Social Interactions and Social Divisions in Program Intervention

Hector H. Sandoval*

April 2022

Abstract

Using micro-level data from the Progresa program in Mexico, this paper estimates a behavioral model of school attendance and social interactions to study whether social divisions appeared as a result of the program targeting, as well as the impact of such divisions on the complementarities arising from social interactions and on program outcomes. Although the qualitative evaluation described how social divisions manifested between beneficiaries and non-beneficiaries due to the program targeting, the empirical results suggest that the distinction introduced by the program did not alter the endogenous-effects network and hence the complementarities. The evidence is consistent with a positive spillover program effect on school attendance and an endogenous effect that account for 25 percent of the overall treatment effect.

JEL codes: I21, I25, I28, Z13.
Keywords: social interactions, peer effects, social divisions, education, Progresa.

*Address: University of Florida, Department of Economics, PO Box 117140, 325 MAT, Gainesville, FL 32611-7140. Telephone: +1 (352) 392-0475, e-mail: hsandoval@ufl.edu.
I am grateful to my advisors Steven Durlauf, James Walker, and Jesse Gregory for their support, encouragement, and guidance, for comments from Richard Romano, Áureo de Paula, Robert Haveman, Jean-Paul Chavas, David Denslow, the participants in the 2014 Public Workshop in UW-Madison, 2015 Seminar in University of Florida, 2015 Southern Economic Association Annual Meeting, 2016 Seminar in the University of Florida, Food and Resource Economics Department, and 2016 Eastern Economic Association Annual Meeting. Earlier drafts were circulated under the title "Social Interactions in Rural Mexico."

Introduction

Defining oneself at least in part by membership in a group is part of human nature. Individuals belong to groups in which they share a common identity or similar characteristics, for example, the beneficiary status of a social program. Importantly, group membership can shape individual choices when social interactions are present, that is, when an individual's payoff from a given action depends directly on the actions of others in that individual's group. A fundamental aspect of social interactions is the presence of complementarities among individuals which can result in spillover program effects in the context of a program intervention. However, the distinction introduced by social programs between beneficiaries and non-beneficiaries can itself disrupt group membership and alter such complementarities. This makes program targeting particularly relevant because it can lead to social divisions, which are undoubtedly an unintended effect of the program that can ultimately limit its impact. These issues are studied in this paper.

Using micro-level data from the conditional cash transfer (CCT) program Progresa in Mexico, this paper estimates a behavioral model of school attendance and social interactions to study how social divisions resulting from program targeting can affect the complementarities arising from social interactions and their subsequent effects on program outcomes. CCT programs like Progresa have now been implemented in over sixty countries on five continents (Parker and Todd, 2017), however, the potential effects of program intervention on social relationships and program outcomes have remained understudied. This paper studies this issue and makes several contributions to the literature. It is the first paper to study social divisions due to program targeting. Specifically, it provides a decomposition of the treatment effect based on the behavioral model, which shows how social divisions can have an impact on spillover program effects. Second, it examines other peer-effects networks and shows that, in the context of the Progresa program, complementarities are stronger at the school-grade level. Finally, it departs from the *linear-in-means model* by emphasizing that contextual- and endogenous-effect networks may differ and that their selection should be determined by the context under study.

In 1997, the Mexican government launched the Progresa program to foster the human capital development of children in poor rural households.[1] A major goal of the program was to encourage school attendance and allow children to complete basic education. The quantitative evaluation and other studies that assessed the impacts of the program showed positive results on schooling and other outcomes, as well as positive spillover effects among non-beneficiaries.[2] However, the

[1]Seeing the positive results associated with the program, several Latin American countries adopted similar programs in the following years: *Bolsa Familia* in Brazil, *Familias en Acción* in Colombia, *Chile Solidario* in Chile, *Mi Familia Progresa* in Guatemala, and *Juntos* in Peru, among many others.

[2]See, for example, Behrman, Parker, and Todd (2011), Behrman, Sengupta, and Todd (2005), Schultz (2004), and Skoufias (2001). Spillover effects are studied in Bobonis and Finan (2009), Lalive and Cattaneo (2009), and Angelucci and De Giorgi (2009). For a more general summary of what is known about Progresa program effects, see Parker and

qualitative evaluation showed that the distinction introduced by the program between beneficiaries and non-beneficiaries led to confrontations, social tensions, and social divisions that affected the relationships within the treated communities. These disruptions to the social structure can alter the complementarities arising from social interactions, which in turn can affect the spillover effects on non-beneficiaries.[3]

The rural localities where Progresa was initially implemented as a randomized experiment were small in size and had one school facility at most. Moreover, community relations were also longstanding characteristics of these localities. These features of the localities make them suitable for studying social interactions. This paper matches the baseline and the first follow-up surveys from the Progresa experiment with administrative records to construct a dataset of 34,117 children in primary (elementary) and secondary (middle high) school to estimate a model of school attendance and social interactions. In the sample, around 68 percent of the children are identified as beneficiaries of the program.

The model is based on the Bayesian game in Blume *et al.* (2015) which provides a behavioral interpretation of the parameters. In the model, school attendance decisions not only depend on a student's own individual characteristics but also on peers' characteristics and peers' attendance behavior, both of which can produce a spillover program effect. However, in the presence of social divisions, the influence of peers may be diminished or even cease to exist. In particular, the parameter of interest is the marginal rate of substitution (*MRS*) between social and private utility, which measures the private utility needed to compensate for the cost incurred when deviating from peers' behavior. In the reduced-form version of the model, this parameter corresponds to the endogenous (peer) effect found in the empirical literature.

