Essays on social interactions

Gioia de Melo

Supervisor: Giulio Zanella

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics

Department of Economics and Statistics

University of Siena

August 2012

Acknowledgements

First and foremost, I would like to express my gratitude to my advisor, Giulio Zanella, for his guidance and encouragement during my graduate studies. Most of the empirical strategies applied in the first two chapters were suggested by him. I am also extremely thankful to Sam Bowles for his constant support and especially for his advice and discussions during the elaboration of the third chapter. This work has also benefited from discussions with my colleagues in the Ph.D program and several academics from University of Siena. I am also grateful to the Instituto de Economía, Universidad de la República, for providing the work flexibility that enabled me to complete my graduate studies.

The essay constituting the first chapter of the Thesis, "Peer effects identified through social networks. Evidence from Uruguayan schools", was partly written while I was visiting researcher at University of California, Berkeley, whose hospitality I gratefully acknowledge. I am particularly thankful to Edward Miguel for sponsoring my visit and to Patrick Kline for his invaluable advice. I am also grateful for advice received from Yann Bramoullé, Pamela Campa, Giacomo De Giorgi, Francesco Drago, Fred Finan, Bernard Fortin, Tiziano Razzolini and Jesse Rothstein as well as to seminar participants at UC Berkeley, the University of Siena and the IZA European Summer School in Labor Economics for useful comments. I am also very thankful to Andrés Peri from División de Investigación, Evaluación y Estadística, Administración Nacional de Educación Pública, for providing access to the data used in the first two chapters of the dissertation.

The third chapter of the thesis is co-authored with Matías Piaggio, whom I consider an excellent teammate for such a demanding project. I am very grateful to the Latin American and the Caribbean Environmental Economics Program for funding the project and to participants at LACEEP's XII and XIII workshops for the comments received. I am particularly thankful to Francisco Alpizar, Andres Rius, Jordi Brandts, and Fredik Carlsson for their invaluable advice. I am also very thankful for the comments received from Gabriele Camera, Luigi Luini, Shachar Kariv, Jorge Maldonado, Matías Brum and Paolo Pin. The project would have never been possible to implement without the commitment and support of Hugo Bai, Gabriel Burdín, María Dolores Carnevale, Paula Carrasco, Marco Colafranceschi, Andres Dean, Inés Erro, Sandra García, Federico Méndez, Sybille Mertens, Cecilia Oreiro, Sylvina Porras, Lorena Repetto, Andrés Rius, Nicolás Schmidt and Andrea Vigorito during the experimental sessions. I am especially grateful to Cecilia Toledo, for her support and patience helping us coordinate the field work. I would also like to thank Javier Vitancurt, (DINAMA), and Graciela Fabiano (DINARA), for their cooperation while collecting technical information. Finally, I would like to thank the fishers from the five communities who participated in the experiment.

I would like to thank my parents Alfredo and Emilia and my grandmother Anna for their endless understanding and support. This thesis is dedicated to them.

Abstract

The thesis focuses on the role of social interactions in two different settings: education and the management of common pool resources. The main aim is to assess whether the presence of peers influences behavior. A second goal is to explore which are the effects of group heterogeneity. The underlying motivation is to try to understand which is the most desirable composition of individuals in terms of social welfare in both education and local environmental management. The dissertation combines observational studies exploiting a unique data set of students with an experimental study performed in the field.

The thesis is composed of three chapters, corresponding to three independent essays with the following titles:

1. Peer effects identified through social networks. Evidence from Uruguayan schools
2. The social determinants of junior high dropout in Uruguay
3. The perils of peer punishment. Evidence from a common pool resource framed field experiment.

The first chapter "Peer effects identified through social networks. Evidence from Uruguayan schools", analyzes the presence of peer effects in test scores using a novel identification strategy that enables to estimate separately behavioral (endogenous) effects from contextual effects. The study exploits a unique data set with information on individual specific reference groups. Estimates suggest that there are strong positive endogenous peer effects and therefore that the social multiplier is large. The study illustrates that in a context of socioeconomic stratification in which schools are increasingly tending to be more homogeneous in student characteristics; social interactions will amplify educational disparities.

The second chapter "The social determinants of junior high dropout in Uruguay", analyzes the influence of classmates in the last grade of primary school on the decision to attend junior high. The study applies a completely different identification strategy to the first chapter. While the first chapter considered as the relevant reference group peers specifically named by the individual, the second chapter considers the whole class (excluding the individual). In addition, the main identification strategy applied in the second chapter can only estimate a composite social interaction effect that encompasses both endogenous and contextual effects. The findings of the second chapter suggest that while having classmates with a better socioeconomic background increases the likelihood of attending junior high, the presence of classmates with higher scores may discourage low achieving students to continue studying. Indeed, this negative peer effect is only observed at the lowest quartile of the score distribution and could suggest that having high achieving classmates may deteriorate low achieving students' confidence regarding schooling aspirations.

Overall it seems that on average students benefit from learning in environments mixed with high achieving students (chapter 1). However, this conclusion seems not to apply for the schooling aspirations of the lowest achieving students (chapter 2). It is debatable which reference group among those considered in the two chapters is more appropriate. While it could be argued that one is influenced the most by the individuals with whom one is closest, it is also arguable that the characteristics of the whole class can affect students. Furthermore, while disentangling endogenous from contextual effects is crucial in terms of the policy implications endogenous effects have (and this highlights the importance of the strategy followed in the first chapter), policies can only influence the composition of classrooms and indirectly through classrooms affect the composition of friends at school. This, in turn, justifies studying the relevance of classmates as a whole. These studies reflect how complex is to generalize the relevance of peers to different educational outcomes and to assume peer effects are homogeneous. It also suggests different identification strategies may be capturing different effects.

The third chapter "The perils of peer punishment: Evidence from a common pool resource framed field experiment", studies whether social preferences differ in a context in which individuals exploiting a common pool resource belong to different communities. The study performs a common pool resource experiment that evaluates the effects of non-monetary punishment by peers among communities of Uruguayan fishers.

Assessing the relevance of non-monetary punishment as a tool to enhance cooperation is of particular importance in regard to community management of common pool resources, because informal sanctions typically take place within a community. We find no evidence of in-group favoritism. That is, subjects do not exhibit greater levels of cooperation when interacting exclusively with subjects from their own community. On the contrary, interacting with fishers from other communities has a positive effect on cooperation when punishment is available. We interpret this result as follows: in a context in which individuals do not know each other (or hardly know each other) but are aware that there is a slight chance they might see each other again, being publicly punished provides the only information others have about oneself and in this sense it may be important to avoid being flagged in such a way. In other words, the relationship between the sensitivity to peer punishment in in-group/out-group contexts may be non-monotonic.

In addition, we find strong conformity effects: individuals adjust their period-by-period decisions in order to converge with their peers' average in a previous period. These results highlight the potential relevance of social comparisons as a form of non-pecuniary policy seeking changes in behavior. Finally, we observe that community membership has an influence over individuals' decisions, a finding not explained by observable socioeconomic factors. This may suggest that social norms differ among communities.

Contents

1 Peer effects identified through social networks. Evidence from Uruguayan schools

This chapter provides evidence on peer effects in standardized tests by exploiting a unique data set on social networks in Uruguayan primary schools. The identification method enables one to disentangle endogenous from contextual effects via instrumental variables that emerge naturally from the network structure. Correlated effects are controlled for via classroom fixed effects. I find significant endogenous effects in reading and math: a one--standard deviation increase in peers' scores increases own scores by about 40 percent of a standard deviation. Simulation exercises show that, when schools are stratified by socioeconomic status, peer effects may amplify educational inequalities.

1.1 Introduction

Because peer effects constitute a form of externality, they are of particular relevance to welfare-enhancing policies (Durlauf, 1998; Hoxby, 2000; Glaeser and Scheinkman, 2001). Significant levels of peer influence can have policy implications not only in terms of efficiency but also of inequality. In fact, educational policies ranging from tracking to desegregation programs have been justified in terms of presumed peer effects.[1]

The dependence of individual behavior on peers' behavior can generate a social multiplier or feedback loop and can also lead to multiple equilibria (Manski, 1993; Glaeser, Sacerdote and Scheinkman, 2003; Soetevent, 2006). Since social interactions are likely to influence schooling decisions, study habits, and individual aspirations, it follows that socioeconomic stratification in the establishment of social networks has serious implications for the persistence of educational disparities and of broader social inequalities across generations (Benabou, 1996; Durlauf, 1996,

[1] In the United States desegregation plans were prompted by the 1954 Supreme Court decision (Brown v. Board of Education) that declared it illegal to segregate schools by race---and later by the Coleman (1966) report that concluded racial segregation has a negative effect on the educational achievement of minority children. Some more recent studies (Guryan, 2004; Card and Rothstein, 2007) have provided some evidence in favor of this hypothesis. Today, there are many countries implementing desegregation programs; most notable is India's nationwide program, the Right to Education Act, which reserves one fourth of private schools placements for disadvantaged children. In turn, tracking has been promoted under the assumption that a high-achieving peer has more effect on another high-achieving student than on a low-achieving student and under the assumption that more homogeneous levels in classrooms allow teachers to target instruction accordingly with students' needs.

2004; Bowles, Loury, and Sethi, 2007; Graham, 2011). Moreover, the search for valuable social interactions can lead to inefficient stratification (Benabou, 1993, 1996; Zanella, 2007).

That being said, much debate has addressed the actual relevance of peer effects especially given the identification challenges posed by any study of social interactions and there is still no consensus on their magnitude. This chapter assesses the impact of peer effects in test scores by applying an identification strategy recently developed in three independent papers: Bramoullé, Djebbari and Fortin (2009); De Giorgi, Pellizzari, and Redaelli (2010); and Lin (2010). This strategy exploits information on individual-specific peer groups in which the existence of partially overlapping peers allows for using the characteristics of peers' peers (and of peers' peers' peers) as instrumental variables to obtain an exogenous source of variation in peer behavior. In this way, the strategy enables one to isolate the endogenous peer effect and thus solving the so-called reflection problem. This is especially important because only endogenous effects can generate a social multiplier, and most previous studies have estimated a composite social effect that includes both endogenous and contextual effects.[2]

The intuition behind this framework is that peers' peers, who are not also the students' peers, can only have an impact on that student's outcomes indirectly by influencing the outcomes of her peers. Including classroom fixed effects allows me to control for the self-selection of students into schools and for unobserved shocks at the class level. The chapter also shows that, within a given class, there seems to be no self-selection into groups of peers with similar socioeconomic background.

I use a data set of primary schools in Uruguay (not previously employed for research purposes) that provides information on reference groups. Students self-report whom they would like to invite to their house to play and whom they would like to work with for a school assignment. To the best of my knowledge, the only other data set with similar characteristics is the National Longitudinal Study of Adolescent Health (Add Health). Calvó-Armengol, Patacchini, and Zenou (2009) and also Lin (2010) use the information in Add Health's social networks to study peer effects in education.[3] De Giorgi, Pellizzari and Redaelli (2010) apply a similar strategy to study the influence of classmates on a student's choice of college major at Bocconi University.

I find strong evidence of endogenous effects for both reading and math, although peer effects are not significant for science. A one--standard deviation increase in peers' scores increases the student's scores by 40 percent of a standard deviation in reading (and 37 percent of a standard deviation in math). This effect is smaller than, but comparable to that of having a mother who

[2] A social multiplier or feedback loop occurs when the direct effect of an improvement in one characteristic of an individual has an indirect effect on other individuals through social interactions (Soetevent, 2006).

[3] Bramoullé et al. (2009) also use the Add Health data set to study peer effects on the consumption of recreational services; Fortin and Yazbeck (2011) use it to study peer effects in fast-food consumption.

completed college. In contrast, contextual effects seem not to be significant. I then employ a simulation exercise to assess the extent to which peer effects amplify educational inequality in a context of schools stratified by socioeconomic status. I estimate that if peers were assigned randomly, then the standard deviations of reading and math scores would decrease by 4.5 percent and 10 percent, respectively.

The main contribution of this chapter is to apply a recently developed identification strategy to a new comprehensive data set that is representative at the country level for students in their last year of primary school. A significant advantage of this data set---compared with those used in most studies that analyze peer effects in test scores---is that here the tests on reading, math, and science were devised and scored by the national educational authority and so are not biased by teachers' perceptions and/or preferences. In this way, each student took the same three tests.[4] Moreover, the data set used in this paper give a very precise idea of what the real peer group is and yield individual-level information not available elsewhere about network formation in different activities (leisure and study). They also present a heterogeneous scenario of schools and students and, most importantly, provide enough variability to allow drawing inferences. The paper's second contribution is to analyze, by means of a simulation exercise, the possibility that peer effects act to amplify educational inequality. The findings reported here do not directly support any particular policy intervention but do demonstrate that peer effects in learning should be taken into account when designing any educational policy ranging from the decision of where to build a new school---in a system in which students are assigned to the nearest school from their house---to more complex policies.

The chapter is organized as follows. Section 1.2 reviews the main empirical literature on peer effects in education, Section 1.3 discusses the identification strategy, and Section 1.4 describes the data. Section 1.5 reports the main results; Section 1.6 provides some alternative specifications. Section 1.7 analyzes the implications of peer effects in a context of socioeconomic segregation. Section 1.8 compares endogenous and contextual effects estimated using network data with social interaction effects derived from exploiting variations in peer characteristics between classrooms within a school. Section 1.9 concludes.

1.2 Related literature

Although peer effects in education have been studied since the 1960s, there is still no consensus on their relevance (Soetevent, 2006). In the last two decades, the empirical literature on peer

[4] Add Health contains information on students' grade-points averages (GPAs).

effects has been subjected to powerful criticisms regarding identification issues raised by Manski (1993, 2000), Moffitt (2001), and Brock and Durlauf (2001). Several studies have attempted to address these econometric challenges, but the evidence on the relevance of peer effects remains mixed.

A first challenge is to isolate peer effects from correlated effects that arise from sorting and/or unobserved omitted variables.[5] In addition, the study of social interactions involves a simultaneity problem or reflection problem: the presence of exogenous effects implies that characteristics affect not only each individual's outcome but also each peer's outcome, but the researcher observes only the equilibrium outcome in which all the individuals' outcomes are jointly determined (Soetevent, 2006). Hence it is extremely hard to find an exclusion restriction (i.e., an explanatory variable of individual outcomes that does not affect indirectly peers' outcomes) that would enable one to separate endogenous from contextual effects in a linear-in-means model (Manski, 1993).[6] In other words, the structural parameters cannot be recovered from the reduced form owing to collinearities between individual and contextual variables. Another challenge is that identifying social interactions is impossible unless the group composition is known (Manski, 1993, 2000). In what follows, I review the main strategies for overcoming these challenges that have been pursued in previous studies.

Correlated effects

Sacerdote (2001) and Zimmerman (2003) study peer effects in education by exploiting data on randomly assigned college roommates, where the random assignment allows them to separate social interactions from correlated effects. Graham (2008) suggests a novel method for identifying social interactions using conditional variance restrictions. By using experimental data from project STAR, he distinguishes the excess variance due to peer effects from that due to

[5] As was initially pointed out by Manski (1993), there are three possible effects that can account for similar behavior within a group. First, children may act similarly because they are influenced by their peers' behavior (proxied by outcomes); according to Manski's typology, these are *endogenous* effects. Second, children may attain similar outcomes also because they are influenced by their peers' characteristics. For instance, children may perceive their peers' parents as role models and the involvement of parents in their children's education may also indirectly benefit the children's peers; these are viewed as *exogenous* (or contextual) effects. Finally, children in a class may exhibit similar outcomes owing to the presence of *correlated* effects---as when, for example they are taught by the same teacher or have the same socioeconomic background or are equally motivated to study. Whereas endogenous and exogenous effects reflect the impact of social interactions, correlated effects do not.

[6] In this model (which is standard in the literature) the outcome of an individual is linearly related to his own characteristics, the corresponding mean characteristics of his peers, and their mean outcome.

group-level heterogeneity and/or sorting.[7] Graham's estimates suggest a substantial impact of peer quality on kindergarten achievement.

Hoxby (2000) identifies social interactions by exploiting the variation in gender and racial composition of a grade within schools during adjacent years. Lavy and Schlosser (2011) also rely on variation in gender composition across adjacent cohorts, and Ammermueller and Pischke (2009) use changes in composition across classrooms within the same grade. This strategy is useful for isolating correlated effects provided the changes yield sufficient variation (Nechyba, 2006). Other studies use school-by-grade effects (Calvó-Armengol et al., 2009; Lin, 2010) or school-by-grade and student effects (Hanushek, 2003).

The reflection problem

Many studies do not disentangle endogenous and exogenous effects and therefore estimate a composite social interaction effect (or assume there is but one form of interaction). This is the case in Hoxby (2000), Sacerdote (2001), Zimmerman (2003), Graham (2008), and Ammermueller and Pischke (2009). Yet, it is especially important to isolate endogenous effects because only they can generate a social multiplier. Hanushek et al. (2003) estimate endogenous and exogenous effects separately by instrumenting the peers' score with their lagged achievement (though they acknowledge the downward bias inherent in that approach). The reflection problem can also be circumvented by specifying a model in which behavior varies either nonlinearly with group mean behavior or linearly with some characteristic of group behavior other than the mean (Manski, 2000; Brock and Durlauf, 2001).

Another possibility is to find an instrumental variable that directly affects the behavior of some but not all the group members. In this way, endogenous and exogenous effects can be disentangled under a partial-population experimental setting whereby the outcome variable of some randomly chosen members of the group is modified exogenously (Moffitt, 2001). That strategy is applied by Bobonis and Finan (2009), who study neighborhood spillovers from induced school participation of children eligible for the PROGRESA program. Cooley (2010) disentangles endogenous and exogenous effects by utilizing the introduction of student accountability policies in North Carolina public schools. These policies imposed an additional cost on low performance and thus affected the effort only of those who perceived themselves to be in danger of failing. Cooley identifies peer spillovers by comparing classrooms that contain varying percentages of ``accountable" students with classrooms of otherwise similar composition but in which students were not held accountable. A novel strategy for disentangling endogenous from exogenous effects involves the use of partially overlapping reference groups (Calvó-

[7] The experimental aspect of project STAR enables Graham (2008) to assume that teacher quality is distributed randomly across classrooms.

Armengol et al., 2009; Laschever, 2009; De Giorgi et al., 2010; Lin, 2010). I detail this strategy in Section 1.3.

Reference groups

Data constraints often require the reference group to be defined arbitrarily (Nechyba, 2006). Most papers that study peer effects in education assume that individuals interact within broad groups and are affected by an average intragroup externality that identically affects all the members of a grade within a school or classroom. Given the information on social networks available from the Add Health data set, some studies have considered individual-specific reference groups. Lin (2010) assumes that the individuals named by a student as friends are his reference group and Calvó-Armengol et al. (2009) concentrate on the position of each individual named in a social network (the Katz--Bonacich index[8]).

1.3 Identification Strategy

Bramoullé et al. (2009) determine the conditions under which endogenous and contextual effects can be identified when individuals interact through social networks known by the researcher and when correlated effects are assumed to be fixed within groups. In this chapter, I follow their identification strategy. The model developed here is an extension of the linear-in-means model of Manski (1993) and Moffitt (2001), but now each individual has his own specific reference group. Let the structural model for any student i belonging to classroom c be as follows:

$$y_{ci} = \alpha_c + \beta \frac{\sum_{j \in P_i} y_{cj}}{p_i} + \gamma x_{ci} + \delta \frac{\sum_{j \in P_i} x_{cj}}{p_i} + \varepsilon_{ci}, \quad E[\varepsilon_{ci} \mid x_{ci}, \alpha_c] = 0 \tag{1}$$

Here y_{ci} is the test score of student i and x_{ci} is a *1* × *K* vector of individual characteristics (for simplicity, hereafter we assume that there is only one characteristic). Each student i may have a specific peer group or set of nominated friends P_i of size p_i. The term β captures the endogenous or behavioral effect, and δ captures the exogenous effect of peers' predetermined characteristics. I address the problem of correlated effects by introducing classroom fixed effects that capture unobserved variables common to students in the same classroom. This approach allows for correlation between the classroom's unobserved common characteristics (e.g., teacher quality) and observed characteristics such as parental education. However, individual

[8] This measure counts, for each node in a given network, the total number of direct and indirect network paths of any length stemming from that node. Paths are weighted by a factor that decays geometrically with path length.

characteristics are assumed to be strictly exogenous after conditioning on the classroom fixed effect.

Let I_c be the identity matrix for classroom c and let ι be the corresponding vector of 1s. Let G be an $n \times n$ interaction matrix for the n students in classroom c, with $G_{ij} = \frac{1}{p_i}$ if j was named by i and $G_{ij} = 0$ otherwise. Note that G is row-normalized. The model can be written in matrix notation as:

$$y_c = \alpha_c \iota_c + \beta G_c y_c + \gamma x_c + \delta G_c x_c + \varepsilon_c, \quad E[\varepsilon_c \mid x_c, G_c, \alpha_c] = 0 \tag{2}$$

Then to eliminate classroom fixed effects, I apply a ``within" transformation via pre-multiplying equation (2) by $D_c = I_c - \frac{1}{n_c}\iota_c \iota_c'$. That is, I average equation (1) over all the students in i's classroom and then subtract it from i's equation. The structural model can now be written as:

$$D_c y_c = \beta D_c G_c y_c + \gamma D_c x_c + \delta D_c G_c x_c + D_c \varepsilon_c \tag{3}$$

where the reduced form is:

$$D_c y_c = D_c (I_c - \beta G_c)^{-1} (\gamma I_c + \delta G_c) x_c + D_c (I_c - \beta G_c)^{-1} \varepsilon_c \tag{4}$$

Bramoullé et al. (2009) show that if the matrices I, G, G^2 and G^3 are linearly independent, then social interactions can be identified. This implies that $E[DGy \mid x]$ is not perfectly collinear with (Dx, DGx), which means that $(DG^2 x, DG^3 x, \ldots)$ are valid instruments for the outcomes of ones' peers.[9] In other words, the characteristics of a student's peers' peers (and of his peers' peers' peers, etc.) who are *not* her peers serve as instruments for the outcomes of her own peers, thus resolving the reflection problem. The intuition behind this framework is that the characteristics of peers' peers who are not the student's peers can have only an indirect impact on the student's behavior by influencing her peers' behaviors. Bramoullé et al. (2009) note that a sufficient condition for identification is that the network's diameter (i.e., the maximal distance between any two peers in the student network) be no less than 3. This, in turn, requires that there be at least one case in which: i named j, j named k, and k named l; but i named neither k nor l and j did not name l. Nevertheless, the authors demonstrate that identification often holds also in transitive networks as well, in which case it derives from the directed nature of the network. In more general terms, social effects can be disentangled as long as there is some variation in reference groups. In this chapter, identification is based on both the existence of partially

[9] These variables have been previously transformed as deviations from their corresponding classroom mean.

overlapping groups (links of distance 3 or more) and on the network's directed nature (i.e., the direction of influence from one node to another).[10]

A crucial identification assumption is that there are no unobserved characteristics that differ among children in a classroom and that also affect both achievement and the likelihood of becoming friends. For instance, if the most able children become friends among themselves and attain better scores than the rest of the class, then the networks will not be exogenous conditional on α_c and x_c and so estimates of social interactions will be inconsistent. Alternatively, if highly disruptive children tend to interact mostly with other disruptive children and also score poorly (owing to this unobserved characteristic and not due to their peers' influence), then inconsitent estimates would again result. Of course, it is not feasible to test whether there is self-selection in terms of unobservables. The following section presents evidence suggesting that at least there is no selection in terms of observables related to parental background.

1.4 Data

The analysis is based on a unique data set: the fifth Evaluación Nacional de Aprendizajes, which took place in October 2009 and comprises a 322-school sample (24 percent of Uruguay's schools) in which approximately 8,600 students were evaluated. The sample is representative of sixth-grade students (the last grade in primary school, students 11--12 years old) and covers children in both private and public schools. The evaluation consists of math, science, and reading tests which were created and scored by ANEP, the central authority responsible for education in Uruguay.[11] This is a major advantage compared to data sets in which students are graded by their teachers because teachers' expectations of (or preferences for) their students could distort grading within a class. Every student who was evaluated took the same reading, math, and science test. The data set also includes questionnaire answers from students and their families as well as from teachers and the school principals.

Two questions on the students' questionnaire are of particular importance for this study because they provide information on reference groups:

- If you were to invite two classmates to play at your house, whom would you invite?
- If you were to invite two classmates to work on an assignment for school, whom would you invite?

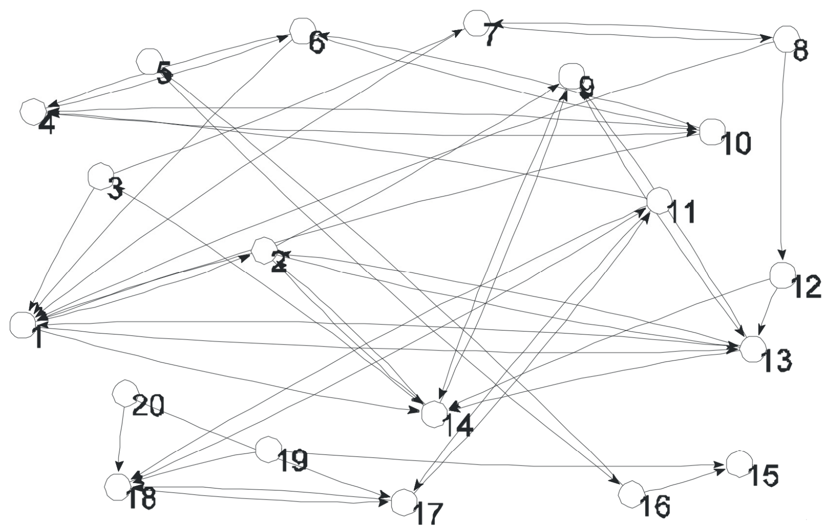
Figure 1.1 depicts the network structure resulting from the information provided by answers to these two questions from one actual classroom. Links of distance at least 3 (i.e., that satisfy the

[10] If student A names B but B does not name A, then B is viewed as A's peer but A is not viewed as B's peer.

[11] Administración Nacional de Educación Pública (ANEP).

identification condition) can be observed.[12] Also, I checked that the matrices I, G, G^2 and G^3 are linearly independent (where G is a matrix that contains all the classroom networks), which is another way of verifying that the identification condition established by Bramoullé et al. (2009) is satisfied.[13]

Figure 1.1 – A classroom viewed as a network



The reference-group questions mentioned previously dictate that a student name at most 4 peers. This does constitute a limitation, since reference groups exceeding that number are thus not adequately captured. However, it should be taken into account that the problem is not as severe as in studies where nodes are sampled because in this study students name their closest peers first. Considering both questions (party and work), 13 percent named 4 distinct peers who can be

[12] For example, individual 7 named 8 who named 12 who named 13, 7 did not name either 12 or 13 and 8 did not name 13. In turn, 13 named 9, 14, 2 and 1, none of whom were named by the previous individuals.

[13] This was checked by vectorizing matrices I, G, G^2 and G^3 and verifying that the matrix formed by these four vectors is of rank 4.

identified in the data set (on average they named 2.4 distinct peers).[14] One might expect that students name their closest friends in the ``play" question but not necessarily in the ``work" one but, 65 percent of students repeated at least one peer in the two questions (40 percent repeated the name of one peer and 25 percent repeated the two peers named in the party question in the assignment question, see Table A.1.1 in the Appendix).[15]

On average children were named (i.e., were considered part of others' reference-group) 1.7 times in both the play and work question. Students that were named between 1 and 4 times amount to 69 percent in the play question and 66 percent in the work question while 14 percent of students were not named by anyone in either question. This general pattern suggests that children who were named by others as peers are distributed quite uniformly within classrooms---in other words, the whole class did not name the same student. This contributes to identification by increasing the distance in terms of number of links between individuals (since the likelihood of finding links of ≥ 3 would be lower if most of the arrows were pointing toward just a few students). As mentioned before, most children who are named in the work question are also named in the play question; also it is uncommon to be named many times in one question and not at all in the other. Another interesting feature is that the mean of the peer score is higher than that of the individual score. This relation holds even when only the play network is considered, which suggests that being a good student in primary school increases popularity (see Table A.1.2 in the Appendix).

Table 1.1 presents descriptive statistics on the original data set and the final samples for the variables to be used when performing estimates. Even though the family survey provides a wide range of socioeconomic information, there is incomplete information on some variables for some students. Naturally, this deficiency complicates the calculation of peer variables. In order to minimize the number of omitted observations, the regressions include only a few variables ---all of which have a low percentage of missings and are commonly used in studies on education.[16]

[14] It may happen that students named children that either were absent on the date of the evaluation or for whom there is no information on family characteristics. Taking into account those students (who cannot be considered in the final estimations), children on average named 2.7 distinct peers, 15% named only one peer in the play question and 14.6% named only one peer in the work question. There are also 249 individuals who are isolated (i.e., who did not name anybody in the two questions).

[15] Note that student i 's naming of student j does not imply that the two are actually friends. It might instead be the case that i would like to be friends with j (say, because he admires or likes j even if they are not currently close friends. What matters, however, is that j will likely to exert influence on i for no reason other than i considers j as part of his reference group. The study's identification strategy assumes that children are influenced only by those classmates whom they name.

