMCE points for worst	-4.056	-6.323	-2.468
Equivalence in Schooling		More than	
	-3 years	-3 years	-2 years
Change in Mean Math score, Percentage of Standard Deviation			
For the classes:	-5.01%SD**	-10.16%SD**	-4.23%SD**
	(0.363)	(0.726)	(0.287)
For benefited students:	-7.96%SD**	-10.16%SD**	-7.25%SD**
	(0.584)	(0.726)	(0.502)
For worst students:	-7.96%SD**	-12.42%SD**	-4.84%SD**
	(0.584)	(0.9)	(0.339)
For the rest of their class:	-3.48%SD**	-8.98%SD**	-1.7%SD**
	(0.249)	(0.634)	(0.121)
Percentual Change in Mean cost of effort			
For the classes:	3.625%**	7.448%**	3.054%**
	(0.094)	(0.192)	(0.092)
For benefited students:	6.356%**	7.448%**	5.488%**
	(0.166)	(0.192)	(0.165)
For worst students:	6.356%**	9.667%**	3.461%**
	(0.166)	(0.253)	(0.096)
For the rest of their class:	2.264%**	6.343%**	1.08%**
	(0.059)	(0.162)	(0.065)
Percentual Change in Mean effort			
For the classes:	-2.269%**	-4.521%**	-1.921%**
	(0.062)	(0.119)	(0.056)
For benefited students:	-3.883%**	-4.521%**	-3.421%**
	(0.105)	(0.119)	(0.096)
For worst students:	-3.883%**	-5.797%**	-2.204%**
	(0.105)	(0.153)	(0.061)
For the rest of their class:	-1.439%**	-3.865%**	-0.696%**
	(0.039)	(0.101)	(0.039)

Standard errors are in parenthesis below.

** These results are significant at a 95% confidence level.

Table 11:
Counterfactual 3: Same gender classes

	% Change	
Class Size	20	10
Mean Math score - Percentual Change of SD		
For all students	0.558%SD**	0.501%SD**
	(0.082)	(0.078)
For males	-2.133%SD**	-2.267%SD**
	(0.48)	(0.456)
For females	3.249%SD**	3.107%SD**
	(0.559)	(0.54)
Mean Cost of Effort - Percentual Change		
For all students	0%	0%
	(-)	(-)
For males	-0.139%**	-0.197%**
	(0.007)	(0.009)
For females	0.157%**	0.122%**
	(0.009)	(0.008)
Mean Effort - Percentual Change		
For all students	0.62%**	0.577%**
	(0.073)	(0.074)
For males	0.171%	0.183%
	(0.3)	(0.277)
For females	1.051%**	1.014%**
	(0.317)	(0.296)

Standard errors are in parenthesis below.

** These results are significant at a 95% confidence level.

Table 12:
Counterfactual 4: Tracking inside the school

	% Change	
Class Size	20	10
Mean Math score - Percentual Change of SD		
For all students	2.412%SD**	3.029%SD**
	(0.417)	(0.465)
For worst students	-3.706%SD	-1.978%SD**
	(3.242)	(0.388)
For best students	8.531%SD**	11.628%SD**
	(3.833)	(1.656)
Mean Cost of Effort - Percentual Change		
For all students	0%	0%
	(-)	(-)
For worst students	4.489%**	4.425%**
	(1.93)	(0.295)
For best students	-4.217%**	-6.441%**
	(1.909)	(0.804)
Mean Effort - Percentual Change		
For all students	2.139%**	2.796%**
	(0.269)	(0.369)
For worst students	-0.646%	0.162%
	(1.391)	(0.447)
For best students	4.932%**	7.065%**
	(1.518)	(0.706)

Standard errors are in parenthesis below.

** These results are significant at a 95% confidence level.

Table 13:
Counterfactual 4: Replicating the CSW 2011 experiment

	% Change
Mean Math score - % Change of SD	
For middle students	0.432%SD
	(0.643)
For worst students	-0.273%SD
	(0.682)
For best students	-0.767%SD
	(0.737)
Mean Cost of Effort - % Change	
For middle students	1.191%
	(-)
For worst students	-1.515%**
	(0.393)
For best students	-1.108%**
	(0.335)
Mean Effort - % Change	
For middle students	1.528%**
	(0.389)
For worst students	-1.514%**
	(0.37)
For best students	-1.741%**
	(0.352)

Standard errors are in parenthesis below.
** These results are significant at
a 95% confidence level.