In the empirical literature on social interactions, two types of effects are usually distinguished: contextual effects and endogenous (peer) effects. The contextual effects consider the direct influence of peers' characteristics on an individual's choices, while the endogenous (peer) effect considers the direct influence of peers' choices on an individual's behavior. An important aspect of the literature is the use of the *linear-in-means model*, in which both contextual- and peer-effects networks are based on the same reference group (Manski, 1993). However, given the social context in the rural localities where the Progresa program took place, distinguishing between these two networks is important. In particular, this paper considers a more general model that includes two main reference groups, localities and school grades.

Using the behavioral model, this paper provides a decomposition of the treatment effect that shows the potential effect of ruptures in the local social fabrics and adds to a growing literature

Todd (2017).

[3] Adato (2000) explores the effects of the program targeting on community relationships through focus groups and semi-structured interviews.

on spillover effects and experimental designs (Angelucci and De Giorgi, 2009; Angelucci and Di Maro, 2016; and Arduini, Patacchini, and Rainone, 2019). This decomposition shows that the treatment effect on the beneficiaries is comprised of a direct effect from receiving the transfer, an endogenous effect that comes from peers' behavior, and a contextual effect from the transfer at the locality level. The treatment effect on the non-beneficiaries comes exclusively from the endogenous and contextual effects, since they did not receive any transfer. Importantly, the program is not only spilling over to the non-beneficiaries, but also to the beneficiaries through the endogenous and contextual effects. Notably, whenever program targeting results in social divisions between beneficiaries and non-beneficiaries, thus altering the endogenous-effect network, the program will not have any spillover effect on the non-beneficiaries.

Obtaining estimates of endogenous effects confronts a number of problems. In particular, there are three main challenges in estimating social interactions models, namely, simultaneity, endogenous group formation, and a correlated unobservables problem (Moffitt, 2001). To overcome the endogeneity problem, this paper uses the behavioral model to derive instruments for peers' behavior. Specifically, the model shows that the average exogenous cash transfer from the program at the peer-effects network level is a valid instrument. Another challenge is that peers might self-select into groups in ways unobserved by the econometrician. As mentioned before, this paper considers two main reference groups, localities and school grades. Although the formation of the locality might be endogenous, behavioral choices related to school attendance are on a different time scale relative to the locality-level group membership. In other words, the locality can be considered as an exogenous reference group for the purpose of this paper. Nonetheless, self-selection can arise if children decide to repeat a school grade to spend more time with their friends (peers) who have failed, for example. Although grade repetition is common because attendance is not sufficiently regular to fulfill the requirements to complete the grade (Behrman, Sengupta, and Todd, 2005), the exogenous nature of the Progresa cash transfer still permits the identification of the endogenous effect. Finally, to surmount the group level correlated unobservables problem, such as teacher's quality, the panel feature of the data is exploited in some specifications to estimate also a first difference version of the model. Additionally, the empirical specification of the model is comprehensive in describing the full range of individual and family backgrounds and contextual factors that may influence children's attendance.

The benchmark specification considers contextual effects at the locality level and endogenous effects at the school-grade level (completed at baseline). The empirical results from this specification show an endogenous effect of 0.2913 on school attendance, which implies that a 10 percent increase in peers' attendance increases the likelihood of attending by 2.9 percent. This effect is consistent with a positive marginal rate of substitution between private and social utility of order 0.4110. That is, private utility should increase by 0.41 to account for the cost incurred from a marginal deviation from peers' behavior. When considering a structure that accounts for social divisions, in the sense that beneficiary and non-beneficiary children comprised two separate groups making up their own group within the school grade level, the endogenous effect is smaller and insignificant. Although the qualitative evaluation described how social divisions manifested between beneficiaries and non-beneficiaries due to the program targeting, the empirical results suggest that the distinction introduced by the program did not alter the endogenous-effects network and hence the complementarities among individuals. In other words, the empirical evidence does not support the case where social divisions manifested between the children of beneficiary and non-beneficiary families. Overall, the evidence is consistent with a positive spillover program effect on school attendance. In particular, the endogenous effect accounts for 25 percent of the overall treatment effect. Finally, when considering partitions of the endogenous-effects network in terms of exogenous characteristics, such as gender, ethnicity, and socioeconomic status, the results show that complementarities are stronger using the benchmark specification.

An implication of the results is that increased school attendance among the non-beneficiary children in the treated localities is due to social interactions. Alternative explanations for such increase are unobserved changes in the income levels of non-beneficiary families, for example, as a result of the informal risk-sharing activities among the households and/or higher labor earnings.[4] Angelucci and De Giorgi (2009) rule out changes in labor income and find an increase in consumption among the non-beneficiaries through the receipt of transfers and loans after the second year of the program. Importantly, this effect on the non-beneficiaries is not present in the first year of the program, which corresponds to the period analyzed in this paper. In other words, at least for the first year of the program, the evidence is consistent with the social interactions' hypothesis.

The literature on peer effects in education is ample and the evidence leaves little doubt about their existence, and for certain contexts and outcomes peer effects are in fact important factors (Epple and Romano, 2011, and Sacerdote, 2011). In terms of school attendance and in the context of the Progresa program, Bobonis and Finan (2009) estimated an endogenous effect of 0.65 among children in secondary school (grades 7-9) using a *linear-in-means model* that considers both contextual and endogenous effects at the locality level. Similarly, Lalive and Cattaneo (2009) estimated an endogenous effect of 0.505 for children in primary school (grades 3-6). The latter paper decomposes the program effect into a direct and an indirect effect due to social interactions, however, both the model specification and the decomposition ignore contextual effects. As shown in this paper, contextual effects are important because they have a direct effect on beneficiaries and indirect effect on non-beneficiaries. In this spirit, using data from a scholarship program in

[4]Progresa may reduce the labor supply of beneficiary children, as some switch from employment to schooling, and it may also reduce adult labor supply through an income effect. This may translate in higher labor income for non-beneficiary families through higher wages and increases in their labor supply. It may also increase the labor supply of non-beneficiary children, having then a negative impact on their school attendance.

