Social Interactions and Schooling Decisions

Rafael Lalive*, University of Zurich, CESifo, and IZA
Alejandra Cattaneo, University of Zurich

October 1, 2004

Abstract

This paper provides empirical evidence on the relevance of social determinants in schooling. In particular, we study how much a student's school attendance decision is affected by the choices of his or her peers. Identification is based on a randomized intervention that grants a cash subsidy encouraging school attendance to some students in a village but not to others. Results indicate that the ineligible student's decision to attend school is affected strongly by the choices in the peer group. Moreover, social interactions are equally important among boys and girls, and throughout grades 3-6 of primary school.

JEL-Classification: C93, I21, I28
Keywords: peer effects, schooling, field experiment.

*Thanks to Michael Kosfeld for comments on this paper. This research project has been supported by a grant from IEW. Address: Rafael Lalive, IEW, Bluemlisalpstr. 10, CH-8006 Zuerich, rlalive@iew.unizh.ch. Alejandra Cattaneo, SOI, Zuerichbergstrasse 14, CH-8032 Zuerich, alejandra.cattaneo@sts.unizh.ch.

1 Introduction

Perhaps the most important contribution of economics to education science is a model that pins down the economic forces affecting the decision to acquire school knowledge. This *human capital* framework holds that students will attend school if the expected future benefits of attending school outweigh the current (opportunity) costs of spending time in the classroom rather than working (Becker, 1964). The standard human capital model uses the individual return to education and individual foregone earnings to model education decisions (Card, 1999). Yet, schooling decisions may depend in important ways also on the amount of schooling acquired by other individuals – the social determinants of schooling (Akerlof and Kranton, 2002).[1] The idea is that an individual's self image may be enhanced when his or her actions are in line with the behavior in the peer group.[2]

The aim of this paper is to provide evidence on the relevance of the social environment in schooling decisions. There are two important problems in identifying social interactions (Manski, 1993, 1995, 2000). First, individual outcomes and group outcomes are correlated due to endogenous group membership – the "omitted variable" problem. Second, it is difficult to disentangle the effect of group behavior on individual behavior since the individual is also affecting group members – the "reflection problem". This paper relies on a targeted school subsidy that encourages a subgroup of students to attend school in a random selection of villages in rural Mexico – the PROGRESA program. This cash subsidy induces an increase in average schooling in the peer group.[3] We can thus infer whether social interactions are relevant in schooling decisions by studying the ineligible student's response to the program (Moffitt, 2001). The fact that PROGRESA is a targeted intervention addresses directly the "reflection problem" because the schooling subsidy manipulates the school attendance rate in the peer group while leaving unaffected the individual.[4] Moreover, the fact that the program was implemented only in some villages but not in other villages addresses the omitted variable problem because randomization ensures balancing of all determinants of school attendance. This means that the experimental design allows identification of social interactions in schooling.

Results indicate first that there is a positive but statistically insignificant average spillover effect of the program on ineligible individuals. However, we also find that the ineligible stu-

[1] See also Becker (1974) for an early model of social interactions in consumption decisions.

[2] See Coleman (1961) for an ethnographic study of the relevance of peer groups among U.S. High School students. Other reasons for social interaction in schooling include social learning (Manski, 2004) and strategic consideration, i.e. that it pays off to acquire more schooling if other students acquire more schooling due to labor market competition.

[3] See Behrman *et al.* (2001) and Buddelmeyer and Skoufias (2003) for evaluations of this program.

[4] We discuss several concerns with this assumption below.

dents' response is larger the larger is the eligible fraction of students in the peer group – the students in the same grade level with the same gender. Second, when we combine the response of the ineligible student with information on the effect of PROGRESA on peer group schooling, we find that the ineligible students' schooling decisions are strongly, and statistically significantly, affected by their peer's decision. This finding is important for education policy. Strong social interactions in schooling decisions imply that the effects of educational interventions are amplified due to a social multiplier process.[5]

There is a rapidly expanding literature on social interactions in schooling.[6] Case and Katz (1991) use instrumental variables to study neighborhood effects in the Boston area. Hoxby (2000) identifies peer effects from gender and race variation. Sacerdote (2001) studies peer effects among college freshman at Dartmouth college who are assigned to dorms at random. Hanushek *et al.* (2001) study how peer ability affects student achievement. Using PROGRESA data, Bobonis and Finan (2002) document strong spillover effects on the ineligible students. Duflo and Saez (2003) study the role of information and social interactions in retirement plan decisions in a field experiment. Betts and Zau (2004) use administrative data to study peer groups and academic achievement.

This paper contributes to the literature in three important ways. First, the empirical evidence is based on a clean social experiment that manipulates the actions and decisions of a students' peers while leaving unaffected the student. Such an empirical design is essential in solving the "reflection problem". Second, this paper highlights the social component in the spillover effect due to PROGRESA by focusing on the differential impact of the program along the eligible fraction dimension. This is important because spillovers on the aggregate level may be due to a number of reasons that are not related to the peer group. Third, this paper provides empirical evidence on the relevance of social forces in affecting one of the most important decisions – the amount of time to be invested in acquiring knowledge.

The paper is organized as follows. Section 2 presents background on Mexico and PROGRESA. Section 3 discusses the data and presents descriptive evidence. Section 4 explores a simple theoretical framework for social interactions in schooling and discusses the identification strategy. Section 5 presents the main results, and Section 6 concludes.

[5]See Glaeser *et al.* (2003) for a discussion of the social multiplier.

[6]See Glaeser *et al.* (1996) on social interaction in crime, and Lalive (2003), Topa (2001), and Topa and Conley (2002) for social interaction in unemployment.

