

Accounting for Peer Effects in Treatment Response

Rokhaya Dieye Habiba Djebbari Felipe Barrera-Osorio*

July 2014

Abstract

When one's treatment status affects the outcomes of others, experimental data are not sufficient to identify a treatment causal impact. In order to account for peer effects in program response, we use a social network model. We estimate and validate the model on experimental data collected for the evaluation of a scholarship program in Colombia. By design, randomization is at the student-level. Friendship data reveals that treated and untreated students interact together. Besides providing evidence of peer effects in schooling, we find that ignoring peer effects would have led us to overstate the program actual impact.

Keywords: education, social network, impact evaluation.

JEL codes: C31, C93, I22.

***Dieye:** Department of Economics, Université Laval; **Djebbari** (corresponding author): Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS, habiba.djebbari@univ-amu.fr and Department of Economics, Université Laval; **Barrera-Osorio:** Harvard Graduate School of Education, Harvard University. **Acknowledgments:** We thank Yann Bramoullé, Margherita Comola, Bernard Fortin, Nicholas Lawson and seminar participants at DIAL-Université Dauphine, to the GREQAM-AMSE workshop in social and economic networks, to the Statistics and Econometrics of Networks conference in Paris, to the Journées du CIRPEE in Quebec, and to the Société Canadienne de Science Economique in Canada, for many helpful comments. We gratefully acknowledge financial support from le Fonds Québécois de Recherche sur la Société et la Culture and the Social Sciences and Humanities Research Council.

Welfare programs may impose positive or negative effects on individuals who are not directly targeted by them but who interact with beneficiaries (Kremer and Miguel 2007, Oster and Thornton 2011, Bhattacharya et al. 2013). This complicates the evaluation of targeted interventions (Manski 2013a, Manski 2013b). The goal of this paper is to build a framework to evaluate the impact of a program, accounting for possible peer effects. We illustrate the relevance of our framework by assessing an education program targeted to the poor in Colombia.

Peer effects are important in many contexts. Peers share resources, information and learn from each other. Behaviour such as crime, alcohol and drug consumption, failure at school may be reinforced through social interactions. Taking the influence of peers into account is therefore important to evaluate interventions designed to help people escape poverty traps (Durlauf, 2006). Providing an accurate assessment of peer effects is, however, difficult (Manski, 1993). We first need to identify the appropriate peer groups. In the context of schooling decisions, they may include neighbours, schoolmates, classmates, and friends. Secondly, schooling outcomes are often highly correlated among peers. We can expect, for instance, children from similar background to interact together in the same schools. The best schools may also be able to attract the best teachers and unobserved teacher quality may explain average school performance. Besides endogenous sorting and common unobservables, identification is also hindered by simultaneity, as own schooling outcome depends on peers' outcomes, which itself depend on own outcome, a problem also known as the "reflection problem" (Manski, 1993).

In this paper, we provide a bridge between the reduced-form approach that exploits the random assignment of a treatment and the structural approach to estimate peer influence. Our work offers four main contributions. To our knowledge, we are the first to propose a structural model to account for social network effects in treatment response, to discuss identification of the model and to show how to use it to define parameters of interest for policy evaluation. A second contribution of our work is to propose a novel way of using randomized experiments to validate a model. Our third contribution is to show that ignoring peer effects may undermine our assessment of a program's impact. Finally, we show how we can use a model of peer effects to evaluate the overall impact of an intervention. We discuss parameters of interest for policy evaluation. We apply our framework to data from an experiment that randomly assigned school grants to children. To do so, we extend the model and results to account for missing data on nonparticipants.

A small and recent literature on treatment effects exploits exogenous variation in the local intensity of treatment to study social and neighborhood influences in partial population experiments (Kremer and Miguel 2007, Oster and Thornton 2011, Bhattacharya et al. 2013). However, this literature is silent on the nature of the spillover effects that are identified. A model is useful to dissect the process of contamination when treated and untreated student interact. For spillover

effects to materialize, the program must first have a *direct* effect on its recipients. Complementarities in time spent in social activities outside school may then lead children to attend school more when their peers do. This so-called *endogenous* peer effect, translates into a social multiplier, through which the direct effect of the program may be magnified. Finally, in addition to being influenced by peer attendance, attendance to school may change as more peers get treated. This *contextual* effect of the treatment may be positive or negative depending on the mechanism at play. We show that the sign of the error from ignoring peer influence depends on the strength and direction of the direct program impact, resulting from the initial response to changes in incentives, and of indirect peer effects.¹

The structural approach takes peer effects as the primary object of interest. A social network model of peer effects has proved successful on a number of fronts (see e.g., [Bramoullé et al. 2009](#), [Laschever 2009](#), [Lee et al. 2010](#), [Lin 2010](#), [Calvo-Armengol et al. 2009](#)). First, it can convincingly address the problem of identifying the appropriate peer groups. Second, better data on links between agents helps for the identification problems. In particular, a social network approach does not impose that everyone within a group interacts with everyone else and with no one outside, as in a group structure. It recognizes that some links are not present. In a linear-in-means model,² these natural restrictions prove sufficient to address the simultaneity problem in the absence of other sources of spurious correlation ([Bramoullé et al. 2009](#), thereafter BDF 2009).

The social network approach, however, has two important shortcomings. First, although it offers a solution to the reflection problem, it falls short of providing a convincing answer to the problem of correlated unobservables. Most papers in this literature correct for correlated unobservables through network fixed effects. These capture factors that are common to all students in a network. But fixed effects may not correctly address the problem of endogenous sorting within the network. Identification typically fails if social links are not made at random *within* the network.³ Treatment randomization allows us to address the problem of correlated unobservables, without having to assume that they are fixed within network. Indeed, even though friendship links may be endogenously formed, the fraction of treated friends is exogenous because treatment assignment is random. Similarly, even if teachers are not randomly assigned to classrooms or schools, a student's treatment status and the treatment status of his friends are determined by chance.

A downside of relying on any structural model of peer effects is that the framework that de-

¹An intuitive reasoning may mislead us to think that, when treated and untreated individuals interact, the difference in mean outcomes between experimental groups underestimates the actual program impact.

²According to a linear-in-means model, a student's attendance to school may depend on his treatment, the average treatment among his peers, and his peers' average attendance.

³For instance, an unobserved factor may explain both heterogeneity in school outcomes and the structure of friendship within the classroom.

scribes how peer effects are generated may be invalid. For instance, misspecification may occur if the model assumes that all peers have the same impact when, in reality, “bad” peers have a negative effect on performance because of the disruption they produce and “good” peers have a positive or no effect. We exploit the experimental nature of the data to offer a novel way to validate our model.⁴ We show how the naive difference in expected outcome between the treated and untreated is related to the structural parameters of the model, i.e., direct program effect and indirect peer effects. Comparing the experimental to structural estimates allow us to test the social network model.

Typically in randomized experiments, data only covers program participants, i.e., treated and untreated students. Yet a student may well name as a friend a nonparticipant fellow. In a social network model, nonparticipant friends may well matter, for instance, if they transmit program effects from treated to untreated participants. We thus extend our social network model to allow for missing data on nonparticipants.

In our application, we focus on schooling decisions of poor Colombians in Bogota. Poverty affects about a third of the 8.5 million population of the city. Bogota, as many other large cities in Latin America, is highly segregated, with pockets of poverty in certain neighborhoods. Although relatively high for a developing country, average enrolment also hides a lot of heterogeneity, especially at the secondary level. In this context, peer effects may be expected to reinforce educational disparities, trapping young people in a bad equilibrium. In the randomized experiment described in [Barrera-Osorio et al. \(2011\)](#), cash transfers are allocated among poor children using a lottery. Random assignment is done at the individual student-level in two poor neighborhoods of the capital city. The program is still active and has now expanded to cover about 130,000 students.

We focus on the impact on school attendance. Regular school attendance is one condition to continue receiving the subsidy. Data on attendance is collected over a period of 13 weeks during which research staff visited schools unannounced, calling the roll and marking absences. Data are available on school participation of all the children enrolled in all intervention schools, as well as on their friendship network. Children are requested to name their five best friends. We use (preprogram) data on self-reported friendship links collected at the start of the schooling year (attendance data is collected 5 months later). We thus know the treatment status of each child (i.e., whether he was picked up as a lottery recipient), his schooling outcome, the proportion of his friends who are treated, and their schooling outcomes. This makes these data suitable to apply our framework.

We find that the direct program impact and the endogenous peer effect are positive and the contextual effect is negative. The endogenous effect does translate into a social multiplier for

⁴Randomized experiments have been successfully used to validate econometric models ([LaLonde 1986](#), [Todd and Wolpin 2006](#)).

school attendance, which magnifies the direct program impact. But, increases in the fraction of treated peers results into a drop in attendance. Indirect evidence suggest that this effect is consistent with substitution between own and peers labor supply. Although direct impact and indirect peer effects are all statistically significant at 1 percent, the resulting net effect of the program is close to zero and not significant. Ignoring peer effects thus leads to overestimating the actual effect of the program. Our empirical findings are robust to a number of specification checks. Based on our model validation strategy, we find that the model is well specified.

The rest of the paper is organized as follows. In section 1, we present details of the program and describe the experiment. In section 2, we lay out our econometric framework. Data are presented in section 3. Empirical results are discussed in section 4, where we also present some specification tests. We conclude in section 5.

1 The program and the experiment

1.1 *Subsidios Condicionados a la Asistencia Escolar*

Subsidios Condicionados a la Asistencia Escolar (cash transfers conditional on school attendance) is a local program created to reduce dropout rates in Bogota. Cash transfer programs are a popular tool for reducing current poverty and improving human capital investments. Two influential programs (*PROGRESA-Oportunidades* in Mexico and *Bolsa-Familia* in Brazil) serve as a model replicated in various countries (see, e.g., Fiszbein et al., 2009 for a review). Colombia also launched in 2002 a cash transfer program aimed at communities of less than 100,000 inhabitants (*Familias en accion*). In contrast, *Subsidios Condicionados a la Asistencia Escolar* is piloted over two localities of the district of Bogota, a municipality comprised of 20 localities. The rationale is to reach a population who is not eligible to *Familias en accion*. Indeed, most of the population lives in large cities and thus do not qualify for *Familias en accion*. Furthermore, this pilot allows experimenting with various ways of delivering the cash transfers (every two months, partly every two months and partly at the beginning of the next school year, partly every two months and upon graduation).⁵

The program distributes cash transfers to poor families conditional on having children attend school on a regular basis (*i.e.*, at least 80 percent of the time). Its objective is to lower dropout and help limit child labor.⁶ The transfers amount to 30,000 pesos or \$15 per month. They represent about 8 percent of the median income of recipient households. They cover the cost of sending a child to school, including enrollment fees, uniforms and school material. The program was

⁵In our analysis, we do not distinguish between all different ways of delivering cash benefits.

⁶Based on the Colombian Child Labor survey of 2001, about 10 percent of 5-14 years old children are working. Boys are more likely to participate than girls (14.1% of boys *vs.* 6.6% of girls). More than 75 percent of working children combine work (paid or unpaid) and study.

launched in 2005 by the Secretary of Education of the City over two localities of the district of Bogota (Suba and San Cristobal). Children eligibility depends on their parents' income. A national registry classifies households according to their wealth to identify those eligible for social programs. Only the bottom two categories are eligible to these cash transfers. There is also a residency requirement (children had to reside in the localities prior to 2004). The program targets secondary-school children from grade 6 to 11 (age 10 to 17 on average). The intervention lasted for three years.⁷ The pilot study covers the first year, from February 2005 to February 2006. The program is still active, and targeting more than 130,000 students.

1.2 The experiment

1.2.1 Timeline

Subsidios Condicionados a la Asistencia Escolar was advertised at the start of the academic year in January-February 2005 (see timeline in Figure 1). Registration to the program in Suba and San Cristobal was open during 15 days at the end of February-beginning of March 2005. In San Cristobal, the pilot program targets children who completed grade 5 (*i.e.*, who are entering basic secondary school). Suba targets children who completed grade 8 (those in grade 9 to 11). There was oversubscription at the registration stage. A public lottery, selecting beneficiaries among the applicants, was held in each locality on April 4th 2005. This means that two separate randomized experiments were held, one for each locality. By design, randomization is done at the individual level.