Colombia, Dieye *et al.* (2014) employs a social network model to account for peer effects on school attendance. These authors find that the direct program impact and the endogenous peer effect are both positive, while the contextual effect of the treatment is negative. Additionally, they show that ignoring peer effects can lead to biased assessment of the program impacts. The evidence presented in this paper is consistent with Dieye and coauthors', but in contrast to them, the main focus of this paper is on endogenous forms of social divisions that appear as a result of policy intervention, and it's the first to study the role of social divisions in the context of program intervention.

More broadly, this paper contributes to the literature studying how the introduction of programs changes social networks. For instance, Dupas *et al.* (2019) document the effect of expanding access to formal savings accounts on intra and inter-household financial linkages, Comola and Prina (2021) which study the interplay between network changes and treatment effects by looking at how financial interventions also end up affecting those households' networks in Nepal, and Baneerjee *et al.* (2021), which study how the introduction of formal lending institutions changes social networks in rural communities.

The remainder of the paper is organized as follows. Section 1 describes the Progresa program and the randomized experiment. Section 2 describes the Progresa localities, the social relationships, and the social disruptions that arose between beneficiaries and non-beneficiaries. Section 3 describes the behavioral model of school attendance and social interactions. Section 4 presents the data and outlines the empirical strategy, as well as a decomposition of the treatment effect using the behavioral model. Section 5 contains the empirical results. Section 6 presents the main conclusions.

1 Progresa Program

1.1 Overview of the Program

In the year 1997, the Mexican government launched a conditional cash transfer program known as Progresa to give immediate monetary assistance and to foster development of the human capital of poor rural households.[5] The program was intended to reduce long-run poverty by increasing the human capital of poor children. A major goal of the program was to provide poor households with enough resources to allow their children to complete primary and secondary school. To achieve

[5]Progresa stands for *Programa de Educación, Salud y Alimentación* (Program of Education, Health, and Nutrition). During President Fox's administration, from 2000 to 2006, the program was renamed *Oportunidades* (Opportunities). In 2015, President Peña Nieto again renamed the program to *Prosperidades* (Prosperities). Up to year 2013, the program provided benefits to 5.8 million households and continued to be the main antipoverty strategy for the government.

this goal, the program had three different components, education, health, and nutrition, which were provided simultaneously as a package to the eligible (poor) households.[6] The program targeted households and distributed the monetary transfer to the household's mother on a bimonthly basis to increase the likelihood that the benefits have a direct impact on the well-being of children. The idea was that resources controlled by women, instead of men, were more likely to improve child health and nutrition. The program's take-up rate was strikingly high, with 97 percent of the families who were offered the program electing to participate (Parker and Todd, 2017).

The education component targets children in grades 3 to 6 (primary level) and 7 to 9 (lower secondary level). It provides monetary transfers conditional on more than 85 percent of school attendance, which is monitored by the school teacher.[7] The idea is to compensate the household for the value of lost work of the child, reflecting the income they would contribute to their families. The amount of the transfer depends on the child's gender and grade level, and is adjusted every six months for inflation. The grant increases at higher grades and is slightly higher for girls in lower secondary school, since girls tended to have higher dropout rates after completing primary school. During the period from September to December 1998, the transfer for the 6th grade in primary school was $135 Mexican pesos monthly for girls and boys; and for the 7th grade (1st grade in secondary), it was $200 for boys and $210 for girls. This component also adds an extra transfer per year for school materials, $120 for primary school and $240 for secondary school. According to Schultz (2004), the conditional transfer corresponded approximately between one-half and two-thirds of what a child in lower secondary age earned if working full time.

The health component provides basic health care for all family members, with emphasis on the welfare of mothers and children and on preventive health. The health package includes basic sanitation at the family level, family planning, supervision of nutrition and child's growth, vaccinations, anti-parasite treatment, accident prevention and first-aid for injuries, among other health benefits. The nutritional component includes a fixed monetary transfer for consumption, $100 monthly and a nutritional supplement.[8] The supplements are targeted to children between the ages of four months and two years, and pregnant and lactating women. They are also given to children, aged two to four, if any sign of malnutrition is detected. These supplements contain mainly iron, zinc, and vitamins. Similar to the education component, beneficiary families must visit the health center for regular check-ups and preventive care in order to receive the monetary transfer. Moreover, the beneficiaries are required to attend nutrition and health education lectures that pro-

[6]The program followed this approach because the interactions among the components were thought to increase its effectiveness. For example, school attendance and performance are often reduced by poor health and nutritional status.

[7]Failure to comply will lead to the loss of the benefits, first temporarily, then permanently.

[8]The total monthly monetary transfer a family can receive is capped. In the period from July to December 1998, this cap was $625 pesos monthly. Up to November 1999, the transfer averaged $200 pesos (around US $20) per household, Angelucci and De Giorgi (2009).

vide training in using the nutritional supplements and discuss topics related to nutrition, hygiene, infectious disease, immunization, family planning, and chronic diseases detection and prevention. These talks are mainly directed to beneficiary mothers, but everybody is invited to attend, including the non-beneficiaries.

1.2 Randomized Experiment

An evaluation was planned in parallel since the beginning of the program because previous programs and subsidies seemed to be ineffective. Moreover, the evaluation was thought to support the continuation of the program even if the government changed. In the past, it was common for each administration to implement its own particular social programs.[9] Accordingly, the program followed a sequential expansion which made it possible to implement a randomized experiment during the early stages. A control group was selected from the set of poor localities that were eligible for the program, but were not yet planned to be covered due to budget limitations. The control group was no different from the treated one and thus received the benefits a year-and-a-half later.