2 Background

According to the General Education Act from 1993 the educational system in Mexico comprises 3 levels: Basic, which is sub-classified in Pre-primary, Primary and Lower Secondary; Upper secondary and Tertiary. However the only two types that are obligatory for all Mexican citizens according to the Mexican Political Constitution are Primary and Lower Secondary school.

Primary school is imparted for children and adults in three different modalities: general; bilingual-bicultural, provided for indigenous people; and community courses, provided for people in small or isolated communities. Primary school consists of 6 grades, and it can be attended by pupils between 6 and 14 years old. According to data from the Secretariat of Public Education (SEP), 92.8% of children were covered by primary education in the periods 1997-1998 and 1998-1999. The ratio of successfully completed courses was slightly higher for the period 1998-1999, 85.8% against 84.9% in 1998-1999.

The Lower Secondary school lasts for three years and it is provided to the population between 12 and 16 years old who have previously completed the Primary school. According to the SEP, the absorbtion rate for the secondary level was 87.8% and 90% for the periods 1997-1998 and 1998-1999 respectively. The coverage also increased in the same period from 75.7% to 77.7% as well as the ratio of successfully completed courses which increased from 73.8% to 76.1%.

In order to encourage enrollment and permanence in school of children and teenagers under 18 years old who attend grades between third of primary and third of secondary the Mexican government created PROGRESA (Programa de Educacion, Salud y Alimentacion)[7] which is a program aimed at increasing the opportunities and complementing the income of Mexican families living in conditions of extreme poverty. It has three components: education, health and nutrition. However, we will focus only on the educational component which consists of cash transfers which are provided to poor families every two months during the school term (August to June) conditional on sending their children to school. They have two particularities: the sums granted increase as children reach higher grades and in the secondary school the sums awarded to girls are higher than those for boys, due to the higher proportion of girls dropping out of school. The level of the school subsidy approximates what a child would earn in the labor market or what she or he could contribute to the household income. For example, in the period July-December 1998, the amount of the grants increased from 70 to 135 pesos per month between the third and sixth grade of primary school and almost to the double for

[7]The name have been currently changed to Oportunidades

enrolling in first grade of secondary[8]. In order to avoid reducing a household's incentives for self-help the total monthly monetary transfer a family could receive was limited at 525 pesos in July-December 1998. The nominal values of the cash transfers are adjusted every 6 months to take into account changes in cost of living.

These cash subsidies are handed out to the mother because of the belief that the mother is usually better administrating the household resources and because women are disproportionately vulnerable to poverty. The grants are awarded only after confirming the children's attendance to school at least for the 85% of the school calender year. If they fail to fulfill this requirement they loose the grant, at first temporarily and then permanently.

PROGRESA identified localities as being eligible to participate because of a high degree of marginality that was determined primarily on the bases of analysis of data in the 1990 and 1995 population census.[9] The resulting sample contains about 24'000 households from 506 rural communities which were located in seven states: Guerrero, Hidalgo, Michoacan, Puebla, Queretaro, San Luis Potosi y Veracruz. From these 506 localities, 320 were randomly selected as treatment localities and the remaining 186 as control localities. Only the communities chosen as treatment received PROGRESA benefits during the year 1998 (starting in August 1998). Within the treatment and control villages the program was targeted at poor families, that is, those families with per capita income lower than the cost of the standard food basket (roughly 320 pesos). This procedure lead to roughly 52 % of all households being classified as poor. Note that this selection procedure leads to a higher chance of being eligible for households with children compared to single households. By July 1999, PROGRESA added a further 26 % of all household to the list of beneficiaries because it was felt that the initial procedure discriminated against households whose children had already left home. However, in 2000, when the PROGRESA demonstration period had ended, it was discovered that many of the households re-classified in July 1999 as poor did not receive any benefits.

An important issue arises due to the strict targeting of the program to the poor. If children from poor households only interact with other children from poor households, there could be important peer effects that can not be detected with this experimental design. However, note that this argument implies that any social interactions that we find can be thought to represent a lower bound on within poor / within non-poor social interactions. Second, there is strong

[8]Specifically, the cash subsidy is 70 pesos in grade 3, 80 pesos in grade 4, 95 pesos in grade 5, and 115 in grade 6 of primary school. The cash subsidy then jumps to 200 (210) pesos in grade 1 of secondary school for boys (girls) with marginal increases in grades 2 and 3 averaging again 10-20 pesos. Beneficiary families also receive help for the school material, which equaled 135 pesos for primary school and 170 for secondary school for the period July-December 1998

[9]See Skoufias (2001) for a more detailed discussion.

evidence that the inhabitants of the 506 rural and extremely poor villages in Mexico felt that the selection of poor families was quite arbitrary.[10] This suggests that social relationships exists also between the poor and the non-poor. Moreover, the fact that the program was implemented did not seem to change existing social relationships.[11]

3 Data and Descriptive Evidence

The PROGRESA evaluation database contains information on school attendance, socio-economic characteristics, and localities for 7 waves in the period between October 1997 and November 2000. The empirical analysis uses information on the first four waves, i.e. October 1997 (wave 1), May 1998 (wave 2), October 1998 (wave 3), and May 1999 (wave 4). Thus, we focus on the time period when PROGRESA is a clean targeted intervention – when most household classified eligible for the program did receive benefits and most of the ineligible households did not receive benefits. Second, we concentrate on children living with their mother who are enrolled in grades 3-6 of primary school in October 1997 because all children in these grades are eligible for the subsidy in October 1998.[12] This criterion rules out children enrolled in grades 1-3 of secondary school in the year prior to the start of the program. However, note that (i) for many of these children we have no information on school attendance in October 1998 (this item is only available for children aged 6-16), and (ii) many villages do not have a local secondary school – implying that the children in the local village are only a part of the social network. Third, we concentrate on children aged 6 to 16 years in October 1998 because the school attendance item was only collected for children in this age group. We end up with a sample of 15,653 children of which 9,690 live in "treated" villages – where the program was implemented in August 1998 – and 5,963 live in "control" villages that were denied access to the program in August 1998.