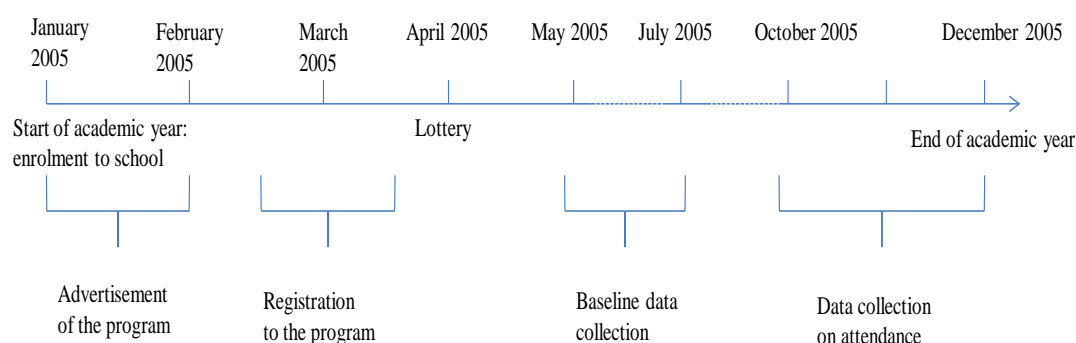


Figure 1: Timeline

Data on school attendance was collected at the end of the school year, about 5 months after the baseline survey that collects data on friendship links.

⁷An academic year runs from end of January to mid-November.

1.2.2 Treatments

Two distinct treatments were offered in San Cristobal. One consists in providing 30,000 pesos per month (equivalent to US\$15), a payment made every two months. The other consists in providing $2/3$ of the annual benefit on a two-month basis and $1/3$ at enrolment at the next academic year. The two treatment groups are about the same size. Children had to attend school at least 80 percent of time each month. The funds are made available in a bank account and students could lose their benefits if one of the three following situations occurred: failure to matriculate to the next grade twice, failure to reach the attendance target in two successive payment periods, or exclusion from school. In Suba, all beneficiary children received $2/3$ of the transfer every two months as well as a lump-sum transfer representing $1/3$ of the amount conditional on graduating from high school. We make use of the experiments separately although we do not distinguish in the case of San Cristobal between the two treatments.

2 Econometric model

We now present our econometric framework. It is motivated by our empirical application on the effect of *Subsidios Condicionados a la Asistencia Escolar* on school attendance. In section 2.1, we introduce a linear model of peer effects in which we allow for student specific peer groups. A student regular attendance to school may be affected by his treatment, his peers' treatments and his peers' attendance. We study identification and show how it depends on random assignment of treatment and the geometric properties of social network structure. In section 2.2, we extend the model to show how to account for missing data on students who did not apply to the program (nonparticipants).⁸ In section 2.3, we discuss how to evaluate a program's impact in the presence of peer effects. We define and discuss direct and total impact. We also show that the naive difference in expected outcomes between treated and untreated participants is generally not equal to the (direct or total) program impact. We relate all the impact parameters to the structural parameters of the model of peer effects.

2.1 A model of treatment and peer effects

Let y_{is} be individual i 's outcome in school s , ($i = 1, \dots, N$) and ($s = 1, \dots, S$) where N is the number of applicants to the program and S the number of participating schools. Let \mathbf{x}_{is} be $(1 \times K)$ vector of characteristics for individual i . For simplicity, we present the model with one characteristic ($K = 1$), but the results hold for any K . Let t_{is} take value 1 if student i in school s is randomly assigned to the treatment group, 0 otherwise. Each individual i has a specific reference group N_i

⁸Non-participants students are either non eligible or eligible but not registered. As explained in the previous section, to be eligible to the program, a student must belong to a poor family and apply to the program.

of size n_i . This group should include all individuals whose outcomes or background characteristics affect i 's outcome.⁹ We develop our analysis in two stages. In this section, we assume that everyone in the participating schools applies to the program. However, in our sample, we only observe information on students who registered to the lottery, *i.e.*, participating students. Data on nonparticipant is missing. In the next section, we extend the analysis to account for missing data on nonparticipants.

We adopt the standard model of peer effects studied, for instance, in BDF (2009) and De Giorgi et al. (2010). We assume that a student's outcome may be affected by his own attributes, by his friends' mean attribute and by their mean outcome:

$$y_{is} = \alpha + \beta(\mathbf{G}\mathbf{y})_{is} + \gamma x_{is} + \gamma_1 t_{is} + \delta(\mathbf{G}\mathbf{x})_{is} + \delta_1(\mathbf{G}\mathbf{t})_{is} + \epsilon_{is}, \quad (1)$$

where \mathbf{G} is a $N \times N$ interaction matrix, with element $G_{ij} = 1/n_i$ if j is a friend of i , and 0 otherwise. For any $(N \times 1)$ vector \mathbf{z} , $(\mathbf{G}\mathbf{z})_{is}$ is the average of z among i 's friends in school s : $(\mathbf{G}\mathbf{z})_{is} = \frac{\sum_{j \in N_i} z_j}{n_i}$. The error term ϵ_{is} reflects unobservables characteristics of student i in school s .

This model captures three types of effects. First, γ_1 and γ simply represent the effects of own treatment t_i and own attribute x_i on outcome y_i . Second, δ_1 and δ capture the impacts of the proportion of treated friends $(\mathbf{G}\mathbf{t})_{is}$ on an individual outcome. In the peer effect literature, they are usually referred to as *contextual* peer effects. Third, β is the endogenous peer effect; that is, the impact of friends' mean outcome $(\mathbf{G}\mathbf{y})_{is}$ on i 's outcome y_{is} .

Since we are mainly interested in the direct and social effects of the program, *i.e.*, in parameters $(\beta, \gamma_1, \delta_1)$, we can further reduce the model to:

$$y_{is} = \alpha + \beta(\mathbf{G}\mathbf{y})_{is} + \gamma_1 t_{is} + \delta_1(\mathbf{G}\mathbf{t})_{is} + u_{is}, \quad (2)$$

where $u_{is} = \epsilon_{is} - \gamma x_{is} - \delta(\mathbf{G}\mathbf{x})_{is}$.

We now discuss identification of model (2). First, observe that randomization balances unobservable effects in the two experimental groups. Randomization also insures that individual attributes, x_{is} , and friends attributes, $(\mathbf{G}\mathbf{x})_{is}$, are balanced between the two experimental groups, implying that $\mathbb{E}[x_{is} | t_{is}] = 0$ and $\mathbb{E}[(\mathbf{G}\mathbf{x})_{is} | t_{is}] = 0$. Treatment randomization is not sufficient, however, to guarantee identification. Observe that model (2) defines a system of simultaneous linear equations. This system is potentially subject of the reflection problem (Manski, 1993), as y_i depends on y_j , which itself may depend on y_j . In matrix notation, model (2) becomes:

$$\mathbf{y} = \alpha \mathbf{1} + \beta \mathbf{G}\mathbf{y} + \gamma_1 \mathbf{t} + \delta_1 \mathbf{G}\mathbf{t} + \mathbf{u}.$$

⁹Individual i is excluded from his reference group, that is, $i \notin N_i$.

Assuming $|\beta| < 1$, matrix $\mathbf{I} - \beta\mathbf{G}$ is invertible and the associated reduced-form model for this system of simultaneous equations is:

$$\mathbf{y} = \alpha(\mathbf{I} - \beta\mathbf{G})^{-1}\mathbf{1} + (\mathbf{I} - \beta\mathbf{G})^{-1}(\gamma_1\mathbf{I} + \delta_1\mathbf{G})\mathbf{t} + (\mathbf{I} - \beta\mathbf{G})^{-1}\mathbf{u}. \quad (3)$$

Expanding the inverse into a series and assuming that no individual is isolated, we obtain:

$$y_{is} = \frac{\alpha}{1 - \beta} + \gamma_1 t_{is} + (\gamma_1\beta + \delta_1)\sum_{k=0}^{\infty}(\beta)^k(\mathbf{G}^{k+1}\mathbf{t})_{is} + v_{is}, \quad (4)$$

where $v_{is} = \sum_{k=0}^{\infty}(\beta)^k(\mathbf{G}^k\mathbf{u})_{is}$.

Thanks to treatment randomization, reduced-form parameters, *i.e.*, the effect of own treatment and peers treatment, are identified from unobserved correlated effects that may arise from sorting and common shocks. Indeed, if friendship links are not affected by the treatment, *i.e.*, $\mathbb{E}[G_{ij} | \mathbf{t}] = \mathbb{E}[G_{ij}]$, then $\mathbb{E}[v_{is} | \mathbf{t}, \mathbf{G}] = 0$. For all k , $(\mathbf{G}^k\mathbf{u})$ is simply a $n \times 1$ vector of weighted averages of distance- k friends' attributes. As a result of treatment randomization, both t_{is} , own treatment, and $(\mathbf{G}^{k+1}\mathbf{t})_{is}$, peers treatment, are exogenous to $(\mathbf{G}^k\mathbf{u})_{is}$, and thus to v_{is} .

From the estimation of the reduced-form equation (4), we can identify the structural parameters in (2), *i.e.*, $(\alpha, \beta, \gamma_1, \delta_1)$, if \mathbf{I} , \mathbf{G} and \mathbf{G}^2 are linearly independent and $\gamma_1\beta + \delta_1 \neq 0$ (see BDF 2009 for a proof of this result). In particular, these three matrices are linearly independent in the presence of intransitivity, *i.e.*, if there exists three individuals i, j, k , such that i is friend to j and j is friend to k but i and k are not friends. Intransitivity has a natural interpretation in terms of instrumental variables: k 's treatment should only affect i 's behavior through its effect on j 's behavior. More generally, it is the heterogeneity in the structure of the network that helps break the simultaneity or reflection problem (Manski, 1993). In practice, it is easy to check that \mathbf{I} , \mathbf{G} and \mathbf{G}^2 are linearly independent with the data at hand. We now summarize these results in Proposition 1 below.

Proposition 1 *Assume that friendship is not affected by the treatment and that the friendship network has some intransitivity. Randomization at the individual level allows identification of direct program effect, indirect endogenous peer effect and indirect contextual peer effect.*

Compared to the identification results in BDF (2009), treatment randomization allows us to relax the assumption that correlated effects must be fixed at the level of the network, but requires us to assume that students do not re-sort on the basis of treatment status, *i.e.*, $\mathbb{E}[G_{ij} | \mathbf{t}] = \mathbb{E}[G_{ij}]$.¹⁰

¹⁰In contrast, Comola and Prina, 2013 propose a framework in which the network may change as a result of an exogenous intervention. They apply it to experimental data collected to assess the effect of a new saving technology in Nepal. They map the network of financial exchanges between all villagers before and after the intervention. A large part of the peer effect on expenditures comes from the recomposition of the network resulting from changes in expenditures. In our application, we do not account for this channel. Our estimated peer effects would be biased downward if, as a result of the subsidy program, children who attend school regularly are more likely to create links with other children who also

In BDF (2009), it is assumed that only unobserved effects that are fixed at the level of schools may cause spurious correlations. This assumption is too strong if, conditional on joining a school, links are not formed at random. In particular, if there is homophily based on a characteristic that is also correlated to the outcome of interest, then unobserved correlated effects due to sorting may still bias the estimates. In the result section, we compare estimates for model (2) with random effects (our preferred specification) and with fixed effects. We expect both estimators to be consistent if the network fixed effects are uncorrelated to the treatment.

2.2 Treatment of missing data

We now adapt and extend the model to account for missing data on nonparticipant friends in the friendship network. In our data, we only have information on program applicants; this is typical of data collected to assess the average effect of a program on participants. In the presence of peer effects, however, nonparticipants may matter. For instance, they may help transmit program impacts from treated to untreated participants. Yet, we do not observe the full network of interactions. In this section, we show how to correct for this incompleteness for the estimation of model (2).¹¹ Our correction makes use of the fact that, in our data, we know the number of nonparticipant friends for each participant.