Table 14:
Counterfactual 5: Variations in Class Size - Short Run

	Effect of Class Size Variation with respect to Class Size=40		
Class size	5	10	20
Mean Math score - Percentual Change of SD			
For all students	9.004%SD**	7.188%SD**	4.725%SD**
	(0.654)	(0.498)	(0.306)
For 5 worst students	6.771%SD**	6.042%SD**	4.166%SD**
	(0.601)	(0.468)	(0.29)
For 5 best students	10.111%SD**	7.708%SD**	5.001%SD**
	(0.695)	(0.509)	(0.311)
Mean Cost of Effort - Percentual Change			
For all students	0%	0%	0%
	(-)	(-)	(-)
For 5 worst students	0.039%**	0.017%**	0.016%**
	(0.002)	(0.001)	(0.001)
For 5 best students	0.022%**	0.014%**	-0.013%**
	(0.005)	(0.001)	(0.001)
Mean Effort - Percentual Change			
For all students	4.57%**	3.38%**	2.104%**
	(0.377)	(0.239)	(0.131)
For 5 worst students	3.621%**	2.991%**	1.975%**
	(0.492)	(0.29)	(0.149)
For 5 best students	3.664%**	2.965%**	1.974%**
	(0.456)	(0.27)	(0.143)

"Short run" means that teacher productivity IS fixed as class size changes.
Standard errors are in parenthesis below.
** These results are significant at a 95% confidence level.

Table 15:
Counterfactual 5: Variations in Class Size - Long Run

Class size	Effect of Class Size Variation with respect to Class Size=40 5	10	20
Mean Math score - Percentual Change of SD			
For all students	-22.267%SD**	-19.522%SD**	-13.064%SD**
	(1.122)	(0.979)	(0.655)
For 5 worst students	-22.408%SD**	-18.982%SD**	-12.502%SD**
	(1.266)	(1.049)	(0.678)
For 5 best students	-23.11%SD**	-20.622%SD**	-14.048%SD**
	(1.131)	(0.993)	(0.664)
Mean Cost of Effort - Percentual Change			
For all students	0%	0%	0%
	(-)	(-)	(-)
For 5 worst students	0.039%**	0.017%**	0.016%**
	(0.002)	(0.001)	(0.001)
For 5 best students	0.022%**	0.014%**	-0.013%**
	(0.005)	(0.001)	(0.001)
Mean Effort - Percentual Change			
For all students	1.417%**	0.713%**	0.355%**
	(0.388)	(0.253)	(0.144)
For 5 worst students	0.504%	0.335%	0.224%
	(0.497)	(0.306)	(0.163)
For 5 best students	0.537%	0.31%	0.214%
	(0.461)	(0.282)	(0.156)

"Long run" means that teacher productivity is NOT fixed as class size changes.
Standard errors are in parenthesis below.
** These results are significant at a 95% confidence level.

Table 16:
Counterfactual 5: Replicating the Lazear 2001 Conclusion

	Mean Math Test Score - Short Run		
	All students	Worst 5 students	Best 5 students
Class size = 5	255.936	243.049	270.271
Class size = 10	256.901**	241.817**	271.709**
Class size = 15	257.529**	240.701**	273.592**
Class size = 20	257.807**	240.035	275.801
Class size = 25	257.682**	239.967	277.804**
Class size = 30	257.097**	240.233	278.952
Class size = 35	256.067**	240.294	279.849

	Mean Math Test Score - Long Run		
	All students	Worst 5 students	Best 5 students
Class size = 5	261.003	236.707	277.317
Class size = 10	259.208**	237.629**	276.409**
Class size = 15	258.131**	238.644	275.918
Class size = 20	257.773**	240.083	275.711
Class size = 25	258.091**	242.123**	275.157
Class size = 30	259.188**	244.502**	273.958
Class size = 35	260.981**	246.667**	272.375**

"Short run" means that teacher productivity IS fixed as class size changes.
"Long run" means that teacher productivity is NOT fixed as class size changes.
** These means are statistically different from the mean
for the previous class size at a 95% confidence level.

Table 17:
Counterfactual 6.1: Effect of Skimming in a District with Magnet School (MS)

	Small District: Lebu		Large District: Puente Alto	
Number of Schools	6		104	
Students with Data	62		4,063	

School Type	Restructured	New	Restructured	New
Mean Math score - Percentual Change of SD				
Old students in MS	-5.39%SD		0.96%SD	
	(1076.2)		(191.5)	
New students in MS	15.37%SD	12.84%SD	0.05%SD	1.32%SD
	(504.2)	(140.2)	(234.6)	(64.3)
Students left behind	1.62%SD	1.62%SD	-1.67%SD	-1.67%SD
	(27.8)	(27.8)	(6.6)	(6.6)
Cost of effort				
Old students in MS	3.251%		-1.047%	
	(648.9)		(209.7)	
New students in MS	3.821%	6.834%	-3.357%	-4.345%
	(593.4)	(80.6)	(185.5)	(38.4)
Students left behind	-1.775%	-1.775%	0.311%	0.311%
	(21.1)	(21.1)	(2.7)	(2.7)
Mean effort				
Old students in MS	-2.787%		0.729%	
	(557.1)		(145.6)	
New students in MS	8.346%	8.322%	1.157%	2.16%
	(55.4)	(55.9)	(190.8)	(32.7)
Students left behind	1.432%	1.432%	-0.149%	-0.149%
	(13.4)	(13.4)	(2.1)	(2.1)

Standard errors are in parenthesis below.