Table 1 provides descriptive statistics on the background characteristics of the children in the sample. Panel A in Table 1 reports statistics for the 10,484 eligible children. About 50 % of all children are enrolled in grades 5-6 in the year prior to the implementation of the program both in treated and control villages (the remaining 50 % being enrolled in grades 3-4). Also, the percentage of girls is identical in treated and control villages. Educational attainment of

[10] "Among beneficiaries, non-beneficiaries and *promotoras*, there was a strong view expressed that 'everyone is poor' – a sense of common identity in poverty." (Adato, 2000, p. vi)

[11] "Many comments were made suggesting that beneficiaries and non-beneficiaries continue to get along with each other fine and 'the same' as before." (Adato, 2000, p. vi)

[12] The sample selection can not be based on the grade attained in October 1998 because grade attainment is endogenous.

parents is low: Roughly 18 % of all mothers, and 19 % of all fathers have completed primary education or higher. Note that for about 10 % of all children, information on the father is missing (in which case we set the educational attainment dummy variable to zero). Thus, information on the average father educational attainment is slightly biased towards zero.[13] The last row in Table 1 reports information on secondary school availability within the village as a proxy for the costs of attending a secondary school which is relevant for children in grades 5-6. Results indicate that roughly 25 % of all eligible children live in villages with a secondary school.

Table 1 about here

Panel B in Table 1 reports the means and treatment contrasts for the 5,169 ineligible children.[14] In contrast to the eligible children, enrolment in grades 5-6 is about 57 %, i.e. slightly higher than in grades 3-4. This is arguably due to the larger fraction of children wanting to complete primary education leading to a larger fraction of grade repeaters. The second difference between eligible and ineligible children is parental educational attainment. Between 25 % and 28 % of eligible children's mothers and fathers have completed primary education or more, whereas the corresponding figures are between 18 % and 19 % for among eligible children. Interestingly, the fraction of children with missing information on the father is 12 % among ineligible children which is slightly higher than in eligible children. The fraction of children living in villages with secondary school is slightly higher in the control sample compared to the treated sample. However, the difference is not statistically significant.[15] The fact that none of the treatment contrasts are significant is strong evidence in favor of randomization.[16]

Table 2 reports descriptive evidence on the effect of the cash subsidy on school attendance. School attendance is a binary indicator variable taking the value 1 if the child attends school at the date of the interview, and zero otherwise.[17] Panel A in Table 2 reports effects for the eligible children. Roughly 77 % of all children in grades 3-6 attend school in control villages

[13] Assuming that father information is missing at random it is possible to reweigh the available information to give the correct father education figure. The corrected father educational attainment figures are .210 in control villages, and .215 in control villages.

[14] Note that 67 % of all children in our dataset are eligible for the cash subsidy. This is a result of the procedure that determines eligibility based on per capita income favoring households with children.

[15] Furthermore to the extent that the availability of a secondary school encourages school progression, imbalance with respect to the availability of secondary schools will tend to bias downward the treatment contrast.

[16] See Behrman and Todd (1999) for an exhaustive discussion of randomization in PROGRESA.

[17] A second important outcome is achievement test scores. However, there is no effect of PROGRESA on test scores (Behrman *et al.*, 2000).

in October 1997. In treated villages, school attendance is slightly higher, 78 %, one year prior to the start of the program. However, the treatment contrast is not significant at any conventional level of significance. One year later, in October 1998, school attendance is 69 % in control villages – 8 percentage points lower than the year before. This means that a substantial fraction of children in our sample have dropped out of school in control villages. In contrast, in treated villages school attendance is 76 % – only 2 percentage points lower than the year before. This means that the program increased school attendance by 6 percentage points – a significant impact both in the economic and statistical sense.

Table 2 about here

Panel B in Table 2 reports the effect of the cash subsidy on the ineligible children. School attendance is about 76 % in treated villages, and about 78 % in control villages about one year before the program was introduced – the treatment contrast being insignificant. By October 1998, school attendance has dropped by 7 percentage points in control villages but only by 5 percentage points for ineligible children in treated villages. Thus, the program appears to have increased school attendance sightly, by 2 percentage points, among ineligible children. However, this "spillover" effect is not significantly different from zero at any conventional level of significance.[18] Thus, results in Table 2 indicate that PROGRESA strongly increases schooling for the eligible children but only very weakly for the ineligible children suggesting weak or no social spillovers of the program.

However, note that villages differ strongly with respect to the fraction of households eligible for the schooling subsidy with some villages being characterized by a very low eligible fraction and other villages being characterized by a very high eligible fraction. This means that the ineligible children's response to introducing a cash subsidy in their peer group also must vary in line with the eligible fraction in the peer group. We define the peer group of child i to contain all other children in the same grade *level* and same gender. Grade level refers to grades 3-4 or grades 5-6.Both characteristics are very important in defining the social group that individuals identify with. Furthermore, many schools in rural Mexico do not have separate classrooms for all grades. This means that using grade levels may capture more closely the actual peer groups than using the actual grades. We exclude 28 children for whom the peer group is empty.[19]

[18]This result is in contrast to Bobonis and Finan (2002). However, note that Bobonis and Finan (2002) study secondary school enrolment whereas this paper focuses on school attendance in (predominantly) primary school.