[16] Table A.1.5 shows estimates including a larger number of regressors and hence employing fewer observations. The coefficients are similar to the main estimates.

The final sample for each test (math, reading, and science) consists of all individuals for whom we have valid information not only on their score and family characteristics but also on their peers' scores and characteristics as well as on their peers' peers, and their peers' peers' peers characteristics. The number of observations varies in the final data set for each test because the tests were given on separate dates and some (i.e., absent) children did not take them all. The final samples exhibit slightly better socioeconomic characteristics and test scores, though it makes up a substantial part (between 75 and 80 percent) of the original sample.

Table 1.1 – Descriptive statistics

	Full sample			Reading final sample			Math final sample			Science final sample		
	Obs.	Mean	S.D.	Obs.	Mean	S.D.	Obs.	Mean	S.D.	Obs.	Mean	S.D.
Female	8805	0.49	0.5	6953	0.51	0.50	6593	0.51	0.50	6598	0.51	0.50
Repeated (1 or more ys)	8781	0.31	0.46	6953	0.26	0.44	6593	0.25	0.44	6598	0.25	0.43
Mother: ≤ primary	7722	0.3	0.46	6953	0.28	0.45	6593	0.28	0.45	6598	0.28	0.45
Moth: incompl HS	7722	0.42	0.49	6953	0.42	0.49	6593	0.42	0.49	6598	0.42	0.49
Moth: HS-incompl college	7722	0.15	0.36	6953	0.16	0.37	6593	0.16	0.37	6598	0.16	0.37
Moth: compl college	7722	0.13	0.33	6953	0.14	0.34	6593	0.14	0.35	6598	0.14	0.35
Reading score	8605	501.6	101.9	6953	511.6	99.00						
Math score	8371	501.6	102.4				6593	512.53	100.08			
Science score	8402	501.1	101.1							6598	512.00	94.98
Number of peers named	8623	2.42	1.04	6953	2.38	0.91	6593	2.33	0.91	6598	2.33	0.91
Other variables in the data set not included in the final samples to minimize loss of observations												
Father: ≤ primary	7259	0.32	0.47									
Fath: incompl HS	7259	0.45	0.5									
Fath: HS-incompl college	7259	0.14	0.35									
Fath: compl college	7259	0.09	0.29									
Numb. persons in house	7862	4.92	1.85									
Books: less 10	6979	0.28	0.45									
Books: btw 10 & 50	6979	0.35	0.48									
Books: more than 50	6979	0.37	0.48									
Slum	7862	0.12	0.32									
Wealth	5823	2.41	1.32									

Notes: Dummies for the number of books at home are defined based on the question: Approximately how many books are there in the household? With the following options: (1) there are no books, (2) there are less than 10, (3) there are between 10 and 50, (4) there are more than 50. The wealth index considers different durable goods a household may own.

As mentioned in Section 1.3, the identification strategy would be invalidated if children sort with children who are similar in an unobserved way that is correlated with their academic achievement. In line with Sacerdote (2001), Bayer, Ross, and Topa (2008) and with Drago and Galbiati (2012), I analyze whether there is sorting in terms of observables within a class and conclude that there is none in terms of socioeconomic background. Bayer et al. (2008) remark that this does not prove that there is no sorting on unobservables but it does suggest the assumption is a reasonable one. In one of these tests, I run two OLS regressions for own socioeconomic characteristics (mother's education index and wealth index) as a function of the corresponding peer characteristic and control for classroom fixed effects. Table 1.2 shows that neither of the two coefficients turn out to be significant. Table A.1.3 shows that although 45.5 percent of students whose wealth index is above the class median named only peers whose wealth index is also above that median, also 43.3 percent of students whose wealth index is *below* the class median also named only peers whose wealth index is above the class median.[17] It can also be seen that students whose mother's education is or above the class median have peers similar to those of students whose mother's education is below the class median. In this sense, there does not seem to be self-selection into peers of similar socioeconomic characteristics. There is a preference for interacting with individuals of socioeconomic characteristics above the class median but this applies for both those whose own characteristics are above and below the class median.

Table 1.2 – Own socioeconomic characteristics regressed on peer characteristics. Evidence of no selection on observables

	Mother educ. index	Wealth index
Same variable for peers	-0.01	0.03
	(0.02)	(0.03)
Obs.	6,953	4,928
Number of clusters	318	309
Classroom fixed effects	Yes	Yes

Notes : The mother's education index ranges from 1 to 9 and summarizes different levels of education (years of education cannot be reconstructed precisely).

Standard errors (in parentheses) are clustered at the school level. Both peer scores and own score have been normalized.

The wealth index weights different durable goods a household may own through factor analysis. The durables considered are: boiler, washing machine, phone, car, microwave and computer.

*** p<0.01, ** p<0.05, * p<0.1

1.5 Results

In this section I present estimates of peer effects in achievement for reading, math and science standardized tests while following the strategy outlined in Section 1.3. For computation of the reference group, all distinct peers named in the two questions (play and work) were weighted equally.[18] Table 1.3 reports OLS estimates both with and without classroom fixed effects.[19] When classroom fixed effects are included, the OLS estimates indicate that endogenous effects are significant only for math (and are very small). Table 1.4 presents 2SLS estimates with

[17] For instance, 16 percent of students with wealth at or above the class median did not name any peer at or above the class median (ie. only named peers below the class median) while 19 percent of students with wealth below the class median did not name any peer at or above the class median.

[18] Table A.1.9 presents other reference-group specifications.

[19] In the final sample there are 395 classrooms or groups in the reading estimates, 392 in the math data set, and 394 for science.

standard errors clustered at the school level.[20] Observe that the F-tests of the excluded instruments in the first stage for the math, reading, and science test indicate that weak instruments are not a concern.

Table 1.3 – OLS estimates

	Reading	Math	Science	Reading	Math	Science
Endogenous effect	0.15***	0.29***	0.25***	-0.02	0.04**	0.01
	(0.02)	(0.03)	(0.05)	(0.02)	(0.02)	(0.02)
Own characteristics						
Female	0.12**	0.00	-0.03	0.11**	0.01	-0.02
	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)
Repeat	-0.45***	-0.51***	-0.36***	-0.48***	-0.54***	-0.37***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Mother: incompl HS	0.14***	0.10***	0.15***	0.11***	0.07**	0.13***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)
Mother: comp HS-incomp colllege	0.45***	0.31***	0.40***	0.37***	0.25***	0.35***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Mother: compl college	0.67***	0.54***	0.54***	0.58***	0.49***	0.52***
	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)
Contextual effects						
Female	0.00	0.01	0.01	0.04	-0.03	-0.01
	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)
Repeat	-0.05	0.10***	-0.01	-0.17***	-0.11***	-0.12***
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)
Mother: incompl HS	0.14***	0.03	0.06	0.09**	0.01	0.06
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
Mother: comp HS-incomp colllege	0.30***	0.25***	0.26***	0.21***	0.20***	0.22***
	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
Mother: compl college	0.40***	0.28***	0.20***	0.28***	0.26***	0.25***
	(0.06)	(0.06)	(0.07)	(0.07)	(0.06)	(0.06)
Obs.	6953	6593	6598	6953	6593	6598
R-squared	0.26	0.31	0.23	0.11	0.11	0.07
Classroom fixed effects	No	No	No	Yes	Yes	Yes

Notes : Standard errors (in parentheses) are clustered at the school level. Both peer scores and own scores have been normalized.

*** p<0.01, ** p<0.05, * p<0.1

[20] Clustering at the classroom level does not alter the significance of the estimates. It seemed more reasonable to cluster at the school level because clustering at the classroom level would imply assuming zero correlation between classrooms within a school.

Table 1.4 – 2SLS estimates

	Reading	Math	Science
Endogenous effect	0.40***	0.37***	0.22
	(0.11)	(0.13)	(0.16)
Own characteristics			
Female	0.11*	0.02	-0.01
	(0.06)	(0.05)	(0.05)
Repeat	-0.45***	-0.51***	-0.36***
	(0.03)	(0.03)	(0.03)
Mother: incompl HS	0.08***	0.05**	0.12***
	(0.03)	(0.02)	(0.03)
Mother: comp HS-	0.33***	0.22***	0.32***
incomp colllege	(0.04)	(0.04)	(0.04)
Mother: compl college	0.51***	0.43***	0.48***
	(0.05)	(0.04)	(0.05)
Contextual effects			
Female	-0.04	-0.02	-0.01
	(0.07)	(0.05)	(0.06)
Repeat	0.08	0.12	-0.02
	(0.08)	(0.10)	(0.08)
Mother: incompl HS	0.04	-0.04	0.02
	(0.04)	(0.05)	(0.06)
Mother: comp HS-	0.02	0.1	0.12
incomp colllege	(0.09)	(0.08)	(0.10)
Mother: compl college	-0.07	0.06	0.10
	(0.14)	(0.11)	(0.15)
Excluded instruments (first stage)			
Peers' peers motheduc	0.07***	0.06***	0.08***
	(0.02)	(0.02)	(0.02)
Peers' peers' peers motheduc	0.08***	0.07***	0.03
	(0.02)	(0.02)	(0.03)
Obs.	6953	6593	6598
F-test excluded inst	13.89	11.91	10.38
p-val overidentification test	0.81	0.37	0.94
Number of clusters	318	316	318
Classroom fixed effects	Yes	Yes	Yes

Notes : Standard errors (in parentheses) are clustered at the school level. Both peer scores and own score have been normalized.

*** p<0.01, ** p<0.05, * p<0.1

According to the estimates in Table 1.4, endogenous effects are large and highly significant for reading and math but are not significant for science.[21] A one--standard deviation increase in peers' score increases own performance by 40 percent of a standard deviation for reading and by 37 percent of a standard deviation for math.[22] This is smaller than but still comparable to the effect of having a mother who completed college. It is also similar in magnitude to the impact of having been held back in school at least one year. These estimates lie between those obtained by Graham (2008) for kindergarten students and those reported by Lin (2010) for adolescents, suggesting that peers' influence on academic achievement decreases with age. A straightforward measure of the social multiplier cannot be computed within this framework: some children are named more often than others, so the aggregate sum of peers' scores is not directly comparable to the sum of individual scores.

Exogenous effects are never significant, suggesting that social interactions operate mainly through peers' actions. This finding confirms the same results reported by Laschever (2009) and De Giorgi et al. (2010).[23] Cooley (2010) obtains some counterintuitive results regarding the impact of contextual effects and argues that, after conditioning on peer achievement the expected sign of contextual effects is ambiguous. In turn, Lin (2010) finds that many peers' characteristics are significant determinants of GPA performance.

The higher 2SLS than OLS estimates may come as a surprise. The OLS estimates may be biased downward because of classical measurement error in peers' scores. Also, it could be due to the presence of heterogeneous peer effects on students' scores. In the latter case, (consistent) OLS estimates would report an average effect across all students whereas the 2SLS estimand is a weighted average of responses to a unit change in treatment for subjects whose treatment is affected by the instrument (Angrist and Imbens, 1995).[24] The weighting function might reflect

[21] The correlation among the tests is around 0.6. The reason why peer effects seem to be not significant for science is a question that should be further explored. One possible explanation is that math and reading tests assess cognitive skills that may improve in response to class interaction with one's peers whereas the science test is likely to contain more questions whose answers require more memory. An interesting fact is that there seems to be somewhat higher levels of motivation toward science, which also is perceived to be less difficult than math or reading. Table A.1.4 shows how often children believe that they almost always understand what they have been taught; the percentage is higher in science than in math or reading. Also, the percentage of children who enjoy what they are taught ``almost always'' is higher in science than in math or reading.

[22] In contrast, Carrell et al. (2009) find strong effects in math and science but not significant effects in foreign language courses among students at the US Air Force Academy.

[23] Laschever (2009) examines how social ties formed during World War~I affected a veteran's likelihood of having a job in 1930.

[24] Two-stage least squares can estimate a local average treatment effect in the presence of heterogeneous treatment effects provided the monotonicity condition is satisfied. This additional restriction requires that the instrumental

how the compliers (i.e., peers whose scores are affected via endogenous or exogenous social interactions) are distributed over the range of scores.[25] That 2SLS estimates are higher than OLS estimates could result from peers effects being greater for subjects whose peers are themselves positively affected by other peers (instrument compliers). Note that De Giorgi et al. (2010) also find a negative bias in the OLS estimates; their explanation applied to this context would suggest the presence of network-specific shocks that work in different directions.

1.6 Alternative specifications

In this section I provide some alternative specifications for the previously reported results. Table A.1.5 reports estimates following the same specification as in Table 1.4 but including additional individual and peer characteristics. This reduces the sample size significantly since for an individual to be included in the estimation her socioeconomic characteristics, her peers' peers' socioeconomic characteristics (and her peers' peers' peers' characteristics) need to be complete. The estimates in Table A.1.5 are similar to those reported in Table 1.4. Table A.1.6 presents results including the information provided by approximately 700 students who are not included in the estimates. For these students, there is complete information on their scores and characteristics but not on their peers (either because they did not name any or, more often, because the peers they named were absent on testing day or did not provide information on socioeconomic characteristics); hence they cannot be included in the regression. Nonetheless, these observations provide valuable information that can be used to compute---for other students---their peers' peers' characteristics and their peers' peers' peers characteristics.[26] The estimated endogenous coefficients are slightly larger than those reported in Table 1.4.

variable affects treatment intensity in the same direction for everyone (Angrist and Imbens, 1995). There may be heterogeneous effects due to observable characteristics (i.e., treatment effects could be homogeneous after conditioning for observable characteristics); alternatively, individuals with the same characteristics may respond differently to the treatment.

[25] Angrist and Imbens (1995) show that 2SLS in a framework of variable treatment intensity yields an average of the derivative, with weight given to each possible value of the treatment variable in proportion to the instrument-induced change in the variable's cumulative distribution function at that point. In addition, 2SLS with covariates generates an average of covariate-specific average causal responses, and 2SLS with multiple instruments generates a weighted average of average causal responses for each instrument. Because the model estimated here involves variable treatment intensity as well as multiple instruments and covariates, the resulting weights are a combination of all these factors.

[26] I correct peers' peers' characteristics and peers' peers' peers characteristics when these observations were named as direct peers by multiplying by a factor that weights peers without considering them. For example: if A named B who named C and D and if D did not name anybody (or named someone who was absent), I use D's information to compute A's peers' peers' characteristics but then correct by a factor that---instead of weighting D's peers and C's peers equally when computing B's peers' peers' characteristics---assigns all the weight to C, who is the only one with valid information on his friends.

Table A.1.7 replicates the estimates of Table 1.4 but while considering only those classrooms in which, among peers, selection on observables (as measured by the correlation between an individual's characteristic and her peers' characteristic at the classroom level) is relatively low. The first three columns of the table present the estimates for individuals for whom the within-classroom correlation between the student's mother education and their peers' mother's education is lower than 0.3. For the reading test, endogenous effects remain significant and large in magnitude while they are not longer significant for math. The next three columns show the estimates for individuals for whom the correlation between being a repeater and having peers who are repeaters is lower than 0.3. In this case, estimates are significant and large in magnitude for all three tests: reading, math, and science.

Table 1.4 included school-level dummies for mother's education and for peers' mothers' education while using as instruments an index of peers' peers' mothers' education and peers' peers' peers' mothers' education. The instruments are variables with values that range from 1 to 9 and reflect different levels of education. A variable indicating years of education cannot be precisely reconstructed.[27] In Table A.1.8 I perform an additional estimation in which---instead of including dummies for different levels of mother education---I attempt to reconstruct years of schooling with some measurement error.[28] In this way, I express covariates and instruments in exactly the same way. The results are quite similar to those in Table 1.4: endogenous peer effects are large for reading and math but not significant for science, and exogenous effects are never significant.

Table A.1.9 reports the endogenous coefficient estimates obtained when considering alternative reference groups. When using the network information contained in only one question (play or work), the test of the null hypothesis loses some power because in that case there are fewer valid observations (fewer students have information on their peers and their peers' peers) and the remaining network is also weakened (many individuals have fewer peers). [29] Overall, the endogenous coefficient estimates do not differ substantially across the different specifications,

[27] In the survey mothers were asked to mark yes/no to the following options: (1) did not attend primary, (2) incomplete primary, (3) complete primary, (4) 1 or 2 years of secondary school, (5) 3 years of secondary school, (6) 4 or 5 years of secondary school, (7) complete high school (6 years), (8) incomplete college, (9) complete college.

[28] This variable ranges from 0 to 16. For instance, I assigned 16 years of schooling to mothers who have completed college even though college in Uruguay may take more than 4 years. For answers indicating 1 or 2 years of secondary school, I assumed it was only 1 (i.e., 7 years of schooling).

[29] Recall that a maximum of two peers could be named in each question.

but they are larger and more significant when considering only the peers named in the work question than in the play question. This result could be due to children choosing better students as their reference group for study purposes. The mean of peer scores is higher in the work than in the play network. However, most children are named in both questions (only 11 percent were named at least once in the play question and not named in the work question). I also estimate a specification in which a peer who is named in both questions is weighted more than a peer who is named in only one.[30] In this case, the *F*-tests of the excluded instruments for reading, math, and science always reach acceptable levels, and the estimates are only slightly smaller in magnitude than those reported in Table 1.4.

The estimated model is an extension of the standard linear-in-means social interaction model in which student specific reference groups are allowed. This model constrains peer effects to have distributional consequences but no efficiency consequences. As a first attempt to see whether peer effects are heterogeneous among different kinds of students, I estimate peer effects for children with different levels of mother's education separately. Unfortunately, this reduces the significance of most of the estimates (see Table A.1.10). The only endogenous effect that is significant for both reading and math is the one of children whose mothers have finished primary school but did not complete high school. This could be explained by that category being the largest category in the sample (42 percent of children in the sample share this characteristic). It is interesting that, in reading, the peers' mothers' education (contextual effect) is positive and significant only for children whose own mothers have the lowest education levels. Endogenous peer effects are significant for both females and males. The endogenous effect seems to be larger for females in reading but smaller for females in math.

1.7 Potential impact on educational inequality

Social interactions are likely to influence schooling decisions, study habits, and individual aspirations. For this reason, socioeconomic stratification as social networks are forming has a strong influence on the persistence of educational disparities and on broader social inequalities across generations (Benabou, 1996; Durlauf, 1996, 2004; Bowles et al., 2007; Graham, 2011). In this section, I assess the extent to which inequalities in educational outcomes are amplified by peer effects operating in a context of socioeconomic stratification.

In terms of income distribution, Uruguay is the least unequal country in Latin America; however, inequalities in the Uruguayan educational system are large even when compared to other Latin American countries. In the PISA 2009 math tests, Uruguay achieved the highest mean and the

[30] For instance, if a student names A and B in the play question and names A and C in the work question, then the peer score and characteristics are computed while assigning weights of 0.25 to B and C and 0.5 to A.

highest scores at the 95th percentile of all the Latin American countries that participated in the tests. But the scores achieved by the 5th percentile of the distribution were lower than those achieved in Chile and Mexico. Furthermore, Uruguay's dropout rates at age 15 are significantly higher than those in Chile.[31] If the same percentage of 15-year-olds attended high school in both countries, then the observed differences between Uruguayan and Chilean test score distributions could be even larger (this is particularly important when one considers that educational inequalities are likely to translate, through wages, into future socioeconomic inequalities). One possible explanation for the larger disparities in test scores in Uruguay is that socioeconomic segregation may be amplifying educational inequality through peer effects. In the Uruguayan public school system, students are assigned to schools according to their neighborhood of residence. This is a critical factor in determining how neighborhood socioeconomic stratification affects education. To illustrate the level of such stratification present in the data set, I computed some simple ANOVA estimates: 42 percent of the variation in the variable that summarizes students' mother's education is due to between-school variance, and 45 percent of the variation in a wealth index (that considers different durable goods a household may own) can also be attributed to differences between schools.

In order to quantify the potential impact of peers on inequality in a context of socioeconomic segregation, I compare the distribution of the actual reading and math scores with the one resulting from reshuffling peers among the sample of children who have the same number of peers.[32] In other words, if an individual originally named 3 peers, then I assign her randomly 3 new peers that had been named by individuals who in total had named 3 peers (each of these 3 new peers was named by a different student). In this sense, I maintain the degree of ``popularity'' (number of times a child is named by others) and the degree of ``sociability'' (number of peers the child identified) that individuals exhibit in the actual sample. The logic here is that a hypothetical social planner could reassign children to different schools but could not alter how popular and/or sociable they are.[33] I then multiply all the individual characteristics and peer scores and characteristics by the coefficients from the original regressions and add the residuals from the original predicted reading and math scores, respectively. Figure 1.2 compares the actual scores' distributions with the resulting distributions averaged over 100 simulations. As expected,

[31] In Uruguay, only 82 percent of 15-year-olds attended the educational system; in Chile 97 percent did so (2006 data).

[32] I do not reshuffle among the total data set because the distribution of the number of peers named is not uniformly distributed along socioeconomic characteristics. In particular, children belonging to higher socioeconomic strata tend to name slightly more peers. Since children from higher socioeconomic neighborhoods tend to have better scores, it follows that if peers are reshuffled among all individuals in the data set then the mean of the variable for peers' score will increase slightly (given the lower number of peers named by children in poorer neighborhoods) complicating distributional comparisons.

[33] The estimation does rely on the (fairly extreme) assumption that these randomly matched peers would become friends.

changing actual peers into random peers concentrates the distribution more about its mean and reduces its mass in the high and low achieving tails. The actual reading score has a mean of 512 and a standard deviation of 99, whereas the simulated distribution has (the same mean and) a standard deviation of 94.6; the absolute gap between the 95th and 5th percentiles is reduced from 309.4 to 302.6. The distribution of math scores exhibits a reduced standard deviation (from 100 to 90), and the gap between 95th and 5th percentiles is reduced from 313.1 to 286.7 (see Table 1.5). A possible explanation for the lack of a greater reduction in inequality is that actual within-school friendship ties are not driven by socioeconomic background (this was shown in Tables 1.2 and A.1.3). Observe also, that these estimations assume peer effects are homogeneous for all students, the impact of reshuffling students randomly could be much greater if treatment effects were instead heterogeneous among children with different socioeconomic background, in particular, if lower socioeconomic students benefited more from social interactions.

Figure 1.2 – Distributional impact: comparison with random peers

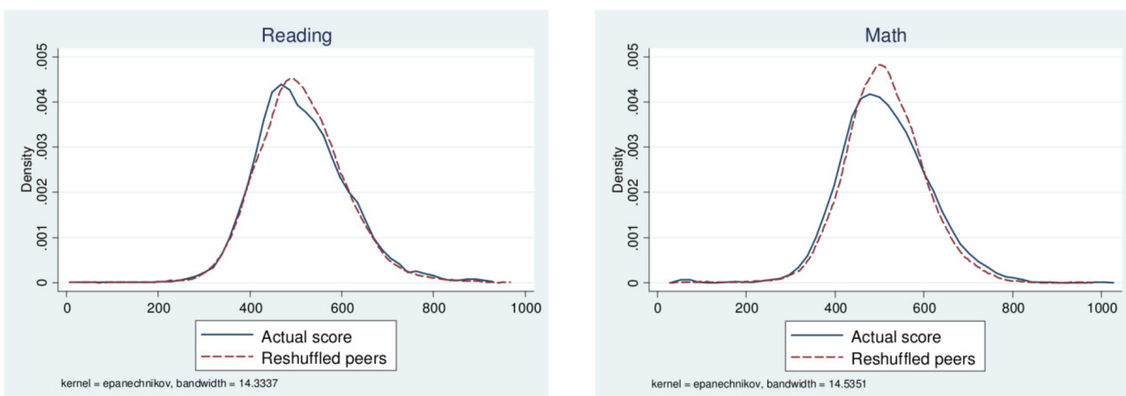


Table 1.5 – Changes in the distribution of reading and math scores

	Reading		Math	
Percentile	Actual score	After reshuffling	Actual score	After reshuffling
5th	369.4	368.6	367.5	376.2
10th	395	397.5	396	406.3
15th	414.2	417.3	418.5	427.2
20th	428.7	434	432.1	442.4
25th	446.3	448.8	447.2	454.9
30th	453.9	461.5	458.4	466.7
35th	468.4	473.1	472.5	478.3
40th	479.5	484.2	480.4	488.5
45th	488.5	494.9	493.9	498.8
50th	501.5	506	505.5	509.1
55th	515.2	517.1	518.7	519.2
60th	528.8	528.9	531.6	530.1
65th	541.1	541.8	544.9	541.8
70th	556.8	555.2	558	555.3
75th	572.4	569.1	573.6	568.7
80th	588.9	586.2	592	582.4
85th	613	606.2	614.4	601.8
90th	642.3	631.4	639	625.4
95th	678.8	671.3	680.7	662.9
95th – 5th	309.4	302.6	313.1	286.7

This is an out-of-sample computational experiment that seeks to proxy (in an extreme way) the possible distributional impact of policies that intervene in the determination of socioeconomic interaction environments for individuals. Durlauf (1996) refers to this type of policies as *associational redistribution*: ``an interactions-based perspective alters the redistributive focus away from policies designed to equalize per-student expenditure to those that attempt to equalize the total school environment" (Durlauf, 1996, p.267). I regard this exercise as useful but am aware of its limitations. First, as Piketty (2000) notes, these policies can provoke controversy because most people consider the choice of one's peers to be an area not within the purview of public policy. Second, evidence regarding the impact of desegregation plans is mixed. Rivkin and Welch (2006, p.1043) review several studies that assess the impact of school desegregation and conclude that the ``effects of integration on black students remains largely unsettled. If there is a marginal consensus, it is that effects are probably small, but beneficial." Third, if peer effects operate mainly via friendship networks, then it will be difficult to determine the impact of moving a child from a school whose average student is from low socioeconomic background to a school whose average student is from a higher average background (or vice versa), since it is not certain that the relocated child would establish any links with children of different characteristics. Evidence from the Add Health data set suggests that mere exposure to more heterogeneous schools does not promote interracial integration per se (Moody, 2001). Also, Carrell, Sacerdote and West (2012) find that grouping low ability students with high ability ones has a negative impact on low ability students. Carrell et al. (2012) interpret this result as grouping low ability with high ability students may have provided more opportunities (relative to random assignment) for increased homophily with low ability students becoming friends among low ability students.[34] Finally, this exercise abstracts from teacher behavior changing in response to student reassignment. Duflo, Dupas and Kremer (2011) conclude that tracking could favor both high- and low-achieving students because it facilitates teachers' adaptation of their instruction level especially when teachers are incentivized to instruct to the top of the distribution. However, the wages of public school teachers in Uruguay are not linked to their students' achievement.

1.8 Comparison with classroom peer effects

In this section I compare the peer effects estimated using network data with social interaction effects (endogenous and exogenous effects altogether) obtained from a within classroom transformation. In line with Ammermueller and Pischke (2009), I exploit the random variation of students' characteristics across classrooms within 6th grade in public schools. That is, I use the same data set but only include in the estimation those public schools for which more than one classroom was evaluated. Peer variables in this case are the average characteristics of the classroom excluding individual *i*. This strategy allows me to only estimate an overall social interaction effect because in this strategy there is no exclusion restriction that enables to solve the reflection problem. A crucial identifying assumption for this strategy is that groups within a school are not conformed in any systematic way (for instance that there is no tracking). For this reason, I only consider public schools where tracking is not allowed in Uruguay. Estimates include school fixed effects and observable teacher characteristics which represent another potential determinant of differential academic attainment between groups within a school.

Table 1.5 reports peer effects following the outlined strategy. While the percentage of students with different levels of mother education and the percentage of repeaters in the classroom do not seem to be significant the percentage of girls in the classroom has a positive and substantial coefficient for all the three tests (reading, math and science). Estimates including an interaction term suggest that the positive effect of a greater percentage of females in the class is only

[34] An alternative hypothesis the authors mention refers to the potential relevance of the presence of middle ability students in order to generate positive peer effects for the lower ability students.

observed for males.[35] In turn, Hoxby (2000) and Lavy and Schlosser (2011) find that both males and females perform better in classrooms where there are more females.

A potential reason why this gender effect may not have been captured in the contextual effects estimates using network data is that within classrooms there seem to be isolated networks in terms of gender. Indeed, substantial signs of homophily in terms of gender can be observed: 92 percent of the peers that girls name are girls and 91 percent of the peers that boys name are boys.

For a subsample of approximately 1100 students there are available not only their scores in 2009 but also scores in reading and math from 2006 when they were in 3rd grade. These tests are also standardized tests set and graded by SERCE in every school.[36] Both for reading and math it appears that own test scores in 3rd grade are highly significant in explaining 6th grade test scores. However, the rest of the classroom's average test scores in 3rd grade do not seem to matter.[37] Note that for this subsample, the percentage of females in the class is no longer significant while the percentage of classmates who have mothers that completed college is highly significant.

[35] Estimates on the impact of being female and of the percentage of females in the class (excluding i) increase substantially when including an interaction between both, probably because including the interaction raises multicollinearity problems.

[36] Segundo Estudio Regional Comparativo y Explicativo. This evaluation was performed in sixteen Latin American countries plus a state in Mexico.

[37] I am considering 3rd grade test score of classmates in 6th grade. These classmates need not have been in the same group when they were in 3rd grade but that is not an issue because all students sat for the same test and I am concerned with the effect of 6th grade classmates.