Even in our extended model, however, we ignore external effects of the program on the group of nonapplicants. One argument for focussing on the set of applicants is that *it is* the population of interest to the policy-maker who is deciding whether to continue or extend the program. If the program does not pass the cost-effectiveness test when excluding benefits to noneligible people, the recommendation will likely be to stop funding it, even if there is evidence that nonapplicants benefit from it.¹²

Our treatment of missing data is based on two intuitive elements. First, randomness in the program selection process implies that we can expect treated and untreated applicants to have the same number of nonapplicant friends, and friends who are similar in terms of observed and unobserved characteristics.¹³ Another and more technical argument relates to the conditions for identification. Intuitively, identification relies on the fact that the friend of a friend may not be attend school regularly and drop links with those who do not. This change in friendship structure should have taken place within the same school year, in a 5-month interval. In the absence of longitudinal data on friendship links, we cannot test this hypothesis.

¹¹Chandrasekhar and Lewis (2011) also study how missing data may affect estimation of a social network model in the special case where nodes are selected randomly, *i.e.*, where data on social contacts are only collected for a random subset of individuals.

¹²Ignoring external effects to non-participants may not always be justified, see Miguel and Kremer (2004). Yet, it is very plausible that, in our context, policy-makers do not care much about external effects to children from well-off families.

¹³We test the first statement using the data at hand, but cannot test whether characteristics of nonapplicant friends are similar, since we do not have any information of nonapplicants.

a direct friend. Not observing nonparticipants implies that we will not observe some of these intransitive triads. But all the intransitive triads that we observe between applicants *are* indeed intransitive. As long as we observe enough of these intransitive triads, we should be able to identify the structural parameters of the model. More formally, in our setting, we know the total number of friends for each participant ($i = 1, \dots, N$) from school ($s = 1, \dots, S$), n_{is} , and the number of friends among participants, m_{is} . We can thus decompose the proportion of treated among friends for individual i as follows:

$$\frac{1}{n_{is}} \sum_{j=1}^{n_{is}} t_{js} = \frac{1}{n_{is}} \sum_{j=1}^{m_{is}} t_{js} = \frac{m_{is}}{n_{is}} \left(\frac{1}{m_{is}} \sum_{j=1}^{m_{is}} t_{js} \right) = \frac{m_{is}}{n_{is}} (\mathbf{G}^1 \mathbf{t})_{is},$$

where \mathbf{G}^1 is a square matrix of size N defined such that, for all (i, j) in the set of participants, $g_{ij}^1 = \frac{1}{m_{is}}$ if i and j are friends, 0 otherwise. The first equality is implied by the fact that none of the nonparticipants are treated ($D_{is} = 0 \Rightarrow t_{is} = 0$). Consequently, the treatment vector is only defined for participants (size $N \times 1$). Similarly, consider the following expression for the average outcome among friends for individual i :

$$\frac{1}{n_{is}} \sum_{j=1}^{n_{is}} y_{js} = \frac{1}{n_{is}} \left(\sum_{j=1}^{m_{is}} y_{js} + \sum_{j=m_{is}+1}^{n_{is}} y_{js} \right) = \frac{m_{is}}{n_{is}} \left(\frac{1}{m_{is}} \sum_{j=1}^{m_{is}} y_{js} \right) + \frac{n_{is} - m_{is}}{n_{is}} \left(\frac{1}{n_{is} - m_{is}} \sum_{j=m_{is}+1}^{n_{is}} y_{js} \right).$$

It becomes:

$$\frac{1}{n_{is}} \sum_{j=1}^{n_{is}} y_{js} = \frac{m_{is}}{n_{is}} (\mathbf{G}^1 \mathbf{y})_{is} + \frac{n_{is} - m_{is}}{n_{is}} (\mathbf{G}^0 \mathbf{y}^0)_{is},$$

where, \mathbf{G}^1 is defined as previously. The observed \mathbf{y} is of size N , and the unobserved \mathbf{y}^0 is of size J , where J is the total number of nonparticipants who are friends to participants. The $N \times J$ matrix \mathbf{G}^0 represents social interactions between each participant $i \in [1, N]$ and his nonparticipants friends. Observe that $\frac{n_{is} - m_{is}}{n_{is}} (\mathbf{G}^0 \mathbf{y}^0)_{is}$ is unobserved because missing.

We can now transform model (2) to account for missing data on nonparticipants:

$$y_{is} = \alpha + \beta \frac{m_{is}}{n_{is}} (\mathbf{G}^1 \mathbf{y})_{is} + \gamma_1 t_{is} + \delta_1 \frac{m_{is}}{n_{is}} (\mathbf{G}^1 \mathbf{t})_{is} + u_{is}. \quad (5)$$

Compared to model (2) in the previous section, we now weight the observed network interaction matrix to account for missing data on nonparticipants using the proportion of participants among friends, $\frac{m_{is}}{n_{is}}$. In addition, the error term u_{is} now includes as an additional term the unobservable $\frac{n_{is} - m_{is}}{n_{is}} (\mathbf{G}^0 \mathbf{y}^0)_{is}$. Importantly, the proportion of friends who are nonparticipants, $\frac{n_{is} - m_{is}}{n_{is}}$, is uncorrelated with treatment, and so is $(\mathbf{G}^0 \mathbf{y}^0)_{is}$. Consequently, as in the general setting, the reduced-form of this model is identified if students do not re-sort on the basis of treatment status. Similarly, structural parameters $(\alpha, \beta, \gamma_1, \delta_1)$ are identified if the friendship network between participants is intransitive. In Appendix A, we write model (5) in matrix notation.

2.3 Parameters of interest

How can program impacts be evaluated in the presence of peer effects? In the literature on treatment effects, most studies rely on the assumption that individual outcomes are invariant

with respect to other individuals' treatment assignment. This assumption, also known in the statistical literature as the Stable Unit Value Assumption (SUTVA), does not hold in the presence of social interactions between treated and untreated participants, however.¹⁴ In the presence of social interactions, the program may affect participants in two ways. There may be a direct effect due to changes in individuals' incentives and also an indirect effect caused by social interaction. In this section, we study program impacts in the presence of peer effects structured as in the general model (2).¹⁵

We now define the direct and total effects of the program, and relate them to the structural parameters of model (2). The average total effect of the treatment in the population, Δ^{tot} , which corresponds to the intervention scaled up to include all participants, is defined as follows:

$$\Delta^{tot} = \mathbb{E}(y|\mathbf{t} = \mathbf{1}) - \mathbb{E}(y|\mathbf{t} = \mathbf{0}). \quad (6)$$

The first term is the expected outcome when all applicants are treated, and the second one is the expected outcome when no one receives the program. Intuitively, the larger the fraction treated, the larger the indirect effects, so that a scaled-up program treating all participants generates the highest indirect effect in absolute term. Thus, the total effect of pilot program for which coverage is only partial is different from the total effect for a program offering benefits to all eligible individuals (Philipson 2000). This is why we focus on the latter, as defined in (6).

From reduced-form (3), observe that:

$$\begin{aligned} \mathbb{E}(y|\mathbf{t} = \mathbf{1}) &= \alpha(\mathbf{I} - \beta\mathbf{G})^{-1}\mathbf{1} + (\mathbf{I} - \beta\mathbf{G})^{-1}(\gamma_1\mathbf{I} + \delta_1\mathbf{G})\mathbf{1}, \\ \mathbb{E}(y|\mathbf{t} = \mathbf{0}) &= \alpha(\mathbf{I} - \beta\mathbf{G})^{-1}\mathbf{1}. \end{aligned}$$

If there is no isolated individual, this yields:¹⁶

$$\mathbb{E}(y|\mathbf{t} = \mathbf{1}) - \mathbb{E}(y|\mathbf{t} = \mathbf{0}) = \frac{\gamma_1 + \delta_1}{1 - \beta}\mathbf{1}, \quad (7)$$

and hence

$$\Delta^{tot} = \frac{\gamma_1 + \delta_1}{1 - \beta}.$$

This is the *composite* social effect in Manski's seminal paper (1993). If there are no endogenous and no contextual effects, $\delta_1 = \beta = 0$, or no interactions between agents, then Δ^{tot} reduces to γ_1 , and randomization of treatment is sufficient to identify the average total effect of the treatment in the population, which is equal to γ_1 in model (5). If $\gamma_1 > 0$ (a positive direct effect of the treatment), $\delta_1 > 0$ (a positive contextual effect of the treatment) and $0 < \beta < 1$ (a positive, but

¹⁴In the following, we do not attempt to provide the full specification of potential outcomes. Rather, and to be precise, we call untreated individuals ($t_i = 0$) those who have not received the treatment, although they may be affected by the treatment status of their peers. Treated individuals can be directly and indirectly affected by the treatment.

¹⁵The results also hold for the special case model (5) with missing data on nonparticipants.

¹⁶Replace matrix \mathbf{G} by matrix \mathbf{H} as defined in Appendix A to obtain a similar result for model (5).

not too large, endogenous effect), then the direct effect of the program on the participants, γ_1 , underestimates the total effect of the program, Δ^{tot} . If $\gamma_1 > 0$, $\delta_1 < 0$ and the contextual effect is larger in magnitude than the direct effect, ignoring peer effects may lead to overestimating the effect of the program. In our empirical analysis below, we provide estimates for the direct effect, γ_1 , and the total program effect, Δ^{tot} .

For our last result in this section, let us look at the difference in expected outcomes between treated and untreated participants. This is a natural benchmark parameter: in the absence of peer effects, this difference identifies the average effect of the treatment on the treated. This does not hold in the presence of peer effects, however. In the following, we express this difference as a function of the structural parameters and of the network:¹⁷

Proposition 2

$$\mathbb{E}[y_{is} | t_{is} = 1] - \mathbb{E}[y_{is} | t_{is} = 0] = \gamma_1 + (\gamma_1\beta + \delta_1) \frac{1}{N} \text{Tr}((\mathbf{G}(\mathbf{I} - \beta\mathbf{G})^{-1})) \quad (8)$$

Randomization is key for this result to hold. The proof is derived in Appendix B and relies on three arguments: (1) the expected fraction of treated friends at distance k (excluding oneself) is the same for treated and untreated individuals, (2) the unobserved effects are also balanced between the treated and the untreated, (3) treated units may experience an effect of own treatment that feeds back from interactions with others.

The difference in expected outcomes between treated and untreated participants is the sum of two terms. The first term, γ_1 , is simply the direct effect due to changes in students' incentives resulting from the treatment. This is the program impact when there are no peer effects ($\beta = \delta_1 = 0$ or no interactions between students). The second term captures how the direct treatment effect gets modified through interactions with others. The i^{th} diagonal element of matrix $\mathbf{G}(\mathbf{I} - \beta\mathbf{G})^{-1}$ is equal to a weighted sum of the number of cycles from i to i in the graph.¹⁸ When i is treated, this affects his direct friends both through contextual effects (δ_1) and endogenous effects ($\gamma_1\beta$). In turn, they affect friends at distance 2, 3, and so on, and all these indirect effects eventually find their way back to i through cycles in the graph. Thus, the difference in expected outcomes between treated and untreated participants accounts for both the direct effect from own treatment and the indirect effects back to one-self. We label this effect the individual mirror effect:

$$\Delta^{mirror} = \gamma_1 + (\gamma_1\beta + \delta_1) \frac{1}{N} \text{Tr}((\mathbf{G}(\mathbf{I} - \beta\mathbf{G})^{-1})).$$

In general, it is not equal to the direct (γ_1) or total program impact ($\frac{\gamma_1 + \delta_1}{1 - \beta}$).

¹⁷Again, simply replace matrix \mathbf{G} by matrix \mathbf{H} as defined in Appendix A to obtain a similar result for model (5).