[19]In defining peer groups we balance the cost of loosing observations due to empty peer groups with the benefits of tightly defined peer groups. While defining the peer group to contain the children in the same grade does not change results in important respects, this results in a substantially stronger loss in terms of the number of observations with non-empty reference groups.

Table 3 studies the differential response to the program with respect to the eligible fraction in the peer group. Table 3 distinguishes four levels of impact of the program on peer group schooling. School attendance in peer groups with an eligible fraction between 0 percent and 25 percent is not expected to change much because most children in the peer group are ineligible. The average response of ineligible children in peer groups with this very low eligible fraction is -0.8 percentage point (not significantly different from zero). The second level of social impact arises in peer groups with an eligible fraction between 26 percent and 50 percent. In these groups, the school subsidy has a larger impact on average schooling in the peer group compared to the lowest level of social influence. Interestingly, the average response of ineligible children in these groups is now positive – school attendance increases by 1.4 percentage points due to the program (not significantly different from zero). The third level of social impact arises in peer groups with an eligible fraction of between 51 to 75 percent. The ineligible children's average response to the program is, again, positive and amounts to 1.3 percentage points (not significantly different from zero). The highest level of social influence is exerted in peer groups with eligible fraction exceeding 75 percent. Clearly, the school subsidy has a tremendous impact in these peer group. It turns out that the response of the ineligible children in these peer groups is also strongest. School attendance increases by almost 5 percentage points due to the school subsidy.[20] This effect is economically and statistically significant. Thus, descriptive evidence in Table 3 is consistent with an interpretation that the ineligible students' decision to attend school are affected by the corresponding decisions of other students in the same grade level and with the same gender.

Table 3 about here

A number of concerns with the empirical evidence in Table 3 arise. First, an increase in school attendance among ineligible children can be due to an expansion of school resources in treated villages that did not occur in control villages. However, note that the detailed location information available does not indicate any change in the number of classrooms available in primary and secondary school. Second, beneficiaries may have shared some of the cash benefits with non-beneficiaries. Note that this concern is only relevant if beneficiaries share PROGRESA benefits *conditional* on attending school. Moreover, there is no evidence that sharing between beneficiaries and non-beneficiaries is quantitatively important (Bobonis and Finan, 2002). Third, the treatment was not assigned at random conditional on the eligible fraction. Thus, the eligible fraction in treated villages might be correlated with other deter-

[20]Note that a stronger response of ineligible children could also indicate that the treatment effect is stronger in peer groups with a high eligible fraction. The analysis in section 5 accounts for this fact.

minants of school attendance in a way that differs across treatment and control villages. We discuss this point at the end of the following section.

4 Theoretical Framework and Identification

It is useful to capture the salient forces affecting the decision to acquire schooling in a very simple and stripped down model. The standard human capital framework identifies the individual return to schooling as the primary marginal benefit, and the (out of pocket and / or opportunity) costs plus foregone income as the primary marginal costs of schooling. These aspects can be captured in the following utility function $U \equiv bS - rS - 1/2kS^2$ where b is the return to schooling, r is the subjective discount rate, and k captures the fact that potential earnings of children dropping out of school are increasing more than one for one in schooling.[21] The first order condition for utility maximization implies that $S = (b - r)/k$. This pins down the individual forces that shape schooling. Schooling is higher, the higher the return to schooling, b, the lower the subjective discount rate, r, and the less convex potential earnings are, i.e. the lower k is.

To illustrate the importance of the social forces we introduce a notion of utility that that captures both *individual* benefits and costs and *social* gains and losses student i enjoys from schooling is as follows[22]

$$U_i = bS_i - rS_i - \frac{1}{2}kS_i^2 - \alpha \sum_{j=1, j\neq i}^{N} (S_i - S_j)^2/N \tag{1}$$

where N is the number of students in the peer group, and α captures the importance of the social gains and losses of schooling relative to the individual costs and benefits of schooling. Thus, we capture the importance of schooling decisions in the peer group by adding a quadratic loss function to the individual utility function. The social utility term is maximized when student i conforms to schooling in the peer group.[23] The first order condition for utility maximization is

[21] Note that the PROGRESA subsidy scheme that approximates potential earnings for drop outs is convex. Moreover, Card (1999) uses this basic model of schooling to organize the literature on the causal effect of education on earnings.

[22] Note that it is straightforward to allow for heterogeneity in the return to schooling b and the discount rate r. For the econometric analysis, however, it is crucial that the curvature of the schooling cost function k and the importance of social forces in schooling α do not vary strongly across individuals. We discuss heterogeneity in these parameters in the following section.

[23] See Akerlof (1980) for a similar way of capturing conformity. Bernheim (1994) shows that conformity also arises even if it is not directly embedded in the utility function.

$$S_i = \frac{b-r}{\alpha+k} + \frac{\alpha}{\alpha+k} \sum_{j=1, j\neq i}^{N} S_j/N \tag{2}$$

This shows that the average schooling level in the peer group affects the level of schooling of individual i in addition to the individual returns and costs of schooling.[24] The higher the level of schooling is in the peer group, the more schooling will be acquired by the individual. Moreover, schooling decisions that recognize individual costs and benefits differ from the decisions that recognize social gains on top of individual costs and benefits. As the importance of social gains, α, increases, individual schooling increases provided that the individual net benefit $b-r$ is not too large.