Table 1.6 – Classroom peer effects

	Reading	Math	Science	Reading	Math	Science	Reading	Math
Own characteristics								
Female	0.15***	0.17***	0.16***	0.48**	0.52**	0.49**	0.06	0.14**
	(0.05)	(0.05)	(0.05)	(0.21)	(0.21)	(0.21)	(0.05)	(0.06)
Repeat	-0.58***	-0.59***	-0.60***	-0.58***	-0.59***	-0.60***	-0.43***	-0.40***
	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)
Mother: incompl HS	0.17***	0.16***	0.17***	0.17***	0.16***	0.18***	0.05	0.09
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)
Mother: comp HS-incomp colllege	0.51***	0.51***	0.52***	0.51***	0.51***	0.52***	0.30***	0.35***
	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)
Mother: compl college	0.67***	0.68***	0.70***	0.68***	0.69***	0.70***	0.52***	0.47***
	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.11)	(0.12)
Reading score (3rd grade primary)							0.47***	
							(0.03)	
Math score (3rd grade primary)								0.48***
								(0.04)
Classroom characteristics								
Females in the class	0.48*	0.49*	0.50*	0.84***	0.88***	0.86***	0.13	0.62
	(0.27)	(0.28)	(0.28)	(0.30)	(0.32)	(0.31)	(0.41)	(0.43)
Female × Females in the class				-0.65	-0.71*	-0.67		
				(0.39)	(0.39)	(0.40)		
Repeaters in the class	-0.46	-0.41	-0.47	-0.45	-0.41	-0.47	-0.03	0.01
	(0.36)	(0.38)	(0.37)	(0.36)	(0.39)	(0.37)	(0.42)	(0.38)
Mother: incompl HS (percentage in class)	-0.31	-0.30	-0.32	-0.32	-0.32	-0.33	0.08	0.27
	(0.33)	(0.33)	(0.34)	(0.33)	(0.33)	(0.34)	(0.31)	(0.34)
Mother: comp HS-incomp colllege (percentage in class)	-0.25	-0.31	-0.26	-0.25	-0.32	-0.26	0.10	0.52
	(0.55)	(0.58)	(0.57)	(0.55)	(0.58)	(0.57)	(0.63)	(0.60)
Mother: compl college (percentage in class)	0.35	0.33	0.36	0.32	0.29	0.33	1.18**	1.16**
	(0.45)	(0.47)	(0.46)	(0.45)	(0.47)	(0.46)	(0.44)	(0.53)
Reading score (3rd grade primary) class average							-0.11	
							(0.14)	
Math score (3rd grade primary) class average								-0.11
								(0.10)
Teacher characteristics								
Female teacher	0.01	0.02	0.01	0.01	0.02	0.01	0.01	-0.09
	(0.07)	(0.05)	(0.07)	(0.07)	(0.05)	(0.07)	(0.04)	(0.06)
Teacher experience (years)	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.02	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Tearcher experience squared	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Obs.	2,114	2,048	2,114	2,114	2,048	2,114	1,122	1,135
R-squared	0.13	0.13	0.13	0.13	0.14	0.13	0.33	0.34
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors (in parentheses) are clustered at the school level. Both peer scores and own score have been normalized. All peer variables refer to averages for the 6th grade class excluding the individual.
*** p<0.01, ** p<0.05, * p<0.1

To sum up, to estimate classroom peer effects I used a subsample (only public schools where more than one group was evaluated) of the data set for which I had estimated network peer effects. The identification strategy followed in this section does not enable to identify endogenous and contextual effects separately so estimates are not directly comparable. The estimates commented above suggest that there are significant peer effects at the class level captured by the percentage of females in a larger sample and then by classmates' mother education in the smaller one. However, behavioral effects cannot be disentangled.

1.9 Conclusions

The findings reported in this chapter indicate significant peer effects in academic achievement at the primary school level. Estimates suggest that there are strong endogenous peer effects and thereby a large social multiplier. A one--standard deviation increase in a student's peers' score increases the focal student's scores by 40 percent of a standard deviation in reading and 37 percent of a standard deviation in math. This magnitude is smaller yet comparable to that of one's mother having completed college. In contrast, contextual effects do not seem to be significant suggesting that social interactions operate mainly through peers' actions.

In making these estimates I apply a recently developed identification strategy to a unique data set of primary schools in Uruguay. This strategy enables me to solve the reflection problem and thus disentangle endogenous from contextual effects, two social interaction effects with distinct policy implications. The intuition behind this framework is that peers' peers who are not the focal student's peers can only affect that student's behavior indirectly by influencing the behavior of her peers. In other words, it is assumed that peers' peers' characteristics can be excluded from the structural equation explaining a student's scores and thus can serve as instrumental variables that help explain the peers' scores. Correlated effects are handled with by including classroom fixed effects. Standard errors are clustered at the school level. Also, I estimated peer effects using an alternative identification strategy that exploits variations in peer characteristics between classrooms within a school. Estimates suggest that peer effects are also substantial at the classroom level but the latter strategy does not allow one to identify endogenous from contextual effects.

The high significance of peer effects signals their potential importance as amplifiers of educational inequalities in socioeconomically stratified environments. That is, if it matters whom one interacts with at school, then differences in social environment will contribute to polarized outcomes. According to the exercise performed in Section 1.7, if peers were assigned randomly then the standard deviation in scores would decrease by roughly 5--10 percent.

Social interactions can be viewed as affecting individuals' preferences, constraints and expectations (Manski, 2000). However, research on specific mechanisms remains scarce. Some of the most notable contributions in this respect are Akerlof and Kranton (2002), Kremer and Miguel (2007), Lazear (2001), Austen-Smith and Fryer (2005) and Lavy and Schlosser (2011). There is also relevant evidence from other disciplines, including anthropology and social psychology.[38] In further research it would be particularly interesting to explore the mechanisms through which peer effects operate.

Appendix 1

Table A.1.1 – Distribution of students according to number of peers named

Reading final sample				
	Peers in party question	**Peers in work question**	**Total distinct peers**	
	Obs	Obs	Obs	%
0	333	265	–-–	–-–
1	1920	1887	1147	16.5
2	4700	4801	2930	42.1
3	–-–	–-–	1973	28.4
4	–-–	–-–	903	13.0
Total	6,953	6,953	6,953	100
Percentage that named one peer twice				39.5
Percentage that named two peers twice				25.4
Math final sample				
	Peers in party question	**Peers in work question**	**Total distinct peers**	
	Obs	Obs	Obs	%
0	352	274	–-–	–-–
1	1997	1962	1204	18.3
2	4244	4357	2799	42.5
3	–-–	–-–	1806	27.4
4	–-–	–-–	784	11.9
Total	6,593	6,593	6,593	6593
Percentage that named one peer twice				39.3%
Percentage that named two peers twice				24.1%

Note : Students name peers only once but the three samples (reading, math and science) have different number of observations because the tests took place at different dates. Reported values for final samples (ie. after dropping observations with incomplete information on own or peer scores and characteristics).

[38] Doise and Mugny (1984) documented that children can solve problems more effectively when working in pairs or small groups than when working alone: the resulting conflict of views enables one child's perceptions and strategy to stimulate the other child to develop new strategies. A widely studied case of peer pressure in the context of educational attainment involves black students discouraging their black peers from excelling academically, which is viewed as 'acting white' behavior (Fordham and Ogbu, 1986). Individuals exposed to such social interactions are discouraged from investing in education because they fear being rejected by their social peer group.

Table A.1.2 – Mean individual and peer scores by network

Network	Mean individual score	Mean peer score
Reading		
Play and work	511.6	525.9
Play	514.2	522.7
Work	513.8	534.5
Math		
Play and work	512.5	528.0
Play	515.3	524.3
Work	514.9	537.8
Science		
Play and work	512.0	523.8
Play	514.1	520.9
Work	513.9	531.1
School type (reading final sample)		
Private schools	577.1	591.2
Ordinary public schools	516.9	530.0
Full-time (public)	488.4	505.3
Critical social context (public)	463.6	478.2
Rural (public)	476.9	477.9

Note: Reported values for final samples (ie. after dropping observations with incomplete information on own or peer scores and characteristics). Scores in the original sample were standardized to a mean of 500 and a standard deviation of 100.

Table A.1.3 – Distribution of students' and their peers' characteristics relative to the class median

% of peers named whose wealth is above or equal to the class median	Distribution of students whose wealth is above or equal to the class median	Distribution of students whose wealth is below the class median
0%	15.89%	19.10%
25%	0.54%	0.56%
33%	5.90%	4.23%
50%	19.12%	21.50%
67%	9.44%	9.17%
75%	3.54%	2.19%
100%	45.56%	43.25%
Total	100%	100%
% of peers named whose mothers' education is above or equal to the class median	**Distribution of students whose mothers' education is above or equal to the class median**	**Distribution of students whose mother's education is below the class median**
0%	10.55%	13.82%
25%	0.84%	1.11%
33%	5.65%	6.18%
50%	21.52%	21.78%
67%	12.31%	13.19%
75%	5.17%	5.11%
100%	43.97%	38.81%
Total	100%	100%
	Students above or equal to the class median	**Students below the class median**
Average % of peers with wealth above the median	66.18%	63.30%
Average % of peers with mothers' education above the median	68.90%	64.66%

Table A.1.4: Degree of difficulty and preferences for reading, math, and science.

Can you easily understand what is taught in class?			
	Reading	Math	Science
Almost always	40.00%	35.70%	44.00%
Sometimes	50.70%	54.10%	47.60%
Almost never	9.40%	10.20%	8.40%
Do you like what is taught in class?			
	Reading	Math	Science
Almost always	59.20%	65.00%	67.60%
Sometimes	33.50%	30.10%	25.80%
Almost never	7.30%	4.90%	6.60%

Note: The two questions (degree of difficulty and preferences) were asked for the three subjects separately. Frequencies reported.

Table A.1.5 – 2SLS estimates including other regressors

	Reading	Math	Science
Endogenous effect	0.35**	0.37*	0.10
	(0.15)	(0.22)	(0.17)
Own characteristics			
Female	0.12*	0.02	0.06
	(0.06)	(0.05)	(0.05)
Repeat	-0.41***	-0.50***	-0.32***
	(0.03)	(0.03)	(0.03)
Mother: incompl HS	0.04	0.03	0.07**
	(0.03)	(0.03)	(0.03)
Mother: comp HS-	0.22***	0.13***	0.23***
incomp colllege	(0.04)	(0.04)	(0.05)
Mother: compl college	0.37***	0.30***	0.36***
	(0.06)	(0.05)	(0.05)
Numb. persons in house	-0.03***	-0.01**	-0.03***
	(0.01)	(0.01)	(0.01)
Books: btw 10 & 50	0.11***	0.10***	0.10***
	(0.03)	(0.04)	(0.03)
Books: more than 50	0.24***	0.23***	0.27***
	(0.03)	(0.04)	(0.04)
Preschool age 2, 3 or 4	-0.02	0.03	0.03
	(0.04)	(0.04)	(0.04)
Preschool age 5 or never	-0.09*	-0.02	0.04
	(0.05)	(0.05)	(0.05)
Slum	-0.08*	-0.05	-0.12***
	(0.04)	(0.04)	(0.05)
Contextual effects			
Female	-0.03	-0.02	-0.08
	(0.07)	(0.05)	(0.06)
Repeat	0.08	0.12	-0.07
	(0.08)	(0.13)	(0.08)
Mother: incompl HS	0.04	-0.08	0.01
	(0.04)	(0.06)	(0.06)
Mother: comp HS-	0.06	0.06	0.15*
incomp colllege	(0.09)	(0.09)	(0.08)
Mother: compl college	-0.01	-0.00	0.10
	(0.14)	(0.13)	(0.12)
Numb. persons in house	0.02*	0.01	-0.01
	(0.01)	(0.01)	(0.01)
Books: btw 10 & 50	-0.06	-0.04	-0.01
	(0.06)	(0.07)	(0.05)
Books: more than 50	-0.08	-0.08	0.08
	(0.08)	(0.09)	(0.08)
Preschool age 2, 3 or 4	0.05	0.01	0.03
	(0.06)	(0.05)	(0.06)
Preschool age 5 or never	-0.06	-0.10	-0.06
	(0.07)	(0.07)	(0.08)
Slum	-0.03	-0.03	0.07
	(0.06)	(0.06)	(0.06)
Obs.	5,674	5,369	5,375
F-test excluded inst	7.61	4.3	10.92
Classroom fixed effects	Yes	Yes	Yes

Notes: Standard errors (in parentheses) are clustered at the school level. Both peer scores and own score have been normalized.

*** p<0.01, ** p<0.05, * p<0.1

Table A.1.6 – 2SLS estimates using additional information

	Reading	Math	Science
Endogenous effect	0.43***	0.40***	0.25
	(0.12)	(0.13)	(0.17)
Own characteristics			
Female	0.10*	0.01	-0.01
	(0.06)	(0.05)	(0.05)
Repeat	-0.44***	-0.50***	-0.35***
	(0.03)	(0.03)	(0.03)
Mother: incompl HS	0.08***	0.06**	0.12***
	(0.03)	(0.02)	(0.03)
Mother: comp HS-	0.33***	0.22***	0.31***
incomp colllege	(0.04)	(0.04)	(0.05)
Mother: compl college	0.50***	0.43***	0.48***
	(0.05)	(0.04)	(0.05)
Contextual effects			
Female	-0.03	0.00	0.00
	(0.07)	(0.05)	(0.06)
Repeat	0.1	0.15	0.01
	(0.08)	(0.1)	(0.09)
Mother: incompl HS	0.04	-0.05	0.02
	(0.04)	(0.05)	(0.06)
Mother: comp HS-	0.01	0.09	0.11
incomp colllege	(0.1)	(0.08)	(0.11)
Mother: compl college	-0.09	0.05	0.08
	(0.14)	(0.11)	(0.15)
Obs.	6953	6593	6598
F-test excluded inst	13.46	11.62	10.62
p-val overidentification test	0.75	0.37	0.91
Number of clusters	319	320	322
Classroom fixed effects	Yes	Yes	Yes

Notes: Standard errors (in parentheses) are clustered at the school level. Both peer scores and own score have been normalized.

*** p<0.01, ** p<0.05, * p<0.1

Table A.1.7 – Estimations *excluding* classrooms that exhibit some selection on observables among peers

	Classrooms with low correlation among individual's and peers' mothers' education			Classrooms with low correlation among individual's and peers' being repeaters		
	Reading	Math	Science	Reading	Math	Science
Endogenous effect	0.34**	0.28	0.18	0.42***	0.38***	0.36**
	(0.14)	(0.18)	(0.19)	(0.12)	(0.15)	(0.16)
Own characteristics						
Female	0.11*	0.02	-0.01	0.1	0.01	0.08
	(0.06)	(0.05)	(0.06)	(0.07)	(0.06)	(0.07)
Repeat	-0.46***	-0.50***	-0.36***	-0.44***	-0.50***	-0.35***
	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Mother: incompl HS	0.09***	0.07***	0.12***	0.09***	0.05*	0.13***
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Mother: comp HS-	0.35***	0.22***	0.32***	0.31***	0.20***	0.27***
incomp colllege	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)
Mother: compl college	0.52***	0.44***	0.49***	0.47***	0.45***	0.47***
	(0.06)	(0.05)	(0.05)	(0.07)	(0.06)	(0.06)
Contextual effects						
Female	-0.02	-0.02	-0.02	-0.02	0.01	-0.06
	(0.08)	(0.05)	(0.06)	(0.08)	(0.06)	(0.07)
Repeat	0.04	0.05	-0.06	0.07	0.14	0.01
	(0.09)	(0.13)	(0.09)	(0.08)	(0.11)	(0.09)
Mother: incompl HS	0.04	-0.04	0.03	0.06	-0.07	-0.03
	(0.05)	(0.05)	(0.07)	(0.05)	(0.06)	(0.07)
Mother: comp HS-	0.04	0.1	0.11	-0.03	-0.02	-0.04
incomp colllege	(0.1)	(0.08)	(0.11)	(0.1)	(0.09)	(0.11)
Mother: compl college	0.02	0.07	0.12	-0.02	0.04	-0.04
	(0.15)	(0.13)	(0.16)	(0.16)	(0.14)	(0.17)
F-test excluded inst	7.97	5.92	7.22	9.34	9.51	7.34
p-val overid. test	0.67	0.56	0.85	0.83	0.2	0.82
Obs.	6095	5680	5690	4426	4127	4098
Classroom fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes : Standard errors (in parentheses) are clustered at the school level. Both peer scores and own score have been normalized.

*** p<0.01, ** p<0.05, * p<0.1

Table A.1.8 – Years of schooling instead of school dummies

	Reading	Math	Science
Endogenous effect	0.34**	0.37*	0.09
	(0.17)	(0.21)	(0.2)
Own characteristics			
Female	0.12**	0.01	-0.02
	(0.06)	(0.05)	(0.05)
Repeat	-0.44***	-0.49***	-0.35***
	(0.03)	(0.03)	(0.03)
Moth. years of schooling	0.05***	0.04***	0.05***
	(0)	(0)	(0)
Contextual effects			
Female	-0.03	-0.01	-0.01
	(0.08)	(0.05)	(0.06)
Repeat	0.04	0.13	-0.06
	(0.09)	(0.14)	(0.09)
Moth. years of schooling	0	0.01	0.02
	(0.02)	(0.01)	(0.02)
Excluded instruments (first stage)			
Peers' peers moth. yrsch	0.05***	0.04***	0.04***
	(0.01)	(0.01)	(0.01)
F-test excluded inst	17.5	15.5	15.25
Obs.	6953	6593	6598
Classroom fixed effects	Yes	Yes	Yes

Table A.1.9 – Other reference group specifications (endogenous effects)

	Reading	Math	Science
Play network	0.37	0.30**	0.31*
	(0.27)	(0.14)	(0.17)
F-test excluded inst	3.21	8.3	8.12
Obs.	6458	6057	6054
Work network	0.56***	0.42**	0.13
	(0.11)	(0.21)	(0.15)
F-test excluded inst	13.69	6.32	14.55
Obs.	6529	6160	6141
Weighting peers named twice more	0.37***	0.34**	0.2
	(0.11)	(0.13)	(0.15)
F-test excluded inst	13.96	11.79	12.02
Obs.	6953	6953	6598

Notes: Standard errors (in parentheses) are clustered at the school level. Both peer scores and own score have been normalized.

*** p<0.01, ** p<0.05, * p<0.1

Table A.1.10 – Heterogeneous effects

	Mother's education				Gender	
	≤ Primary	Incompl HS	HS - incompl College	Compl College	Females	Males
Reading						
Endogenous effect	-0.2	0.33**	1.49	-0.14	0.59***	0.44***
	(0.23)	(0.14)	(0.89)	(0.61)	(0.16)	(0.17)
Contextual effects						
Female	0.01	-0.04	0.04	0.07	-0.02	-0.08
	(0.11)	(0.09)	(0.29)	(0.28)	(0.11)	(0.11)
Repeat	-0.29*	0.07	0.77	-0.34	0.12	0.1
	(0.16)	(0.1)	(0.79)	(0.54)	(0.12)	(0.12)
Mother: incompl HS	0.21***	-0.02	-0.2	0.32	-0.03	0.11
	(0.07)	(0.06)	(0.31)	(0.33)	(0.07)	(0.07)
Mother: comp HS-	0.39**	-0.02	-0.61	0.54	-0.1	0.08
incomp colllege	(0.17)	(0.11)	(0.36)	(0.39)	(0.15)	(0.13)
Mother: compl college	0.44	0.04	-1.2	0.41	-0.37	0.03
	(0.3)	(0.16)	(0.74)	(0.48)	(0.24)	(0.17)
F -test excluded inst	6.4	14.13	2.04	1.2	8.94	7.95
Obs.	1924	2919	1038	868	3549	3397
Math						
Endogenous effect	0.18	0.42***	-0.44	0.42	0.35**	0.49***
	(0.21)	(0.23)	(0.54)	(0.97)	(0.17)	(0.17)
Contextual effects						
Female	-0.05	0.03	0.03	-0.22	-0.08	-0.05
	(0.09)	(0.09)	(0.17)	(0.24)	(0.08)	(0.08)
Repeat	-0.01	0.21	-0.83	0.04	0.12	0.16
	(0.15)	(0.17)	(0.58)	(0.9)	(0.13)	(0.14)
Mother: incompl HS	0.07	-0.11	0.1	0.14	-0.05	-0.09
	(0.08)	(0.07)	(0.22)	(0.32)	(0.06)	(0.08)
Mother: comp HS-	0.08	0.04	0.33	0.2	0.18*	0.05
incomp colllege	(0.15)	(0.1)	(0.27)	(0.29)	(0.1)	(0.12)
Mother: compl college	0.18	0.05	0.66	0.03	0.16	0.02
	(0.22)	(0.17)	(0.36)	(0.29)	(0.17)	(0.16)
F -test excluded inst	11.31	6.31	2.29	0.63	9.09	8.01
Obs.	1791	2761	997	844	3363	3222

Notes : Standard errors (in parentheses) are clustered at the school level. Both peer scores and own score have been normalized.

*** p$<$0.01, ** p$<$0.05, * p$<$0.1

2 The social determinants of junior high dropout in Uruguay

This chapter analyzes the determinants of junior high dropouts in Uruguay with particular emphasis on peer effects from classmates in the last year of primary school. I exploit presumably random variations in student composition among 6th grade groups within primary schools and control for students' individual characteristics and teacher's characteristics. While I find positive peer effects of classmates' socioeconomic background, there seems to be a negative effect of classmates' scores on junior high attendance. The latter result could suggest that having high achieving classmates may deteriorate low achieving students' confidence regarding schooling goals.

2.1 Introduction

Dropout rates at secondary education are substantial in developing countries. According to the World Bank, while in 2010 gross enrollment rates at secondary school in High Income countries were situated at 101 percent, in Middle Income and Low Income countries they were at 71 and 41 percent, respectively.[39] Although Uruguay exhibits universal attendance at primary school, the gross enrollment rate at secondary school is around 88 percent, a lower rate than that of Brazil and Chile.[40] Dropout rates in Uruguay are worrisome not only at senior high (grades 10-12) but also at junior high (grades 7 to 9) despite the latter being compulsory.[41] Indeed, according to household survey data in 2011, 11.4% of adolescents aged from 13 to 15 years old did not attend either primary, junior high or vocational schools.

Not completing compulsory education severely impacts on future earnings and wealth. Oreopoulos (2007) argues that lifetime wealth increases by about 15% with an extra year of compulsory schooling based on data for the US, Canada and the UK. Oreopoulos (2007) argues that youths ignore or heavily discount long term wealth consequences when deciding to drop out of school. In turn, Eckstein and Wolpin (1999) conclude that adolescents who drop out of high school have very different traits (ability, motivation, expectations about the rewards from graduation and preferences for leisure) than those who graduate and that if they were forced to remain in school for more years the increase in graduation rates would not be substantial.

Studying the determinants of dropout decisions is particularly relevant at least for developing countries. Along with family background and the institutional features of the educational system, peers appear as a major potential influence in dropout decisions. This chapter looks at peer effects from classmates in the last year of primary school on junior high dropout decisions.

[39] The gross enrollment rate is defined as total enrollment in secondary education, regardless of age, expressed as a percentage of the population of official secondary education age.

[40] Based on data from Anuario Estadístico de Educación 2009. Ministerio de Educación y Cultura.

[41] Junior and senior high can be accomplished either at non vocational schools (liceos) or at vocational schools (Universidad del Trabajo del Uruguay, UTU).

Among the economics literature there are several studies that attempt to identify peer effects in educational outcomes (mainly test scores) and suggest peers may be of particular importance.[42] However, there is scarce literature on the relevance of peer effects in drop out choices. Gaviria and Raphael (2001) find that attending schools where children drop out increases substantially the likelihood of dropping out. Gould, Lavy and Paserman (2009) observe adverse long term peer effects of immigrants on natives' educational outcomes (drop out and passing rates in matriculation exams). In turn, Mora and Oreopoulos (2011) do not find an impact of non reciprocating friends' intentions to drop out on students' intention to drop out while they find a small effect of reciprocating friends.[43]

In this chapter I study the influence of peers in the last grade of primary school on the decision to attend junior high two and three years later. I use unique data on the current status of attendance (2011 and 2012) of adolescents who had participated in standardized evaluations while they were in 3rd and 6th grade of primary school in 2006 and 2009, respectively. The identification strategy follows that of Ammermueller and Pischke (2009). I exploit the random variation of students' characteristics across classrooms within 6th grade groups in primary public schools. Given that student tracking into classrooms is strictly not allowed in Uruguayan public schools, this can be thought of as a natural experiment. There are several studies that apply similar identification strategies in order to identify peer effects. Hoxby (2000) exploits the variation in gender and racial composition of a grade within schools during adjacent years. Hanushek et al. (2003) also exploit small differences in peer characteristics across cohorts within a school. Gould, Lavy and Paserman (2009) consider the variation in the percentage of immigrants across grades within schools. By performing a within school analysis, I address selection into schools (and neighborhoods). I also control for observable teacher characteristics which represent another potential determinant of differential choices of adolescents who attended the same school and grade in different classrooms.

The chapter is organized as follows. Section 2.2 characterizes junior high dropouts using household survey data. Section 2.3 describes the identification strategy used to estimate peer effects. Section 2.4 describes the data and section 2.5 analyzes de results. Finally, section 2.6 concludes.

[42] See Graham (2008), Carrell, Fullerton and West (2009), and De Giorgi, Pellizzari and Redaelli (2010) among the most recent studies.

[43] The authors define a non reciprocating friend to the cases in which one student lists another as a friend but the latter does not reciprocate.

2.2 Characteristics of dropouts in Uruguay

In this section, I characterize junior high drop outs by analyzing Household Survey data. Table 2.1 indicates that among those adolescents who did not attend junior high or vocational schools in 2011, almost half had completed primary school but never attended secondary school.

Table 2.1 – Characteristics of 13-15 year olds that do not attend either primary or secondary school

	%
Dropped out during primary school	7.0
Completed primary school but never attended secondary school	47.5
Started attending secondary school and later dropped out	45.4
Total	100
Source: National Household Survey 2011.	

Table 2.2 provides information on the main reason stated for dropping out among 13-15 year olds. Even if more than half did not respond to the question, there are a significant percentage of dropouts who argue that they did not complete junior high because they were not interested or wanted to learn other things.

Table 2.2 – Main reason for dropping out among 13-15 year olds

	%
No response	54.6
Started to work	1.2
Was not interested/Was interested in learning other things	28.1
Got pregnant or spouse got pregnant	1.3
Studying was too hard	5.2
Had to deal with family issues	2.4
Other reasons	7.2
Total	100.0
Source: National Household Survey 2011.	

Table 2.3 shows the marginal effects of individual socioeconomic characteristics on the probability of attending junior high. As could be expected the probability of attendance is higher among those who attended a private school and is increasing in per capita household income. It is in turn, lower for adolescents who live in slums and for those who live in rural areas. The probability of attendance is higher for girls and for 13 and 14 year olds relative to 15 year olds. It is worth noting that having attended pre-school increases the probability of attendance to junior high by 12%.

Table 2.3 – Probability of attending junior high or vocational schools (marginal effects)

Towns with more than 5000 inhabitants (excluding Montevideo)	-0.00
	(0.01)
Towns with less than 5000 inhabitants	-0.00
	(0.01)
Rural areas	-0.08***
	(0.02)
Per capita household income (logs)	0.08***
	(0.01)
Attended a private primary school	0.04***
	(0.01)
Lives in a slum	-0.03**
	(0.01)
Female	0.05***
	(0.01)
13 year old	0.09***
	(0.01)
14 year old	0.06***
	(0.01)
Black origins	-0.00
	(0.01)
Lives with mother	-0.04
	(0.03)
Lives with father	-0.01
	(0.02)
Attended pre-school	0.12***
	(0.03)
Obs	6,193
Pseudo r-squared	0.16

Robust standard errors in parentheses with marginal effects computed at the mean of the independent variables.
Source: own calculations base on National Household Survey 2011.
*** p<0.01, ** p<0.05, * p<0.1

2.3 Identification strategy

The identification strategy follows Ammermueller and Pischke (2009) and relies on variations across classes within schools. Classrooms are an important unit of analysis as students interact with their classmates several hours per day during a whole academic year. In what follows, I examine the relevance of classroom peers during the last year of primary in public schools.

Let the reduced form specification of the decision to attend junior high be:

$$y_{ics} = \alpha_s + \beta X_{ics} + \gamma C_s + \lambda X_{(-i)cs} + \mu_{cs} + \varepsilon_{ics}$$

Where y_{ics} is the choice to either continue studying or drop out for the adolescent i who attended classroom c and school s while he was in the last year of primary school, α_s are school fixed effects, X_{ics} are individual or family characteristics, C_s are teacher characteristics and $X_{(-i)cs}$ are the average characteristics of the classmates of student i. Finally, μ_{cs} reflects correlated effects and ε_{ics} individual error terms. By estimating a reduced form only a composite social interaction effect can be obtained. That is, λ captures both endogenous and contextual effects (Manski, 1993).

In this study, I control for selection into schools by including school fixed effects and only exploit variation within schools and across classrooms. The crucial identification assumption is that groups within schools are not formed under the basis of achievement or parental background and teachers are not assigned into classrooms with any specific criteria. If classroom composition was actually affected by tracking then some classes would have systematically less probability of producing dropouts than others and the estimates would be biased by these correlated effects. In Uruguay tracking is not allowed in public schools. Thereby, it can be argued that differences in composition between classrooms within a school are random.