Table 18:
Counterfactual 6.2: Implementation of Magnet Schools Policy
Students are random conditional on observables

		Score 2009		School Product. 2009		Score 2011		School Product. 2011	
Region	District	In Sample	Predicted	In Sample	Predicted	In Sample	Predicted	In Sample	Predicted
XV	Arica	255.88	256.6	1.037	1.035	252.85	250.28	1.047	1.043
VI	Rengo					281.92	246.32	1.062	1.15
RM	Puente Alto					332.07	251.2	1.077	1.286
RM	Santiago								
IX	Purén	221.72	242.56	1.024	0.955	292.49	282.67	1.054	1.178
V	San Antonio					327.42	253.19	1.048	1.262
IX	Villarrica					332.21	228.9	1.031	1.277
V	Villa Alemana	255.14	252.85	1.052	1.06	250.56	252.69	1.047	1.053
VIII	Lota								
VII	Linares	321.89	246.3	1.018	1.255				
XII	Punta Arenas					223.99	246.58	1.048	0.957
II	Calama					284.39	277.33	1.054	1.14
RM	Renca					264.42	239.54	1.055	1.118
V	San Felipe	256.6	254.76	1.044	1.05	264.19	251.78	1.05	1.09
RM	Lo Barnechea	274.63	241.46	1.036	1.146				
X	Puerto Montt	299.59	273.91	1.043	1.118	299.77	309.54	1.053	1.118
XI	Coyhaique					287.54	238.98	1.041	1.197
II	Antofagasta	245.88	264.08	1.061	1.004	257.82	233.39	1.22	1.086
I	Iquique	244.51	249.7	1.022	1.005	235.97	234.78	1.049	1.009
X	Ancud	235.36	244.64	1.039	1.008	212.58	230.46	1.061	0.946
XI	Aisén	277.86	265.73	1.014	1.05	293.08	267.85	1.029	1.109
VIII	Concepción	287.49	262.61	1.068	1.147	280.86	326.56	1.059	1.116
RM	Conchalí	253.91	257.66	1.044	1.032	237.8	250.12	1.055	1.009
V	Los Andes	285.08	260.61	1.066	1.144				
V	Viña del Mar					247.26	284.1	1.057	1.026

Table 19:
Counterfactual 6.2: Implementation of Magnet Schools Policy
Students belong to the top 10% conditional on observables

		Score 2009		School Product. 2009		Score 2011		School Product. 2011	
Region	District	In Sample	Predicted	In Sample	Predicted	In Sample	Predicted	In Sample	Predicted
XV	Arica	255.88	334.64	1.037	0.844	252.85	336.08	1.047	0.838
VI	Rengo					281.92	345.01	1.062	0.907
RM	Puente Alto					332.07	365.19	1.077	1.002
RM	Santiago								
IX	Purén	221.72	306.88	1.024	0.796	292.49	341.59	1.054	0.935
V	San Antonio					327.42	335.94	1.048	1.027
IX	Villarrica					332.21	316.56	1.031	1.072
V	Villa Alemana	255.14	338.5	1.052	0.842	250.56	330.3	1.047	0.843
VIII	Lota								
VII	Linares	321.89	297.97	1.018	1.081				
XII	Punta Arenas					223.99	333.35	1.048	0.767
II	Calama					284.39	342.14	1.054	0.912
RM	Renca					264.42	333.95	1.055	0.878
V	San Felipe	256.6	335.12	1.044	0.847	264.19	333.74	1.05	0.872
RM	Lo Barnechea	274.63	316.46	1.036	0.929				
X	Puerto Montt	299.59	365.32	1.043	0.896	299.77	372.03	1.053	0.888
XI	Coyhaique					287.54	321.69	1.041	0.952
II	Antofagasta	245.88	355.37	1.061	0.795	257.82	432.4	1.22	0.805
I	Iquique	244.51	306.91	1.022	0.857	235.97	325.07	1.049	0.814
X	Ancud	235.36	320.86	1.039	0.817	212.58	331.35	1.061	0.746
XI	Aisén	277.86	337.67	1.014	0.873	293.08	349.02	1.029	0.899
VIII	Concepción	287.49	357.29	1.068	0.9	280.86	350.25	1.059	0.89
RM	Conchalí	253.91	335.96	1.044	0.838	237.8	336.56	1.055	0.804
V	Los Andes	285.08	354.03	1.066	0.9				
V	Viña del Mar					247.26	343.3	1.057	0.815