The aim of this paper is to provide reliable evidence on the importance of social gains in shaping schooling decisions. The importance of social forces in schooling is nicely captured by $\gamma \equiv \alpha/(\alpha+k)$. This parameter is non-zero if and only if α is non-zero. Moreover, γ captures nicely the relative importance of convexity in foregone child income, k, and of social gains due to conformity to peer group schooling, α. The basic econometric model that is used in identifying the importance of social interactions in schooling is[25]

$$S_i = x_i\beta + \gamma \sum_{j=1, j\neq i}^{N} S_j/N + \epsilon_i \tag{3}$$

Cross section evidence on the magnitude of the parameter γ suffers from two important problems – omitted variables, and simultaneity (Manski, 1993). This paper addresses these problems in the context of a randomized, targeted intervention. Let $e_i = 1$ if child i is eligible for the school subsidy, and $e_i = 0$ otherwise. Furthermore, let $T_i = 1$ if the village is a treated village, and $T_i = 0$ otherwise. PROGRESA hands out a cash subsidy that is slightly convex in the grade level attained. This means that PROGRESA decreases the magnitude of the parameter k for eligible individuals. This directly increases the desired level of schooling among eligible children leading to an increase in school attendance (since school attendance is a pre-requisite for grade passing).[26] Moreover, the program also affects the average desired level of schooling at the peer group level – only weakly in peer groups with a low eligible fraction ($e_{gi} \equiv \sum_{j=1, j\neq i}^{N} e_i/N$) but quite strongly in peer groups with a high eligible fraction. Thus, the interaction term treated village * eligible fraction ($T_i e_{gi}$) is an instrument that is expected

[24] Note that a more general discussion of the relevance of social forces in schooling recognizes not only the average level of schooling but also dispersion in schooling in the peer group. However, we focus on average schooling decisions for simplicity.

[25] This model is similar to the canonical model of social interactions discussed in Manski (1993) and identical to the model used in Glaeser *et al.* (2003).

[26] See Attanasio *et al.* (2003) for an exhaustive structural analysis of the effects of PROGRESA.

to affect average change in schooling in the peer group in treated villages. Is this instrument orthogonal to omitted variables – the second condition for a valid instrument? This condition can be violated among eligible children because PROGRESA affects both, the individual costs and benefits to schooling, and the conformity gains. However, the individual costs and benefits remain unaffected by PROGRESA among ineligible children. This issue can be explored by regressing each of the observed characteristics in the dataset on the eligible fraction, separately for treated and control villages. Results indicate that the correlation between the control variables and the eligible fraction is *not* statistically different between ineligible children living in control villages and treated villages. This is in line with randomization of the program balancing all observed and unobserved determinants of schooling. This means that the salience of PROGRESA in the peer group is a valid instrument for the average change in schooling in the peer group for ineligible children.

5 Results

Table 4 reports the main results of the empirical analysis. The first Column in Table 4 provides the IV estimates. The dependent variable is the change in school attendance between October 1997 and October 1998 for ineligible children. Results indicate that individual schooling decisions are strongly affected by the corresponding decisions in the peer group. The marginal effect of peer group schooling on individual schooling of the ineligible student, γ, is .581. The parameter estimate is significantly different from zero at the 5 % level. This implies that social forces are about equally important as individual costs and benefits in human capital investment. Moreover, the strong response of ineligible individuals to the change in average peer group schooling implies that the effects of this educational intervention is amplified. The idea is that the intervention generates a first direct increase in the desired level of schooling that is limited to eligible students. By increasing the average level of schooling in the peer group, PROGRESA generates a second, social shock to the desired level of schooling for all students. Thus, this intervention triggers a social multiplier leading to an increase in schooling that is about twice $(1/(1-\gamma))$ as strong as the initial direct shock of the school subsidy.

Table 4 about here

Table 4 also shows that school attendance drop significantly more strongly for children in grades 5-6 of primary school compared to children enrolled in grades 3-4. This is evidence of children dropping out of primary school before completion. Furthermore, estimates suggest that educational attainment of the father seems to matter. School attendance increases sig-

nificantly more strongly for children whose father has completed primary education compared to children whose father has not completed primary school (or with missing information on the father's education). School attendance of children with no information on the father drops significantly more strongly than for children with information on the father. The remaining control variables do not significantly affect the change in school attendance. Note that this is due to the fact that the dependent variable reflects the change in school attendance rather than the level of school attendance.[27] This implies that all time invariant determinants of schooling have already been eliminated.

The second Column in Table 4 reports the effect of the subsidy on the change in peer group average schooling between October 1997 and October 1998 – the first stage results. The effect of PROGRESA on peer group schooling is captured by the two parameters "Treated" and "Treated * Eligible Fraction in the Peer Group". The latter parameter is significantly different from zero. Increasing the eligible fraction in the peer group from 0 to 1 leads to an 8 percentage points higher change in school attendance. It is instructive to compare the effect of the intervention on peer group average schooling to the direct effect of the school subsidy on eligible individuals. This effect can be measured by comparing eligible students to ineligible students *living in the same village* if both groups of students are affected equally strongly by their social environment. Descriptive evidence in Table 2 suggests that this direct effect is on the order of 4 percentage points. Interestingly, the effect of the educational intervention on average peer group schooling is 8 percentage points – roughly twice as high as the direct effect of PROGRESA conditional on the social environment. The "Treated" parameter that measures differences between treated and control villages when the eligible fraction is 0 is not significantly different from zero. This suggests that possibly unobserved determinants of schooling are balanced with respect to treatment. The last row in Column 2 of Table 4 gives the value of the F-statistic for the first stage regression. Clearly, the fact that this statistic exceeds the crucial threshold of 1 by 3.4 points indicates that the analysis is not affected by small sample bias due to weak instruments (Bound *et al.*, 1995).