2.4 Data

The data set comprises a panel of approximately 2030 students belonging to 50 public schools who were attending 3rd grade of primary school in 2006 when they participated in the SERCE evaluation[44] and that participated in either the V Evaluación Nacional de Aprendizajes in 2009 or in an evaluation carried out by Universidad de la República in 2009.[45] The SERCE evaluation consisted on math and reading tests together with questionnaires to families and school principals. In 2009, ANEP, the central authority responsible for education in Uruguay, implemented the V Evaluación Nacional de Aprendizajes. This evaluation included math, reading and science tests together with questionnaires to students, their families, teachers and the school principals. ANEP included among the schools evaluated, schools that had participated in the SERCE evaluation in order to build a panel. In 2009, Universidad de la República performed a parallel evaluation (including questionnaires to families, teachers and school principals) on a sample of schools who had participated in the SERCE evaluation in 2006.[46] In 2011 and 2012 students IDs from all students who had been evaluated both in 2006 and 2009 were matched to

[44] Segundo Estudio Regional Comparativo y Explicativo. This evaluation was performed in sixteen Latin American countries plus a state in Mexico.

[45] The original panel consists of 3500 students but given the identification strategy outlined, schools in which only one classroom was evaluated were not considered in this chapter..

[46] The Project "Aprendiendo con las XO: el impacto del Plan Ceibal en el aprendizaje 2009" aimed at determining a preliminary impact of the One Lapltop Per Child program (Plan Ceibal).

administrative data of public junior high and vocational schools in order to determine the current status of the 3000 students.[47] In the V Evaluación Nacional de Aprendizajes, students had been asked to which junior high school or vocational school they intended to attend, this allowed to identify students who in 2009 attended public schools and now attend junior high in private institutions. Also, most students' telephone numbers were made available as well as contacts at the primary schools, hence it was possible to check whether those not found in the official data sets had actually dropped out.

In order to perform within school estimates I only utilize schools with two or more classes (50 schools). Table 2.4 shows the descriptive statistics of the data set which include socioeconomic characteristics of the students, test scores from the SERCE evaluation in 2006 and from the V Evaluación Nacional de Aprendizajes in 2009 as well as characteristics from 6th grade teachers in 2009 (last year of primary school) and the students' current attendance status. In this data set 17% of students who attended 6th grade of primary school in 2009 are currently not attending junior high.[48]

[47] These checks were performed in the context of a project Gioia de Melo and Alina Machado are carrying out in order to analyze the long term effects of the One Laptop Per Child program.

[48] Low socioeconomic strata schools are overrepresented in the sample because the educational authorities in the 2009 evaluation selected the sample in such a way to be able to study this type of schools more in depth.

Table 2.4 Descriptive statistics

	Obs	Mean	Std. Dev.	Min	Max
Attends junior high	2029	0.83	0.38	0	1
Female	2029	0.49	0.50	0	1
Mother educ: < = primary	1982	0.36	0.48	0	1
Mother educ: incompl. highschool	1982	0.48	0.50	0	1
Mother educ: highschool or incompl. college	1982	0.10	0.31	0	1
Mother educ: compl. college	1982	0.06	0.23	0	1
Books: < 10	1984	0.27	0.45	0	1
Books: btw 10 & 50	1984	0.39	0.49	0	1
Books: > 50	1984	0.27	0.45	0	1
Number of persons in house	1831	5.07	1.85	2	16
Preschool (years)	1910	2.71	1.48	0	6
Never repeated in primary school	2005	0.72	0.45	0	1
Repeated only once	2005	0.21	0.41	0	1
Repeated more than once	2005	0.07	0.26	0	1
Slum	1796	0.19	0.39	0	1
Reading score (3rd grade primary)	1835	0.00	1.00	-4.3	3.6
Math score (3rd grade primary)	1931	0.00	1.00	-4.4	4.4
Reading score (6th grade primary)	1587	0.00	1.00	-4.9	3.9
Math score (6th grade primary)	1551	0.00	1.00	-4.7	3.4
Belief keep studying after finishing primary	1517	4.56	0.95	1.0	5.0
Belief will attend junior high	1466	4.34	1.23	1.0	5.0
Belief will attend vocational school	1473	2.69	1.59	1.0	5.0
Belief will attend college	1493	3.26	1.54	1.0	5.0
Female teacher	1949	0.95	0.22	0.0	1.0
Teacher experience (years)	1949	16.89	9.11	1.0	35.0
Tearcher experience squared	1949	368.28	340.43	1	1225
Teacher has good relationship with students	1949	0.62	0.49	0	1

Source: V Evaluacion Nacional de Aprendizajes 2009, SERCE 2006 and Aprendiendo con las XO: el impacto del Plan Ceibal en el aprendizaje 2009.

Notes: Dummies for the number of books at home are defined based on the question: Approximately how many books are there in the household? With the following options: (1) there are no books, (2) there are less than 10, (3) there are between 10 and 50, (4) there are more than 50. Questions on beliefs required to answer to what extent the student agreed with each statment. The options were: (1) completely disagree, (2) disagree, (3) do not agree or disagree, (4) agree, (5) completely agree. Teacher characteristics were answered in the teachers survey. The dummy "Teacher has a good relationship with students" takes the value of 1 if the teacher answered option 1 in the question: In general are you satisfied with your relationship with the students? with the following options: (1) very satisfied, (2) satisfied, (3) unsatisfied, (4) very unsatisfied.

As in Ammermueller and Pischke (2009), the socioeconomic characteristics used as peer measures are self reported by students' parents. Ammermueller and Pischke (2009) note this may lead to measurement error in the individual and peer variables. I also include two student variables which are not self reported: students' scores in the SERCE math and reading tests (2006). These tests were created and scored by SERCE. This is a major advantage compared to data sets in which students are graded by their teachers because teachers' expectations of (or preferences for) their students could distort grading within a class. Contrary to the data used in Ammermueller and Pischke (2009), all students in the classroom were surveyed and not just a sample. Still, some were absent the date of the survey or their parents decided not to fill it. The average number of students per classroom in this sample is of 19.

It should be noted that students that repeated any grade in primary school after 2006 were excluded from the analysis even though they were evaluated in 2009. The reason being that during the 2009 evaluations, those students who had participated in the 2006 evaluation and were not in 6th grade primary (because they had repeated) were evaluated but the rest of their class was not. Only 6th grade classes were evaluated as a whole. Overall 330 students who were in 5th or 4th grade in 2009 were not considered in the estimates because of not having complete information on their classroom peers. In 2011 among these 330 students, approximately 72 percent attended junior high or vocational schools with a minority still attending primary school. To sum up, I only consider students who either never repeated or that repeated one or more times in the first three years of primary school but did not repeat any grade between 2006 and 2009. Potentially they may have repeated 6th grade in 2009.[49]

Table 2.5 reports the decomposition of variance in peer variables within and between schools. Most of the variance in these variables is between schools, which suggests selection into schools is pervasive.[50] However, there is still some variance left within classrooms in a school.

Table 2.5 – Decomposition of variance of peer variables

	Sum of squares		
	Bewteen schools	Within schools	Total
Mother education (index)	4084.78	188.56	4273.34
Books at home (index)	422.17	43.51	465.68
Reading score	403.02	60.69	463.71
Math score	409.78	61.49	471.27

2.5 Results

The peer averages are calculated using information for all students who report a value for this specific variable in the data set, not just the students in the final sample. In each case, the peer variable for student *i* is the classroom mean omitting the value of the variable for student *i*. The sample size conditioning on non missing student background in all the variables included in the estimation shrinks to 81%, significantly above the corresponding percentages per country in Ammermueller and Pischke (2009) which, range from 40 to 75%.

Table 2.6 shows marginal effects for the probability of attending junior high with and without including school fixed effects to control for selection into schools. Due to the potential for error correlation across individuals within a given school, I cluster all standard errors at the school

[49] Chances that some students who are considered drop outs in the data set continue attending primary school are extremely slim: educational regulations determine that when children turn 15 they must be redirected to other educational institutions (special educational facilities) but cannot remain in primary school; thereby the only case in which a student who attended 6th grade in 2009 could still be attending primary school in 2012 is that he never repeated any grade before 6th grade and then repeated 6th grade three times.

[50] Fernandez (2009) confirms that dropouts are concentrated in high schools where most students come from low and extremely low socioeconomic background.

level. Own characteristics have the expected signs. Not having repeated any year in primary school increases the probability of attending junior high by approximately 25%. It is interesting to note that having repeated one year in primary school still increases the probability of attending junior high relative to having repeated more than one year. The relevance of having attended preschool that had been found when analyzing household survey data is no longer observed. In line with many studies, teachers' observable characteristics (6th grade primary) do not explain educational attainment and in this case junior high school attendance. However, there could be relevant unobservable teacher characteristics that could be influencing junior high attendance choices and in that way could be confounding peer estimates.

As for peer effects, both the OLS estimates and the within school estimates suggest that there peer effects are present (see Table 2.6). In what follows I discuss only the within school estimates, which control for selection into schools. Increasing the percentage of students who own between 10 and 50 books in one standard deviation (departing from the average percentage in the sample) increases the probability of attendance by 3 percent.[51] The effect is observed solely for households who own between 10 and 50 books as only 27 percent of adolescents in the sample live in households that own more than 50 books and these adolescents are likely to have attended schools where no classmate dropped out ex post. The number of books at home is the sole indicator selected by Ammermueller and Pischke (2009) to study peer effects. The authors study peer effects in scores and use an index of books at home instead of separate dummies. They find peer effects slightly smaller than those found in this chapter.

Table 2.6 also reports the impact of classmates test scores in primary on the probability of junior high attendance (see columns 5 and 6 for the within school reading and math estimates, respectively). I am considering the 3rd grade test scores of classmates in 6th grade. These classmates need not have been in the same group when they were in 3rd grade but that is not an issue because all students sat for the same test and I am concerned with the effect of 6th grade classmates. There is a negative effect from classmates test scores. A one standard deviation increase in classmates' average reading test scores leads to a 6 percent decrease in the probability of attendance (and 9 percent decrease in the probability of attendance for the case of math). These results are in contrast with those observed in Table 1.6 in chapter 1 which suggested that classmates scores in third grade had no significant incidence on students' 6th grade scores.

Both the positive peer effects observed regarding the number of books at home and the negative effects from test scores are confirmed in Table A.2.1 where each variable and its corresponding peer variable are the only regressors. However, when including 6th grade standardized test scores for a smaller sample of schools (only for those who participated in the V Evaluación Nacional de Aprendizajes in 2009) instead of 3rd grade test scores for the same reference groups (6th grade

[51] As mentioned before, the identification strategy here applied cannot disentangle contextual from exogenous effects. The reduced form estimation yields a composite social interaction effect.

classmates), peer effects do not seem to be significant (see Table A.2.1). The correlation between scores in 2006 and 2009 is 0.56 and 0.61 for reading and math, respectively.

Table 2.6 – Probability of attending junior high or vocational schools. Marginal effects

	(1)	(2)	(3)	(4)	(5)	(6)
Own characteristics						
Female	**0.05*** **	**0.04*** **	**0.04*** **	**0.05*** **	**0.03** **	**0.04*** **
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Mother educ: incompl. highschool	**0.05*** **	**0.05*** **	**0.05*** **	**0.04** **	**0.04** **	**0.03* **
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Mother educ: highschool or incompl. college	0.03	0.03	0.03	0.02	0.01	0.00
	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)
Mother educ: compl. college	0.03	0.02	0.02	0.02	0.01	-0.00
	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
Books: btw 10 & 50	**0.05*** **	**0.03** **	**0.05*** **	**0.06*** **	**0.04*** **	**0.06*** **
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Books: > 50	**0.04*** **	**0.03* **	**0.04*** **	**0.05*** **	**0.03* **	**0.04*** **
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Number of persons in house	**-0.01*** **	**-0.01** **	**-0.01*** **	**-0.01** **	**-0.01* **	**-0.01*** **
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Slum	**-0.04* **	-0.03	-0.03	-0.03	-0.03	-0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Preschool (years)	0.00	0.00	-0.00	0.01	0.01	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Never repeated in primary school	**0.25*** **	**0.24*** **	**0.20*** **	**0.26*** **	**0.25*** **	**0.20*** **
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Repeated only once	**0.07*** **	**0.07*** **	**0.05*** **	**0.07*** **	**0.07*** **	**0.04* **
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)
Reading score (3^{rd} grade primary)		**0.03*** **			**0.04*** **	
		(0.01)			(0.01)	
Math score (3^{rd} grade primary)			**0.03*** **			**0.03*** **
			(0.01)			(0.01)
Peer characteristics						
Female	0.11	**0.15** **	0.12	0.13	0.09	0.11
	(0.08)	(0.07)	(0.08)	(0.13)	(0.12)	(0.10)
Mother educ: incompl. highschool	0.07	0.11	0.11	-0.06	-0.01	-0.08
	(0.08)	(0.07)	(0.07)	(0.12)	(0.10)	(0.11)
Mother educ: highschool or incompl. college	0.07	0.12	0.20	-0.09	0.03	0.00
	(0.11)	(0.12)	(0.12)	(0.17)	(0.16)	(0.16)
Mother educ: compl. college	**0.27* **	**0.28* **	**0.36** **	0.06	0.09	0.10
	(0.15)	(0.14)	(0.17)	(0.26)	(0.24)	(0.25)
Books: btw 10 & 50	0.10	**0.13** **	**0.15** **	**0.22** **	**0.24** **	**0.36*** **
	(0.07)	(0.06)	(0.07)	(0.11)	(0.11)	(0.10)
Books: > 50	0.04	0.05	0.09	0.08	0.09	0.21
	(0.08)	(0.08)	(0.08)	(0.16)	(0.16)	(0.14)
Number of persons in house	**-0.04* **	**-0.04** **	**-0.03* **	-0.03	-0.03	-0.03
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)
Slum	-0.04	-0.01	-0.03	-0.02	-0.02	-0.05
	(0.06)	(0.05)	(0.06)	(0.12)	(0.10)	(0.11)
Preschool (years)	-0.02	**-0.03* **	-0.02	0.03	**0.04* **	**0.03** **
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
Never repeated in primary school	-0.10	-0.03	-0.10	-0.16	-0.13	-0.18
	(0.11)	(0.09)	(0.11)	(0.22)	(0.20)	(0.18)
Repeated only once	0.01	0.05	-0.03	-0.12	-0.13	-0.19
	(0.13)	(0.11)	(0.13)	(0.27)	(0.23)	(0.22)
Reading score (3^{rd} grade primary)		**-0.06** **			**-0.13*** **	
		(0.02)			(0.04)	
Math score (3^{rd} grade primary)			**-0.11*** **			**-0.21*** **
			(0.03)			(0.03)
Teacher characteristics						
Female teacher	**-0.05*** **	**-0.04** **	**-0.04* **	-0.03	0.00	0.02
	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
Teacher experience (years)	-0.00	0.00	0.00	-0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Tearcher experience squared	0.00	-0.00	0.00	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Teacher has good relationship with students	0.01	0.02	0.02	0.01	0.01	0.03
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)
Obs	1,641	1,511	1,577	1,515	1,389	1,442
Pseudo r-squared	0.19	0.21	0.21	0.24	0.27	0.27
School fixed effects	No	No	No	50	50	50

Notes: Estimates of a probit model with marginal effects computed at the mean of the independent variables. Standard errors (in parentheses) are clustered at the school level.

*** p<0.01, ** p<0.05, * p<0.1

Tables 2.7 and 2.8 report peer effects of the percentage of classmates from different quartiles of the distribution of reading scores and math scores, respectively. While the negative effect observed in Table 2.6 for reading scores is not significant any longer, Table 2.8 suggests that an increase in the percentage of classmates whose scores are in the second, third or fourth quartile of the distribution decreases the probability of attendance to junior high. Columns 3 and 4 of Tables 2.7 and 2.8 explore interaction effects but results do not seem very conclusive. Table A.2.2 in the Appendix reports peer effects on attendance from classmates' scores for each quartile of the distribution separately. According to results in Table A.2.2, students whose scores in 3rd grade were in the first and second quartile of the distribution are negatively affected by an increase in classmates' average scores. This exercise is also performed for a subsample of schools that participated in the evaluation in 6th grade. While Table A.2.1 suggested there were no significant effects if these tests were considered instead of 3rd grade tests, when splitting the sample in quartiles a negative effect of classmates' math scores on attendance is observed for students in the second quartile of the distribution.

Table 2.7 – Probability of attending junior high or vocational schools. Own and peer reading scores. Marginal effects

	(1)	(2)	(3)	(4)
Own characteristics[1]				
First quartile		**-0.11*****		-0.22
		(0.03)		(0.21)
Second quartile	**0.05****	-0.04	0.01	**0.14****
	(0.02)	(0.03)	(0.10)	(0.06)
Third quartile	**0.07*****	-0.01	-0.00	0.06
	(0.02)	(0.02)	(0.11)	(0.08)
Fourth quartile	**0.08*****		**0.15*****	
	(0.02)		(0.04)	
Peer characteristics[1]				
Fraction first quartile		0.23		
		(0.15)		
Fraction second quartile	-0.13	0.10		
	(0.16)	(0.14)		
Fraction third quartile	-0.16	0.07		
	(0.13)	(0.13)		
Fraction fourth quartile	-0.23			
	(0.15)			
First quartile × fraction first quartile				0.07
				(0.15)
First quartile × fraction second quartile				0.12
				(0.22)
First quartile × fraction third quartile				-0.05
				(0.20)
Second quartile × fraction first quartile				**-0.31***
				(0.16)
Second quartile × fraction second quartile			-0.00	-0.28
			(0.19)	(0.19)
Second quartile × fraction third quartile			-0.03	**-0.37****
			(0.17)	(0.19)
Second quartile × fraction fourth quartile			**0.31****	
			(0.16)	
Third quartile × fraction first quartile				-0.20
				(0.16)
Third quartile × fraction second quartile			0.01	-0.17
			(0.18)	(0.18)
Third quartile × fraction third quartile			0.22	0.00
			(0.18)	(0.20)
Third quartile × fraction fourth quartile			0.18	
			(0.15)	
Fourth quartile × fraction second quartile			-0.17	
			(0.19)	
Fourth quartile × fraction third quartile			-0.01	
			(0.15)	
Fourth quartile × fraction fourth quartile			-0.20	
			(0.16)	
Obs.	1,389	1,389	1,389	1,389
Pseudo r-squared	0.27	0.27	0.28	0.28
School fixed effects	50	50	50	50

Notes: Estimates of a probit model with marginal effects computed at the mean of the independent variables. Standard errors (in parentheses) are clustered at the school level.

1) Own characteristics and peer characteristics except for quartiles in reading and teacher characteristics are included in the estimation but not reported in order to reduce the table size.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.8 – Probability of attending junior high or vocational schools. Own and peer math scores. Marginal effects

	(1)	(2)	(3)	(4)
Own characteristics[1]				
First quartile		**-0.05***		-0.26
		(0.03)		(0.18)
Second quartile	0.03	-0.02	0.10	0.03
	(0.02)	(0.03)	(0.06)	(0.07)
Third quartile	**0.06*****	0.03	0.08	0.01
	(0.02)	(0.03)	(0.06)	(0.11)
Fourth quartile	**0.04****		**0.16*****	
	(0.02)		(0.06)	
Peer characteristics[1]				
Fraction first quartile		**0.56*****		
		(0.11)		
Fraction second quartile	**-0.35****	**0.21***		
	(0.15)	(0.12)		
Fraction third quartile	**-0.51*****	0.05		
	(0.18)	(0.12)		
Fraction fourth quartile	**-0.56*****			
	(0.11)			
First quartile × fraction first quartile				**0.28****
				(0.13)
First quartile × fraction second quartile				0.12
				(0.18)
First quartile × fraction third quartile				0.08
				(0.20)
Second quartile × fraction first quartile				0.10
				(0.16)
Second quartile × fraction second quartile			-0.10	-0.05
			(0.20)	(0.17)
Second quartile × fraction third quartile			-0.25	-0.22
			(0.18)	(0.16)
Second quartile × fraction fourth quartile			-0.03	
			(0.15)	
Third quartile × fraction first quartile				0.07
				(0.16)
Third quartile × fraction second quartile			0.16	0.21
			(0.16)	(0.19)
Third quartile × fraction third quartile			-0.18	-0.21
			(0.24)	(0.25)
Third quartile × fraction fourth quartile			-0.04	
			(0.14)	
Fourth quartile × fraction second quartile			-0.26	
			(0.19)	
Fourth quartile × fraction third quartile			-0.09	
			(0.17)	
Fourth quartile × fraction fourth quartile			**-0.32****	
			(0.15)	
Obs.	1,442	1,442	1,442	1,442
Pseudo r-squared	0.27	0.27	0.27	0.27
School fixed effects	50	50	50	50

Notes: Estimates of a probit model with marginal effects computed at the mean of the independent variables. Standard errors (in parentheses) are clustered at the school level.

1) Own characteristics and peer characteristics except for quartiles in reading and teacher characteristics are included in the estimation but not reported in order to reduce the table size.

*** p<0.01, ** p<0.05, * p<0.1

For a subsample of primary schools, (those that participated in the V Evaluación Nacional de Aprendizajes in 2009) children answered questions regarding their beliefs about future studies. For each statement (I will keep on studying after primary, I will attend junior high, I will attend vocational school and I will attend college), students had to rank their degree of approval of those statements ranging from 1 (completely disagree) to 5 (completely agree). Table 2.9 shows that while own beliefs are relevant for explaining attendance in junior high (after controlling for other individual characteristics), peers' beliefs do not seem to influence own behavior. The only individual belief which is not significant is the belief of attending vocational school. This can be explained because most adolescents attend junior high (approximately 80 percent).

Table 2.9 – Probability of attending junior high or vocational schools including own and classmates' beliefs about future studies. Marginal effects

Own characteristics				
Female	**0.04****	**0.04****	**0.04***	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
Mother educ: incompl. highschool	**0.05****	**0.04****	**0.04****	**0.04****
	(0.02)	(0.02)	(0.02)	(0.02)
Mother educ: highschool or incompl. college	-0.02	-0.02	-0.02	-0.05
	(0.06)	(0.06)	(0.06)	(0.06)
Mother educ: compl. college	-0.01	-0.01	-0.00	-0.03
	(0.07)	(0.07)	(0.07)	(0.08)
Books: btw 10 & 50	**0.06*****	**0.06*****	**0.06*****	**0.05*****
	(0.02)	(0.02)	(0.02)	(0.02)
Books: > 50	0.04*	0.04	0.03	0.03
	(0.02)	(0.02)	(0.02)	(0.02)
Number of persons in house	**-0.01****	**-0.01****	**-0.01****	**-0.01****
	(0.01)	(0.01)	(0.01)	(0.01)
Slum	-0.05*	-0.05*	-0.04	-0.05*
	(0.03)	(0.03)	(0.03)	(0.03)
Preschool (years)	-0.00	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Never repeated in primary school	**0.27*****	**0.25*****	**0.29*****	**0.28*****
	(0.06)	(0.06)	(0.07)	(0.07)
Repeated only once	**0.06*****	**0.06*****	**0.07*****	**0.07*****
	(0.02)	(0.02)	(0.02)	(0.02)
Belief keep studying after finishing primary	**0.02****			
	(0.01)			
Belief will attend junior high		**0.02****		
		(0.01)		
Belief will attend vocational school			-0.01	
			(0.01)	
Belief will attend college				**0.02*****
				(0.00)
Peer characteristics				
Female	**0.39***	**0.37****	0.30	0.29
	(0.21)	(0.17)	(0.20)	(0.20)
Mother educ: incompl. highschool	**-0.20***	-0.17	-0.17	-0.18
	(0.11)	(0.11)	(0.13)	(0.12)
Mother educ: highschool or incompl. college	-0.29	-0.30	-0.26	-0.28
	(0.23)	(0.23)	(0.24)	(0.24)
Mother educ: compl. college	-0.34	-0.26	-0.32	-0.28
	(0.28)	(0.28)	(0.30)	(0.27)
Books: btw 10 & 50	**0.23***	**0.21***	0.15	0.17
	(0.12)	(0.12)	(0.14)	(0.14)
Books: > 50	0.17	0.17	0.15	0.15
	(0.18)	(0.17)	(0.18)	(0.18)
Number of persons in house	-0.05	-0.04	-0.05	-0.05
	(0.04)	(0.04)	(0.04)	(0.03)
Slum	-0.20	-0.21	-0.14	-0.13
	(0.13)	(0.13)	(0.14)	(0.14)
Preschool (years)	0.02	0.02	0.03	0.02
	(0.03)	(0.03)	(0.04)	(0.03)
Never repeated in primary school	-0.28	-0.23	-0.28	-0.35
	(0.26)	(0.25)	(0.27)	(0.26)
Repeated only once	-0.43	-0.40	-0.41	-0.46
	(0.31)	(0.29)	(0.34)	(0.31)
Belief keep studying after finishing primary	0.02			
	(0.06)			
Belief will attend junior high		-0.04		
		(0.04)		
Belief will attend vocational school			0.00	
			(0.03)	
Belief will attend college				0.00
				(0.03)
Teacher characteristics				
Female teacher	-0.00	0.00	-0.03	-0.02
	(0.05)	(0.05)	(0.05)	(0.04)
Teacher experience (years)	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Tearcher experience squared	0.00	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Teacher has good relationship with students	-0.03	-0.03	-0.04	-0.03
	(0.02)	(0.02)	(0.03)	(0.02)
Obs	1,067	1,067	1,009	1,026
Pseudo r-squared	0.27	0.27	0.26	0.28
School fixed effects	41	41	41	41

Notes: Estimates of a probit model with marginal effects computed at the mean of the independent variables. Standard errors (in parentheses) are clustered at the school level.

*** p<0.01, ** p<0.05, * p<0.1

For a subsample of schools there are data available on nominated friends. A strategy exploiting information on networks was attempted but no significant results were obtained (see Table A.2.3). The no result applying this strategy may have two possible explanations. The first one is that friends at school may be less important for ex post dropouts. That is, ex post dropouts may be influenced by the whole classroom as shown in the tables before but not necessarily by the peers they named potentially either because they have greater difficulty socializing at school or because they have closer peers outside school. Indeed, on average adolescents who attend junior high listed more friends in 6th grade primary than those who ex post dropped out (see Table A.2.4). However, it may also be due to having greater difficulty completing the questionnaire.[52] The second possible explanation is that there are not sufficient observations to apply this strategy. Indeed, while peer effects in reading and math scores are highly significant for a sample of approximately 7000 observations as shown in Chapter 1, when I replicate these estimates for the subsample of schools used in this chapter to estimate classroom peer effects, network peer effects in test scores are no longer significant for reading (see Table A.2.3).

2.6 Conclusion

In this chapter I exploited a unique data set of adolescents who participated in evaluations while they were in 3rd and 6th grade primary school and currently should be attending junior high or vocational schools and for whom there is data available on their status of attendance. I study whether classmates from 6th grade primary school (either through contextual or endogenous social interactions) influence adolescents' decisions to attend junior high. To do so, I exploit presumably random variations in student composition between 6th grade groups within primary schools and control for students' individual characteristics and teacher's characteristics.

Nechyba, (2006) points out that this strategy is useful for identifying peer effects provided the changes across grades or in this case classrooms yield sufficient variation. Indeed, in this data set most variation in variables is between schools. However, I still find significant positive peer effects from increasing the fraction of classmates who live in households that own books.[53] This is the same variable for which Ammermueller and Pischke (2009) find positive peer effects in

[52] Dropouts listed significantly more friends that ex post dropped out than high school attendants. However, this should not be interpreted as causal as it is influenced by the fact that there is strong selection into schools due to residential segregation. Thereby, high school dropouts tend to concentrate in poorer schools. Instead, if we compare the percentage of friends that scored above the class mean in math and reading between dropouts and high schools attendants, we observe that it is not conclusive that ex post dropouts have friends with a lower academic performance (at least when considering their friends from school).

[53] This is captured by an indicator of the percentage of classmates that live in households who own between 10 and 50 books.

test scores.[54] In turn, I find some evidence that increasing the level of achievement of an individual's classmates has a negative impact on own junior high attendance decisions, particularly among those in the lowest quartile of the score distribution (who are potentially more likely to dropout). A possible explanation is that low achieving students may feel less confident regarding schooling goals when they compare themselves with high achieving peers. Alternatively, it could be the case that low achieving students achieve better schooling outcomes in more homogeneous classes if this facilitates teachers' adaptation of their instruction level to specially focus on low achieving students. Nonetheless, these results should be taken with caution given that peer estimates using test scores in 6th grade rather than 3rd grade for a subsample of schools do not show any significant peer effects.

At the individual level grade retention seems to be the most important factor related to junior high attendance choices. Manacorda (2012) exploits the discontinuity generated by a rule that determines repetition for students with more than three failed subjects and concludes that repetition has a substantial impact on dropout. This suggests the costs from grade retention may outweigh the benefits. Also, the fact that own beliefs about future studies at the last year of primary school are significant in explaining attendance may suggest that in most cases the decision of attendance to junior high is already made while in the last year of primary school. Indeed, according to household survey data a significant percentage of junior high dropouts declare to have completed primary school but never have attended junior high. These issues point to the relevance of implementing policies that aim at promoting junior high attendance at the last year of primary school.