The program's selection process followed three stages. First, the program geographically selected highly marginalized localities.[10] Within this set of localities, the program randomly selected localities into treatment. In total, the experiment involved 506 localities in seven states in southern and central Mexico.[11] Of the 506 localities, 320 were randomly assigned to the treatment group and the 186 remaining localities were randomized out to form the control group. The random assignment was performed at the locality level because of the broader geographic nature of the program and because randomization at the household level was considered infeasible.

Second, a baseline census was collected in October 1997 in those 506 localities to construct a "poverty index" and identify the beneficiaries (poor) and non-beneficiaries (non-poor) households. The households were evaluated to be in poverty not just by income, but also by other factors, such as the dwelling characteristics and ownership of durable goods, livestock, and land, among others. Only the households below a certain poverty threshold were eligible beneficiaries of the program. Importantly, poverty status was determined in both treatment and control localities.

Finally, a community assembly was arranged in each treated locality to make public the list of selected families and to provide information about the objective of the program. The assemblies were conducted to improve the accuracy of the targeting and to achieve transparency, but in prac-

[9] Notably, during the Fox administration, the program was expanded to urban areas and the educational grants extended to the high school level.

[10] At this stage, the program used a "marginality index" that was constructed using principal components analysis based on literacy, dwelling characteristics, water, drainage, electricity, and household size using census data from 1990 and 1995. This index is stratified in five categories, from very low to very high levels of marginality.

[11] The seven states are: Guerrero, Hidalgo, Michoacán, Puebla, Querétaro, San Luis Potosí, and Veracruz.

tice this step barely resulted in significant changes to the list. At the assembly, households were informed that their eligibility status won't change until November 1999 despite any change in their income. The beneficiaries in the treatment locality began to receive the benefits by May 1998, whereas the ones in the control localities received them by December 1999.

2 Social Interactions in Rural Mexico

In the context of rural Mexico, site of the Progresa experiment, social relationships were well-established within the rural localities where people manifested a group identity. In particular, there are four important features that make these localities suitable for studying social interactions. First, the size of these localities are the smallest among the rural localities in Mexico.[12] Table 1 contains the demographic composition of the localities by treatment status. On average, there are around 48 households per locality, with a mean household size of 5.2 members. This corresponds to 250 inhabitants, where one-third of the population are school-agers (aged 6 - 17), another one-third are young adults (aged 18 - 44), and less than one-third are identified as indigenous.[13] On average, around 60 percent of the residents in each locality are identified as poor.

[12]The National Institute of Statistics and Geography in Mexico considers a locality to be rural if the population size is below 14,999 inhabitants.

[13]The indigenous population are those aged 5 and over who report speaking an indigenous language.

Table 1: Demographic Composition of the Localities by Treatment Status

	Control mean	**Control** s.d.	**Treatment** mean	**Treatment** s.d.	diff.	(t-test) p-value
Population	260	171	241	169	19	0.2175
Babies (0-2)	0.0742	0.0270	0.0746	0.0260	-0.0004	0.8741
Preschoolers (3-5)	0.0859	0.0256	0.0862	0.0264	-0.0003	0.9133
School-agers (6-17	0.3279	0.0485	0.3287	0.0484	-0.0008	0.8618
Young adults (18-44)	0.3262	0.0458	0.3240	0.0469	0.0022	0.6099
Adults (45-64)	0.1268	0.0408	0.1288	0.0389	-0.0020	0.5827
Seniors (65+)	0.0575	0.0440	0.0564	0.0327	0.0012	0.7368
Indigenous	0.2873	0.4116	0.2639	0.3909	0.0235	0.5237
Monolingual	0.0591	0.1422	0.0383	0.1128	0.0208	0.0704
Bilingual	0.2266	0.3311	0.2240	0.3349	0.0026	0.9320
Males	0.5007	0.0314	0.5084	0.3910	0.0235	0.5237
Households	49	30	46	31	3	0.2695
Household size	5.24	0.7657	5.21	0.7917	0.0331	0.6470
Dependency ratio	0.9156	0.1954	0.9122	0.1898	0.0034	0.8476
Poor	0.5780	0.2225	0.5953	0.2185	-0.0174	0.3922
Localities	186		320		Total: 506	

Source: ENCASEH 1997. The sample corresponds to the population in the 506 localities at baseline (October, 1997). The p-value comes from a t-test for the equality of means between the treatment and control samples. The 320 localities from the treatment group were randomly selected.

Second, a longstanding characteristic of the localities (treated and control) is the participation of locality members in community activities in which they come together to perform volunteer work to benefit the entire community.[14] The activities include cleaning and painting schools, cleaning clinics, streets, and surroundings, picking up garbage, building fences and gardens, cutting grass, and tending community orchards.

A third important characteristic is that most of the localities have an association of parents that make contributions for school resources and take school and educational decisions.[15] Table 2 shows the proportion of localities with a school parent association, the proportion of localities that had community activities in the past year, and the proportion of school facilities. It also includes the average number of distinct community activities for both control and treatment localities. Almost all of the localities had community activities where most residents (men and women) participate. Considering data from a locality survey gathered in 1998, around 89 percent of the localities reported community activities, and the median number was 10 in the year preceding the

[14]Some of the community activities are known as *faenas*.