Results for the remaining characteristics in the First Stage Regression indicate that there is a weakly significant negative correlation between the change in school attendance and the eligible fraction in control villages. Average peer group school attendance drops significantly more strongly in grades 5-6 compared to grades 3-4. Villages with a local secondary school enjoy an increase in school attendance compared to villages without a local secondary school. Interestingly, characteristics that refer to the ineligible *individual* have no explanatory power

[27] All characteristics except for "father information missing" are correlated with the level of school attendance in a significant manner.

for the average change in school attendance in the peer group. This is due to the fact that the peer group can be quite heterogeneous with respect to these characteristics.

Table 5 discusses two important concerns with the identification strategy. Column A in Table 5 reports the baseline result from Table 4 for convenience. Column B in Table 5 addresses the concern that there could be changes in the local labor market that affect schooling decisions differently in treated locations compared to control locations. This concern can be addressed by controlling for state * local region (*municipalidad*) interactions. This procedure identifies 191 local regions (compared to only 7 states used in all analyses so far). This means that the effect of group schooling on individual schooling is only identified from the variance in the salience of the program within very narrowly defined regions. Results indicate that social forces are important in schooling decision. The parameter estimate on the coefficient that captures the importance of social interactions is larger than the corresponding baseline estimate. This means that baseline estimates tend to give a lower bound on the actual relevance of social forces in schooling.

Table 5 about here

Column C in Table 5 discusses the primary identifying assumption that time trends in schooling would have been identical in the absence of treatment for ineligible students in treated and control villages. Recall that PROGRESA collects information at bi-annal frequency. This means that in May 1998 (wave 2) and in May 1999 (wave 4) there is, again, information on school attendance before and after introducing the program. Furthermore, whereas the situation in October (baseline estimates) captures school attendance in the beginning of the school year, the waves 2 and 4 which take place in May reflect school attendance in the end of the school year. Thus, using waves 2 and 4 in identifying social interactions is a sensitivity analysis that provides information on the extent to which within school year *trends* are comparable with respect to treatment status of the village. There is an important data issue with waves 2 and 4. Whereas waves 1 and 3 contain information on all children, such information is missing for about 20 % of all children in waves 2 and 4. It is not clear why this is the case. Nevertheless, we can address this issue by setting school attendance information to zero if it is not available. To correct for missing school attendance information we weigh each observation with the inverse of the probability of being observed in waves 2 and 4 conditional on village, grade level and gender. This procedure is valid if the data is missing at random conditional on village, grade level, and gender. It turns out that for 85 children, the probability of being observed is zero leading to a slightly smaller sample of 5,056 observations. Results indicate that social interactions are about equally important in data from waves 2 and 4

compared to the baseline estimates. This is consistent with similar time trends in treated and control variables in the absence of treatment.

So far the empirical analysis has assumed that the social interactions parameter $\gamma = \alpha/(\alpha + k)$ is identical across individuals. This assumption motivates a standard constant coefficient regression model for identification. Table 6 explores the extent to which this assumption is true along the two important dimensions gender and grade level. Results for girls suggest that the social interactions parameter point estimate is slightly lower than in the baseline estimate failing the threshold for statistical significance at the 10 % level (Column A in Table 6). Results for boys are significantly different from zero and slightly larger than the baseline estimate. Thus, results by gender do not suggest any important differences in the extent to which social forces affect human capital decisions.[28] Interestingly, the change in school attendance for boys appears to be significantly related to whether or not the father is present in the household. School attendance drops significantly more strongly for boys without information on the father.

Table 6 about here

Columns C and D in Table 6 report the social interactions coefficient by grade level. Results indicate that social interactions are significantly different from zero and almost equally important for younger and older children. Interestingly, father educational attainment appears to matter strongly for children in grades 5-6 but not for children enrolled in grades 3-4 prior to the start of the program. Children whose father has completed primary school or more are significantly more likely to continue to attend school compared to children whose father has not completed primary education. Furthermore, the presence of the father is associated with a significant increase in school attendance in grades 5-6. This suggests that fathers are important in encouraging their children to attempt to reach secondary school level. The overall impression from the results for subgroups is that social gains to schooling are very important. Furthermore, these social gains appear to affect schooling of all students in roughly the same manner, i.e. irrespective of the actual identity of the student.

[28] Our results of equally important social interactions for boys and girls are in contrast to Handa *et al.* (2000) finding strong effects on ineligible girls; and our results are in contrast to Bobonis and Finan (2002) reporting strong effects for ineligible boys. Note, however, that both papers are based on the average response among ineligible children. As this paper demonstrates, the average response may not be informative on the relevance of social interactions.

6 Conclusions

This paper argues that conformity of individual schooling decisions to peer group schooling decisions may be an important determinant of human capital formation. The relevance of the resulting conformity effects can be studied in the context of an experiment that grants a cash subsidy to a subgroup of students in villages across rural Mexico. This subsidy encourages the eligible students to remain in school rather than to start working. Moreover, if conformity effects are relevant, the ineligible students may also decide to acquire more schooling provided that a salient fraction of their peer are eligible for the subsidy.

Results indicate first that there is a positive but statistically insignificant average effect of the program on ineligible individuals. However, we also find that the ineligible students' response to the school subsidy is larger the larger is the eligible fraction of students in the peer group – the students in the same grade level with the same gender. Second, when we combine the response of the ineligible student with information on the effect of PROGRESA on peer group schooling, we find that the ineligible students' schooling decisions are strongly, and statistically significantly, affected by their peer's decision. This finding is important for education policy. Strong social interactions in schooling decisions imply that the effects of educational interventions are amplified due to a social multiplier process.

References

Adato, M. (2000). The Impact of PROGRESA on Community Social Relationships. Report, IFPRI.

Akerlof, G. A. (1980). A Theory of Social Custom, of Which Unemployment May be One Consequence. *Quarterly Journal of Economics*, **90**(4), 749–775.