¹⁸A cycle starting at i is a set of individuals j_1, \dots, j_l such that i is friend with j_1 , j_1 is friend with j_2, \dots , and j_l is friend with i .

An interesting implication of (8) is that it can be used to validate the model. Indeed, if model (2) is well-specified, then we should not reject the null that $\mathbb{E}[y_{is} | t_{is} = 1] - \mathbb{E}[y_{is} | t_{is} = 0] = \gamma_1 + (\gamma_1\beta + \delta_1)\frac{1}{N}\text{Tr}(\mathbf{G}(\mathbf{I} - \beta\mathbf{G})^{-1})$. We implement this test in the empirical section. We also provide evidence on the size of the error we would make by ignoring social interaction effects in our assessment of the program.

3 Data

3.1 Description of the data and sample

During two weeks at the end of February-beginning of March 2005, all eligible parents residing in Suba and San Cristobal were invited to register their children to the program. On April 4th 2005, a lottery selecting beneficiaries among applicants was held (one in each locality). In total, the lotteries were run over more than 17,000 applications, of which 10,000 received the subsidy and close to 7,000 did not. However, owing to budget limitations, only a subset of schools was included in the study sample.¹⁹ The study focuses on 68 of the 251 schools in Suba and San Cristobal with the largest number of applicants. Every applicant in the selected schools is included in the study sample. The resulting sample size is 6,886, including 1,146 pupils in Suba (527 treated and 619 untreated applicants) and 5,740 in San Cristobal (3,722 treated and 2,018 untreated applicants). The distribution of the sample by locality and treatment status is provided in Table 1.²⁰ Since the sample is larger for San Cristobal, our preferred estimates are based on this sample. However, we also present and discuss estimates for Suba.

Baseline data was collected in May-July 2005, a month after random assignment. The baseline survey instrument was administered to applicant children. The follow-up was collected at the beginning of the next academic year, in February-March 2006.²¹ For the follow-up, households to which applicant children belong were visited. More than 98 percent of the children surveyed in the first round were found.²² In addition, between baseline and follow-up data collections, teams of research assistants were sent to gather direct observations on attendance of children to schools

¹⁹Survey data was only collected for the study sample.

²⁰All applicants in the 68 schools are included in the attendance data, but only 91% of these are surveyed at baseline. Because we construct social networks based on information provided in the baseline survey, our working sample is restricted to those who responded to the survey (6,886 children out of 7,569).

²¹Baseline data include information on family structure and school related expenses, as well as children labor force participation, their education, aspirations, and friendship network. Follow-up data gathers the same type of information as the baseline, except for the friendship network. In addition, they include grades in Math, Sciences and Spanish.

²²This includes children who did not re-enrolled in the next academic year. However, we do not observe grades for these children. Since the program affects dropout, then “test-takers” in the treatment group and “test-takers” in the control group are not drawn from the same population, as in Angrist et al. (2006). Thus, we do not consider grades as an outcome variable for our model.

for a period of 13 weeks around the end of the 2005 school year. The visits were not announced in advance and the assistants called the roll of all students and marked absences.

For this study, we use data from the baseline and postprogram data on school attendance. The baseline survey includes a question on friendship relationships, worded as follows: *“Make a list of 5 best friends who study in your school (or classmates with whom you spend most of your time), including their full names and last names”*. Children could nominate friends who did not apply to the program, but the network is limited to the school and only five nominations are permitted. In practice, less than 1 percent of students friends are not in the same grade level, and less than 1 percent nominate 5 friends, so the framing of the question is not binding. In San Cristobal, there are 15.1 percent of student for whom we do not have information on any friend. There are two reasons for that. We find that 3 percent of them do not actually report having any friend and are not designated as friends by other. We consider them as socially isolated. The other reason is that the study is restricted to applicants. There is no information collected on nonapplicants. These pupils only report friendship links to nonapplicants and are not designated as friends by any pupil in the study sample. We consider them as socially isolated from the group of applicants.

Our main outcome of interest is school attendance. Regular school attendance is one condition to continue receiving the scholarship. Using the observed attendance data, we construct an aggregate that measures the percentage of time the student is found present in class during the school visits (*i.e.*, number of presences divided by number of visits). This measure of school attendance is considered to be more accurate than the self-reported measure from survey data ([Barrera-Osorio et al., 2011](#)).

3.2 Descriptive statistics

In Table 2, we present baseline descriptive statistics for pupils in the treatment and control groups in San Cristobal. We present baseline network descriptive statistics for children from San Cristobal in Table 3.²³ In Table 4 and 5, we report the same statistics for Suba.

In San Cristobal, children are 13.5 years of age on average. Male students represent 48 percent of the sample. Average family wealth index (official SISBEN score) is 11.6 over 100, meaning that the average family is in the first decile of the wealth distribution. Most student are enrolled in grades 6 to 8 (20 percent in each of these grades). About 16 percent are enrolled in each of grade 9 and 10, and only 8 percent in grade 11. Around 11 percent of the currently enrolled students had already dropped out of school in the past. For 27 percent of them, the main reason for dropping out is “lack of money”. Overall, 38 percent had already repeated a grade. At baseline, 22 percent of boys (17 percent of girls) work for pay in addition to studying. Those who work for pay work

²³See Appendix C for how we construct the social network matrix.

for an average of 7 to 8 hours a week. Children enrolled in grades 9 – 11 in Suba are older by 2 years, 38 percent of boys work and study (21 percent of girls) and they work an hour more on average. They are otherwise similar to pupils in San Cristobal (Table 4).

Looking at network characteristics for San Cristobal (Table 3), we find that children cite on average 4.37 friends, of which 2.58 are among applicants. About 12 percent are not friends to any other applicant, and 0.6 percent have no friends at all. Including those who are friendless, average characteristics of direct friends (distance 1) and indirect friends (distance 2 and more) are quite similar.

In Tables 2 and 3, we also present balancing tests on individual characteristics and characteristics of direct and indirect friends. We do not find any significant difference in average between experimental groups for age, gender, SISBEN poverty score, number of friends, attributes of direct friends and attributes of friends at distance 2 and 3, consistent with what we expect with randomization. As we also expect, the fraction of friends treated at distance 2 and at distance 4 is very different in the two experimental groups. This is because with nondirected networks, an individual is always friend with himself, except for isolated individuals. When we ignore cycles back to one-self, fraction of friends treated at distance 2 and 4 is similar in the two groups.²⁴ The results follow the same pattern for Suba.

Finally, in Table 6, we show attendance levels in both localities. There is a 87 percent attendance rate in San Cristobal, and 86 percent in Suba (higher grades). Note that average attendance is lower in the analysis carried out by Barrera-Osorio et al. (2011), respectively 80 percent and 78 percent. This difference comes from the fact that we restrict the sample to children who are selected for the baseline survey (for which there are data on friendship links). Our sample restriction does not alter the validity of the experiment (see Appendix tables 1 and 2).

4 Empirical results

We estimate equation (5) presented in section 2 using maximum likelihood (see Appendix D for the expression of the likelihood).²⁵ According to the model, attendance depends on own and peers treatment status, as well as peers attendance choices.

²⁴Results available upon request.

²⁵Alternatively, one can estimate the model by Generalized 2SLS. An advantage of ML over G-2SLS is that with ML we can exploit all the heterogeneity in the social structure to help with identification. By comparison, Generalized 2SLS only exploits exclusion restrictions related to the presence of intransitive triads in the network.

4.1 Peer effect parametric estimation

4.1.1 Estimation results for San Cristobal

Accounting for peer effects lowers the effect of the program on school attendance in San Cristobal (Table 7). Looking at column 1, we find that lottery winners show higher attendance than lottery losers (a significant 1.1 percentage point increase from a base mean of 87 percent). This suggests that incentives directed to scholarship recipients work, in the sense that they do lead them to privately increase their attendance. An increase in peers attendance results in an increase in own attendance. The effect is small (0.03) and statistically significant. This finding is consistent with positive complementarities between peers in time spent in social activities outside school. In contrast, the local intensity of treatment among peers negatively impacts school attendance: an increase in the fraction of peers treated leads to a significant drop in attendance. What may explain that a higher fraction of treated peers is associated with lower school attendance? It may be that treated children who dropped out of their jobs to comply with the program requirements are freeing up labor that their peers take-up.²⁶ When taking into account the positive endogenous peer effect and the negative contextual effect of the treatment through peers, we find that the overall program effect (panel 5, Table 7) is close to zero and not significant.

We find that the naive difference in mean attendance between experimental groups, (γ^{ols}) is positive and significantly different from zero (column 3). Comparing it to the estimated total program effect (panel 5, Table 7), we find that ignoring peer effects would lead us to wrongly conclude that a full-scale program would be effective. It is often argued that ignoring peer effects when these are present implies that the naive experimental estimator can be interpreted as a lower-bound to the actual program effect. Intuitively, in the presence of social interactions, treated and untreated outcomes should be closer than in the absence of interactions. Here, this is not what we find. The reason is the following. We do indeed find evidence of a positive endogenous effect, consistent with the rest of the literature. But we also allow the treatment status of others to affect attendance directly beyond the effect it has through behavior, and this contextual effect of treatment is found to be negative. Results are robust to adding controls (age, gender, grade and SISBEN score); see Columns 2 and 4.

Moreover, we find that the naive experimental estimator (γ^{ols}) is only capturing the individual mirror effect (Δ^{mirror}), that is, the effect from private incentives to attend school on program beneficiaries and social influences on attendance that feedback to the group of program beneficiaries.

²⁶That beneficiaries and peers are substitutes on the child labor market is one possible explanation for the negative contextual effect. A behavioural explanation is also plausible. For instance, a student, who in the absence of the program has intrinsic motivations for attending school, may respond negatively to the introduction of financial incentives for attendance, and this effect may be stronger when more of his friends benefit from them. We provide some suggestive evidence that the former explanation is valid (see section 4.2.6) but cannot test the behavioral explanation with the data at hand.

We find that the individual mirror effect (panel 3, Table 7) is positive, statistically significant and of similar magnitude as $\hat{\gamma}^{ols}$. We actually cannot reject the hypothesis that the individual mirror effect is equal to the difference in expected outcome between experimental groups (panel 4, Table 7).²⁷ Because the individual mirror effect is a function of the model parameters, this constitutes a strong test of the model. We find that the model is validated on these data.

To sum up, our results suggest that private incentives to attend school lead to a small increase in school attendance among program beneficiaries. This effect is amplified through social interaction: peers reinforce each other's behavior. Treatment status of peers also affect attendance directly: exogenously higher intensity of treatment among peers leads to a reduction in attendance. This effect is consistent with substitution between own and peers labor supply. Taken together, treatment and peer effects on attendance implies that the program is not effective in increasing school attendance. In other words, the naive experimental estimator overestimates the actual program impact. The econometric model is tested and validated for San Cristobal.

4.1.2 Estimation results for Suba

The results for Suba (Table 8) are essentially similar to the ones obtained for San Cristobal. Accounting for peer effects in treatment response implies that the program is not an effective mean to improve school attendance in Suba, where the program targets higher-grade students. While the naive experimental effect is positive, the total program effect (panel 5) is negative in Suba and not significant. This is due to the fact that the contextual effect from the treatment more than offsets the direct treatment effect to beneficiaries.

As in San Cristobal, winning the lottery results in a 1.2 percentage point higher attendance (from a base mean of 86 percent). The sign of the effect is consistent with the private incentive embedded in the program, but the effect is not statistically significant. The endogenous peer effect is positive and statistically significant, and the contextual effect of peers treatment is negative and statistically significant. A 1 percentage point (p.p.) higher intensity of treatment among peers is associated with a 1.6 p.p. drop in attendance while a 1 p.p. increase in school attendance among peers is associated with a 4.7 p.p. rise in attendance. These effects are in the same order of magnitude as in San Cristobal. The model is also validated using data from Suba. As shown in panel 3, the estimate of the individual mirror effect on scholarship beneficiaries, $\hat{\Delta}^{mirror}$, which sums the effect from private incentives to attend school and social influences on attendance that feedback to the group of program beneficiaries, is very close in value to the difference in means between experimental groups ($\hat{\gamma}^{ols}$). As a result, we cannot reject that the two are equal (panel 4).