[54] Ammermueller and Pischke (2009) use the number of books at home as their sole indicator of family background because it has the best item response but they do not use dummies for each category they use an index that ranges from 1 to 5 which covers the 5 options of the question.

Appendix 2

Table A.2.1 – Probability of attending junior high or vocational schools. Single individual and peer variables. Marginal effects.

	Own characteristics	Peer characteristics
Female	**0.07*** **	-0.06
	(0.01)	(0.11)
Obs	1,914	
Pseudo r-squared	0.12	
School fixed effects	50	
Mother educ: incompl. highschool	**0.12*** **	0.00
	(0.02)	(0.10)
Mother educ: highschool or incompl. college	**0.11*** **	-0.00
	(0.02)	(0.14)
Mother educ: compl. college	**0.11*** **	-0.15
	(0.04)	(0.23)
Obs	1,844	
Pseudo r-squared	0.13	
School fixed effects	50	
Books: btw 10 & 50	**0.10*** **	**0.20*** **
	(0.02)	(0.08)
Books: > 50	**0.12*** **	0.06
	(0.02)	(0.12)
Obs	1,846	
Pseudo r-squared	0.13	
School fixed effects	50	
Number of persons in house	**-0.02*** **	-0.02
	(0.00)	(0.02)
Obs	1,698	
Pseudo r-squared	0.12	
School fixed effects	50	
Slum	**-0.07** **	0.02
	(0.03)	(0.08)
Obs	1,663	
Pseudo r-squared	0.11	
School fixed effects	50	
Preschool (years)	**0.02*** **	0.00
	(0.01)	(0.03)
Obs	1,778	
Pseudo r-squared	0.12	
School fixed effects	50	
Never repeated in primary school	**0.36*** **	-0.14
	(0.04)	(0.17)
Repeated only once	**0.11*** **	-0.05
	(0.02)	(0.19)
Obs	1,890	
Pseudo r-squared	0.18	
School fixed effects	45	
Reading score (3^{rd} grade primary)	**0.09*** **	**-0.05* **
	(0.01)	(0.03)
Obs	1,700	
Pseudo r-squared	0.17	
School fixed effects	50	
Math score (3^{rd} grade primary)	**0.09*** **	**-0.12*** **
	(0.01)	(0.04)
Obs	1,804	
Pseudo r-squared	0.16	
School fixed effects	50	
Reading score (6^{th} grade primary)	**0.10*** **	-0.02
	(0.01)	(0.04)
Obs	1,468	
Pseudo r-squared	0.17	
School fixed effects	41	
Math score (6^{th} grade primary)	**0.10*** **	-0.04
	(0.01)	(0.04)
Obs.	1,430	
Pseudo r-squared	0.15	
School fixed effects	41	

Notes: Estimates of a probit model with marginal effects computed at the mean of the independent variables. Standard errors (in parentheses) are clustered at the school level.

*** p<0.01, ** p<0.05, * p<0.1

Table A.2.2 – Classroom peer effects on the probability of attending junior high or vocational schools by quartile of scores:

	1st quartile	2nd quartile	3rd quartile	4th quartile
Reading score (3rd grade primary)	**-0.45****	-0.03	0.06	0.00
	(0.21)	(0.12)	(0.13)	(0.00)
Obs.	308	238	177	164
Pseudo r-squared	0.33	0.41	0.39	0.52
School fixed effects	Yes	Yes	Yes	Yes
Math score (3rd grade primary)	**-0.34****	**-0.32****	-0.00	-0.05
	(0.17)	(0.15)	(0.00)	(0.07)
Obs.	300	245	166	212
Pseudo r-squared	0.36	0.33	0.60	0.38
School fixed effects	Yes	Yes	Yes	Yes
Reading score (6th grade primary)	0.38	-0.07	-0.00	0.00
	(0.39)	(0.22)	(0.00)	(0.00)
Obs.	253	235	152	129
Pseudo r-squared	0.27	0.36	0.53	0.60
School fixed effects	Yes	Yes	Yes	Yes
Math score (6th grade primary)	-0.54	**-0.58*****	0.00	0.00
	(0.42)	(0.19)	(0.00)	(0.00)
Obs.	238	235	174	125
Pseudo r-squared	0.37	0.34	0.57	0.70
School fixed effects	Yes	Yes	Yes	Yes

Notes: Estimates of a probit model with marginal effects computed at the mean of the independent variables. Standard errors (in parentheses) are clustered at the school level. Quartiles are computed for public schools only. Each entry is the coefficient on the peers' scores from a separate regression.

Own characteristics and peer characteristics (except for peers' scores) and teacher characteristics are included in the estimation but not reported in order to reduce the table size.

*** p<0.01, ** p<0.05, * p<0.1

Table A.2.3 – Peer effects using network data

	Attending junior high	Attending junior high	Attending junior high
Endogenous effects	0.56	0.50	-0.16
	(0.94)	(0.44)	(0.46)
Contextual effects			
Female	0.07	0.11**	0.02
	(0.06)	(0.05)	(0.04)
Repeat	0.06	0.04	-0.06
	(0.17)	(0.07)	(0.11)
Books: btw 10 & 50	-0.00		
	(0.07)		
Books: > 50	-0.07		
	(0.08)		
Reading score (3rd grade primary)		0.00	
		(0.00)	
Math score (3rd grade primary)			0.00
			(0.00)
Excluded instruments (first stage)			
Peers' peers' books: btw 10 & 50	0.03		
	(0.03)		
Peers' peers' reading scores		0.00*	
		(0.00)	
Peers' peers' peers' reading scores		-0.00**	
		(0.00)	
Peers' peers' math scores			0.00
			(0.00)
Peers' peers' peers' math scores			0.00
			(0.00)
Obs.	1,283	1,137	1,229
F-test excluded inst	0.83	3.09	1.31
Number of clusters	41	41	41
Classroom fixed effects	Yes	Yes	Yes

	Reading scores 2009	Math scores 2009	Science scores 2009
Endogenous effects	0.63	0.43**	0.05
	(1.19)	(0.21)	(0.37)
Contextual effects			
Female	0.10	0.02	0.26**
	(0.21)	(0.10)	(0.11)
Repeat	0.36	0.26	-0.06
	(0.86)	(0.20)	(0.14)
Mother: incompl HS	-0.08	-0.02	-0.05
	(0.19)	(0.08)	(0.13)
Mother: comp HS-	-0.19	0.38***	0.11
incomp colllege	(0.41)	(0.14)	(0.14)
Mother: compl college	-0.49	0.21	0.38
	(1.27)	(0.17)	(0.31)
Excluded instruments (first stage)			
Peers' peers motheduc	0.03	0.10***	0.09**
	(0.03)	(0.02)	(0.03)
Peers' peers' peers motheduc	0.00	0.08*	-0.04
	(0.07)	(0.04)	(0.06)
Obs.	1,107	1,063	1,053
F-test excluded inst	0.48	14.76	3.47
Number of clusters	41	41	41
Classroom fixed effects	Yes	Yes	Yes

Notes: Standard errors (in parentheses) are clustered at the school level. Both peer scores and own score have been normalized.

*** p<0.01, ** p<0.05, * p<0.1

Table A.2.4 – Comparison of characteristics of dropouts and junior high attendants' peers

Number of friends						
	Obs	Mean	Std. Err.	Std. Dev.	[95% Con	Interval]
Dropped out	443	1.78	0.06	1.18	1.67	1.89
Attends secondary school	2388	2.39	0.02	1.07	2.34	2.43
Ttest	Ha: diff < 0		Ha: diff != 0		Ha: diff > 0	
P value	0.00		0.00		1.00	

Percentage of friends that attends secondary school						
	Obs	Mean	Std. Err.	Std. Dev.	[95% Con	Interval]
Dropped out	216	0.76	0.03	0.37	0.71	0.81
Attends secondary school	1788	0.93	0.00	0.21	0.92	0.94
Ttest	Ha: diff < 0		Ha: diff != 0		Ha: diff > 0	
P value	0.00		0.00		1.00	

Percentage of friends that scored above the class mean in a reading test						
	Obs	Mean	Std. Err.	Std. Dev.	[95% Con	Interval]
Dropped out	291	0.58	0.02	0.41	0.53	0.62
Attends secondary school	2024	0.61	0.01	0.37	0.60	0.63
Ttest	Ha: diff < 0		Ha: diff != 0		Ha: diff > 0	
Pvalue	0.06		0.13		0.94	

Percentage of friends that scored above the class mean in a math test						
	Obs	Mean	Std. Err.	Std. Dev.	[95% Con	Interval]
Dropped out	238	0.61	0.03	0.41	0.56	0.66
Attends secondary school	1862	0.63	0.01	0.38	0.61	0.65
Ttest	Ha: diff < 0		Ha: diff != 0		Ha: diff > 0	
P value	0.22		0.44		0.78	

3 The perils of peer punishment: Evidence from a common pool resource framed field experiment.[55]

We provide a model and experimental evidence on the effects of non-monetary punishment (NMP) by peers among communities of Uruguayan fishers exploiting a common pool resource (CPR). We find a) experimental groups composed of fishers from different communities (out-groups) who are sometimes in conflict over fishing territories did not overexploit the resource more than groups from a single community (in-groups) and, unlike in-groups, out-groups reduced their exploitation of the resource in response to the threat of punishment; b) cooperative individuals punished free riders while a substantial amount of punishment was targeted by free riders on cooperators, who [in turn] responded by increasing their exploitation of the resource; and c) wealthier individuals practiced greater overexploitation of the resource. Our results suggest that the relevance of in-group favoritism in promoting cooperation due to social preferences may be overrated, and that the effectiveness of peer punishment is greater when individuals are motivated by social preferences and also that coordination is required to prevent anti-social targeting and to enhance the social signal conveyed by the punishment.

3.1 Introduction

The conservation and sustainable use of common pool resources (CPRs) is an important issue worldwide not only because of the relevance that conservation of biodiversity has, but also because local commons are essential for the livelihoods of the world's poorest communities. The exploitation of a CPR poses a typical social dilemma. Hardin (1968) proposed to establish either private or state property rights as a solution to avoid the tragedy of the commons. However, market contracts and governments often fail to prevent overexploitation because the necessary information to design and enforce beneficial exchanges and directives cannot be effectively used by judges and government officials (Bowles and Gintis, 2002). In the last decades, many authors have argued that communal property regimes may enhance cooperation in the preservation of a CPR by enforcing social norms, and in this way fill the gaps of incomplete contracts (Ostrom, 1990; Feeny et al. 1990; Baland and Platteau, 1996; Ostrom et al., 1999 and Ostrom, 2000; Bowles and Gintis, 2002).

Several studies have concentrated on the determinants of successful experiences based on communal property regimes but the issue is far from settled.[56] It has been widely argued that a self-sustaining cooperative equilibrium can be achieved in a context in which self-regarding agents interact in repeated games. But in addition to reasons for cooperating related to the expectation of a subsequent reciprocal benefit sufficient to offset the current cost (Trivers, 1971), social preferences may also constitute a motive for cooperation. Social preferences encompass a

[55] This chapter is the result of a joint project with Matías Piaggio.

[56] For a description of successful cases see Ostrom (1990), Feeney et al. (1990), Ostrom et al. (1999) and Baland et al. (2007).

wide range of motives such as reciprocity, altruism, conformism and also emotions such as shame, guilt and anger (Bowles and Gintis, 2011). In a CPR context, Velez et al. (2009) show that what induces cooperation among Colombian artisanal fishing communities is their desire to conform to or emulate others' behavior in the group. In a similar framework, López et al. (2012) argue that shame is an important channel through which social preferences promote cooperative behavior.

In their seminal paper, Fehr and Gätcher (2000) show that individuals are willing to bear a cost in order to punish free riders in public good games. This pattern has also been observed in the context of a CPR dilemma (van Soest and Vyrastekova, 2006; Noussair et al., 2011). However, while van Soest and Vyrastekova (2006) find that costly punishment is effective in increasing cooperation, Noussair et al. (2011) do not observe significant changes in cooperation. Janssen et al. (2010) conclude that costly punishment is not effective in reducing extraction unless combined with communication. There is also evidence that non-monetary punishment (NMP) (Masclet, 2003; Noussair and Tucker, 2005; Dugar, 2010), social approval (Gächter and Fehr, 1999) and public observability (Barr, 2001; Denant-Boemont, 2011; López, 2012) can be effective for increasing contributions in public good games. Rege and Telle (2004) and Noussair and Tucker (2007) show that initial increases in cooperation as a consequence of public observability tend to fade away in a repeated game context. Due to the absence of monetary incentives, non-monetary punishment allows better isolation of the presence of pro-social emotions when an individual reacts to being punished relative to costly punishment. Assessing the relevance of non-monetary punishment as a tool to enhance cooperation is of particular importance in regard to community management of common pool resources, because informal sanctions typically take place in that setting. Nevertheless, to our knowledge, in the context of a CPR dilemma there is so far no evidence of the effectiveness of non-monetary punishment in promoting cooperation.

Several studies that range from laboratory to field experiments show that individuals may achieve greater levels of cooperation when interacting with members of their own group rather than with outsiders (Bandiera et al., 2005; Miguel and Gugerty, 2005; Ruffle and Sosis, 2006; Goette et al. 2006; Bernhard, et al., 2006; Chen and Xin, 2009).[57] Miguel and Gugerty (2005) argue that in rural western Kenya, social sanctions levied against free riders are more likely to take place within ethnic groups, due to the mutual reciprocity that exists within groups, rather than across groups. Bernhard, et al. (2006) observe in a third-party punishment game with two small native groups from Papua Guinea that the punishment is harsher if the victim of a norm violation belongs to the same group as the punisher. These authors show that in situations in which dictators expect the same level of punishment, they give larger transfers to in-group members than to out-group members. This would point to the relevance of unconditional preferences such as altruism that vary depending on group membership.

[57] In a broader sense, Akerlof and Kranton (2000; 2005) and Bowles and Gintis (2002) have highlighted the relevance that social identity and group affiliation have on individuals' behavior in most economic organizations.

In this study, we assess whether non-monetary punishment (NMP) is effective in promoting cooperation via social preferences in a CPR dilemma.[58] We are also interested in determining whether social preferences, such as altruism, reciprocity, shame and anger differ in a context in which individuals exploiting a common pool resource belong to different communities relative to the case in which only individuals from the same community are allowed to exploit the resource. In accord with Bernhard et al. (2006), Goette et al. (2006), and Tanaka et al. (2008), we test in-group favoritism in naturally-occurring groups in field settings. That is, we test whether fishermen that live in different communities have a greater propensity to cooperate when they interact among themselves, than when interacting with fishers who do not belong to their community, and we explore whether they differ in their sensitivity to NMP in these two scenarios.

We perform a framed field experiment (Harrison and List, 2004) where the subject pool is fishers from the Uruguayan sea coast who fish in two coastal lagoons and live in nearby villages. We concentrate on coastal lagoons because, unlike the open sea where large-scale fishing is widespread, in coastal lagoons the only agents who develop fishing activities are artisanal fishermen. Fishermen from different communities do not interact during their daily life, but they are used to encountering each other while fishing, as they tend to move from one lagoon to the other depending on fish availability. We implement both an in-group/out-group treatment and a NMP treatment. During the in-group treatment subjects played a CPR game only with members of their own community, while during the out-group treatment we required that they play the game with members of another community. The NMP implied that by facing a monetary cost, individuals could express their disapproval of others' extraction decisions. Disapproval was reflected by receiving flags that vary in color in accordance with the level of disapproval achieved among the rest of the group members.

To our knowledge, this is the first study that examines the relevance of informal sanctions in a CPR game. Additionally, our study combines three innovative features (that have not been implemented at the same time before). First, instead of inducing artificial in-group/out-group differences we enable individuals from different communities meeting each other. Second, groups are reshuffled after each period in order to avoid repeated game effects that could lead to a self-sustaining cooperative equilibrium. Third, individuals are charged a monetary cost for punishing others even if those socially punished do not face any monetary cost. This step was implemented in order to avoid subjects punishing the others carelessly.

We find no evidence of in-group favoritism. On the contrary, interacting with fishers from other communities has a positive effect on cooperation when punishment is available. That is, during the out-group treatment, individuals reduce their extraction level when NMP is available irrespective of whether they are effectively punished. We observe that the effectiveness of

[58] For the purpose of this study we define cooperation in a narrow sense as the behavior through which one agent internalizes some of the externalities he or she imposes on other users, and maintains his own use below what would maximize his individual profits (Baland et al., 2007). Baland et al. (2007) note that cooperation often requires coordination. That is, the creation of institutions is needed in order to regulate the use of the resource. In this study we concentrate on the simplest form of cooperation, as the experiment does not allow for communication or the introduction of any institutional form.

informal sanctions deteriorates by the fact that not all individuals' types are sensitive to NMP, and that these types of sanctions can be used to punish both free riders and cooperators. We argue that for peer punishment to be effective it should require coordination to prevent anti-social targeting and to enhance the social signal conveyed by the punishment. Finally, we observe that individuals adjust their extraction levels period by period according to their deviation, with respect to the group's average in a previous period, as if following an implicit social norm.

The chapter is organized as follows: Section 2 develops a theoretical model; Section 3 describes the experimental design; Section 4 reports results; finally, Section 5 concludes.

3.2 Theoretical framework

In this section we suggest a model that disentangles different motives of behavior that individuals may face when interacting in the context of a common pool resource dilemma. The model is an adaptation of a model developed by Bowles (2004), in which peer monitoring and forms of social disapproval enable individuals to achieve agreed levels of effort.

In this model individuals not only care about their own payoffs but also value (either positively or negatively) the material payoffs of their peers. Individuals may experience spite or altruism (which are independent of the others' actions), as well as reciprocity (the value they place on others' payoffs depends on the others´ past behavior or in their beliefs about the others' types). Individuals share a social norm regarding how much extraction is admitted and may experience shame if they are publically sanctioned for violating it. Besides, they face motives for punishing others socially when the others deviate from the social norm. However, punishing others may be costly, as it can deteriorate the relationship with ones' peers.

Consider a common pool resource exploited by two individuals i and j (the model can easily be generalized to n members). We define individual i's utility function as:

$$(1) \qquad u_i = \Pi_i + \beta_{ij}\Pi_j - s_i + d_i$$

Utility is the sum of the individual's own material payoffs (Π_i) plus the valuation of the others' material payoffs ($\beta_{ij}\Pi_j$), minus the subjective valuation of shame (s_i), and subjective utility or relief experienced from expressing NMP (d_i). The individual chooses an extraction level a_i and a level of NMP μ_{ij} toward his peer in order to maximize equation (1).

Let i´s material payoff be:

$$(2) \qquad \Pi_i = p_1 a_i + p_2 \left[\sum_{k=i}^{k=j} (a^{Max} - a_k) \right] \quad \forall\, k = i, j \text{ and } i \neq j$$

where both i and j have the same maximum endowment of a^{Max} to use for extracting a particular resource (i.e. fishnets). As in Cardenas (2004), the aggregate extraction by the two players $\left[\sum_{k=i}^{k=j} a_k\right] \forall\, k = i, j \text{ and } i \neq j$ reduces i's material payoffs. Alternatively, the externality can also be described as a public good benefit from conservation, i.e. lack of extraction.

Individuals have preferences as to the other's payoffs. This is reflected in i's utility by the coefficient β, which depends on both unconditional preferences (altruism or spite) and on reciprocity. Member i's degree of other regarding preferences towards j is:

$$(3) \qquad \beta_{ij} = \alpha_i + \lambda_i (b - a_j)$$

where $\alpha_i \in (-1,1)$ is i's unconditional spite or altruism, and $\lambda_i \in (0,1)$ is her degree of reciprocity. The level of reciprocal motivation therefore depends on the extent to which j has deviated from the social extraction norm (b): If j has extracted less than b, and $\lambda_i > 0$ then i experiences good will toward j and positively values his payoffs. But if j extracted more than the social extraction norm or if i feels spite towards j, then i may experience malevolence toward j ($\beta_{ij} < 0$) and enhance his utility by disapproving of j's performance.

Punishment for deviating from the social norm is expressed through non-monetary mechanisms. Socially punishing defectors enhances one's utility through the relief of expressing emotions.[59] But it does not come without cost. Individuals who express their disapproval in relation to others' actions face a cost of $c(\mu) = \frac{\gamma_i}{2}\mu_{ij}^2$ due to a deterioration in the relationship with their peers. Therefore, disapproval motives (d_i) are a function of the benefits that punishing provides ($-\beta_{ij}\mu_{ij}$), which is a function of the punishment provided as well as altruism (spite), reciprocity, how the other deviates from the social norm, and the cost of punishing:

$$(4) \qquad d_i = \mu_{ij} \left[-\beta_{ij} - \frac{\gamma_i}{2} \mu_{ij} \right]$$

Finally, being publicly punished by others may cause shame,

$$(5) \qquad s_i = \sigma_i (a_i - b) \mu_{ji}$$

[59] We are ruling out dynamic effects of punishment, that is, punishing j in order to get him to cooperate in the future.

where σ_i is a measure of one´s susceptibility to social punishment. The level of disutility experienced from being socially punished depends on one's susceptibility to punishment, the degree of divergence of the individual's extraction relative to the social norm (how fair the individual evaluates the punishment as being) and the severity of the punishment received.

The utility-maximizing level of punishment is:

$$(6)\ \mu_{ij} = -\frac{\beta_{ij}}{\gamma_i} = -\frac{1}{\gamma_i}\left[\alpha_i + \lambda_i(b - a_j)\right]$$

In other words, *i* chooses the level of disapproval that equates the marginal cost of punishment ($\gamma_i\mu_{ij}$) with the marginal benefit of punishment, $-\beta_{ij}$, the negative of the valuation placed on the payoff of the other (as long as $\beta_{ij} < 0$, and chooses zero punishment otherwise). Where punishment is positive, it is clearly increasing in λ_i and decreasing in α_i, as one would expect.

The extraction level of *i* will be a function of *j*´s level of extraction, as well as of the parameters related to the other regarding preferences.

$$(7)\ a_i = \frac{\gamma_j\left(p_1 - p_2\left(1 + \beta_{ij}\right)\right)}{2\lambda_j\sigma_i} + b + \frac{\alpha_j}{2\lambda_j}$$

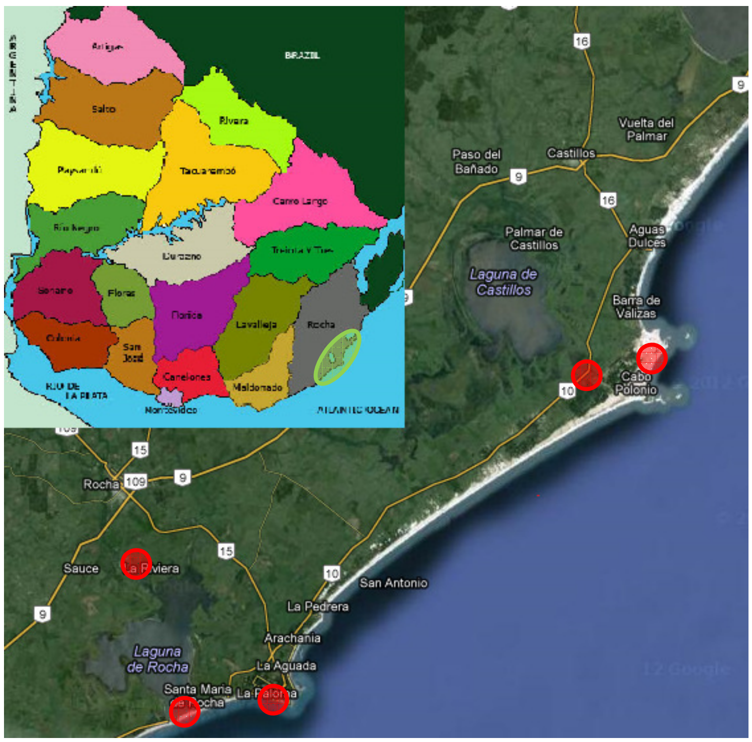
Eq. 7 suggests that *i*'s extraction level varies positively the more altruist *j* is, the higher *j*'s marginal cost of disapproving is and the higher *j*'s extraction levels are. In turn, *i*'s level of extraction will diminish if he is very sensitive to shame, or if *j* is a reciprocator. In this way, social preferences (other than spite) may induce individuals to behave in ways that diverge from the Nash equilibrium in a social dilemma. That is, by valuing other players' payoffs, altruism and reciprocity can make individuals behave closer to the social optimum. Reciprocity motives may also induce an individual to express NMP to norm violators. If individuals feel shame when punished by others, this may also help avert coordination failures in terms of resource extraction.

3.3 Experimental design

3.3.1 Subject pool

Fishermen from five communities that fish either in the Laguna de Rocha or in Laguna de Castillos (two coastal lagoons 50 kilometers away from each other on the Uruguayan sea coast) or in both were recruited (Figure 3.1). We consider a community to be a group of people that live in the same settlement and constantly interact among each other. Individuals from different communities do not differ in terms of ethnicity, while they show some differences in socioeconomic characteristics. These communities differ in terms of how connected they are to the rest of society and the exit options they face. While some live very isolated and fishing is their main source of income (Laguna de Rocha, Puerto los Botes and to less extent El Puente), others are more connected to more densely populated areas and can exploit other exit options (Valizas and Barrio Parque). Facing other exit options as therefore reflected in their income and wealth (see Table A.3.1 in the Appendix).

Figure 3.1 – Location of field experiment (the five communities marked by red dots)



Fishermen from different communities are not used to meeting each other in their daily lives, but they do so when they move across lagoons during fishing high seasons. This is particularly important during the shrimp high season, which usually takes place once a year in the Laguna de Castillos, but rarely in the Laguna de Rocha, because of geographical reasons. PROBIDES (2002) reports that fishermen complained about fishermen from other communities coming in the high season to fish in the lagoon where they, the complainants, fish all year round. We believe that the place of residence is one of the main dividing factors among fishermen from different communities.

3.3.2 The experiment

The experiment consisted of a CPR game of 20 periods, structured in two stages of 10 periods each. In each stage, players interacted either only with members of their own community (in-group treatment) or in groups mixed with fishermen from another community (out-group treatment). This was not explained to the subjects. By this we mean that we did not mention during the in-group stage that all members from the same community were going to play together. Individuals were simply told in which group they would play based on their identifier.

The CPR game was framed around the decision of how many nets to use when fishing. Subjects made their decision in subgroups of four subjects. During the first five periods of each stage subjects played a regular CPR game, in which they consider a common pool resource exploited by individuals who have the same maximum endowment (eight nets) to fish. Individual benefits increase in the number of nets one uses and decrease with the aggregate level of nets used (see Table A.3.2 in the Appendix). Player-i's earnings in each period during the first five periods of each stage were given by the payoff function:

$$\pi_i = 18a_i + 12\sum_{j=1}^{4}(8 - a_j)$$

During the last five periods of each stage a NMP treatment was conducted. In this treatment subjects were allowed to express disapproval of others' fishnet choices. As a consequence, subjects who were punished by other players were assigned a flag, and its color (yellow, orange, or red) indicated how much their peers disapproved of the number of nets they had decided to throw. After making the usual decision on how many nets to use and being informed of the total number of nets used by the subgroup (and therefore being able to determine the others' average extraction), they were able to allocate 0 to 10 disapproval points to each possible choice of fishnets the others may have made (see Table A.3.3 in Appendix).[60]

Punishment points implied no monetary cost to the punished but did imply a monetary cost to the punisher. The cost of each punishment point for the punisher was equivalent to one point in his earnings account. The subject was charged for the total number of disapproval points he used to punish others, irrespective of whether someone had actually chosen the number of nets the punisher decided to punish.[61] In this way, player i's payoff function during the last five periods of each stage is:

$$\pi_i = 18a_i + 12\sum_{j=1}^{4}(8 - a_j) - \sum_{k=1}^{8}\mu_{ij,k}$$

The cost of punishing was set quite low compared to the points a subject could earn during one period. For instance, in one period if all subjects played the Nash equilibrium, each would earn 144 points, whereas if the social optimum was achieved each would earn 354 points. If during the NMP treatment the subject decided to disapprove of all possible fishnet choices with the maximum number of disapproval points, his cost would amount to 80 points (0.5 US). The aim

[60] Subjects could also choose to punish a choice of a number of nets identical to their own. In that case, the punishment would be directed solely to others and not to themselves. This was only explained in case someone asked. Potentially they could disapprove of the eight extraction alternatives at the same time.

[61] The reason why the punisher was charged by the total disapproval points and not just for the ones that corresponded to effective fishnet choices is that it was much simpler to explain and it enabled the subject to calculate the cost by himself. We consider that simplifying mechanisms is particularly important in a framed field environment like ours.

of this treatment was to recreate the state of being socially punished in the field (gossip, direct criticism, etc.) and evaluate its effects on the next periods' extractive decisions. We consider that punishing others socially may also have a social cost to the punisher but we were not particularly interested in studying it; we just intended to show that NMP was not for free for the punisher.

Punishment points for actual choices were added up and yellow, orange and red flags were assigned in accordance with the ranges shown in Table A.3.4 (see Appendix). It was not possible for someone to receive a red flag with just one subject disapproving his fishnet choice.