[15]For example, in an interview with school principals from Progresa localities, they mentioned that parents took the responsibility to raise funds to build a provisional classroom and that parents and community members donated their manual labor (Adato *et al.*, 2000).

survey. These community activities are not part of Progresa and the data show that the program did not increase the number of them. Furthermore, the survey shows that around 85 percent of the control and 87 of the treated localities have a school parent association. Although these numbers correspond to the post-treatment period (1998), data from the control localities are informative about the situation that prevailed before the introduction of the program.

Finally, it is worth noting that around 95 percent of the treatment and control localities have one primary school facility, however, this is not true for secondary school facilities. Around 23 percent of the control localities have a secondary school, while only 20 percent of the treated have one, but the difference is not statistically significant.[16] These schools had limited resources and were generally poorly equipped (Adato *et al.*, 2000). Overall, these particular features of the Progresa localities provides an appropriate context to study social interactions.

Table 2: Proportion of Localities with School Parent Association, Community Activities, and Education Facilities by Treatment Status

	Control		**Treatment**			(t-test)
	mean	s.d.	mean	s.d.	diff.	p-value
Associations						
School Parent Assoc.	0.8548	0.3532	0.8746	0.3317	-0.0198	0.5285
Community activities	0.8979	0.3037	0.8934	0.3091	0.0044	0.8757
# of activities last year	16.50	18.14	17.86	19.26	-1.35	0.4647
School facilities						
Kinder	0.8602	0.3477	0.8307	0.3756	0.0295	0.3822
Primary	0.9409	0.2365	0.9624	0.1907	-0.0215	0.2641
Secondary	0.2283	0.4209	0.1924	0.3948	0.0358	0.3397
Localities	186		320		Total: 506	

Source: Locality Survey 1998. The information corresponds to a locality survey gathered in October 1998, after the treatment started. The p-value comes from a t-test for the equality of means between the treatment and control samples. The 320 localities from the treatment group were randomly selected.

2.1 Social Disruptions

As part of Progresa's external evaluation, a qualitative evaluation was carried out to learn about the social impacts and social costs of the program.[17] This project explored the effects of the distinction introduced by the program on the community relationships through a focus group with *promotoras*,[18] beneficiaries, and non-beneficiaries; and through semi-structured interviews with

[16] These numbers for secondary school facilities consider the TV-Secondary and Non-TV-Secondary.

[17] See Adato (2000) for the complete report.

[18] *Promotoras* are beneficiary women selected by their fellow beneficiaries to voluntarily serve as a link between themselves and the Progresa personnel in each locality. They help the program to convey information to beneficiaries and monitor program operation.

the doctors and school principals in the communities.[19]

Beneficiaries and non-beneficiaries in the focus groups were both described as poor, in need of assistance, sharing a common poverty identity, and perceiving each other as a whole 'family'.[20] When asked who do not need the benefits, some said the teachers, professionals, and government workers. Importantly, they did not agree with the distinction introduced by the program between beneficiaries (poor) and non-beneficiaries (non-poor), and their comments indicated that this distinction led to confrontations, social tensions and divisions that affected the participation of non-beneficiaries in the community activities, in the program health talks, and in their involvement with the school parent association.

Comments among the participants in the focus groups described resentment and frustration of non-beneficiaries created by their exclusion from the program and their lack of understanding for such differentiation. Some comments described direct conflict between beneficiaries and non-beneficiaries, but most of the social tensions reported involved resentment, envy and gossip, or comments stating that the community was not well. A small group of comments indicated tensions between children of beneficiaries and non-beneficiaries, mentioning envy about school utensils, clothing, and the scholarship. Furthermore, doctors in the semi-structured interviews said that the program targeting damaged the interactions within the communities; and their comments, similarly, accounted resentment, gossip, and the unwillingness to participate in community activities from the non-beneficiaries.[21,22]

More concretely, social divisions materialized in at least three different ways within the Progresa localities. First, although community activities were a longstanding tradition, non-beneficiaries stopped participating because they felt unrecognized and because they were 'not being paid' to do the work (not receiving the program benefits). Second, the health talks were for everyone, but comments in the focus groups mentioned that non-beneficiaries did not attend because they were not invited, felt unwelcome, or did not want to go because they were non-beneficiaries. Third, parents typically payed fees to the school association and participated in the cleaning of the school facilities among other activities. These associations have no connection with the program, but non-beneficiaries were reluctant to contribute towards the school resources and no longer wanted to pay the voluntary fees, arguing that the demands on them for school supplies should be reduced. In

[19] A total of 23 focus groups were conducted involving 230 participants (beneficiaries and non-beneficiaries), representing 70 communities from 7 regions, and a total of 17 doctors and 18 secondary school principals were interviewed.

[20] Angelucci *et al.* (2010) identified the extended family network within the same locality. They found that 80 percent of the couple-headed households are embedded in an extended family network within the same locality. The average household has family ties to over five other households in the locality and there are an equal number of eligible and non-eligible in the average family.

[21] 13 out of 17 doctors interviewed said there were problems in their communities related to the program's designation of beneficiaries and non-beneficiaries.

[22] Some comments suggested also that beneficiaries and non-beneficiaries continued to get along in the same way as before.

the interviews, the school principals mentioned that beneficiaries contributed more to the school's fees, while non-beneficiaries asked beneficiaries to pay instead.[23]

In addition, beneficiary women participated together in monthly meetings, health talks, community activities, and went together to collect their program transfer. These collective activities gave them the opportunity to interact more with each other, strengthening their relations. However, the creation of this group of beneficiary women might had indirectly reinforced the social divisions between beneficiaries and non-beneficiaries in the localities.

Overall, the evidence from the focus groups and interviews describes how the targeting of the program affected the social relationships in Progresa localities and created social divisions between beneficiaries and non-beneficiaries. This points out an important unintended effect regarding program intervention, which can in turn alter spillover program effects through changes in the peer group. This issue is described and studied empirically in the following sections.