²⁷Standard errors for Δ^{mirror} are estimated using the delta method, see Appendix E.

4.2 Robustness checks

In this section, we discuss findings from alternative specifications of the model. Robustness checks are only performed for San Cristobal, for which the sample size is larger. For each of these robustness checks, we discuss how we expect results to change compared to the benchmark findings in Table 7. Overall, results from this section comfort us in our modelling choices. In section 4.2.1, we include unobserved effects fixed at the level of the network. In section 4.2.2, we consider the reduced-form equation in which only the treatment contextual peer effects matters. Then, in section 4.2.3, we present estimation results treating isolated individuals differently from others. In section 4.2.4, we also consider an alternative specification of the social interaction matrix where links are not reciprocated. In section 4.2.5, we present semiparametric estimates that do not impose linearity of the effect from peers treatment status. Finally, in section 4.2.6, we offer some additional insights on the program contextual effect.

4.2.1 Fixed effects

In Table 9, we report estimation results for San Cristobal adding network fixed effects. School fixed effects are included in order to control for all unobserved correlated effect that are fixed at the level of the network. With randomization that ensures that everyone’s treatment is uncorrelated with school (unobserved) attributes, both maximum likelihood estimator (estimates in Table 7) and conditional maximum likelihood estimator (estimates in Table 9) should be consistent. As expected, we find that our three parameters of interest (the direct program effect on beneficiaries and the two peer effects) are similar in Table 7 and Table 9. These results are robust to the inclusion of additional covariates.

4.2.2 Contextual effects only

In Table 10, we show estimates of a model with contextual effects only. We allow the fraction of friends at distance-1 and the fraction of friends at distance-2 to enter the equation. This specification can be considered as an approximation of the reduced-form equation (4).²⁸ Neither the direct program effect, nor the two contextual effects are statistically significant using data for San Cristobal. This result could lead us to dismiss peer effects as not being important. But, it is actually consistent with our result in Table 7 showing that endogenous and contextual peer effects are of opposite signs. This finding also stresses the relevance of separately estimating endogenous and contextual peer effects.

²⁸It is an approximation in the sense that we stop at friends at distance-2.

4.2.3 Isolated individuals

In Table 11, we present estimates for San Cristobal of model (2) restricting the sample to exclude isolated individuals. We find that the direct program effect is similar to the benchmark from Table 7. The magnitude of the treatment contextual effect and of the endogenous peer effect are also similar. The total program effect is not significantly different from zero.

In model (2), the intercept should not be the same for isolated individuals and individuals who are not isolated. But in our estimation we impose that intercepts be the same. Allowing for different intercepts should provide additional variation to identify the endogenous effect. When we allow intercepts to differ for isolated and nonisolated students, the coefficient of the treatment contextual effect is unaffected (Table 12). The magnitude of the endogenous effect is slightly higher, but the total effect of the program is still not significantly different from zero. Socially isolated students show lower attendance than students reporting having friends.

4.2.4 Directed network

We now estimate a model for which we use a directed social interaction matrix. Indeed, reported friendship links may not be reciprocal. We relax our hypothesis that all friendship links are reciprocated. Our findings are quantitatively robust to this assumption: direct and indirect effects are of the same size, sign and significance level with directed links (Table 13) than with undirected links (Table 7). This is not surprising given that identification relies on intransitive triads: if the triad is intransitive when links are undirected, it is also with directed links. Similarly, if the triad is not intransitive when links are reciprocal, then it is not intransitive with unreciprocated links.

4.2.5 Semi-parametric model

A potential concern is that the specification of the model may be driving our results. We now consider a semi-parametric specification. Do parametric estimates have limited support in the data? Is linearity too strong an assumption?

We estimate the following semiparametric model with contextual effects only:

$$y_i = \pi_1 t_i + h((\mathbf{Gt})_i) + u_i \quad (9)$$

We provide details of the estimation in Appendix F.

We find that the estimated function is globally flat (Figure 2). The confidence interval is extremely wide at the right-end of the data range (dotted lines). The function is estimated with good statistical precision over the range of points from $[0; 0.8]$. In Figure 3, we can see that most of the data support is for values of $(\mathbf{Gt})_i$ between $[0; 0.8]$, and very few individuals have more than 80 percent of their friends treated. Taken together, these findings suggest that the linear-in-means specification used in our parametric estimation is consistent with the data.

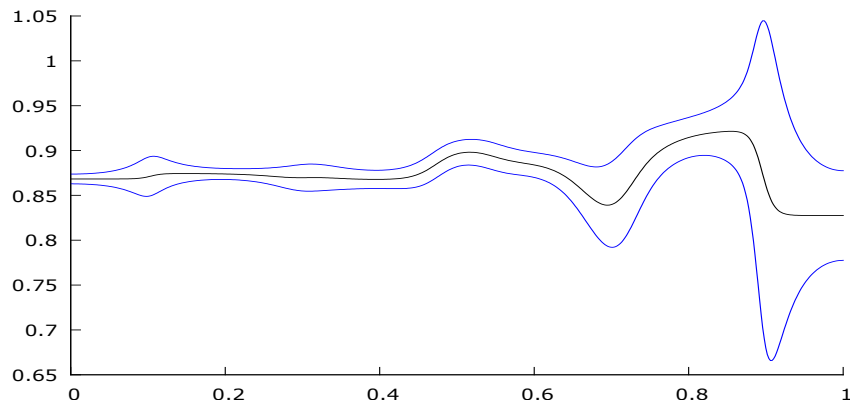


Figure 2: Semi-parametric estimation in San Cristobal

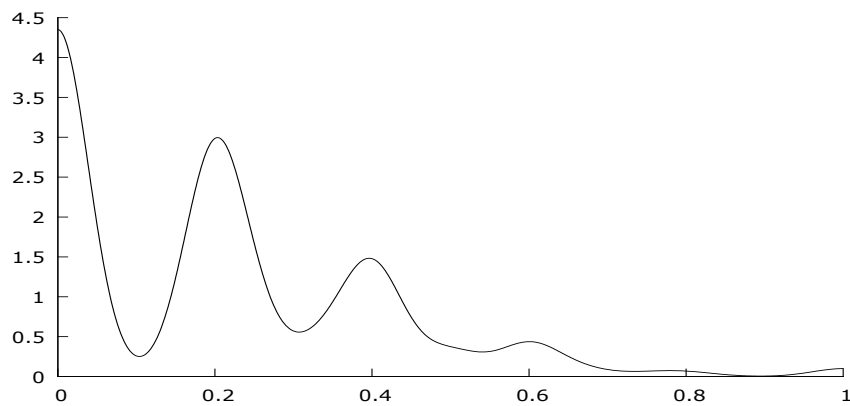


Figure 3: Density of the fraction of friends treated in San Cristobal

4.2.6 Program contextual effect

Another potential concern is that the treatment contextual effect is a mere by-product of the experiment: with a scaled-up program, should we expect peers treatment to affect individual choice directly, beyond its effect through peers choices? To answer this question, one would need to unfold the mechanisms through which the treatment contextual effect is generated. *E.g.*, if it is caused by the substitution in peers labor supply, should we still expect substitution when the program is at-scale?

We first show that substitution in peer labor supply is a plausible explanation for why, holding peers' attendance fixed, own attendance decreases as the fraction of treated peers increases. To do so, we interact the program contextual variable (fraction of treated friends) with a dummy that identifies students aged 13 and above (students enrolled in higher grades). The idea is that older children are more likely to participate in the labor market in the absence of the program. Recall that, in order to keep receiving monetary benefits, students are required to attend school on a regular basis. So, we can expect that treated students enrolled in higher grades are the most affected

by the conditionality. For substitution in peer labor supply to be a plausible explanation, the program contextual effect should be even more negative for the group of students in higher grades. Our findings provide support for this explanation. In San Cristobal (Table 14), the program contextual effect is not statistically different from zero for younger students. For older students, it is negative, statistically significant and more than three times the magnitude of the effect reported in the main estimation result pooling all students together (Table 7).²⁹

Even though the evidence is consistent with a substitution in peers labor supply, the extent of substitution may be smaller for a program at-scale. Indeed, with a program at-scale, all applicants receive cash benefits upon regular attendance, freeing up labor for their peers from more well-off families who may not be interested in substituting for them. Then, a conservative approach is to consider that expression (7) gives a bound to the total effect. The other bound is obtained by setting $\delta_1 = 0$.

5 Conclusion

In this paper, we account for peer effects in order to assess a scholarship program intended to limit child labor and improve progress through school in two poor neighborhoods of Bogota. Our setting is one where assignment to the treatment is at the individual level, so that treated and untreated individuals may be influencing each other. We use a model that exploits information on the fine structure of interactions. Besides providing evidence of peer effects in schooling, we find that ignoring peer effects leads to a biased evaluation of the program actual impact. Through the endogenous peer effect, the positive effect on school attendance from receiving program benefits is magnified. But the social multiplier on school attendance is small and this effect is more than offset by the negative effect through peers treatment status. This unintended negative contextual effect of the treatment is found to be consistent with substitution between peers' labor supply.

We also investigate whether the results may be driven by our modelling choices. We propose a way to test the model specification which relates the model parameters to the difference in expected outcome between experimental groups. We validate the model based on data from both localities (each locality features its own randomized experiment). We also estimate a semiparametric version of the model and find that the linearity-in-means assumption holds on the data support.

One limitation of our work, driven by the data at hand, is that we cannot account for peer effects between program applicants and nonapplicants. Our randomized experiment is limited to the set of applicants. Although we argue that decision-makers may not be interested in benefits to nonapplicants in the case of this scholarship program, accounting for benefits to nonapplicants may be important in other contexts.

²⁹Results for Suba are very similar (Table 15).

REFERENCES

- Angrist, J., Bettinger, E., and Kremer, M. (2006). Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia. *American Economic Review*, 96(3):847–862.
- Barrera-Osorio, F., Bertrand, M., Linden, L. L., and Perez-Calle, F. (2011). Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia. *American Economic Journal: Applied Economics*, 3(2):167–95.
- Benabou, R. (1996). Inequality and Growth. Working Papers 96-22, C.V. Starr Center for Applied Economics, New York University.
- Bhattacharya, D., Dupas, P., and Kanaya, S. (2013). Estimating the Impact of Means-tested Subsidies under Treatment Externalities with Application to Anti-Malarial Bednets. Economics Series Working Papers 646, University of Oxford, Department of Economics.
- Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55.
- Calvo-Armengol, A., Patacchini, E., and Zenou, Y. (2009). Peer Effects and Social Networks in Education. *Review of Economic Studies*, 76(4):1239–1267.
- Chandrasekhar, A. G. and Lewis, R. (2011). Econometrics of Sampled Networks. Mimeo, Stanford University.
- Comola, M. and Prina, S. (2013). Do Interventions Change the Network? A Dynamic Peer Effect Model Accounting for Network Changes. Mimeo, Paris School of Economics.
- De Giorgi, G., Pellizzari, M., and Redaelli, S. (2010). Identification of Social Interactions through Partially Overlapping Peer Groups. *American Economic Journal: Applied Economics*, 2(2):241–75.
- Durlauf, S. (2006). *Groups, Social Influences and Inequality*. Poverty Traps. Princeton University Press.
- Evans, W. N., Oates, W. E., and Schwab, R. M. (1992). Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy*, 100(5):966–91.
- Fiszbein, A., Schady, N., Ferreira, F. H. G., Grosh, M., Keleher, N., Olinto, P., and Skoufias, E. (2009). *Conditional Cash Transfers : Reducing Present and Future Poverty*. Number 2597 in World Bank Publications. The World Bank.
- Heckman, J. J. (2000). Microdata, Heterogeneity and the Evaluation of Public Policy. Nobel Prize in Economics documents 2000-4, Nobel Prize Committee.
- Katz, L. F., Kling, J. R., and Liebman, J. B. (2001). Moving to opportunity in boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics*, 116(2):607–654.