We employed a hybrid strategy method to implement this treatment. Punishment points were assigned after the subjects had been informed of the total number of nets used by the subgroup and therefore subjects could determine the average number of nets used by the others. It is a "hybrid" strategy method because individuals made decisions in two stages (and not as in the classical strategy method, where both decisions [extraction and punishment] are made at the same time). Brandts and Charness (2010) argue that following a strategy method instead of a direct punishment treatment can lead to lower disapproval among individuals. Also, Blount and Bazerman (1996) argue that individuals are less concerned with fairness when simultaneously choosing between two outcomes than when considering each outcome separately. For this reason, we chose a hybrid strategy method, one that is more similar to assigning punishment based on knowing the effective fishnet choices of each of the other members of the subgroup, but that still preserves anonymity. We discarded the alternative of disclosing actual individual levels of extraction in a random order because we considered there was a risk that anonymity would be violated.[62]

3.3.3 The structure of the experiment

Subjects were recruited during a survey that took place in March 2011. The aim of the survey was to gather data on socioeconomic characteristics and environmental perception among the resource users of artisanal fisher communities in Rocha's coastal lagoons. At the end of the questionnaire, the interviewee was asked whether he would be interested in participating in an activity where he could earn on average 2 daily wages (30 US dollars), depending on the decisions he would make. A week before the experiment we visited the communities where we delivered flyers in person to people from the five communities, and we made phone calls to those who had already been surveyed but could not be located while we visited the communities.

The experiment was conducted in two sessions in November 2011. Both sessions took place at La Paloma, a town in the province of Rocha, Uruguay. The communities that participated in each session were determined randomly (Table 3.1). Contrary to most framed field experiments, in this study subjects were transported from the place where they lived to the town where the

[62] Keeping anonymity both in individuals' extraction decisions and in NMP was a priority. Indeed, as Anderies et al. (2011) point out, working with communities in field experiments requires developing this task with responsibility, because the game may not end when experimenters leave, and this may have spillover consequences in their daily life.

experiment took place.[63] The aim of this design was to make subjects from different communities meet. This required that fishermen leave their community to attend the activity. It was particularly cumbersome to convince subjects to travel, and we believe it was the main reason why the number of participants was not as high as desired.

Table 3.1 – Characteristics of the experimental sessions

In-group: "Groups and subgroups with individuals belonging to the same community".
Out-group: "Groups and subgroups with subjects belonging to two communities".
NMP: "Expressing disapproval of others' extraction levels. Those punished receive flags".

	Subjects		Treatments by period			
	Included in analysis	Discarded [a]	1-5	6-10	11-15	16-20
Session 1						
Laguna de Rocha	8	3	ingroup	ingroup punishment	outgroup	outgroup punishment
Valizas	8	3	ingroup	ingroup punishment	outgroup	outgroup punishment
Session 2						
El Puente	12		outgroup	outgroup punishment	ingroup	outgroup punishment
Puerto los Botes	8		outgroup	outgroup-punishment	ingroup	outgroup punishment
Barrio Parque	8		outgoup	outgroup punishment	ingroup	outgroup punishment
Total	44	6				

[a] During session 1 the subjects who turned up from Laguna de Rocha and Valizas were not multiples of four so three subjects from each community were selected randomly to play in subgroups of three and were reshuffled solely among the six all the periods. They were not considered in the analysis.

When subjects arrived at the venue, they drew a number from a bag (one bag per community). This number represented their identifier, and assigned each subject into a group of either eight or twelve members for each stage. Within these groups, subjects would play a CPR game in subgroups of four. The out-group treatment implied subgroups in which two subjects belonged to one community and two to the other.[64] In order to avoid repeated game type of behavior, after each period subjects were reshuffled among all subjects in a group of eight or twelve. The subgroups they would play in the 20 periods were predetermined by the identifier number. It was common knowledge that the matching procedure between periods was random and had been determined by the initial draw of participants' identifier numbers. After each period, the experimenters indicated to the participants which subgroup of four they would play in the next period; at the end of the first 10 periods, participants were told in which group they would then

[63] Buses for each community were hired to pick up participants and transport them to the venue.
[64] In session 2, as there was one community in which there were twelve subjects (El Puente) and in the other two there were eight, during the out-group treatment, subgroups were composed of two subjects from el Puente and two from one of the other two communities or three from El Puente and one from the other community. In all cases the out-group treatment implied mixing just two communities.

play in (this implied a change in treatment from in-group to out-group or vice versa). During session 1, subjects played in an in-group treatment during the first stage, while in session 2 we reversed this order (see Table 3.1). This design enabled us to control for order effects.

Once in subgroups of four members, subjects were asked to sit with their backs facing each other so that they could not see the others' choices. Each group was conducted by a moderator who gave the instructions throughout the game, plus a monitor for every subgroup of four. This ensured that subjects did not interact during the game, and that an experimenter was always available to explain them how to use the material.

Subjects received a payoff table and an earnings sheet where they kept a record of their decisions and points gained. The payoff table summarized the pay-off consequences of all combinations of own nets used and the total number of nets used by the other three members of a subgroup (see Table A.3.2 in the Appendix). The exchange rate was set at 100 points for 0.62 US dollars. When looking at the payoff table, subjects had to make a decision as to how many nets to use (minimum one, maximum eight), which they wrote on a slip of paper and handed it in to the experimenter. Once the four subjects had written out their decisions, the total number of nets used by the subgroup was announced so that each subject could calculate the number of points they had earned and write that figure on their earnings sheet. The explanation of the game followed Cardenas (2003). The actual experiment began once the moderator had conducted three rehearsal periods and once all questions from participants had been clarified. All decisions were made privately and individually and only the total extraction by the four players was publicly announced.

Before the punishing treatment started an example was provided. The example showed three subjects' disapproval cards: one punishing without any criteria, one punishing those who used many nets and one not punishing at all. The choice of nets and the disapproval points assigned were private information; the only public information was the flag received in case the subject was punished by the rest by more than one point. Subjects had to hold the flag so that others could see it during the next period of the game.

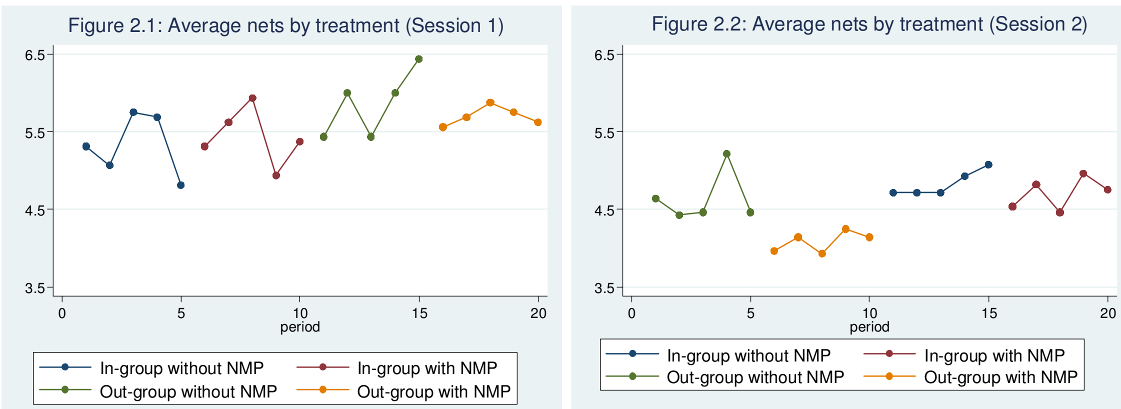
At the end of each experimental session we conducted a post-experiment survey which contained questions about reasons for disapproval, and feelings when being disapproved of. Each session of the experiment lasted about three hours and participants earned on average 30 US dollars (including a 5 US dollar show-up fee), a figure which amounts to 10% of a monthly minimum wage. [65]

[65] The experimental design excluding the in-group out-group treatment was tested with 36 undergraduate students.

3.4 Results

Figure 3.2 shows average extraction levels by period and treatment for session 1 and 2, respectively. At first glance, it suggests that the in-group/out-group treatment does not seem to induce significant changes in behavior when NMP is not available. Subjects' extraction levels in session 1 were higher during the out-group treatment without NMP but did not change substantially for subjects in session 2. The NMP treatment seems to have had a slight positive effect in terms of cooperation especially during the out-group treatment. It lowered average extraction levels in the second stage of session 1 and in both stages in session 2. It should also be noted that the three communities that participated in session 2 exhibited extraction levels significantly below those of the two communities that participated in session 1.

Figure 3.2 – Average nets by treatment



3.4.1 Testing treatment effects

Next we study players' extractive decisions in a dynamic analysis. This allows us to test the in-group out-group and NMP treatments. For this purpose, we estimate a dynamic model following a specification similar to that of Hayo and Vollan (2012), such that:

$$a_{it} = \alpha_i + \beta_1 \pi_{i;t-1} + \beta_2 \pi_{-i;t-1} + \beta_3 out-group\ w/NMP_{\ i} + \beta_4 in-group\ \ w/\ NMP_{\ i} \\ + \beta_5 in-group\ w/out\ NMP_{\ i} + \beta_6 stage_i + e_{it}$$

where:

a_{it} is i's extraction level in period t, $\pi_{i;t-1}$ is individual $i's$ payoff in a previous round. High payoffs in the previous round can be achieved either because there is cooperation (high group payoff and high individual payoff) or because of self-interested behavior (low group payoff and high individual payoff). Controlling for the group's payoff allows us to distinguish which of the two strategies is reinforced over time.

$\pi_{-i;t-1}$ is the payoff of the rest of individual i's subgroup (excluding individual i) in the previous round. Even if the game is a series of one shots and members of a subgroup change in every period, subjects may use information on the behavior of other subjects as a guide for future behavior. A negative relation between the group's payoffs in the previous period and the individual's extraction levels may suggest the existence of social preferences.

Treatments are tested in two ways. First, two dummy variables were included in the model: *in-group* that equals 1 if the players are under an in-group treatment (and 0 if they are playing the out-group treatment) and *NMP*, that equals 1 when the extraction decision is taken during a round that allows for NMP [rounds 6 to 10 and 16 to 20] and 0 otherwise). We also included a third dummy variable (*second stage*), which equals 1 for rounds 11 to 20. Second, we tested the interaction between treatments (see equation above). For this purpose, three dummy variables were included: *out-group with NMP*, *in-group with NMP*, and *in-group without NMP* (*out-group without NMP is the base scenario*). Each of them equal 1 during the periods that they describe, and 0 otherwise. A fixed effects model was performed to control for individuals' time invariant characteristics. Time fixed effects were not included because they show high correlation with treatment variables (treatment dummy variables are time fixed effects).

Columns (1) to (6) in Table 3.2 show that while the in-group treatment has no effect on individuals' decisions, players chose lower extraction levels when playing during the NMP. Column (2) shows that the NMP treatment effect is significant independently of the additional variables included.

Table 3.2 – Dynamic fishnet decisions. Dependent variable: fishnets_{it}

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$earnings_{i;t-1}$				0.004*	0.004*	0.004*		0.004*
				(0.002)	(0.002)	(0.002)		(0.002)
$earnings_{-i;t-1}$				-0.002*	-0.002*	-0.002*		-0.002*
				(0.001)	(0.001)	(0.001)		(0.001)
in-group	0.002		0.002	0.023		0.023		
	(0.139)		(0.139)	(0.128)		(0.128)		
NMP		-0.225*	-0.225*		-0.227**	-0.227**		
		(0.113)	(0.113)		(0.107)	(0.107)		
out-group with NMP							-0.414**	-0.401**
							(0.160)	(0.153)
in-group without NMP							-0.187	-0.159
							(0.199)	(0.187)
in-group with NMP							-0.223	-0.215
							(0.172)	(0.158)
second stage	0.402***	0.402**	0.402***	0.392***	0.383**	0.376***	0.402***	0.379***
	(0.139)	(0.161)	(0.139)	(0.130)	(0.149)	(0.132)	(0.139)	(0.131)
_cons	4.733***	4.847***	4.846***	5.165***	5.358***	5.352***	4.940***	5.387***
	(0.120)	(0.106)	(0.137)	(0.526)	(0.539)	(0.540)	(0.164)	(0.534)
Obs.	880	880	880	836	836	836	880	836
Subjects	44	44	44	44	44	44	44	44
r2 within	0.019	0.025	0.025	0.031	0.037	0.037	0.029	0.041
r2 overall	0.009	0.012	0.012	0.180	0.180	0.180	0.014	0.171
r2 between	.	.	0.108	0.927	0.940	0.939	0.069	0.935

*** p<0.01;** p<0.05; * p<0.1

Robust standard errors in parenthesis

Heterogeneous treatment effects of NMP between in-group and out-group settings are shown in columns (7) and (8) of Table 3.2. On the one hand, it can be seen that the level of nets chosen under the out-group without NMP are not significantly different from the ones under the in-group, both with and without NMP. On the other hand, subjects under the out-group with NMP treatments extract lower levels than when the NMP is not allowed (the -0.4 coefficient amounts to 20% of a standard deviation in nets). Finally, the behavior of individuals under the in-group treatment is not significantly affected by the NMP treatment. The *second stage* dummy variable is positive and significant in all models. That is, subjects increase the average extraction level during the second stage, independently of the treatment they played first. The fact that cooperation decays throughout the game follows previous literature.

Regarding earnings in previous rounds, Models (4), (5), (6) and (8) shows that $\beta_1 > 0$, and $\beta_2 < 0$. This result is consistent with Hayo and Vollan (2012), and suggests that social preferences mechanisms are influencing players' decisions. As stated before, $\beta_1 > 0$, jointly with $\beta_2 < 0$ implies that individuals behave more cooperatively if their group in the past round performed well. This implies that their recent past experience influences their decisions, despite changing partners after each round.

3.4.2 Determinants of extraction decisions

In this section we analyze whether there are socio-demographic determinants of individual choices regarding extraction decisions. We do this for three variables of interest: number of nets chosen in the first period (columns 1 and 2), total number of nets chosen throughout the 20 periods (columns 3 to 5) and average nets (columns 6 to 8). Table 3.3 reports for each of these variables the general and reduced estimations.

Almost no individual-level economic and demographic variable seems to explain extraction choices, as Heinrich (2001) and Hayo and Vollan (2012) found. Wealth and age are the only observable individual determinants of choices which are significant. The magnitude of the wealth coefficient is worth noting: a one standard deviation increase in the wealth index increases the average choice of nets in 44% of a standard deviation. The wealth index was elaborated by means of factor analysis. The index considers different durable goods a household may own.[66] Cardenas (2003) also finds a positive relation between wealth and choices of extraction, and hypothesizes that low wealth status may reflect greater experience in managing a common pool resource. However, in our study this does not seem to be the case. Being a subject whose main activity is fishing is not related to extraction levels (see Table 3.3). Cardenas also provides an alternative explanation, which in our case can be understood if wealthier participants showed smaller marginal utilities from the cash earned in the experiment, thereby having less incentives to cooperate because the marginal value of potential gains is smaller than for the poorer participants. Hayo and Vollan (2012) report a positive coefficient on the upper middle and highest quartiles of income and also argue that high income might reveal a person's stronger preference for consumption, risk and competition.

The other significant determinant of fishnet choices is community membership. El Puente (the baseline in the regression) extracted significantly less than the other four communities. Also, the Wilcoxon-Mann-Whitney ranksum (WMW ranksum) tests reject median and mean extraction levels equality between places of residence, two-by-two, at 10% level of confidence (Table A.3.5 in Appendix). This hypothesis is not rejected only in the case between Barra de Valizas and Barrio Parque, with reference to average nets thrown, and between Laguna de Rocha and Barra de Valizas and Barrio Parque, with reference to average earnings during the experiment. However, median average earnings equality between the last two is rejected. These results, together with the non-significance of individual characteristics, strongly support the hypothesis that group level institutions or social norms influence individuals' behavior.

[66] The variables the index includes are the following: water heater, fridge, TV, radio, cable TV, DVD, washing machine, microwave, computer, Internet, phone, motorbike, car and horse.

Table 3.3 – Determinants of subjects' extraction decisions

	Nets first period		Total nets			Average nets		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Laguna de Rocha	0.84	1.79**	31.01**	40.33***	39.94***	1.55**	2.02***	2.00***
	(1.13)	(0.87)	(13.48)	(11.44)	(12.15)	(0.67)	(0.57)	(0.61)
Valizas	2.53**	1.17	52.92***	51.88***	50.16***	2.65***	2.59***	2.51***
	(1.08)	(0.87)	(12.93)	(11.60)	(12.08)	(0.65)	(0.58)	(0.60)
Botes	2.97**	1.42	25.05	24.22**	25.25**	1.25	1.21**	1.26**
	(1.26)	(0.87)	(15.06)	(11.15)	(11.73)	(0.75)	(0.56)	(0.59)
Barrio Parque	3.23**	1.42	38.35**	29.81**	31.23**	1.92**	1.49**	1.56**
	(1.29)	(0.87)	(15.35)	(11.49)	(12.13)	(0.77)	(0.57)	(0.61)
female	0.64		-5.46			-0.27		
	(0.86)		(10.28)			(0.51)		
age	-0.02		-1.07***	-0.53*	-0.58*	-0.05***	-0.03*	-0.03*
	(0.03)		(0.36)	(0.27)	(0.29)	(0.02)	(0.01)	(0.01)
years of schooling	-0.02		-3.50*			-0.17*		
	(0.17)		(1.98)			(0.1)		
drinkable water	-1.8		-6.39			-0.32		
	(1.08)		(12.83)			(0.64)		
electricity	-1.04		-17.02			-0.85		
	(1.06)		(12.61)			(0.63)		
wealth	0.49*		11.94***	8.04***		0.60***	0.40***	
	(0.29)		(3.4)	(2.90)		(0.17)	(0.14)	
per capita income (logs)	-0.99**		1.5			0.07		
	(0.48)		(5.7)			(0.29)		
fishing main activity	1.11		-2.29			-0.11		
	(0.77)		(9.17)			(0.46)		
perception[a]	-0.28		-3.85			-0.19		
	(0.77)		(9.17)			(0.46)		
trust[b]	-0.16		-13.11			-0.66		
	(1.21)		(14.4)			(0.72)		
second quartile (wealth)					7.00			0.35
					(11.13)			(0.56)
third quartile (wealth)					25.35**			1.27**
					(11.82)			(0.59)
fourth quartile (wealth)					27.02**			1.35**
					(12.72)			(0.64)
Constant	11.85**	3.83***	114.58**	72.59***	81.56***	5.73**	3.63***	4.08***
	(4.46)	(0.55)	(53.1)	(15.38)	(16.34)	(2.65)	(0.77)	(0.82)
Obs.	43	44	43	44	44	43	44	44
R-squared	0.35	0.12	0.61	0.46	0.45	0.61	0.46	0.45

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a] Believes that preserving the environment in coastal lagoons is mainly a responsibility of the people rather than the government.

[b] Believes one can trust most people.

3.4.3 Types

In this section we classify subjects with reference to their actual cost of deviating from self-interested behavior, and their relationship with their subgroup partners into four categories: free riders, cooperators, conditional cooperators and others. The subject's actual cost of deviating from self-interested behavior is computed as the difference between the payoff the subject would have obtained in that period if he had extracted the maximum level (given the others' extraction choices) and the actual payoff he obtained. The classification was determined by analyzing individuals' behavior in the periods in which NMP was not available and subjects had already played one period and therefore had a reference point (periods 2-5 and 11-15).

Following the algorithm used by Kurzban and Houser (2005), we define a cost of deviating threshold equal to 18 points.[67] We consider a free rider to be a subject whose cost of deviating is below or equal to the 18 points threshold during all 9 periods. A cooperator is a subject whose cost of deviating is always above or equal to the 18 points threshold during these periods. In turn, a conditional cooperator is defined as a subject whose cost of deviating is both above and below the 18 points threshold and; who shows a positive slope in an ordinary-least-squares regression of own cost of deviating on the average cost of deviating from other members of his subgroup in a previous period. The slope measures the subject's responsiveness to others' behavior and could be interpreted as a proxy of λ mentioned in Section 2. Subjects that exhibit any other behavior in terms of their cost of deviating are classified as "other." The scatter plot of each subject's cost of deviation and others' average cost of deviation in the previous period are shown in Figure A.3.1 of the Appendix.

Table 3.4 – Distribution and main characteristics by behavioral types

Type	Frequency	Average nets in 20 periods	Average total earnings (US)
Free rider	18%	6.9	28.31
Conditional cooperator	25%	4.9	28.22
Cooperator	20%	3.2	28.82
Other	36%	4.9	29.26

Table 3.4 reports the frequency of each type, average extraction levels and average earnings by type. There is a similar fraction of free riders and cooperators and a slight predominance of conditional cooperators. However, there is a substantial fraction of the subject pool which cannot be classified under any of these three types. It is worth noting that cooperators achieved higher earnings than free riders. This is due to a significant concentration of cooperators within some communities. Table A.3.6 in the Appendix shows that 50% of subjects coming from El Puente were classified as cooperators, enabling them to achieve greater earnings when playing together during the in-group treatment. In turn, Barrio Parque exhibits a high concentration of free riders, which lowered their earnings during the in-group treatment. Subjects classified as "Other"

[67] The threshold is set at 18 points because it is the median cost of deviating during the 9 periods.

achieved the highest average earnings. In general, these subjects behaved very similarly to free riders. However, in some periods they increased their cost of deviating above the 18 point threshold, simultaneous to the rest of their partners. This allowed them to benefit from occasional synchronized cooperation and to thus obtain greater earnings.

3.4.4 Punishing behavior

In this section we analyze punishers' behavior. On average, 71% of subjects chose to punish in each period in which punishment was allowed. Disapproval was substantial throughout the game and was surprisingly quite high in the last period, even if subjects knew the experiment would be over after that period. Figure 3.3 presents average punishing points by period for the two sessions separately. It should be noted that in the out-group treatment, subjects were mixed among in-group and out-group members and did not know the extraction levels of each of them, so punishment could not be directly specifically to out-group members with certainty. Session1 exhibited higher levels of punishment during the out-group treatment, though this did not occur in session 2 in which the average disapproval levels are not significantly different in the out-group and in-group treatments. Considering the two sessions together, the amount of punishment is not significantly different in the out-group and in-group treatments.[68]

Figure 3.3 – Average punishing points by period

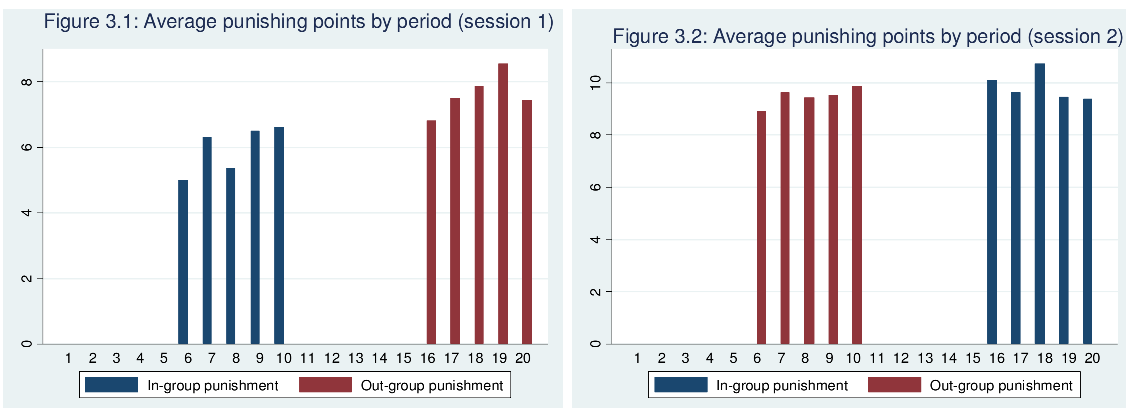


Figure 3.4.1 reports total disapproval points by the number of nets the subject chose to use in that period (horizontal axis) and to which choice of number of nets he decided to punish (bars). Those who use less than 6 nets disapprove of those who use more nets. This punishment of free riding could also be considered altruistic punishment as individuals incur material costs when punishing and reap no material benefits from punishing, because after punishing players are reshuffled before playing the next period. Also, we can observe antisocial punishment (punishment to cooperators): those who use 6 or more nets choose to disapprove of those who used fewer nets. Figure 3.4.2 shows per subject disapproval points instead of total points. Most subjects using 2 nets focused their disapproval on those throwing 6 nets and more, and only few subjects were spending a lot of disapproval points on those throwing 4 and 5 nets. There are only three subjects that used 6 or more nets and they spent a large number of disapproval points in

[68] A Wilcoxon-Mann-Whitney ranksum does not reject the equality between punishment directed during the out-group and in-group treatments for the two sessions together (*p-value*: 0.54).

lower extraction levels. This misdirected punishment is also observed by Falk et al. (2000), Masclet et al. (2003) and Gätcher and Herrmann (2011). Figures A.3.2.1 and A3.2.2 in the Appendix show average extraction levels in both sessions excluding sub-groups in which these three subjects participated. There, it can be observed that the effectiveness of the NMP treatment is greater than that observed in Figure 3.2.

Figure 3.4 – Punishment points by receiver and sender's nets

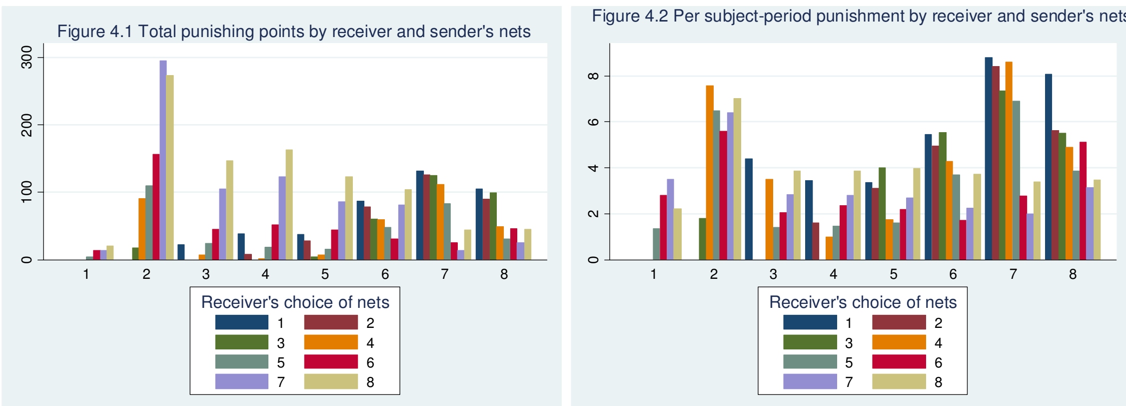


Figure 3.4 also indicates that there is some punishment from senders toward receivers using the same number of nets as themselves, especially when using a large number of nets. This could be interpreted as trying to discourage others from free riding while not sticking to the social norm in their actions (i.e., "do as I say and not as I do"). NMP was quite intense, when translating disapproval points into flags effectively received. We observe that on average there were 1.7 flags delivered per subgroup of four per period. Table 3.5 shows the distribution of flags received depending on whether the subject had chosen an extraction level below or above the subgroup's mean. Although the majority of the flags (60%) were awarded to subjects with extraction levels above the subgroups' mean, the remainder 40% were awarded to individuals with extraction levels below their subgroup's mean.

Table 3.5 – Total flags shown by round

		Negative deviation max{0;āt-1-ai,t-1}				Positive deviation max{0;ai,t-1-āt-1}			
Period	Obs.	Total flags	Yellow	Orange	Red	Total flags	Yellow	Orange	Red
6	44	2	0	2	0	10	8	2	0
7	44	6	4	1	1	6	4	2	0
8	44	7	4	3	0	11	7	4	0
9	44	7	3	3	1	8	5	3	0
10	44	7	5	2	0	16	12	3	1
16	44	9	5	4	0	15	11	1	3
17	44	9	8	1	0	15	9	2	4
18	44	14	7	7	0	9	3	4	2
19	44	7	4	3	0	13	6	7	0
20	44	8	2	6	0	11	10	0	1
Total flags	190	76	42	32	2	114	75	28	11
%	100%	40.0%	22.1%	16.8%	1.1%	60.0%	39.5%	14.7%	5.8%

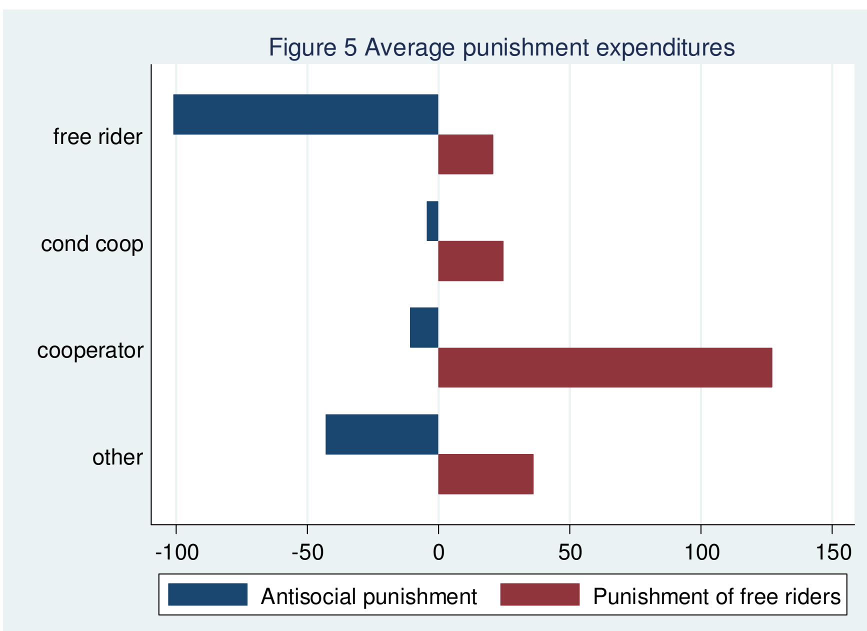
Cooperators were the ones who spent on average most points on disapproving others' behavior followed by free riders (see Table 3.6). As expected, cooperators were the least punished and free riders the most.