3 Social Interactions Bayesian Game

This paper uses the social interactions Bayesian game in Blume *et al.* (2015) as a framework for the empirical analysis. The main features of the model are the following. N (finite) number of agents are members of a network. An agent i's type is a vector, (x_i, x_{-i}, z_i), consisting of his public (observable) characteristic x_i, his peers' public characteristic values, x_{-i}, and his private (unobservable) characteristic, z_i. The vector of agents' types has an exogenous probability distribution known by the agents. Each agent chooses an action ω_i, a school attendance decision, to maximize the following utility function,

$$U_i(\omega_i, \omega_{-i}) = \underbrace{(\gamma x_i + z_i + \delta \sum\nolimits_j c_{ij} x_j)\omega_i - \frac{1}{2}\omega_i^2}_{\text{Private component}} - \underbrace{\frac{\phi}{2}(\omega_i - \sum\nolimits_j a_{ij}\omega_j)^2}_{\text{Social component}} \tag{1}$$

The individual utility consists of two components. First, the private component includes the effect of own individual characteristics, public and private, x_i and z_i, as well as the contextual effects, that is the effect of peers characteristics, $\sum_j c_{ij} x_j$. The second component corresponds to the social utility, which captures the endogenous social interactions (peer) effect as the squared distance between an individual i's own choices with respect to peers' choices. This social component is a cost, which is minimized whenever individual i's behavior conforms to peers' behavior. In other words, the endogenous effect comes from the pressure to conform. The term ϕ is the marginal rate of substitution (*MRS*) between social and private utility, that is, the private utility needed to compensate a marginal deviation from the social norm. Importantly, this parameter is an assessment of

[23] 8 out of 18 secondary school principals interviewed addressed the problems between beneficiaries and non-beneficiaries.

the relative importance of the social and private incentives behind individual choices. The terms a_{ij} and c_{ij} are the elements of sociomatrices A and C, which determine endogenous (peer) and contextual effects, respectively. Each sociomatrix has dimension $N \times N$, and each entry describes the strength of the relationship between the pair (i, j). These entries are known values, which are explicitly specified in the estimation section below. Importantly, the model specification allows the endogenous- and contextual-effects networks to differ. The first order conditions from the expected utility maximization is the following.[24]

$$\omega_i = \frac{\gamma}{1+\phi} x_i + \frac{\delta}{1+\phi} \sum_j c_{ij} x_j + \frac{\phi}{1+\phi} \sum_j a_{ij} \mathbb{E}(\omega_j | x) + \frac{1}{1+\phi} z_i \tag{2}$$

That is, individual i's school attendance depends on his public (observable) characteristics, peers' public characteristics, and peers' expected behavior. In particular, the term $\phi/(1+\phi)$ corresponds to the endogenous (peer) effect in the literature, which gives raise to spillover program effects. When social divisions occur in the endogenous-effects network, the peer effect might have a reduced impact or cease to exist since some of the links are broken, that is, $a_{ij} = 0$ for some i.

The static incomplete information model chosen considers social interactions as emerging from social norms. Other models in the education literature, like Calvó-Armengol, Pattachini, and Zenou (2009), consider strategic complementarities in production as the source of social interactions. In that paper, the work of other students spills over to increase the marginal product of a student's effort. Although in their model contextual- and endogenous-effects networks are the same, the estimating equation that come from the first order conditions (equation 2) can potentially be rationalized by Calvó-Armengol *et al.* (2009) model too. However, this paper focuses on school attendance rather than achievement, and taking into account the social context described in Section 2, a model that considers social pressure through conformity effects, rather than complementarities in production, as the source of endogenous social interactions is more suitable to study social disruptions and school attendance.

4 Data, Estimation Strategy, and Treatment Effects

4.1 Data and School Attendance

Progresa's short-run evaluation consists of two pre-program surveys, October 1997 (baseline, known as ENCASEH) and March 1998; and five follow-up rounds every six months between

[24] The game has a unique Bayesian Nash equilibrium with linear strategy profiles. Two axioms guarantee the existence of the equilibrium. For completeness, they are: 1) $\phi \geq 0$, A and C are non-negative matrices, such that for each i $\sum_j a_{ij}$ and $\sum_j c_{ij}$ are either 0 or 1, and for all i, $a_{ii} = c_{ii} = 0$; and 2) second moments of the prior distribution exist. The reader is referred to the paper by Blume *et al.* (2015) for the technical details.

October 1998 and November 2000 (each known as ENCEL). This paper uses the baseline and the first follow up surveys only, that is October 1997 and 1998.[25] Each survey is a locality-level census from around 22,000 households with detailed information on income, household characteristics, individual socioeconomic status, geographic location, and school attendance.

The program authorities made public the administrative records that contain the bimonthly program cash transfer given to each household since the beginning of the program.[26] The empirical analysis considers the average cash transfer of three consecutive periods, from July to December 1998. These dates match the collection dates of the evaluation survey.

The sample considered consists of children aged between 6 and 16 in grades 0 to 9 at baseline (October 1997). The dataset consists of 34,117 children, who were matched in both surveys, 21,126 children in treatment localities and 12,991 in control localities. Around 68 percent of all the children are identified as poor, and thus beneficiaries of the program. The average bimonthly transfer received by the children's household in the sample is around $146 pesos. Half of the children are girls and less than one-third are indigenous. Around 74 percent of the corresponding household's heads are literate, with 3 years of schooling on average. Table 7 in the appendix contains the descriptive statistics of the whole sample comparing treatment and control groups.