- Kremer, M. and Holla, A. (2009). Improving education in the developing world: What have we learned from randomized evaluations? *Annual Review of Economics*, 1:513–542.
- Kremer, M. and Miguel, E. (2007). The illusion of sustainability. *The Quarterly Journal of Economics*, 122(3):1007–1065.
- LaLonde, R. J. (1986). Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *American Economic Review*, 76(4):604–20.
- Laschever, R. (2009). The Doughboys Network: Social Interactions and the Employment of World War I Veterans. Working paper, Purdue University.
- Lee, L.-f., Xiaodong, L., and Xu, L. (2010). Specification and Estimation of Social Interaction Models with Network Structures. *The Econometrics Journal*, 13(2):145–176.
- Lin, X. (2010). Identifying Peer Effects in Student Academic Achievement by Spatial Autoregressive Models with Group Unobservables.
- Lise, J., Seitz, S., and Smith, J. (2005). Equilibrium Policy Experiments and the Evaluation of Social Programs. Working Papers 1076, Queen’s University, Department of Economics.
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542.
- Manski, C. F. (2013a). Comment. *Journal of Business & Economic Statistics*, 31(3):273–275.
- Manski, C. F. (2013b). Identification of treatment response with social interactions. *Econometrics Journal*, 16(1):S1–S23.
- Miguel, E. and Kremer, M. (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217.
- Moffitt, R. A. (2001). *Policy Interventions, Low-Level Equilibria, and Social Interactions*, pages 45–82. MIT Press.
- Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and its Applications*, 9:141–142.
- Oster, E. and Thornton, R. (2011). Menstruation, Sanitary Products, and School Attendance: Evidence from a Randomized Evaluation. *American Economic Journal: Applied Economics*, 3(1):91–100.
- Philipson, T. J. (2000). External Treatment Effects and Program Implementation Bias. NBER Technical Working Papers 0250, National Bureau of Economic Research, Inc.
- Robinson, P. M. (1988). Root-n-consistent semiparametric regression. *Econometrica*, 56(4):931–954.
- Sacerdote, B. (2011). *Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?*, volume 3 of *Handbook of the Economics of Education*, chapter 4, pages 249–277. Elsevier.
- Schultz, T. P. (2001). School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program. Working Papers 834, Economic Growth Center, Yale University.

- Schultz, T. P. (2004). School subsidies for the poor: evaluating the Mexican Progresa poverty program. *Journal of Development Economics*, 74(1):199–250.
- Todd, P. E. and Wolpin, K. I. (2006). Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review*, 96(5):1384–1417.
- Watson, G. S. (1964). Smooth regression analysis. *Sankhya, Series A*, 26(359-372).

Table 1: Distribution of research subjects by experimental group

	Suba			San Cristobal		
	T =0	T=1	Total	T=0	T=1	Total
Grades 6-8	0	0	0	1,196	2,240	3,436
Grades 9-10	455	400	855	643	1,147	1,790
Grades 9-11	619	527	1,146	822	1,482	2,304
Total	619	527	1,146	2,018	3,722	5,740

Table 2: Pupils descriptive statistics at baseline, San Cristobal

	Sample (N=5,740)	Control (N=2,018)	Difference	T-stat
Age	13.53 (1.93)	13.57 (1.92)	-0.05 (0.03)	-0.94
Male	0.48 (0.50)	0.48 (0.50)	0.009 (0.009)	0.65
SISBEN score	11.63 (4.54)	11.75 (4.64)	-0.17 (0.08)	-1.34
Grade 6	0.21 (0.41)	0.20 (0.40)	0.007 (0.0107)	0.68
Grade 7	0.21 (0.40)	0.21 (0.40)	-0.0008 (0.007)	-0.07619
Grade 8	0.18 (0.38)	0.18 (0.38)	0.0024 (0.007)	0.22
Grade 9	0.16 (0.37)	0.16 (0.37)	0.004 (0.007)	0.38
Grade 10	0.14 (0.35)	0.15 (0.36)	-0.014 (0.006)	-1.46
Grade 11	0.089 (0.28)	0.088 (0.28)	0.0013 (0.005)	0.16
Ever dropped out	0.11 (0.31)	0.11 (0.31)	-0.0027 (0.006)	-0.32
Reason for dropping out: lack of money	0.27 (0.44)	0.25 (0.44)	0.018 (0.025)	0.51
Repeated a grade	0.38 (0.48)	0.38 (0.48)	-0.002 (0.009)	-0.15
Paid work	0.17 (0.38)	0.18 (0.38)	-0.0056 (0.007)	-0.53
Hours worked	7.51 (10.34)	7.95 (11.31)	-0.67 (0.50)	-0.95

Note: mean and standard deviation in parenthesis; difference and standard error in parenthesis.

Table 3: Network descriptive statistics at baseline, San Cristobal.

	Sample	Control	Difference	T-statistics
Individual characteristics				
Number of friends	4.37 (1.12)	4.371 (1.13)	-0.0011 (0.021)	-0.037
Number of friends among applicants	2.58 (2.01)	2.53 (1.95)	0.076 (0.036)	1.39
No friends	0.006 (0.08)	0.008 (0.09)	-0.0017 (0.0016)	-0.74
No friends among applicants	0.127 (0.33)	0.13 (0.33)	-0.005 (0.006)	-0.52
Characteristics of friends at distance-1				
Treated	0.54 (0.42)	0.54 (0.42)	0.002 (0.007)	0.20
Age	11.19 (5.3)	11.31 (5.24)	-0.19 (0.098)	-1.30
Male	0.41 (0.45)	0.42 (0.45)	-0.006 (0.008)	-0.52
SISBEN score	9.72 (5.52)	9.87 (5.54)	-0.22 (0.10)	-1.44
Characteristics of friends at distance-2				
Treated	0.54 (0.37)	0.27 (0.25)	0.42 (0.004)	52.3
Age	11.19 (5.29)	11.32 (5.22)	-0.19 (0.097)	-1.36
Male	0.41 (0.43)	0.41 (0.43)	-0.0006 (0.008)	-0.053
SISBEN score	9.71 (5.27)	9.80 (5.27)	-0.11 (0.1)	-0.77

Note: mean and standard deviation in parenthesis; difference and standard error in parenthesis.

Continued on next page

Table 3 – continued from previous page

	Sample	Control	Difference	T-statistics
Characteristics of friends at distance-3				
Treated	0.54 (0.37)	0.51 (0.37)	0.04 (0.007)	4.41
Age	11.18 (5.28)	11.31 (5.22)	-0.19 (0.097)	-1.29
Male	0.41 (0.42)	0.42 (0.43)	-0.008 (0.008)	-0.70
SISBEN score	9.73 (5.26)	9.87 (5.28)	-0.21 (0.14)	-1.45
Characteristics of friends at distance-4				
Treated	0.54 (0.36)	0.32 (0.27)	0.34 (0.005)	40.56
Age	11.19 (5.28)	11.31 (5.21)	-0.19 (0.09)	-1.33
Male	0.41 (0.41)	0.41 (0.41)	-0.003 (0.007)	-0.25
SISBEN score	9.71 (5.17)	9.78 (5.16)	-0.10 (0.09)	-0.70

Note: mean and standard deviation in parenthesis; difference and standard error in parenthesis.

Table 4: Pupils descriptive statistics at baseline, Suba

	Sample (N=1,146)	Control (N=619)	Difference	T-stat
Age	15.25 (1.23)	15.21 (1.22)	0.082 (0.051)	1.13
Male	0.40 (0.49)	0.40 (0.49)	0.004 (0.02)	0.14
SISBEN score	13.54 (4.23)	13.43 (4.23)	0.23 (0.17)	0.90
Grade 9	0.39 (0.49)	0.39 (0.48)	0.008 (0.02)	0.29
Grade 10	0.35 (0.47)	0.34 (0.47)	0.015 (0.019)	0.55
Grade 11	0.25 (0.43)	0.26 (0.44)	-0.023 (0.018)	-0.93
Ever dropped out	0.12 (0.32)	0.12 (0.32)	0.003 (0.013)	0.18
Reason for dropping out: lack of money	0.26 (0.44)	0.3 (0.46)	-0.08 (0.055)	-1.10
Repeated a grade	0.37 (0.48)	0.35 (0.47)	0.04 (0.019)	1.41
Paid work	0.28 (0.45)	0.29 (0.45)	-0.033 (0.019)	-1.26
Hours worked	8.26 (8.82)	7.81 (6.57)	1.033 (0.51)	0.97

Note: mean and standard deviation in parenthesis; difference and standard error in parenthesis.

Table 5: Network descriptive statistics at baseline, Suba

	Sample	Control	Difference	T-statistics
Individual characteristics				
Number of friends	4.37 (1.11)	4.32 (1.19)	0.11 (0.049)	1.66
Number of participating friends	1.81 (1.76)	1.76 (1.71)	0.12 (0.07)	1.18
No friends	0.008 (0.093)	0.009 (0.098)	-0.002 (0.004)	-0.38
No friends among applicants	0.22 (0.41)	0.22 (0.42)	-0.008 (0.017)	-0.33
Characteristics of friends at distance-1				
Treated	0.33 (0.41)	0.34 (0.41)	-0.012 (0.017)	-0.51
Age	10.72 (7.00)	10.48 (7.09)	0.53 (0.29)	1.27
Male	0.28 (0.42)	0.27 (0.42)	0.022 (0.017)	0.87
SISBEN score	9.67 (6.95)	9.34 (6.92)	0.71 (0.29)	1.71
Characteristics of friends at distance-2				
Treated	0.33 (0.38)	0.13 (0.19)	0.44 (0.008)	22.97
Age	10.71 (7.00)	10.49 (7.09)	0.47 (0.29)	1.14
Male	0.28 (0.40)	0.27 (0.40)	0.02 (0.016)	0.81
SISBEN score	9.65 (6.80)	9.34 (6.81)	0.68 (0.28)	1.69

Note: mean and standard deviation in parenthesis; difference and standard error in parenthesis.

Continued on next page

Table 5 – continued from previous page

	Sample	Control	Difference	T-statistics
Characteristics of friends at distance-3				
Treated	0.33 (0.38)	0.32 (0.38)	0.025 (0.015)	1.11
Age	10.72 (7.00)	10.47 (7.08)	0.53 (0.29)	1.28
Male	0.28 (0.40)	0.27 (0.40)	0.02 (0.016)	0.85
SISBEN score	9.66 (6.82)	9.36 (6.82)	0.65 (0.28)	1.61
Characteristics of friends at distance-4				
Treated	0.33 (0.37)	0.15 (0.21)	0.39 (0.009)	20.26
Age	10.71 (6.98)	10.49 (7.09)	0.47 (0.29)	1.13
Male	0.28 (0.39)	0.27 (0.39)	0.02 (0.016)	0.95
SISBEN score	9.66 (6.77)	9.35 (6.78)	0.66 (0.28)	1.65

Note: mean and standard deviation in parenthesis; difference and standard error in parenthesis.

Table 6: School attendance by experimental group.