Akerlof, G. A. and Kranton, R. E. (2002). Identity and Schooling: Some Lessons for the Economics of Education. *Journal of Economic Literature*, **150**(December 2002), 1167–1201.

Attanasio, O., Meghir, C., and Szekely, M. (2003). Using Randomised Experiments And Structural Models For 'Scaling Up': Evidence From The PROGRESA Evaluation. Working paper no ewp03/05, Edepo.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. NY: Columbia University Press.

Becker, G. S. (1974). A Theory of Social Interactions. *Journal of Political Economy*, **82**(6), 1063–1093.

Behrman, J. R. and Todd, P. E. (1999). Randomness in the Experimental Samples of PROGRESA. Report, IFPRI.

Behrman, J. R., Sengupta, P., and Todd, P. E. (2000). The Impact of PROGRESA on Achievement Test Scores in the First Year. Report, IFPRI.

Behrman, J. R., Sengupta, P., and Todd, P. E. (2001). Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment. Working paper, University of Pennsylvania.

Bernheim, D. B. (1994). A Theory of Conformity. *Journal of Political Economomy*, **102**(5), 841–877.

Betts, J. R. and Zau, A. (2004). Peer Groups and Academic Achievement: Panel Evidence from Administrative Data. Working paper, University of California.

Bobonis, G. J. and Finan, F. S. (2002). Transfers to the Poor increase the Schooling of the Non-Poor: The Case of Mexico's Progresa Program. Working paper, University of California at Berkeley.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak. *Journal of the American Statistical Association*, **90**(430), 443–450.

Buddelmeyer, H. and Skoufias, E. (2003). An Evaluation of the Performance of Regression Discontinuity Design on PROGRESA. Working paper, University of Melbourne.

Card, D. (1999). The Causal Effect of Education on Earnings. In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics, Vol. 3*, pages 1801– 1863. North-Holland.

Case, A. C. and Katz, L. F. (1991). The Company You Keep: The Effect of Family and Neighborhood on Disadvantaged Youth. Working paper 3708, NBER.

Coleman, J. S. (1961). *The Adolescent Society: The Social Life of the Teenagers and Its Impact on Education*. NY: The Free Press.

Duflo, E. and Saez, E. (2003). The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment. *Quarterly Journal of Economics*, **118**(3), 815–842.

Glaeser, E. L., Sacerdote, B., and Scheinkman, J. A. (1996). Crime and Social Interactions. *Quarterly Journal of Economics*, **111**(2), 507–548.

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

Handa, S., Huerta, M.-C., Pere, R., and Straffon, B. (2000). Poverty, Inequality, and Spill-Over in Mexico's Educational, Health, and Nutrition Program. Report, IFPRI.

Hanushek, E. A., Kain, J. F., Markman, J. M., and Rivkin, S. G. (2001). Does Peer Ability Affect Student Achievement. Working paper 8502, NBER.

Hoxby, C. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation. Working paper 7867, NBER.

Lalive, R. (2003). Social Interactions in Unemployment. Discussion paper 803, IZA.

Manski, C. F. (1993). Identification of Endogenous Social Interactions: The Reflection Problem. *Review of Economic Studies*, **60**(3), 531–542.

Manski, C. F. (1995). *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press.

Manski, C. F. (2000). Economic Analysis of Social Interactions. *Journal of Economic Perspectives*, **14**(3), 115–136.

Manski, C. F. (2004). Social Learning from Private Experiences: The Dynamics of the Selection Problem. *Review of Economic Studies*, **71**(2), 443–458.

Moffitt, R. A. (2001). Policy Interventions, Low Level Equilibria, and Social Interactions. In S. N. Durlauf and H. P. Young, editors, *Social Dynamics*, pages 45–82. Cambridge, Massachusetts: MIT Press.

Sacerdote, B. (2001). Peer Effects with Random Assignment. *Quarterly Journal of Economics*, **116**(2), 681–704.

Skoufias, E. (2001). PROGRESA and its Impacts on the Human Capital and Welfare of Households in Rural Mexico: A Synthesis of the Results of an Evaluation by IFPRI. Report, IFPRI.

Topa, G. (2001). Social Interactions, Local Spillovers and Unemployment. *Review of Economic Studies*, **68**(2), 261–295.

Topa, G. and Conley, T. G. (2002). Socio-Economic Distance and Spatial Patterns in Unemployment. *Journal of Applied Econometrics*, **17**(4), 303–327.

Table 1:
Descriptive Statistics

	A. Eligible			*B. Ineligible*		
	Control	Treated	Difference (abs. z-Value)	Control	Treated	Difference (abs. z-Value)
Grades 5-6	0.496	0.494	-0.002 (0.137)	0.564	0.576	0.012 (0.890)
Girl	0.502	0.492	-0.010 (0.868)	0.50	0.51	0.010 (0.650)
Mother: Primary Ed. or More	0.184	0.178	-0.007 (0.345)	0.256	0.280	0.024 (0.874)
Father: Primary Ed. or More	0.189	0.195	0.006 (0.284)	0.253	0.266	0.014 (0.452)
Father Information Missing	0.099	0.095	-0.004 (0.413)	0.117	0.119	0.001 (0.129)
Secondary School in Village	0.260	0.256	-0.004 (0.069)	0.340	0.265	-0.075 (1.146)
N(obs.)	3880	6604		2083	3086	

Notes: Eligible means Household is eligible for Cash Subsidy.
Sample considers children in grades 3-6 in October 1997, aged 6-16 in October 1998, living with mother.
Source: Own Calculation, Based on Progresa Evaluation Data.