Table 3 contains the school attendance rate before (1997) and after (1998) the program for the children in the sample. At baseline, school attendance was around 85.7 and 86.7 percent for the control and treatment localities, respectively. This difference is significant, potentially as a result of the sample size.[27] After the treatment, school attendance in 1998 reflected significant differences, where higher school attendance is associated with the Progresa program. School attendance declined because children tend to dropout at older ages, however, the program was effective in reducing the dropout rates for both beneficiary and non-beneficiary children. Using a difference-in-differences approach to account for pre-program differences, the program is associated with a significant 2.8 percentage points increase in school attendance for children in the sample. These results are in line with the ones found in the official external evaluations of the program. The short-run experimental treatment effects reported there ranged from 5 to 15 percentage points depending

[25] Only the October 1998 follow up survey is used for several reasons. First, in March 1999, the program carried out a revision of the eligible households, extending the program's coverage. Although the new beneficiary population, known as *densificados*, didn't cash their transfer according to the records, their inclusion can potentially bias the results. Second, the control localities pressured the government to receive the benefits a year earlier than originally planned. Finally, March surveys are excluded because school attendance is low in the summer due to the harvest.

[26] The first transfer registered corresponds to the months of March and April 1998. However, at the beginning of the program, few households received the transfer due to implementation problems. As a consequence, many households started to receive the subsidy later, and those payments included the amount due from past periods.

[27] Behrman and Todd (1999) compared the treatment and control groups for a wide variety of indicators at baseline, and concluded that at the community level, the treatment and control groups appeared to be random. At the individual level, they found small significant differences in some characteristics, but argued that the large sample size could lead to the rejection of minor differences.

on age, gender, and school level.[28]

Table 3: School Attendance by Treatment Status

	Control		**Treatment**			(t-test)
	mean	s.d.	mean	s.d.	diff.	p-value
Individual school attendance						
Baseline 1997	0.8569	0.3502	0.8668	0.3398	-0.0099	0.0101
Post-treatment 1998	0.8140	0.3891	0.8512	0.3559	-0.0372	0.0000
Difference (1998-1997)	-0.0550	0.3163	-0.0267	0.3024	-0.0282	0.0000
Observations	12,991		21,126		Total: 34,117	

Source: ENCASEH 1997 and ENCEL 1998. The sample corresponds to children who have completed any primary or secondary school grade at baseline (October, 1997). Control and treatment refer to the locality where the children live. School attendance 1997 comes from the baseline survey, October 1997. School attendance 1998 comes from the first post-intervention round, November 1998. The item in the survey is the same for both years, *"Does your child attend to school?"*.

4.2 Estimation Strategy

Replacing the expectation in expression (2) by the observable average in the sample, a self-consistency assumption, and adjusting the expression to accommodate the features of the Progresa program give rise to the following econometric specification, which is estimated in the next section.

$$\omega_{ipc} = \alpha + \frac{\beta}{1+\phi} t_{ipc} + \frac{\gamma_0}{1+\phi} x_{ipc} + \frac{\gamma_1}{1+\phi} p_{ipc} + \frac{\gamma_2}{1+\phi} z_{ipc} + \frac{\gamma_3}{1+\phi} y_{ipc} + \frac{\delta_0}{1+\phi} \bar{x}_c + \frac{\delta_1}{1+\phi} \bar{p}_c + \frac{\delta_2}{1+\phi} \bar{z}_c + \frac{\delta_3}{1+\phi} \bar{y}_c + \frac{\phi}{1+\phi} \bar{\omega}_p + \eta_{ipc} \tag{3}$$

The variable ω_{ipc} denotes the school attendance of child i, with peer-effects network p and contextual-effects network c. t_{ipc} is a dummy that denotes the treatment status; x_{ipc} is the amount per child transfer received by i's household,[29] which is positive if the household belongs to a treated locality and it's identified as a beneficiary (poor) of the program. $\bar{x}_c$ denotes the corresponding average transfer at the contextual-effects network, which is positive only in treated localities. The dummy variable p_{ipc} denotes the beneficiary (poverty) status of the child, which is known for treated and control localities; $\bar{p}_c$ denotes the fraction of beneficiary (poor) children in the contextual-effects network; and $\bar{\omega}_p$ denotes the average school attendance in the peer-effects

[28] See Skoufias (2001) for a complete summary of the results, Schultz (2004) for the results on school enrollment, and Behrman, Sengupta, and Todd (2005) for the results on grade repetition, dropout rates, and school reentry rates among dropouts.

[29] This quantity is in Mexican pesos (divided by \$100) and includes the monetary transfer from the scholarship, school supplies, and consumption received by the household.

network, thus $\phi/(1+\phi)$ corresponds to the endogenous social interaction effect. z_{ipc} is a vector of variables correlated with the transfer, that is, school grade and gender; y_{ipc} denotes other individual characteristics used in the literature to study school attendance, such as, age, ethnicity, disabilities, household's head characteristics, the demographic structure of the household, and dwelling characteristics; and $\bar{z}_c$ and $\bar{y}_c$ are the corresponding contextual-level variables. All the averages in the contextual- and endogenous-effects networks are calculated excluding individual i since $a_{ii} = c_{ii} = 0$. Individual and peer characteristics are measured at baseline, while outcome are based on the first follow-up. The model specification is comprehensive in describing the full range of individual and family backgrounds and contextual factors that may influence children's attendance.