Table 3.6 – Punishment sent and received by subject's type

	Average cost of punishment	Average number of flags received
Free rider	122	5.4
Conditional Cooperator	29	4.2
Cooperator	138	3.0
Other	79	4.6

Following Herrmann et al. (2008), we consider punishment for extraction levels greater than one's own as punishment of free riding, and antisocial punishment to punishment directed to extraction levels equal to or smaller than one's own. Figure 3.5 shows average punishment expenditures following this categorization by subjects' types. Cooperators performed most punishment on free riders. In turn, free riders did most of the antisocial punishment although it was also observed to some extent in all the types. As with what was observed for overall punishment, antisocial punishment does not significantly differ between in-group and out-group treatments.

Figure 3.5 – Average punishment expenditures



Among the post-experiment questions, we included one question regarding reasons for disapproving. Table 3.7 confirms that most disapproval points were directed to those who used many nets. However, from the three subjects who significantly disapproved of those who used few nets, two argued that they were disapproving those who threw few nets because they were missing out on chances to fish. They were the ones receiving the most punishment from others. The majority of subjects chose to disapprove of others' behavior because they were using too many nets (55%). It is worth noting that the average total cost of disapproval of the three subjects who disapproved of those who chose a low number of nets, was particularly high.

Table 3.7 – Reasons for punishing

Reasons for punishing	Average cost of disapproval	Average nets in punishment periods	Average number of flags received	Percentage of total subjects
Those who play different	144	6.8	6.5	5%
Those who threw many nets	84	4.3	3.7	55%
Did not disapprove	0	6.2	5.4	11%
Without any criteria	46	5.3	4.5	14%
Did not understand	21	3.3	3.0	5%
Those who threw few nets	376	6.4	6.7	7%
Part of the game	22	4.4	4.0	5%

Following Masclet et al. (2003) we estimated the following model:

$$P_{ik}^{t} = \beta_0 + \beta_1(\max\{0, a_i - a_k\}) + \beta_2(\max\{0, a_{av} - a_k\}) + \beta_3(\max\{0, a_k - a_i\}) + \beta_4(\max\{0, a_k - a_{av}\})$$

Where P_{ik}^{t} is the number of disapproval points that *i* assigns to *k* in round *t*, the coefficient β_1 is associated with positive deviations from the punisher's fishnet choice, that is, cases in which the punished chose fewer nets than the punisher, β_2 reflects the impact of positive deviations from the subgroup's average. In turn, β_3 reflects the relevance of negative deviations from the punisher's fishnet choice, that being situations in which the punished subject chose more fishnets than the punisher. Finally, β_4 is associated with the negative deviation from the subgroup's average. We included individual fixed effects to control for individuals' time invariant characteristics. We estimated the following model for each fishnet choice that could be punished. For instance, the first column in Table 8 reflects the determinants of punishing those subjects who chose 1 fishnet. As Table 3.8 shows, both positive (antisocial punishment) and negative (punishment of free riding) deviations from the punisher's fishnet choice are significant. But as Masclet et al. (2003) showed, there is an additional effect regarding deviations of the punished subject from the subgroup's average. [69]

[69] Estimates from a Tobit model point to the same conclusions but in that model, coefficients are slightly smaller in magnitude.

Table 3.8 – Determinants of disapproval points directed to each of the fishnet options

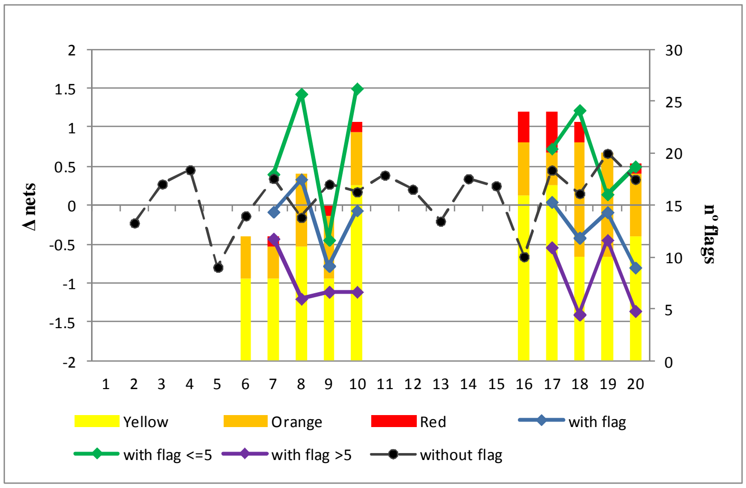
	1_net	2_nets	3_nets	4_nets	5_nets	6_nets	7_nets	8_nets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Positive deviation from i's own extraction ($\max\{0, nets_i - nets_k\}$)	0.79***	0.53***	0.89***	0.48***	0.67***	2.22***	1.46***	
	(0.07)	(0.07)	(0.10)	(0.10)	(0.12)	(0.15)	(0.42)	
Negative deviation from i's own extraction ($\max\{0, nets_k - nets_i\}$)				1.33***	0.35***	0.49***	0.33***	0.04
				(0.28)	(0.09)	(0.06)	(0.06)	(0.06)
Positive deviation from average ($\max\{0, nets_{av} - nets_k\}$)	0.13	0.31***	0.03	0.83***	0.70***	0.54	0.31	
	(0.09)	(0.10)	(0.15)	(0.16)	(0.20)	(0.52)	(1.91)	
Negative deviation from average ($\max\{0, nets_k - nets_{av}\}$)			-0.43	-0.16	0.32*	0.26***	0.29***	0.62***
			(0.62)	(0.32)	(0.16)	(0.09)	(0.08)	(0.08)
Constant	0.24***	0.28***	0.30***	0.42***	0.49***	0.38***	0.84***	0.98***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.08)	(0.10)
Obs.	440	440	440	440	440	440	440	440
R-squared	0.55	0.47	0.44	0.32	0.24	0.51	0.29	0.30
Number of id_	44	44	44	44	44	44	44	44

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3.4.5 Reaction to punishment

In this section we analyze whether punishment generated a change in behavior among those who were punished. The descriptive analysis suggests that flags produce variations in individuals' behavior. Figure 3.6 shows that individuals who received a flag in the previous period, on average, changed their behavior in the next period. However, at the individual level this is not always the case. One of the reasons for observing heterogeneity in terms of reaction to punishment is due to the fact that those who are punished are not only the ones who choose a high number of nets, but also those with a low number of nets. Figure 3.6 shows that while those who received a flag when throwing more than five nets diminish their choice in the next period, those who received a flag when throwing five or less nets increase the number of nets chosen in next period.

Figure 3.6 – Total fishnets variations and total number of flags



Also, fishnets variations of those that receive flags seem to be greater, in absolute terms, for those who received a flag compared to those who did not receive one (Table 3.9).

Table 3.9 – Fishnet variations in extraction by flags received

	n	Mean	s.d.	Min	Max
received a flagin t-1	148	-0.26	2.03	-7	5
received a flag when choosing nets >5 in t-1	81	-0.99	1.91	-4	5
received a flag when choosing nets <=5 in t-1	67	0.61	1.83	-7	2
did not receive a flag in t-1 (during NMP)	204	0.25	1.84	-7	7

However, 33% of the individuals who received a flag did not change their behavior in the next round. This proportion is higher between those who decided to throw two (55%) and eight (48%) nets in the previous period (Table A.3.7 in Appendix). Decisions to throw two or eight nets are the modes of nets' distribution. A large number of subjects who chose these values decided not to change their choice, independently of what others think. Noussair et al. (2011) argue that subjects may not view the norm of cooperation as the norm that punishment should enforce, but that other norms such as "try to fish as much as possible" may be the prevailing ones. Therefore, some punished subjects may interpret punishment for using many nets as inappropriate and respond by raising the number of nets or maintaining their choice at the maximum number of nets.

Not all the flag colors produce the same reaction (see flag range in Table A.3.4 in Appendix). Subjects are more indifferent to yellow flags than to the others: 42% of the cases in which a subject received a yellow flag, he did not change his decision in the next period (Table A.3.8 in Appendix). When analyzing subjects' reaction in relation to how they said that they felt when receiving a flag in the post-experiment survey, those that declared indifference did not change their behavior after receiving a flag, or increased their decision in almost 70% of the cases (Table 3.10). Also, those who answered that they felt uncomfortable diminished their decision in the next period 52% of the times they received a flag (while in 28% of the cases, they maintained the decision they had made in the previous period). Table 3.10 also confirms that those subjects who considered they had been punished unfairly raised their extraction levels in the following period because they experienced anger.

Table 3.10 – Nets variations and feelings (%)

	Feeling when receiving a flag - Total						Feeling when receiving a flag If nets in previous round <=5						Feeling when receiving a flag If nets in previous round >5					
nets variation	**Unc.**	**Angry**	**Indif.**	**Fair**	**Other**	**Total**	**Unc.**	**Angry**	**Indif.**	**Fair**	**Other**	**Total**	**Unc.**	**Angry**	**Indif.**	**Fair**	**Other**	**Total**
-	52	0	30.3	36.59	26.7	35.1	25	0	36	20.84	20	26.9	76.9	0	26.8	58.82	30	42.0
=	28	0	42.4	19.51	40	33.1	41.7	0	16	12.5	20	19.4	15.4	0	58.5	29.41	50	44.4
+	20	100	27.3	43.91	33.34	31.8	33.3	100	48	66.67	60	53.7	7.7	0	14.6	11.76	20	13.6
Total	100	100	100	100.01	100.01	100.01	100	100	100	100.01	100	100	100	0	100	99.99	100	100

Regarding behavior by type, when being punished for choosing more than 5 nets, conditional cooperators reacted by lowering extraction levels 50% of the time (as well as others), whereas free riders lowered extraction levels 26% of the time (Table A.3.9 in Appendix).[70] Instead, when being punished for throwing 5 or less nets, free riders reacted by increasing their extraction levels 100% of the time, conditional cooperators 65% and cooperators 38%, respectively.

The next step is to formally test for the behavior depicted above. We adapt the reaction function included in Masclet et al. (2003) and Noussair and Tucker (2005) and test whether player´s *i* decision changes from period *t-1* to period *t* is a function of the punishment received in the previous period, and his extraction deviations from group average decisions:

$$a_i^t - a_i^{t-1} = \beta_0 + \beta_1 * YF_i^{t-1} + \beta_2 * OF_i^{t-1} + \beta_3 * RF_i^{t-1} + \beta_4 * (max\{0, a_i^{t-1} - \bar{a}^{t-1}\}) + \beta_5 * (max\{0, \bar{a}^{t-1} - a_i^{t-1}\})$$

Table 3.11 – Reaction to punishment. Dependent variable: $fishnets_t$-$fishnet_{t-1}$

[70] By definition cooperators never chose more than 5 nets.

Dependent variable: $Fishnets_{i,t} - fishnets_{i,t-1}$												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Periods	**7-10 & 17-20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**	**1-5 & 11-15**	**7-10 & 17-20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**
Positive deviation from average; $max\{0; a_{i,t-1} - \bar{a}_{t-1}\}$					-0.554***	-0.967***	-0.93***	-0.134	-0.684***	-0.967***	-0.156	-0.658***
					(0.119)	(0.179)	(0.17)	(0.456)	(0,173)	(0.163)	(0.448)	(0.159)
Negative deviation from average; $max\{0; \bar{a}_{t-1} - a_{i,t-1}\}$					0.917***	0.832***	0.84***	0.664***	0.068	0.826***	0.633***	0.117
					(0.152)	(0.144)	(0.15)	(0.156)	(0,511)	(0.144)	(0.159)	(0.491)
Flag in t	-0.61*						-0.26	-0.062	-0.159			
	(0.32)						(0.22)	(0.268)	(0,309)			
Yellow flag in t-1		-0.471	-0.033	0.004						-0.256	-0.010	-0.058
		(0.328)	(0.373)	(0.314)						(0.241)	(0.373)	(0.321)
Orange flag in t-1		-0.793*	-0.029	-0.861						-0.329	-0.221	-0.609
		(0.467)	(0.450)	(0.583)						(0.333)	(0.408)	(0.516)
Red flag in t-1		-1.105	2.802***	0.181						0.340	1.824	0.453
		(1.205)	(0.993)	(0.508)						(0.670)	(1.382)	(0.632)
_constant	0.29**	0.300**	0.566***	-0.702***	-0.178	0.132	0.21	-0.132	0.363	0.236	-0.107	0.315
	(0.14)	(0.138)	(0.088)	(0.194)	(0.163)	(0.172)	(0.19)	(0.205)	(0,345)	(0.178)	(0.215)	(0.331)
Nº observations	352	352	213	139	396	352	352	213	139	352	213	139
Nº individuals	44	44	39	32	44	44	44	39	32	44	39	32
r2 within	0.020	0.023	0.033	0.036	0.291	0.350	0.354	0.141	0.159	0.356	0.158	0.178
r2 overall	0.018	0.018	0.020	0.023	0.185	0.196	0.197	0.081	0.099	0.199	0.088	0.110
r2 between	0.013	0.001	0.001	0.002	0.021	0.038	0.039	0.000	0.036	0.040	0.000	0.022

legend: *** p<0.01;** p<0.05; * p<0.1
robust standard errors in parenthesis

Where YF_i^{t-1}, OF_i^{t-1}, and RF_i^{t-1}, are dummy variables that indicate if the individual received a yellow, orange or red flag, respectively, in a previous period, $max\{0, a_i^{t-1} - \bar{a}^{t-1}\}$ is a variable that indicates if the individual extracted more than his subgroup average in a previous period, and the deviation magnitude, while $max\{0, \bar{a}^{t-1} - a_i^{t-1}\}$ is the same but for negative deviations from the subgroup average in a previous period.[71] We test this model for periods where flag reaction could take place (periods 7 to 10 and 17 to 20) and separately for those that chose 5 or less nets and more than 5 nets, respectively. A control model during periods where reaction is not possible is also included in column (5) to compare conformity effects. We included individual fixed effects to control for non-observable factors that may affect individual decisions.

Model (1) in Table 3.11 suggests that being punished generates a downward adjustment in the following periods. When distinguishing by flag colors (Model 2), we observe that receiving an orange flag appears to have a small influence on diminishing individuals' nets choice, but its effect is diluted when splitting the sample between those people who receive a flag when they threw five or less nets, and more than five (Models 3 to 4). It is worth noting the large increase in fishnets choices in *t* when receiving a red flag, having thrown five or less nets in *t-1*. Subjects react strongly, in a non-cooperative way, when they feel they have been unfairly punished.

When conformity effects are allowed (how individuals deviate from the subgroup's average in the previous period), receiving a flag is not significant in any case any longer (Models (5) to (12)). Those who threw fewer nets than the subgroup's average in the previous period increase their decision in the next period, while those who threw more than average in previous period decrease their decision the next period. The presence of conformity effects is consistent with Masclet et al. (2003), and Hayo and Vollan (2012). As might be expected, the second mechanism does not take place if we look only at the reaction of those who received a flag when throwing five or less nets, while the first mechanism does not work when the reaction of those who threw more than five nets is studied. The magnitude of the conformity effect is larger during the NMP periods (especially positive deviations of the subjects' extraction relative to the subgroup's mean), which could indicate that there may be an additional impact of NMP increasing the convergence to the social norm. However, confidence intervals for these effects in periods with and without NMP overlap at the 95% confidence level. Table A.3.11 in the Appendix shows that interactions between deviations from the subgroup in previous periods and having received a flag are not significant, which confirms that the high significance of conformity effects do not seem to be picking the effect of being punished. Also, the conformity effects are not different when individuals are playing solely with people of their own community, relative to the out-group treatment (Table A.3.11 in Appendix).[72]

[71] We also include a specification in which instead of distinguishing different flag colors, there is just a single dummy variable that indicates that the subject received a flag in the previous period.

[72] We also estimated specifications that replicate those of Masclet et al. (2003) and Noussair and Tucker (2005), using the total number of punishment points received in the previous period instead of the flags dummy variables. Results are consistent with the model shown before, but we prefer to keep the former adaptation, because it better reflects the information that individuals had when they made their decision. Also, Masclet et al. (2003) and Noussair and Tucker (2005) included overall deviations from the group's average in the previous period instead of positive and negative deviations when modeling decision changes. We also tested their specification, obtaining similar

To sum up, as shown in section 3.4.1, the NMP treatment has an effect that reduces extraction levels, especially during the out-group treatment. However, when analyzing period-by-period variations in extraction decisions, individuals adjust their choice to the subgroup's average in the previous period rather than react to punishment. This conformity effect is present both when NMP is available and when it is not. Only those who are punished with a red flag and perceive that action as unfair appear to react by raising their extraction levels. This may be explained because they experience anger. However, this effect dilutes when taking conformity into account. The fact that receiving a flag does not have consequences on individuals' decisions can be explained because subjects who are sensitive to NMP lower their extraction levels in advance, to avoid being punished and experiencing shame. Indeed, they correctly anticipate that in order to reduce the probability of being punished the best they can do is lower their extraction levels. Subjects are aware that if they choose high extraction levels they are likely to be punished and if they choose to do so it is because the punishment does not generate a significant disutility. This explains why when individuals are punished they do not react to punishment (unless they did not expect it, as in the case of being punished by antisocial punishers).

3.5 Discussion

In this study we performed a framed field experiment to test the effectiveness of non-monetary punishment (NMP) in the context of a CPR game. We combined this treatment with an in-group/out-group treatment, letting fishermen from different communities play one stage of the experiment solely with members of their own community and the other stage mixed with another community.

Our findings suggest that NMP has an effect diminishing extraction levels only in the out-group treatment. Subjects derive more disutility from being punished when interacting with subjects who do not belong to their own community. Following the theoretical model in Section 3.2, this would imply that the σ_i coefficient (a measure of one´s susceptibility to social punishment) is greater when interacting in an out-group than in an in-group environment. In other words, subjects take the NMP institution more seriously during the out-group treatment. In a context in which individuals do not know each other (or hardly know each other) but are aware that there is a slight chance they might see each other again, being publicly punished would provide the only information others have about oneself and in this sense it may be important to avoid being flagged in such a way. However, the NMP may not be perceived as intimidating when coming from workmates or neighbors. NMP may not matter either if it takes place in a context of complete strangers in which subjects know for sure they will not meet again. In other words, the

results, but we prefer to stick to our model, because as shown before, we would expect different reactions from those punished when throwing low number of nets than those punished when throwing high number of nets. The model above shows consistent results when changing each flag color dummy variable for a unique dummy variable, indicating that the individual received a flag of any color in the previous period (Table A.3.10 in the Appendix).

relationship between the sensitivity to peer punishment in in-group/out-group contexts may be non-monotonic.

Previous literature regarding contributions in public good games finds that non-monetary punishment increases cooperation in a public good game, but its effect is smaller than that of monetary sanctions (Masclet et al., 2003), and it is more effective in increasing cooperation when combined with this kind of sanction (Noussair and Tucker, 2005). Our findings are consistent with these studies in pointing that non-monetary punishment, solely by affecting pro-social emotions, can enhance cooperation. However, its effects are not as strong as those of monetary punishment, which affects not only individuals' pro-social emotions but also their monetary payoffs.

The NMP's effectiveness is diminished by the fact that not only free riders but also cooperators are punished. Indeed, Beckenkamp and Ostmann (1999) and Masclet et al. (2003) report that if subjects perceive the sanctions as unfair, they can react by decreasing cooperation. The latter interpret punishment from non-cooperators as evidence of spiteful preferences. This misdirected punishment is also observed by Falk et al. (2000), and Gächter and Herrmann (2011). Herrmann et al. (2008) point out that one plausible explanation of antisocial punishment is that people might not accept punishment and therefore seek revenge. This is likely, as these subjects were being constantly punished by the rest. In fact, most punishment administered to free riders was performed by cooperators, while most of the antisocial punishment came from free riders. Alternatively, it could also be interpreted as features of their daily lives that subjects bring into the game (Cardenas and Ostrom, 2004). For instance, they may perceive that intensifying current fishing does not have any consequences on the availability of fish in the future (for instance, because they may believe that climate factors or other industries are more important determinants of fish availability). In the same line of thought, Casari and Luini (2009) and Noussair et al. (2011) argue that subjects may not view the norm of cooperation as the norm that punishment should enforce, as other norms such as "try to catch as many fish as possible" may be the prevailing ones. A fourth explanation could be that this behavior is a consequence of bounded rationality, related to cognitive limitations of the game on the part of some players, in line with Simon (1955). Janssen et al. (2010) argue that in a context in which participants can punish back but cannot discuss why they are sanctioned, receiving a sanction does not carry a clear message.

It is particularly interesting to note that subjects are willing to punish others while facing a monetary cost to themselves and may not necessarily expect that this punishment will determine an increase in cooperation. Even if the monetary cost of social punishment was low, subjects were reminded at every period that by socially punishing others they were themselves bearing a cost, as they had to subtract the total cost of punishment from their earnings in their balance sheet. Despite this fact, subjects chose to punish others during the whole experiment, including the last period when no change in others' behavior was possible. In fact, on average per period each subgroup awarded 1.7 disapproval flags to the members of that group. This result is in line with Fehr and Gachter (2000) findings regarding monetary punishment. Following Casari and Luini (2009), Fudenberg and Pathak (2010) and Noussair et al. (2011), we conclude that punishment is not necessarily applied instrumentally to increase cooperation and that subjects have preferences for punishing.

We do not find significant differences in punishing behavior between in-group and out-group treatments. This finding is in contrast to McLeish and Oxoby (2007) and Miguel and Gugerty (2005), who argue that subjects punish free riders more harshly in in-groups than out-groups. On the contrary, Chen and Li (2009) and Currarini and Mengel (2012) find that subjects are less likely to punish in-group members than out-group members.

Subjects do not adjust their period-by-period decisions as a reaction to punishment effectively received. They seem to correctly anticipate that the likelihood of being punished is increasing in extraction levels and those who would experience disutility by being punished reduce their extraction levels beforehand. Those who do not reduce extraction levels do not react to punishment because they are insensitive to it. Instead, those who were unexpectedly punished and who considered the punishment unfair, experienced anger and increased their extraction levels in the subsequent period.

We find strong conformity effects: individuals adjust their period-by-period decisions in order to converge with their peers' average in a previous period. These results are consistent with Hayo and Vollan (2012). The results highlight the potential relevance of social comparisons as a form of non-pecuniary policy seeking changes in behavior (Ferraro and Price, 2011).

Contrary to what has been mostly documented in the literature, we do not find an in-group bias regarding cooperation. That is, individuals do not behave differently when interacting with subjects from their own community than when they are mixed with subjects from other communities, except for being more sensitive to NMP during the out-group treatment. Hewstone et al. (2002) argue that negative feelings toward out-group members tend to occur mostly in circumstances in which belonging to a group draws a strong sense of identity, and that this can be reduced as a consequence of the quantity and the quality of contact between groups. Fishermen from different communities do not interact during their daily life, but they are used to seeing each other while fishing as they tend to move from one lagoon to the other depending on fish availability. The non-relevance of the in-group favoritism may be explained because of the high mobility across lagoons. When surveyed, those fishermen who live by the shores of the coastal lagoons complained about others coming to fish there. But they usually also move to other places to fish, depending on fish availability. Therefore, they are used to seeing others fishing in their own place of residence, and even if they complain they know the same stands for themselves when they go to fish to another location. In other words, everybody acknowledges being an outsider at some point in time. For this reason, as regards to social preferences, granting exclusive access to a common pool resource to a certain community appears not to be a requisite from a resource conservation point of view.

Finally, community membership appears to have an influence over individuals' decisions, a finding not explained by observable socioeconomic factors. This may suggest that social norms regarding extraction levels differ among communities. The importance of community membership has been noted by Henrich et al. (2001) and Hayo and Vollan (2012). In our case it is quite striking to find differential behavior by community, as the communities we studied do not differ in terms of ethnicity or economic organization. Also, in line with other studies (Cardenas, 2003; Hayo and Vollan, 2012), we do find that cooperation is negatively correlated with wealth. This relationship should be studied more in depth, in order to disentangle the causal link between the two.

Overall, our results are consistent with the view that cooperation in a CPR dilemma is determined not only by repeated game behavior but also by social preferences. Subjects are willing to bear costs due to deviations from the self-interested equilibrium, even in one-shot interactions, as has been previously observed in public goods settings. However, previous interactions with other subjects have substantial influence on behavior, reflecting strong preferences for conformism. Individuals with social preferences limit their resource exploitation (cooperate) in response to the threat of punishment, but we do not find evidence of reactions to being effectively punished. We argue that the latter result is due to two reasons. First, subjects anticipate that the probability of being punished increases with their extraction level decision. Therefore, they reduce their extraction decision in advance, avoiding the experience of shame. Second, antisocial punishment was substantial and generated in some cases an increase in extraction among those being unfairly punished. Finally, our results suggest that the relevance of in-group favoritism in promoting cooperation due to social preferences may be overrated, and that for peer punishment to be effective it requires coordination, in order to prevent anti-social targeting and to enhance the social signal conveyed by the punishment.