	Suba			San Cristobal		
	T =0	T=1	Total	T=0	T=1	Total
Grades 6-8	0	0	0	0.872 (0.16)	0.881 (0.15)	0.878 (0.16)
Grades 9-10	0.855 (0.19)	0.873 (0.20)	0.865 (0.20)	0.867 (0.17)	0.878 (0.17)	0.874 (0.17)
Grades 9-11	0.857 (0.18)	0.872 (0.19)	0.863 (0.19)	0.869 (0.18)	0.880 (0.16)	0.876 (0.17)
Observations	619	527	1,146	2,018	3,722	5,740

Table 7: Main estimation results for school attendance in San Cristobal

	Max-Lik.	Max-Lik. with controls	OLS	OLS with controls
<i>Individual characteristics</i>				
Intercept	0.8574 (0.0115)	0.9645 (0.0198)	0.8711 (0.0035)	0.927 (0.0165)
Treatment = 1	0.0112 (0.0042)	0.0111 (0.0041)	0.0098 (0.0044)	0.0100 (0.0044)
Age	—	-0.0146 (0.0017)	—	-0.0135 (0.0017)
Male	—	-0.0038 (0.0039)	—	-0.0023 (0.0042)
Grade	—	0.0104 (0.0022)	—	0.0131 (0.0022)
SISBEN score	—	0.0010 (0.0004)	—	0.0018 (0.0005)
<i>Peer effects</i>				
Contextual	-0.0104 (0.0056)	-0.0107 (0.0056)	—	—
Endogenous	0.0308 (0.0066)	0.0274 (0.0065)	—	—
<i>Individual mirror effect</i>				
Δ^{mirror}	0.0111 (0.0041)	0.0109 (0.0041)	—	—
<i>Model validation</i>				
$H_0: \Delta^{mirror} = \gamma^{ols}$	0.2005	0.163	—	—
<i>Total effect</i>				
	0.0008 (0.0072)	0.00037 (0.0071)	—	—
Observations	5740	5740	5740	5740

Note: Controls are age, gender, grade and SISBEN score.

Standard errors for the individual mirror effect and total effect are calculated using the delta method.

Table 8: Main estimation results for school attendance in Suba.

	Max Lik.	Max Lik. with controls	OLS	OLS with controls
<i>Individual characteristics</i>				
Intercept	0.8308 (0.0323)	1.0009 (0.0736)	0.8575 (0.0075)	0.7873 (0.0800)
Tertiary = 1	0.0129 (0.0089)	0.0136 (0.0089)	0.0148 (0.0111)	0.0154 (0.0111)
Age	—	-0.0077 (0.0044)	—	-0.0049 (0.0055)
Male	—	-0.0059 (0.0091)	—	-0.0090 (0.0113)
Grade	—	-0.0048 (0.0071)	—	0.0134 (0.0086)
SISBEN score	—	-0.0001 (0.0011)	—	0.0013 (0.0013)
<i>Peer effects</i>				
Contextual	-0.0163 (0.0123)	-0.0164 (0.0123)	—	—
Endogenous	0.0472 (0.0124)	0.0448 (0.0125)	—	—
<i>Individual mirror effect</i>				
Δ^{mirror}	0.0126 (0.0089)	0.0133 (0.0089)	—	—
<i>Model validation</i>				
$H_0: \Delta^{mirror} = \gamma^{ols}$	-0.1552	-0.1498	—	—
<i>Total effect</i>				
	-0.0035 (0.0163)	-0.0029 (0.0162)	—	—
Observations	1146	1146	1146	1146

Note: Controls are age, gender, grade and SISBEN score.

Standard errors for the individual mirror effect and total effect are computed using the delta method.

Table 9: Fixed effects model - School attendance in San Cristobal

	Max-Lik.	Max-Lik. with controls	OLS	OLS with controls
Individual characteristics				
Treatment = 1	0.0112 (0.0041)	0.0961 (0.0041)	0.0110 (0.0041)	0.0109 (0.0041)
Age	—	-0.0152 (0.0017)	—	-0.0149 (0.0017)
Male	—	-0.0049 (0.0039)	—	-0.0036 (0.0039)
Grade	—	0.0091 (0.0022)	—	0.0103 (0.0022)
SISBEN score	—	0.0009 (0.0004)	—	0.0011 (0.0004)
Peer effects				
Contextual	-0.0098 (0.0056)	-0.01304 (0.0056)	—	—
Endogenous	0.0298 (0.0066)	0.0233 (0.0065)	—	—
Individual mirror effect				
Δ^{mirror}	0.0111 (0.0041)	0.0095 (0.0041)	—	—
Model validation				
$H_0: \Delta^{mirror} = \gamma^{ols}$	0.0308	-0.2392	—	—
Total effect				
	0.0014 (0.0072)	-0.0035 (0.0071)	—	—
Observations	5740	5740	5740	5740

Note: Controls are age, gender, grade and SISBEN score.

Standard errors for the individual mirror effect and total effect are computed using the delta method.

Table 10: Contextual effects only - School attendance in San Cristobal.

Intercept	0.8721 (0.0113)	0.9803 (0.0196)
Treatment = 1	0.0075 (0.0052)	0.0087 (0.0051)
Contextual peer effects	0.0003 (0.0057)	-0.0001 (0.0056)
Fraction of distance-2 friends treated	0.0082 (0.0074)	0.0049 (0.0074)
Age	—	-0.015 (0.0017)
Male	—	-0.0036 (0.0039)
Grade	—	0.0104 (0.0022)
SISBEN score	—	0.001 (0.0004)
Observations	5740	5740

Table 11: Excluding isolated individuals - School attendance in San Cristobal.

	Max-Lik.	Max-Lik. with controls	OLS	OLS with controls
<i>Individual characteristics</i>				
Intercept	0.8542 (0.0126)	0.9497 (0.0228)	0.8783 (0.0043)	0.9109 (0.0196)
Treatment = 1	0.0084 (0.0049)	0.0077 (0.0049)	0.0051 (0.0053)	0.0047 (0.0052)
Age	—	-0.0123 (0.0021)	—	-0.0119 (0.0022)
Male	—	-0.0038 (0.0047)	—	-0.0036 (0.005)
Grade	—	0.0082 (0.0026)	—	0.0132 (0.0027)
SISBEN score	—	0.0007 (0.0005)	—	0.002 (0.0005)
<i>Peer effects</i>				
Contextual	-0.0104 (0.0064)	-0.0112 (0.0064)	—	—
Endogenous	0.0389 (0.0091)	0.0361 (0.0091)	—	—
<i>Individual mirror effect</i>				
Δ^{mirror}	0.0082 (0.0049)	0.0075 (0.0049)	—	—
<i>Model validation</i>				
$H_0: \Delta^{mirror} = \gamma^{ols}$	0.4350	0.3869	—	—
<i>Total effect</i>				
	-0.0020 (0.0084)	-0.0036 (0.0083)	—	—
Observations	3895	3895	3895	3895

Note: Controls are age, gender, grade and SISBEN score.

Standard errors for the individual mirror effect and the total effect are calculated using the delta method.

Table 12: Dummy for isolated individuals - School attendance in San Cristobal.

	Max-Lik.	Max-Lik. with controls	OLS	OLS with controls
Individual characteristics				
Intercept	0.8601 (0.0124)	0.9676 (0.0204)	0.8625 (0.0046)	0.9198 (0.0167)
Treatment = 1	0.0111 (0.0041)	0.011 (0.0041)	0.0097 (0.0044)	0.0098 (0.0044)
Isolated = 1	-0.0032 (0.0052)	-0.0032 (0.0052)	-0.0128 (0.0045)	-0.0115 (0.0045)
Age	—	-0.0146 (0.0017)	—	-0.0135 (0.0018)
Male	—	-0.0038 (0.0040)	—	-0.0025 (0.0042)
Grade	—	0.0102 (0.0022)	—	0.0131 (0.0022)
SISBEN score	—	0.001 (0.00045)	—	0.0018 (0.0005)
Peer effects				
Contextual	-0.0107 (0.0056)	-0.0109 (0.0056)	—	—
Endogenous	0.0287 (0.0075)	0.0253 (0.0074)	—	—
Individual mirror effect				
Δ^{mirror}	0.0110 (0.0042)	0.0109 (0.0041)	—	—
Model validation				
$H_0: \Delta^{mirror} = \gamma^{ols}$	0.2134	0.1736	—	—
Total effect				
	0.0004 (0.0072)	0.0002 (0.0071)	—	—
Observations	5740	5740	5740	5740

Note: Controls are age, gender, grade and SISBEN score.

Standard errors for the individual mirror effect and the total effect are calculated using the delta method.

Table 13: Directed network - School attendance in San Cristobal.

	Max-Lik.	Max-Lik. with controls	OLS	OLS with controls
<i>Individual characteristics</i>				
Intercept	0.8618 (0.0112)	0.9688 (0.0195)	0.8711 (0.0035)	0.9279 (0.0165)
Treatment = 1	0.0111 (0.0041)	0.0110 (0.0041)	0.0098 (0.0044)	0.099 (0.0044)
Age	—	-0.0145 (0.0017)	—	-0.0135 (0.0018)
Male	—	-0.0038 (0.0039)	—	-0.0023 (0.0042)
Grade	—	0.0100 (0.0022)	—	0.0131 (0.0022)
SISBEN score	—	0.0010 (0.00045)	—	0.0018 (0.00047)
<i>Peer effects</i>				
Contextual	-0.0128 (0.0056)	-0.0124 (0.0056)	—	—
Endogenous	0.0291 (0.0061)	0.0252 (0.0061)	—	—
<i>Individual mirror effect</i>				
Δ^{mirror}	0.0110 (0.0041)	0.0109 (0.0041)	—	—
<i>Model validation</i>				
$H_0: \Delta^{mirror} = \gamma^{ols}$	0.18911	0.1555	—	—
<i>Total effect</i>				
	-0.0017 (0.0071)	-0.0015 (0.0071)	—	—
Observations	5740	5740	5740	5740

Note: Controls are age, gender, grade and SISBEN score.

Standard errors for the individual mirror effect and total effect are computed using the delta method.

Table 14: Dummy for older students - School attendance in San Cristobal

	M-L	M-L with controls	OLS	OLS with controls
Intercept	0.8516 (0.0136)	0.9537 (0.0241)	0.8711 (0.0035)	0.9279 (0.0164)
Treatment = 1	0.0092 (0.0042)	0.0093 (0.0042)	0.0098 (0.0044)	0.0100 (0.0044)
Age	—	-0.0142 (0.002)	—	-0.0135 (0.0017)
Male	—	-0.0011 (0.004)	—	-0.0023 (0.0042)
Grade	—	0.0101 (0.0023)	—	0.0131 (0.0022)
Sisben	—	0.0009 (0.0005)	—	0.0018 (0.0005)
Contextual	0.0065 (0.0070)	-0.0104 (0.0076)	—	—
Older × Contextual	-0.0317 (0.0063)	-0.0049 (0.0079)	—	—
Endogenous	0.0367 (0.0068)	0.0332 (0.0068)	—	—
Observations	5740	5740	5740	5740

Table 15: Dummy for older students - School attendance in Suba

	M-L	M-L with controls	OLS	OLS with controls
Treatment = 1	0.0139 (0.0089)	0.014 (0.0089)	0.0148 (0.0111)	0.0153 (0.0111)
Age	—	-0.0054 (0.0048)	—	-0.0075 (0.0075)
Male	—	-0.0062 (0.0091)	—	-0.009 (0.0113)
Grade	—	-0.004 (0.0071)	—	0.0127 (0.0087)
Sisben	—	-0.0001 (0.0011)	—	0.0013 (0.0013)
Contextual	0.0142 (0.0178)	0.0025 (0.0191)	—	—
Older × Contextual	-0.0434 (0.0183)	-0.0268 (0.0206)	—	—
Endogenous	0.0467 (0.0125)	0.045 (0.0125)	—	—
Observations	1146	1146	1146	1146

Appendix Tables

Table 16: Validation of experimental design in San Cristobal

	Full sample (N = 6, 366)		Restricted sample (N = 5, 740)	
	Control group mean (s.d.)	Difference by treatment status (s.e.)	Control group mean (s.d.)	Difference by treatment status (s.e.)
Age	13.62 (1.96)	0.048 (0.034)	13.56 (1.92)	0.05 (0.035)
Male	0.48 (0.49)	-0.025 (0.008)	0.47 (0.49)	-0.009 (0.009)
SISBEN score	11.77 (4.66)	0.165 (0.082)	11.74 (4.63)	-0.17 (0.08)
Grade 6	0.21 (0.41)	0.023 (0.07)	0.20 (0.40)	-0.007 (0.07)
Grade 7	0.21 (0.40)	-0.009 (0.072)	0.21 (0.40)	0.0008 (0.76)
Grade 8	0.17 (0.37)	-0.0082 (0.006)	0.18 (0.38)	-0.002 (0.007)
Grade 9	0.16 (0.37)	-0.034 (0.006)	0.16 (0.37)	-0.0039 (0.007)
Grade 10	0.15 (0.36)	-0.008 (0.006)	0.15 (0.36)	0.014 (0.006)
Grade 11	0.084 (0.27)	-0.004 (0.005)	0.088 (0.28)	-0.001 (0.005)

Notes: Restricted sample excludes observations with missing identification number, i.e., observations that cannot be matched with baseline network data.