Table 2:
Descriptive Evidence on the Effect of Cash Subsidies on School Attendance

	A. Eligible			*B. Ineligible*		
	Control	Treated	Difference (abs. z-Value)	Control	Treated	Difference (abs. z-Value)
School Attendance (SA), October 1997	0.769	0.782	0.013 (0.897)	0.758	0.782	0.024 (1.305)
School Attendance (SA), October 1998	0.685	0.757	0.072 (3.798)	0.691	0.733	(0.042) (1.907)
Change in School Attendance	-0.084	-0.025	0.059 (5.511)	-0.067	-0.049	0.018 (1.484)
N(obs.)	3880	6604		2083	3086	

Notes: Eligible means Household is eligible for Cash Subsidy. October 1997 is before program, October 1998 is during program.

Source: Own Calculation, Based on Progresa Evaluation Data.

Table 3:

Decomposing the Effect of Cash Subsidies on the Change in School Attendance among the *Ineligible*

Eligible Fraction in Peer Group	Control	Treated	Difference (abs. z-Value)
0.00-0.25	-0.068	-0.077	-0.008 (0.281)
0.26-0.50	-0.057	-0.043	0.014 (0.666)
0.51-0.75	-0.067	-0.053	0.013 (0.668)
0.76-1.00	-0.086	-0.038	0.047 (1.969)
Total	-0.067	-0.049	0.018 (1.484)

Notes: Peer group comprised all children with same gender and grade level. Average number of children in peer group is 12. Number of Observations in second row.

Source: Own Calculation, Based on Progresa Evaluation Data.

Table 4:
The Effect of Peer Group Schooling on Individual Schooling

Dependent Variable	School Attendance (Change)	Avg. SA in Peer Group (Change)
	IV	OLS
Avg. SA in Peer Group (Change)	0.581	-
	(0.240)**	
Treated * Eligible Fraction in Peer Group		0.080
	-	(0.036)**
Treated	-	-0.008
		(0.024)
Eligible Fraction in Peer Group	0.008	-0.047
	(0.018)	(0.027)*
Grade 5-6	-0.024	-0.045
	(0.014)*	(0.008)***
Girl	-0.009	0.008
	(0.009)	(0.008)
Mother: Primary Ed. or More	0.006	0.005
	(0.008)	(0.005)
Father: Primary Ed. or More	0.018	0.004
	(0.009)**	(0.005)
Father Information Missing	-0.040	0.013
	(0.016)**	(0.008)
Secondary School in Village	0.010	0.027
	(0.012)	(0.009)***
State Effects (7)	Yes	Yes
Observations	5141	5141
First Stage F-Statistic	-	4.40***

Notes: Robust standard errors in parentheses (allow for clustering at village level).
* significant at 10%; ** significant at 5%; *** significant at 1%
Peer Group contains children in same grade level (3-4 or 5-6) with same gender
Source: Own Calculation, Based on Progresa Evaluation Data.

Table 5:
Sensitivity Analysis

	A. Baseline	*B. Detailed Regions*	*C. Waves 2 and 4*
Avg. SA in Peer Group (Change)	0.581	0.910	0.641
	(0.240)**	(0.361)**	(0.259)**
Eligible Fraction in Peer Group	0.008	-0.005	0.000
	(0.018)	(0.021)	(0.036)
Grades 5-6	-0.024	-0.009	-0.007
	(0.014)*	(0.019)	(0.029)
Girl	-0.009	-0.012	-0.027
	(0.009)	(0.009)	(0.015)*
Mother: Primary Ed. or More	0.006	-0.000	0.025
	(0.008)	(0.009)	(0.018)
Father: Primary Ed. or More	0.018	0.021	0.011
	(0.009)**	(0.010)**	(0.020)
Father Information Missing	-0.040	-0.047	-0.080
	(0.016)**	(0.018)***	(0.031)**
Secondary School in Village	0.010	-0.018	0.007
	(0.012)	(0.017)	(0.017)
State Effects (7)	Yes	No	Yes
Region Effects (191)	No	Yes	No
Observations	5141	5141	5056

Notes: Robust standard errors in parentheses (allow for clustering at village level).
* significant at 10%; ** significant at 5%; *** significant at 1%
Peer Group contains children in same grade level (3-4 or 5-6) with same gender
Source: Own Calculation, Based on Progresa Evaluation Data.

Table 6:
Effects in Subgroups

	A. Girls	*B. Boys*	*C. Grades 3-4*	*D. Grades 5-6*
Avg. SA in Peer Group (Change)	0.564	0.616	0.564	0.569
	(0.366)	(0.278)**	(0.310)*	(0.313)*
Eligible Fraction in Peer Group	-0.005	0.025	0.038	-0.011
	(0.025)	(0.026)	(0.020)*	(0.028)
Grades 5-6	-0.030	-0.020	-	-
	(0.021)	(0.018)		
Girl	-	-	-0.007	-0.010
			(0.009)	(0.014)
Mother: Primary Ed. or More	-0.002	0.013	0.011	0.003
	(0.012)	(0.011)	(0.008)	(0.014)
Father: Primary Ed. or More	0.017	0.019	0.002	0.034
	(0.013)	(0.012)	(0.009)	(0.015)**
Father Information Missing	-0.031	-0.048	0.009	-0.075
	(0.024)	(0.022)**	(0.021)	(0.023)***
Secondary School in Village	0.010	0.011	0.014	0.006
	(0.015)	(0.016)	(0.011)	(0.020)
State Effects (7)	Yes	Yes	Yes	Yes
Observations	2592	2549	2198	2943

Notes: Robust standard errors in parentheses (allow for clustering at village level).
* significant at 10%; ** significant at 5%; *** significant at 1%
Peer Group contains children in same grade level (3-4 or 5-6) with same gender
Source: Own Calculation, Based on Progresa Evaluation Data.