Appendix 3

Table A.3.1 – Mean socioeconomic characteristics by community

Community	Years of schooling	Electricity	Wealth	Per capita income (US)	Fishing main activity
Laguna de Rocha	6.0	0.13	1.85	149	0.75
Valizas Puente	6.7	0.75	3.06	175	0.67
Barra de Valizas	7.6	0.38	1.68	373	0.63
Puerto los Botes	6.0	1.00	2.52	246	1.00
Barrio Parque	8.0	1.00	4.32	320	0.38

Table A.3.2 – Payoff table

Others' fishnets	Others' total	My fishnets 1	2	3	4	5	6	7	8	Others' average
	3	354	360	366	372	378	384	390	396	1
	4	342	348	354	360	366	372	378	384	1
	5	330	336	342	348	354	360	366	372	2
	6	318	324	330	336	342	348	354	360	2
	7	306	312	318	324	330	336	342	348	2
	8	294	300	306	312	318	324	330	336	3
	9	282	288	294	300	306	312	318	324	3
	10	270	276	282	288	294	300	306	312	3
	11	258	264	270	276	282	288	294	300	4
	12	246	252	258	264	270	276	282	288	4
	13	234	240	246	252	258	264	270	276	4
	14	222	228	234	240	246	252	258	264	5
	15	210	216	222	228	234	240	246	252	5
	16	198	204	210	216	222	228	234	240	5
	17	186	192	198	204	210	216	222	228	6
	18	174	180	186	192	198	204	210	216	6
	19	162	168	174	180	186	192	198	204	6
	20	150	156	162	168	174	180	186	192	7
	21	138	144	150	156	162	168	174	180	7
	22	126	132	138	144	150	156	162	168	7
	23	114	120	126	132	138	144	150	156	8
	24	102	108	114	120	126	132	138	144	8

Table A.3.3 – Punishment card

If the other throws:	I disapprove (0 to 10 points)
1 net	
2 nets	
3 nets	
4 nets	
5 nets	
6 nets	
7 nets	
8 nets	
Total	

Table A.3.4 – Flag range

Flag	Total punishment points received
Yellow	2 - 5
Orange	6 - 10
Red	11 - 30

Table A.3.5 – Wilcoxon-Mann-Whitney ranksum and median equality average nets tests by place of residence

		Mean				
		LR - Barra	Valizas Puente	Barra de Valizas	Puerto Los Botes	Barrio Parque
Median	Laguna de Rocha	x	7.364	-1.764	3.891	-1.711
	p	x	*0*	*0.0777*	*0.0001*	*0.0871*
	El Puente	59.0262	x	-8.998	-4.869	-8.296
	p	*0*	x	*0*	*0*	*0*
	Barra de Valizas	2.6313	57.1846	x	5.714	0.037
	p	*0.105*	*0*	x	*0*	*0.9706*
	Puerto los Botes	4.564	14.7255	9.1414	x	-5.142
	p	*0.033*	*0*	*0.002*	x	*0*
	Barrio Parque	3.4133	59.0262	0.0514	9.8246	x
	p	*0.065*	*0*	*0.821*	*0.002*	x

Table A.3.6 – Distribution of type by community

	Laguna de Rocha	El Puente	Valizas	Los Botes	Barrio Parque	Total
Free rider	25%	0%	25%	13%	38%	7
Conditional cooperator	50%	17%	25%	25%	13%	12
Cooperator	0%	50%	25%	0%	13%	9
Other	25%	33%	25%	63%	38%	16
Total	100% 8	100% 12	100% 8	100% 8	100% 8	44

Table A.3.7 – Nets variations and nets in previous round (in %)

	Total									If nets in previous round <=5						If nets in previous round >5			
	nets in previous round									nets in previous round						nets in previous round			
nets variation	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Total**	**1**	**2**	**3**	**4**	**5**	**Total**	**6**	**7**	**8**	**Total**
-7								3.23	0.68									3.23	1.23
-6							4.55	3.23	1.35								4.55	3.23	2.47
-5						7.14	4.55	3.23	2.7							7.14	4.55	3.23	4.94
-4					5	0	4.55	3.23	2.03					5	1.49	0	4.55	3.23	2.47
-3				9.52	5	7.14	13.64	6.45	6.76				9.52	5	4.48	7.14	13.64	6.45	8.64
-2			0	9.52	0	7.14	4.55	12.9	6.08			0	9.52	0	2.99	7.14	4.55	12.9	8.64
-1		18.18	9.09	9.52	35	14.29	4.55	19.35	15.54		18.18	9.09	9.52	35	17.91	14.29	4.55	19.35	13.58
0	25	54.55	18.18	9.52	10	50	31.82	48.39	33.11	25	54.55	18.18	9.52	10	19.4	50	31.82	48.39	44.44
1	0	0	54.55	28.57	25	7.14	31.82		17.57	0	0	54.55	28.57	25	25.37	7.14	31.82		11.11
2	25	18.18	18.18	14.29	15	7.14			8.78	25	18.18	18.18	14.29	15	16.42	7.14			2.47
3	0	0	0	4.76	5				1.35	0	0	0	4.76	5	2.99				
4	25	9.09	0	14.29					3.38	25	9.09	0	14.29		7.46				
5	25	0	0						0.68	25	0	0			1.49		.		
6	0	0							0	0					0				
7	0								0	0					0				
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Table A.3.8 – Net variations and flag colour in previous round (%)

	Total			If nets in previous round <=5			If nets in previous round >5		
nets variation	Yellow	Orange	Red	Yellow	Orange	Red	Yellow	Orange	Red
-	28.4	44.9	45.5	25.0	32.1	0.0	30.8	61.9	62.5
=	42.1	20.4	18.2	25.0	14.3	0.0	53.9	28.6	25.0
+	29.6	34.7	36.4	50.0	53.6	100.0	15.4	9.5	12.5
Total	100	100	100	100	100	100	100	100	100

Table A.3.9 – net variations by behavioral type (%)

	Type - Total					Type - If nets in previous round <=5					Type - If nets in previous round >5				
nets variation	**free rider**	**cond. coop**	**coop.**	**other**	**Total**	**free rider**	**cond. coop**	**coop.**	**other**	**Total**	**free rider**	**cond. coop**	**coop.**	**other**	**Total**
-	25.01	41.7	19.1	42.4	35.1	0.0	35	19.1	28.0	27	25.8	50	0.0	53	42.0
=	59.4	17	43	25.4	33.1	0.0	0	42.9	16.0	19.4	61.3	37.5	0.0	32	44.4
+	15.6	41.7	38.1	32.2	31.8	100	65	38.1	56.0	53.7	12.9	13	0.0	15	13.6
Total	100	100	100	100	100	100	100	100	100	100	100	100	0	100	100

Table A.3.10 – Reaction to punishment, other specifications. Dependent variable: $fishnets_t$-$fishnet_{t-1}$

	Masclet et al. (2003) & Noussair et al. (2011) model for punishment points received									Masclet et al. (2003) & Noussair et al. (2011) model controlling for flags received				
	Dependent variable: $fishnets_t$ -$fishnets_{t-1}$													
Sample	**1-20**	**1-20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**	**1 - 20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**	**1-20**	**1-20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**
Punishment points in *t-1*	-0.054*	-0.003	-0.010	0.005	-0.046	-0.011	-0.012	-0.011	-0.054					
	(0.032)	(0.030)	(0.041)	(0.039)	(0.055)	(0.028)	(0.042)	(0.037)	(0.056)					
Flag in *t-1*										-0.408**	-0.118	-0.279	0.044	-0.300
										(0.191)	(0.173)	(0.223)	(0.271)	(0.316)
Deviation from average ($n^{t-1}{}_i$-$n^{t-1}{}_{av}$)		-0.781***	-0.910***	-0.655***	-0.682***						-0.778***	-0.903***	-0.655***	-0.711***
		(0.066)	(0.108)	(0.151)	(0.147)						(0.068)	(0.109)	(0.148)	(0.154)
Positive deviation from average ($n^{t-1}{}_i$-$n^{t-1}{}_{av}$)						-0.666***	-0.947***	-0.119	-0.663***					
						(0.098)	(0.166)	(0.450)	(0.165)					
Negative deviation from average ($n^{t-1}{}_{av}$-$n^{t-1}{}_i$)						0.849***	0.841***	0.669***	0.104					
						(0.109)	(0.148)	(0.162)	(0.490)					
_cons	0.076*	0.014	0.063	-0.081	0.402	-0.103	0.144	-0.135	0.418	0.093**	0.034	0.154	-0.085	0.464
	(0.039)	(0.037)	(0.108)	(0.138)	(0.288)	(0.118)	(0.188)	(0.197)	(0.343)	(0.039)	(0.035)	(0.094)	(0.150)	(0.338)
N	836	836	352	213	139	836	352	213	139	836	836	352	213	139
N_g	44	44	44	39	32	44	44	39	32	44	44	44	39	32
r2_w	0.006	0.306	0.370	0.172	0.205	0.288	0.350	0.142	0.165	0.007	0.306	0.374	0.172	0.206
r2_o	0.004	0.185	0.209	0.084	0.111	0.173	0.196	0.081	0.097	0.006	0.185	0.210	0.084	0.113
r2_b	0.188	0.125	0.031	0.000	0.041	0.126	0.037	0.000	0.042	0.071	0.126	0.032	0.000	0.044
legend: *** p<0.01;** p<0.05; * p<0.1														
standard errors in parenthesis														

Table A.3.11 – Specifications including interactions

Dependent variable: $fishnets_t$ -$fishnets_{t-1}$		
Sample	**7-10 & 17-20**	**7-10 & 17-20**
Positive deviation from average ($n^{t-1}{}_i$-$n^{t-1}{}_{av}$)	-0.952***	-0.94***
	(0.183)	(0.20)
Negative deviation from average ($n^{t-1}{}_{av}$-$n^{t-1}{}_i$)	0.876***	0.699***
	(0.166)	(0.169)
Positive deviation from average ($n^{t-1}{}_i$-$n^{t-1}{}_{av}$)*$flag_{t-1}$	-0.010	
	(0.204)	
Negative deviation from average ($n^{t-1}{}_{av}$-$n^{t-1}{}_i$)*flagt-1	-0.095	
	(0.168)	
Positive deviation from average ($n^{t-1}{}_i$-$n^{t-1}{}_{av}$)*outgroup		-0.08
		(0.17)
Negative deviation from average ($n^{t-1}{}_{av}$-$n^{t-1}{}_i$)*outgroup		-0.24
		(0.16)
_cons	0.120	0.14
	(0.165)	(0.17)
N	352	352
N_g	44	44
r2_w	0.351	0.356
r2_o	0.197	0.201
r2_b	0.038	0.040
legend: *** p<0.01;** p<0.05; * p<0.1		
standard errors in parenthesis		

Figure A.3.1 – Own current cost and others' lagged cost of deviating by subject

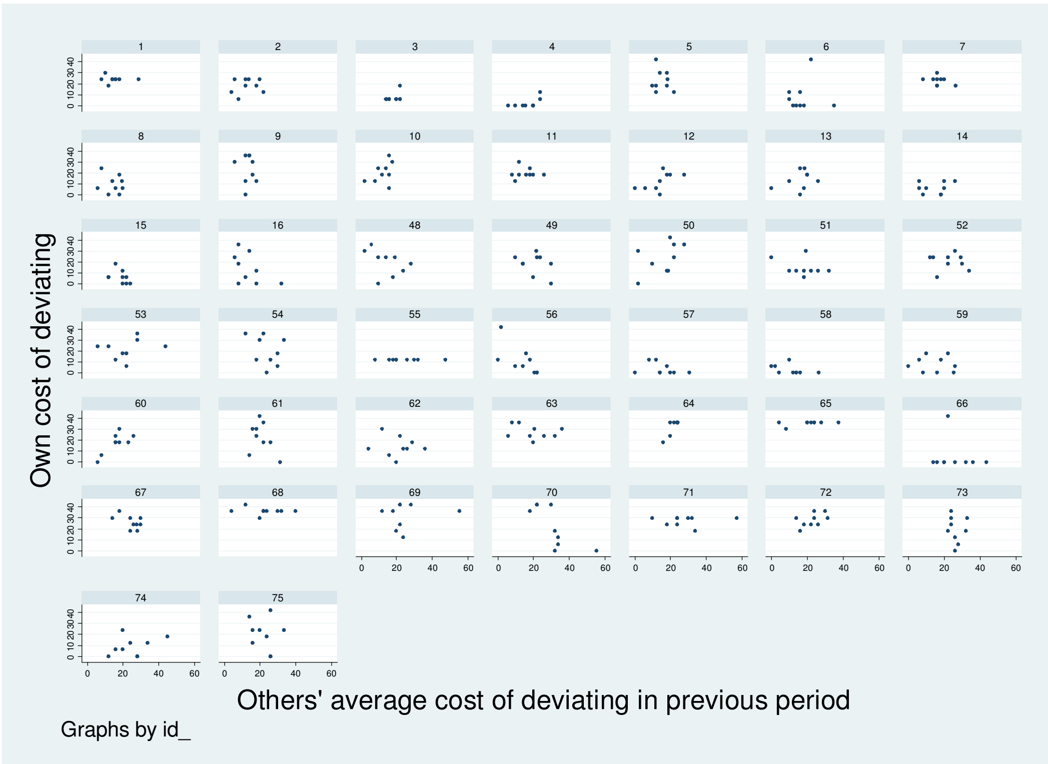
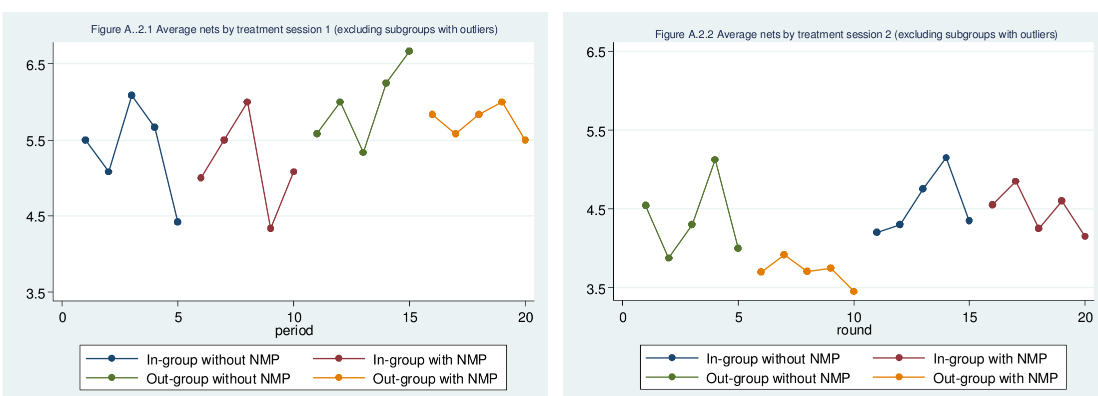


Figure A.3.2 – Average extraction levels by treatment and session excluding subgroups where three subjects performed substantial antisocial punishment



References

Akerlof, G., and Kranton, R. (2000) “Economics and identity”. *Quarterly Journal of Economics*, 115(3): 715–53.

Akerlof, G. and R. Kranton. (2002) ``Identity and Schooling: Some Lessons for the Economics of Education." *Journal of Economic Literature*, 40(4): 1167-1201.

Akerlof, G., and Kranton, R. (2005) “Identity and the economics of organizations”. *Journal of Economic Perspectives*, 19(1): 9–32.

Anderies, J.M., et al. (2011) “The challenge of understanding decisions in experimental studies of common pool resource governance” *Ecological Economics*, 70, pp. 1571 – 1579.

Ammermueller, A. and J.S. Pischke. 2009. ``Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study." *Journal of Labor Economics*, 27(3): 315-348.

Angrist J. and G. Imbens. 1995. ``Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity.", *Journal of the American Statistical Association*, 90(430): 431-442.

Austen-Smith, D. and R. Fryer. 2005. ``An Economic Analysis of 'Acting White'." *Quarterly Journal of Economics*, 120(2): 551-583.

Baland, J.M., Platteau, J.P. (1996) “Halting degradation of natural resources. is there a role for rural communities?”. *Clarendon Press Oxford*.

Baland, J.M., Bardhan, P. and Bowles, S. (2007) *“Inequality, cooperation and environmental sustainability”*. Russell Sage Foundation and Princeton University Press.

Bandiera, O.; Barankay, I. and Rasul, I. (2005). “Social Preferences and the Response to Incentives: Evidence from Personnel Data”. *Quarterly Journal of Economics*, 120(3): 917–62.

Bayer, P., S. Ross and G. Topa. 2008. ``Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." *Journal of Political Economy*, 116(6): 1150-1196.

Barr A. (2001) “Social Dilemmas and Shame-based Sanctions: Experimental results from rural Zimbabwe”. Working Paper 2001.11, Centre for the Study of African Economies University of Oxford.

Beckenkamp, M., and Ostman, A. (1999) “Missing the target? sanctioning as an ambiguous structural solution”. in Foddy, M., Smithson, M., Scheneider, S. and Hogg, M. (eds.) *“Resolving social dilemmas: dynamic, structural, and intergroup aspects”*.. London: Psychology Press, 165-180.

Benabou, R. (1993) ``Workings of a City: Location, Education and Production." *Quarterly Journal of Economics*, 108(3): 619-652.

Benabou, R. (1996) ``Equity and Efficiency in Human Capital Investment: The Local Connection." *The Review of Economic Studies*, 63(2): 237-264.

Bernhard H., Fehr, E. and Fischbacher, U. (2006) "Group affiliation and altruistic norm enforcement". *American Economic Review*, 96(2): 217–21.

Blount, S. and Bazerman, M. (1996), "The inconsistent evaluation of absolute versus comparative payoffs in labor supply and bargaining". *Journal of Economic Behavior and Organization*, 30, 227-240.

Bobonis, G. and Finan, F. (2009) ``Neighborhood Peer Effects in Secondary School Enrollment Decisions." *The Review of Economics and Statistics*, 91(4): 695–716.

Bowles, S. (2004) *"Microeconomics: behavior, institutions and evolution"*. Russell Sage Foundation.

Bowles, S. and Gintis, H. (2002) "Social capital and community governance". *The Economic Journal* Vol. 112 N. 483: 419-436.

Bowles, S. and Gintis, H. (2011) *"A cooperative species"*. Princeton University Press.

Bowles, S., G. Loury, and R. Sethi. (2007) ``Is Equal Opportunity Enough? A Theory of Persistent Group Inequality."

Bramoullé, Y., H. Djebbari and B. Fortin. (2009) ``Identification of peer effects through social networks." *Journal of Econometrics*, 150: 41-55.

Brandts J. and Charness, G. (2010) "the strategy versus the direct-response method: a survey of experimental comparisons".

Brock, W. and S. Durlauf. (2001) ``Interactions-Based Models." in Handbook of Econometrics, Heckman and Leamer (Eds), Elsevier Science B.V.

Calvó-Armengol A., E. Patacchini and Y. Zenou (2009) ``Peer Effects and Social Networks in Education." *Review of Economic Studies*, 76(4): 1239–1267.

Card, D. and J. Rothstein (2007) ``Racial segregation and the black–white test score gap." *Journal of Public Economics*, 91, 2158– 2184

Carrell, S., R. Fullerton and J. West. (2009). ``Does your cohort matter? Measuring peer effects in college achievement." *Journal of Labor Economics*, 27(3): 439–464.

Carrell, S., B. Sacerdote and J. West. (2012) ``From Natural Variation to Optimal Policy? An Unsuccessful Experiment in Using Peer Effects Estimates to Improve Student Outcomes." Working paper.

Cardenas, J.C. (2003) "Real wealth and experimental cooperation: experiments in the field lab". *Journal of Development Economics* 70 (2003) 263– 289.

Cardenas, J.C. (2004) “Norms from outside and from inside: an experimental analysis on the governance of local ecosystems.” *Forest Policy and Economics* (6):229-241

Casari, M. and Luini, L. (2009) “Cooperation under alternative punishment institutions: An experiment”. *Journal of Economic Behavior & Organization*, 71: 273-282.

Chen , Y and Xin S. (2009) “Group identity and social preferences”. *American Economic Review*, 99(1): 431–57.

Coleman, J. (1966) Equality of Educational Opportunity, U.S. GPO, Washington, D.C.

Cooley, J. (2010) ``Desegregation and the Achievement Gap: Do Diverse Peers Help?"

Denant-Boemont, L., Masclet, D. and Noussair, C. (2011) “Announcement, observation and honesty in the voluntary contributions game”. *Pacific Economic Review*, Vol. 16, Nº 2, pp. 207 – 228.

De Giorgi, G., M. Pellizzari and S. Redaelli (2010) ``Identification of Social Interactions through Partially Overlapping Peer Groups." *American Economic Journal: Applied Economics*, 2(2): 241-75.

Denant-Boemont, L., Masclet, D. and Noussair, C. (2011) “Announcement, observation and honesty in the voluntary contributions game”. *Pacific Economic Review*, Vol. 16, Nº 2, pp. 207 – 228.

Doise, W., and G. Mugny. (1984) The social development of the intellect. New York: Pergamon Press.

Drago, F., and R. Galbiati. (2012) ``Indirect Effects of a Policy Altering Criminal Behavior: Evidence from the Italian Prison Experiment." *American Economic Journal: Applied Economics*, 4(2): 199-218.

Duflo, E., P. Dupas and M. Kremer (2011) ``Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review*, 101(5):1739-74.

Dugar, S. (2010) “Nonmonetary sanctions and rewards in an experimental coordination game” Journal of Economic Behavior & Organization 73: 377 – 386.

Durlauf, S. (1996) ``A Theory of Persistent Income Inequality." *Journal of Economic Growth*, 1: 75-93.

Durlauf, S. (1996) ``Associational Redistribution: A Defense," (261-284) in Olin Wright (Ed), Recasting Egalitarianism: new rules for communities, states and markets.

Durlauf, S. (2004) ``Neighborhood Effects." Handbook of Regional and Urban Economics, vol. 4, J. V. Henderson and J.-F. Thisse, eds., Amsterdam: North Holland.

Eckstein, Z. and K. Wolpin (1999) Why youths drop out of highschool: the impact of preferences, opportunities and abilities. *Econometrica*, 67: 1295-1339.

Falk, A., Fehr, E. and Fischbacher, U. (2000) "Informal sanctions". Working paper, University of Zurich.

Feeny, D., Berkes, F., McCay, B. and Acheson, J. (1990) "The tragedy of the commons, twenty-two years laters". *Human Ecology*, Vol. 18, Nº 1.

Fehr, E. and Gächter, S. (2000) "Cooperation and punishment in public good experiments". *American Economic Review*, Vol. 90, Nº 4, pp. 980 – 994.

Fernández, T. (2009) La desafiliación en la educación media en Uruguay. Una aproximación con base en el panel de estudiantes evaluados por PISA 2003. *Revista Iberoamericana sobre Calidad Eficacia y Cambio en Educación*. Vol 7, no. 4: 165-179.

Ferraro, P and Price, M. (2011) "Using non-pecuniary strategies to influence behavior: evidence from a large scale field experiment". NBER working paper series N 17189.

Fordham, S. and J. Ogbu. (1986). ``Black students' school success: coping with the Burden of Acting White." *The Urban Review*, XVIII , 176-206.

Fortin, B. and M. Yazbeck. (2011). ``Peer Effects and Fast Food Consumption and Adolescent Weight Gain." CIRANO Working Papers 2011s-20.

Fudenberg, D. and Pathak, P. (2010) "Unobserved punishment supports cooperation". *Journal of Public Economics* 94: 78-86.

Gächter, S. and Fehr, E. (1999) "Collective action as a social exchange". *Journal of Economic Behavior & Organization*, Vol. 39, pp. 341–369.

Gächter, S. and Herrmann, B. (2011) "The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia". *European Economic Review*, 55, pp. 193 – 210.

Gaviria A. and S. Raphael (2001) School-Based Peer Effects and Juvenile Behavior. *The Review of Economics and Statistics*, Vol 83, no 2: 257-268.

Glaeser, E., B. Sacerdote and J. Scheinkman. (2003) ``The Social Multiplier." *Journal of the European Economic Association* 1(2): 345-353.

Glaeser, E. and J. Scheinkman. (2001) ``Measuring Social Interactions." in Social Economics (Durlauf and Young, eds.), Cambridge: MIT Press, 2001, 83-102.

Goette, L., Huffman, D. and Meier, S. (2006) "The impact of group membership on cooperation and norm enforcement: evidence using random assignment to real social groups". IZA Discussion Paper, Nº 2020.

Gould E, Lavy V. and M. Paserman (2009) Does inmigration affect the long-term educational outcomes of natives? Quasi-experimental evidence. *The Economic Journal*, 119: 1243-1269.

Graham, B. (2008) ``Identifying Social Interactions through Conditional Variance Restrictions." *Econometrica*, 76(3): 643–660.

Graham, B. (2011) ``Econometric methods for the analysis of assignment problems in the presence of complementarity and social spillovers." Handbook of Social Economics 1B: 965 - 1052 (J. Benhabib, M. O. Jacksons & A. Bisin, Eds.). Amsterdam: North-Holland.

Guryan, J. (2004) ``Desegregation and black dropout rates." *American Economic Review*, 94(4), 919–943.

Hanushek, E., J. Kain, J. Markman, and S. Rivkin. (2003) ``Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics*, Vol. 18, Iss. 5, 527-544.

Hayo B. and Vollan, B. (2012) "Group interaction, heterogeneity, rules, and cooperative behaviour: Evidence from a common-pool resource experiment in South Africa and Namibia". *Journal of Economic Behavior and Organization* 81: 9-28.

Hardin, G. (1968) "The tragedy of commons". in Stavins (2000) *"Economics of the environment. Selected Readings"*, W. W. Norton.

Harrison, G. and List, J. (2004) "Field experiments". *Journal of Economic Literature*, Vol. 42, Nº 4. (Dec., 2004), pp. 1009-1055.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. and McElreath, R. (2001) "In search of homo economicus: behavioral experiments in 15 small-scale societies". *American Economic Review*, Vol. 91, Nº 2, pp. 73 – 78.

Herrmann, B; Thoni, C.; and Gachter, S. (2008) "Antisocial punishment across societies". *Science* 319 (5868), 1345-6.

Hewstone, M., Rubin, M. and Willis, H. (2002) "Intergroup bias". *Annual Review of Psychology*, Vol. 53, pp. 575 – 604.

Hoxby, C. (2000) ``Peer Effects in the Classroom: Learning from Gender and Race Variation." NBER working paper no. 7867.

Janssen, M., Holahan R., Lee, A. and Ostrom, E. (2010) "Lab experiments for the study of socio-ecological systems". *Science* 328: 613-617.

Kremer, M. and E. Miguel. (2007) ``The Illusion of Sustainability." *Quarterly Journal of Economics*, 112: 1007–1065.

Kurzban, R. and Houser D. (2005) "Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations". *Proceedings of the National Academy of the United States of America*, 102(5): 1803-1807.

Laschever, R. (2009) ``The Doughboys Network: Social Interactions and the Employment of World War I Veterans."

Lavy, V. and A. Schlosser (2011) ``Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics*, 3(2):1-33.

Lazear, E. (2001) ``Educational Production.'' *Quarterly Journal of Economics*, 116: 777–803.

Lin, X. (2010) ``Identifying Peer Effects in Student Academic Achievement by Spatial Autoregressive Models with Group Unobservables.'' *Journal of Labor Economics*, 28(4): 825-860.

López M.; Murphy, J.; Spraggon, J. and Stranlund, J. (2012) "Comparing the effectiveness of regulation and pro-social emotions to enhance cooperation: experimental evidence from fishing communities in Colombia". *Economic Inquiry*, Vol. 50 (1): 131-142.

Manacorda, M. (2012) "The cost of grade retention." *The Review of Economics and Statistics*. 94(2): 596-606.

Manski, C. (1993) ``Identification of Endogenous Social Effects: The Reflection Problem.'' *Review of Economic Studies*, 60(3): 531-542.

Manski, C. (2000) ``Economic analysis of social interactions.'' *Journal of Economic Perspectives*, 14(3) 115–136.

Masclet, D., Noussair, C., Tucker, S. and Villeval, M.C. (2003) "Monetary and nonmonetary punishment in the voluntary contributions mechanism". *American Economic Review*, Vol. 93, Nº 1, pp. 366 – 380.

Miguel, E. and Gugerty, M.K. (2005) "Ethnic diversity, social sanctions and public goods in Kenya". *Journal of Public Economics*, Nº 89, pp. 2325 – 2368.

Moffitt, R. (2001) ``Policy Interventions, Low-Level Equilibria, and Social Interactions.'' *Social Dynamics*, eds. S. Durlauf and P. Young. MIT Press, 2001.

Moody, J. (2001) ``Race, School Integration, and Friendship Integration in America.'' *American Journal of Sociology*, 107(3): 679-716.

Mora T, and P. Oreopoulos (2011) Peer effects on high-school aspirations: Evidence from a sample of close and not-so-close friends. *Economics of Education Review*, 30: 575-581.

Nechyba, T. (2006) ``Income and Peer Quality Sorting in Public and Private Schools.'' in Handbook of Economics of Education, 2: 1327-1368, Hanushek E. and Welch F. eds, Elsevier.

Noussair, C. and Tucker, S. (2005) "Combining monetary and social sanctions to promote cooperation". *Economic Inquiry* Vol. 43, No. 3 649-660.

Noussair, C. and Tucker, S. (2007) "Public observability of decisions and voluntary contributions in a multiperiod context". *Public Finance Review* Vol. 35 Number 2 176-198.

Noussair, C.; van Soest, D. and Stoop, J. (2011) "Punishment, reward, and cooperation in a framed field experiment". Munich Personal RePEc Archive.

Oreopoulos, P. (2007) Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics*. 91: 2213-2229.

Ostrom, E. (1990) *"Governing the Commons: The Evolution of Institutions for Collective Action*"- New York: Cambridge University Press.

Ostrom, E., Burger, J., Field, C., Norgaard, R., and Policansky, D. (1999) "Revisiting the commons: local lessons, global challenges". *Science* Nº 284, 278 - 282

Ostrom, E. (2000) "Collective action and the evolution of social norms". *The Journal of Economic Perspectives*, Vol. 14, Nº 3.

Piketty, T. (2000) ``Theories of persistent inequality and intergenerational mobility." Handbook of Income Distribution 1, eds. Atkinson A. and Bourguignon F., Amsterdan North-Holland, 430-476.

PROBIDES (2002) "Proyecto de Desarrollo de la Areas Protegidas Lagunas de Garzón y Rocha". Documentos de Trabajo Nº 44.

Rege, M. and Telle K., (2004) "The impact of social approval and framing on cooperation in public good situations". *Journal of Public Economics* 88: 1625– 1644.

Rivkin, S. and F. Welch. (2006) ``Has school desegregation improved academic and economic outcomes for blacks?" Handbook of the Economics of Education 2: 1019 - 1049 (E. Hanushek and F. Welch, Eds.). Amsterdam: North-Holland.

Rodríguez-Gallego, L., Santos, C., Amado S., Gorfinkel, D., González, M.N., Gómez, J., Neme, C., Tommasino, H. and Conde D. (2008) "Costos y beneficios socioeconómicos y ambientales del uso actual de la Laguna de Rocha y su Cuenca. Insumos para la Gestión Integrada de un Área Protegida Costera". Informe Programa de Desarrollo Tecnológico, UdelaR.

Ruffle, B. and Sosis, R. (2006) "Cooperation and the in-group-out-group bias: A field test on Israeli kibbutz members and city residents". *Journal of Economic Behavior & Organization*, Vol. 60, pp 147–163.

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

Soetevent, A. (2006) ``Empirics of the identification of social interactions: An evaluation of the approaches and their results." *Journal of Economic Surveys*, 20(2): 193 - 228.

Tanaka, T.; Camerer, C. and Nguyen, Q. (2008) "Status, ethnicity, and wealth in Vietnam: evidence from experimental games". Unpublished.

Thompson, D. (2007) "Economía e identidad de pescadores laguna de Rocha". Dirección de Cultura, Ministerio de Educación y Cultura.

Trivers, R. (1971) "The evolution of reciprocal altruism". *Quarterly Review of Biology*, 46: 35-57.

Simon, H.A. (1955) "A behavioral model of rational choice". *The Quarterly Journal of Economics*, Vol. 69, Nº. 1, pp. 99-118.

Van Soest, D. and Vyrastekova, J. (2006) "Peer enforcement in CPR experiments: the relative effectiveness of sanctions and transfer rewards, and the role of behavioural types", List, J. (ed.) *"Using Experimental Methods in Environmental and Resource Economics*". Edward Elgar.

Velez, M.A., Stranlund, J.K., and Murphy, J.J. (2009) "What motivates common pool resource users? Experimental evidence from the field". *Journal of Economic Behavior & Organization*, Vol. 70, pp. 485–497.

Zanella, G. (2007) ``Discrete Choice with Social Interactions and Endogenous Memberships." *Journal of the European Economic Association*, 5(1): 122-53.

Zimmerman, D. (2003) ``Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics and Statistics*, 85(1), 9-23.