Table 17: Validation of experimental design in Suba

	Full sample (N = 3, 402)		Restricted sample (N = 1, 146)	
	Control group mean (s.d.)	Difference by treatment status (s.e.)	Control group mean (s.d.)	Difference by treatment status (s.e.)
Age	15.16 (1.84)	-2.13 (0.044)	15.21 (1.21)	0.082 (0.051)
Male	0.46 (0.49)	-0.036 (0.012)	0.40 (0.49)	0.0041 (0.020)
SISBEN score	13.17 (4.41)	0.54 (0.11)	13.43 (4.23)	0.23 (0.18)
Grade 9	0.098 (0.29)	0.31 (0.007)	0.39 (0.49)	0.008 (0.02)
Grade 10	0.08 (0.27)	0.27 (0.006)	0.34 (0.47)	0.015 (0.019)
Grade 11	0.06 (0.23)	0.17 (0.006)	0.26 (0.44)	-0.024 (0.018)

Notes: Restricted sample excludes observations with missing identification number, i.e., observations that cannot be matched with baseline network data.

Appendix A Treatment of missing data

To write the model in matrix form, let

$$\Delta = \begin{pmatrix} \frac{m_1}{n_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{m_N}{n_N} \end{pmatrix}$$

that is, a $(N \times N)$ diagonal matrix with $(\frac{m_1}{n_1}, \dots, \frac{m_N}{n_N})$ on the diagonal and 0 elsewhere. Writing model (5) in matrix notation yields:

$$\mathbf{y} = \alpha \mathbf{1} + \beta(\Delta \mathbf{G}^1 \mathbf{y}) + \gamma_1 \mathbf{t} + \delta_1(\Delta \mathbf{G}^1 \mathbf{t}) + \mathbf{u}.$$

The associated reduced form is thus as follows:

$$\mathbf{y} = \alpha(\mathbf{I} - \beta \Delta \mathbf{G}^1)^{-1} \mathbf{1} + (\mathbf{I} - \beta \Delta \mathbf{G}^1)^{-1} (\gamma_1 \mathbf{I} + \delta_1 \Delta \mathbf{G}^1) \mathbf{t} + (\mathbf{I} - \beta \Delta \mathbf{G}^1)^{-1} \mathbf{u}. \quad (10)$$

Therefore, in order to adapt model (2) from the previous section, one just needs to replace the interaction matrix in that model by $\Delta \mathbf{G}^1$.

Appendix B Proof of Proposition 2

In the following, we derive the proof using the social network matrix restricted to participants. From Equation (5):

$$\begin{cases} \mathbb{E}(y_i | t_i = 1) = \frac{\alpha}{1-\beta} + \gamma_1 + (\gamma_1 \beta + \delta_1) \sum_{k=0}^{\infty} (\beta)^k \mathbb{E}[(\mathbf{H}^{k+1} \mathbf{t})_i | t_i = 1] + \mathbb{E}(v_i | t_i = 1), \\ \mathbb{E}(y_i | t_i = 0) = \frac{\alpha}{1-\beta} + (\gamma_1 \beta + \delta_1) \sum_{k=0}^{\infty} (\beta)^k \mathbb{E}[(\mathbf{H}^{k+1} \mathbf{t})_i | t_i = 0] + \mathbb{E}(v_i | t_i = 0). \end{cases}$$

where $\mathbf{H} = \Delta \mathbf{G}^1$ for ease of notation and $\mathbf{v} = (\mathbf{I} - \beta \mathbf{H}^1)^{-1} \mathbf{u}$.

Let $\mathbf{t} = \mathbf{t}_i + \mathbf{t}_{-i}$, where $\mathbf{t}_i = (0, \dots, 0, 1, 0, \dots, 0)'$. Thus, $\mathbf{H} \mathbf{t} = \mathbf{H}(\mathbf{t}_i + \mathbf{t}_{-i})$. Randomization of the treatment ensures that the expected fraction of treated among friends at any distance k (excluding oneself) is the same for treated and control units: $\mathbb{E}[(\mathbf{H}^k \mathbf{t}_{-i})_i | t_i = 1] = \mathbb{E}[(\mathbf{H}^k \mathbf{t}_{-i})_i | t_i = 0]$. In addition, for all k , $\mathbb{E}[(\mathbf{H}^k \mathbf{t}_i)_i | t_i = 0] = 0$, which implies $\sum_{k=0}^{\infty} (\beta)^k \mathbb{E}[(\mathbf{H}^{k+1} \mathbf{t}_i)_i | t_i = 0] = 0$. But $\mathbb{E}[(\mathbf{H}^k \mathbf{t}_i)_i | t_i = 1]$ may not be equal to zero:

$$\sum_{k=0}^{\infty} (\beta)^k \mathbb{E}[(\mathbf{H}^{k+1} \mathbf{t}_i)_i | t_i = 1] = \mathbb{E}[(\mathbf{H}(\mathbf{I} - \beta \mathbf{H})^{-1} \mathbf{t}_i)_i | t_i = 1] = \mathbb{E}[(\mathbf{H}(\mathbf{I} - \beta \mathbf{H})^{-1}]_{ii} | t_i = 1),$$

where $[\mathbf{H}(\mathbf{I} - \beta \mathbf{H})^{-1}]_{ii}$ is the i^{th} element on the diagonal of $\mathbf{H}(\mathbf{I} - \beta \mathbf{H})^{-1}$. Thus, $\sum_{k=0}^{\infty} (\beta)^k \mathbb{E}[(\mathbf{H}^{k+1} \mathbf{t}_i)_i | t_i = 1] = \frac{1}{N} \text{Tr}(\mathbf{H}(\mathbf{I} - \beta \mathbf{H})^{-1})$. In addition, randomization ensures that: $\mathbb{E}(v_i | t_i = 1) = \mathbb{E}(v_i | t_i = 0)$.

The result follows.

Appendix C Construction of the social interaction matrix

Our sample includes all the applicants to the program in both localities. But, students may name as friends individuals who are not in the sample. We follow a three-steps procedure to construct the matrix of social interaction:

1. All friendship links to a non-applicant (missing) are coded as NA.
2. We build the adjacency matrix. Note that the matrix includes a line for each isolated individual and a line and a column for all non-applicant friends.
3. In the last step, we eliminate the non-applicant friends (line and column).

Appendix D Maximum likelihood

The school-level reduced-form of model (4) can be written in matrix notation, as follows:

$$y_s = (\mathbf{I}_s - \beta \mathbf{H}_s)^{-1} (\mathbf{Z}_s \theta + u_s),$$

where $\mathbf{H}_s = (\Delta \mathbf{G}^1)_s$ for ease of notation. Variables and matrices are written at the school level where each school is of size N_s , and where $\mathbf{Z}_s = (\boldsymbol{\nu}_s, \mathbf{t}_s, \mathbf{H}_s \mathbf{t}_s)$ and $\theta = (\alpha, \gamma_1, \delta_1)$.

We account for correlations between individual error terms within schools by allowing for clustering. Let the error term u_{is} can be written as the sum of a network effect and a idiosyncratic error such that $u_{is} = v_s + e_{is}$. Suppose v_s follows a normal distribution of mean 0 and variance σ_s^2 and e_{is} follows also a normal distribution of mean 0 and variance σ_e^2 . The log-likelihood of the whole sample is then given by the following expression:

$$\begin{aligned} L(y; \theta, \beta, \sigma | \mathbf{H}, \mathbf{t}) = & - \frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma_e^2) - \frac{1}{2} \sum_{s=1}^S \ln(1 + \rho^2 N_s) + \sum_{s=1}^S \ln |\mathbf{I}_s - \beta \mathbf{H}_s| \\ & - \frac{1}{2\sigma_e^2} \sum_{s=1}^S \left[\boldsymbol{\nu}_s' \boldsymbol{\nu}_s - \frac{\rho^2 (N_s \bar{\nu}_s)^2}{1 + \rho^2 N_s} \right], \end{aligned}$$

where $\boldsymbol{\nu}_s = (\mathbf{I}_s - \beta \mathbf{H}_s) \mathbf{y}_s - \mathbf{Z}_s \theta$ and $\rho = \sigma_s^2 / \sigma_e^2$.

Appendix E Model validation

We use the delta method to obtain an estimate of the variance for the right-hand side of expression (8):

$$h(\boldsymbol{\theta}) = \gamma_1 + (\gamma_1\beta + \delta_1) \frac{1}{N} \sum_{i=1}^N \frac{\lambda_i}{1 - \beta\lambda_i}, \quad (11)$$

where λ_i are the eigenvalues of matrix \mathbf{G} and $\boldsymbol{\theta} = (\gamma_1, \delta_1, \beta)'$. Applying the delta method to $y = h(\hat{\boldsymbol{\theta}})$:

$$y = h(\boldsymbol{\theta}) + h' \cdot (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + \text{higher order terms},$$

where $h' = \frac{\partial h}{\partial \boldsymbol{\theta}}|_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}}$. Using equation (11), h' is the following vector:

$$\begin{pmatrix} \frac{\partial h}{\partial \gamma_1} |_{\gamma_1=\hat{\gamma}_1} \\ \frac{\partial h}{\partial \delta_1} |_{\delta_1=\hat{\delta}_1} \\ \frac{\partial h}{\partial \beta} |_{\beta=\hat{\beta}} \end{pmatrix} = \begin{pmatrix} 1 + \hat{\beta} \frac{1}{N} \sum_{i=1}^N \frac{\lambda_i}{1 - \beta\lambda_i} \\ \frac{1}{N} \sum_{i=1}^N \frac{\lambda_i}{1 - \beta\lambda_i} \\ \frac{1}{N} \sum_{i=1}^N \frac{\lambda_i}{1 - \beta\lambda_i} \left(\hat{\gamma} + \frac{(\gamma_1\beta + \delta_1)\lambda_i}{(1 - \beta\lambda_i)} \right) \end{pmatrix}.$$

Let Σ be the variance-covariance matrix of vector $\hat{\boldsymbol{\theta}}$. As $N \rightarrow \infty$, we have:

$$\mathbb{E}(h(\hat{\boldsymbol{\theta}})) = h(\boldsymbol{\theta}),$$

$$V(h(\hat{\boldsymbol{\theta}})) = h' \Sigma h.$$

Appendix F Estimation details for the semi-parametric model

We assume that the error term is independent of own treatment status and mean treatment in the peer group. In the following, to account for missing data on non-participants, simply replace \mathbf{G} by \mathbf{H} where $\mathbf{H} = \Delta \mathbf{G}^\dagger$. Formally, our identifying condition is:

$$\mathbb{E}(u_i | t_i, (\mathbf{G}\mathbf{t})_i) = 0$$

We estimate the parameters of interest, following a three-step procedure (Robinson, 1988):

1. We estimate the following quantities in a non-parametric fashion:

$$\hat{y}_i^* = y_i - \hat{\mathbb{E}}(y_i | (\mathbf{G}\mathbf{t})_i)$$

$$\hat{t}_i^* = t_i - \hat{\mathbb{E}}(t_i | (\mathbf{G}\mathbf{t})_i)$$

2. We estimate π_1 by an OLS regression of y_i^* on t_i^* :

$$\hat{y}_i^* = \pi_1 \hat{t}_i^* + \epsilon_i^*$$

3. We replace the estimated value of π_1 in model (9) and obtain the following expression:

$$y_i - \hat{\pi}_1 t_i = h((\mathbf{G}\mathbf{t})_i) + u_i$$

The final step consists of non-parametrically estimating $h((Gt)_i)$ using the Nadaraya-Watson (Nadaraya 1964 and Watson 1964) estimator.