The rural localities in the sample are small and have at most one school facility. All children in each locality benefit from the school cleaning and painting, which are part of the community activities. Moreover, parents in the localities are actively involved in the school associations and contribute to school resources.[30] These provides evidence that contextual effects are important and that should be considered at the locality level. On the other hand, endogenous effects come from peers facing the same choices. The evidence shows that enrollment rates are higher over the early years of primary education and decline with age and grade progression (Figures 1 and 2 in the appendix). One of the factors behind this pattern is child labor. In fact, child labor is common and labor force participation increases with age (Skoufias and Parker, 2001), reflecting different household opportunity costs and benefits across children ages and grades. In other words, school attendance decisions differ across children in different grades. Therefore the endogenous-effects network is constructed at the school grade level rather than at the locality level. An endogenous-effects networks at the locality level would include schooling decisions of children at different schooling levels who are at a different stage in their career, thus facing different constraints and opportunity costs.

Social interactions model like the one in equation (3) are cumbersome to estimate. In particular, there are three main challenges: simultaneity, endogenous group membership, and correlated unobservables (Moffitt, 2001). Simultaneity arises because $i's$ action affects $j's$ action and vice versa.[31] The endogenous group formation problem arises because self-selection of individuals into groups is pervasive and unobserved to the econometrician. Finally, the correlated unobservables (common shocks) problem arises when there is a group-specific component of the error term that varies across groups and which is correlated with the exogenous characteristics of individuals. These components may represent contextual or environmental unobserved influences on school

[30]The state Ministry of Education is responsible for supplying teachers and classrooms, but other requirements, like maintenance or school supplies, are financed through parental contributions.

[31]This is known in the literature as the reflection problem (Manski, 1993).

attendance, for example, teacher's quality and motivation.

To overcome the simultaneity problem, this paper instruments the average peers' school attendance behavior. Following the logic of the model, a natural instrument arises by taking the expectation of equation (3) over children's peers (over the peer-effects network). Replacing that expectation by the average leads to the following equation and instruments (first-stage),

$$\bar{\omega}_p = \alpha(1+\phi) + \beta\bar{t}_p + \gamma_0\bar{x}_p + \gamma_1\bar{p}_p + \gamma_2\bar{z}_p + \gamma_3\bar{y}_p + \delta_0\bar{x}_c + \delta_1\bar{p}_c + \delta_2\bar{z}_c + \delta_3\bar{y}_c + \eta_p \tag{4}$$

and the corresponding reduced-form is,

$$\begin{aligned}\omega_{ipc} = \alpha(1+\phi) + \beta t_{ipc} + \frac{\gamma_0}{1+\phi}x_{ipc} + \frac{\phi\gamma_0}{1+\phi}\bar{x}_p + \delta_0\bar{x}_c + \frac{\gamma_1}{1+\phi}p_{ipc} + \frac{\phi\gamma_1}{1+\phi}\bar{p}_p + \delta_1\bar{p}_c \\ + \frac{\gamma_2}{1+\phi}z_{ipc} + \frac{\phi\gamma_2}{1+\phi}\bar{z}_p + \delta_2\bar{z}_c + \frac{\gamma_3}{1+\phi}y_{ipc} + \frac{\phi\gamma_3}{1+\phi}\bar{y}_p + \delta_3\bar{y}_c + \tilde{\eta}_{ipc}\end{aligned} \tag{5}$$

where $\bar{x}_p$ corresponds to the average cash transfer in the peer-effects network; $\bar{p}_p$ is the fraction of beneficiaries, and $\bar{z}_p$ and $\bar{y}_p$ are averages of individual characteristics in that network. Table 9 in the appendix shows the summary statistics of these variables.[32]

The endogenous effect in the structural model (3) is identified if equation (4) contains at least one exogenous variable with non-zero coefficient which is excluded from (3). In this case, the average cash transfer in the peer-effects network, $\bar{x}_p$, is a relevant and valid instrument because it is a predictor of peer group average school attendance behavior (if $\gamma_0 \neq 0$) and satisfies the exclusion restriction because of the exogenous nature of the cash transfer, hence uncorrelated with the error term in equation (3). Notice that the average peers' characteristics in equation (4), $\bar{P}_p$, $\bar{z}_p$, and $\bar{y}_p$, can potentially serve as instruments because the peer- and contextual-effects networks differ. However, these variables might be correlated with unobserved determinants of school participation. Importantly, the identification strategy does not rely on them. Identification of the model depends on the distinct configurations of the endogenous- and contextual-effects networks, which is justified given the context discussed before. This identification strategy is based on the results in Blume *et al.* (2015).[33,34]

[32]The variable t_{ipc} equals one for children in treated localities and zero in control localities. The variable $\bar{t}_p$ is the average of this variable considering the peer-effects network and thus it also equals one for children in treated localities and zero in control localities. From (4) to (5) $\bar{t}_p$ is replaced by t_{ipc}.

[33]In the *linear-in-means* model individuals are partitioned into non-overlapping groups, and both peer- and contextual-effects networks are the same. Identification of this model can be achieved by exploiting the partial-population feature of Progresa intervention. This strategy is followed by Lalive and Cattaneo (2009) and Bobonis and Finan (2009). The appendix in the latter shows how identification works in this case.

[34]Although not entirely similar, the identification strategy is related to that provided by Bramoullé, Djebbari, and Fortin (2009). They discuss identification of peer effects considering an extended version of the *linear-in-means model* where each individual has his own specific reference group. They show that identification of endogenous and

Another important issue is that peers might self-select into groups in ways unobserved to the econometrician. In the case of Progresa, for example, self-selection can arise if some children repeat a school grade to spend more time with their peers who have failed. Behrman, Sengupta, and Todd (2005) find that grade repetition is common. Boys tend to lag behind the standard grade progression rate due to grade repetition because attendance at school was not sufficiently regular to fulfill the requirements for completing the grade. Nonetheless, the endogenous effect is identified by the program transfer because the average transfer in the peer-effects network affecting peers' attendance is exogenous and thus uncorrelated with any other variables at the peer-effects network level and the self-selection. The appendix shows formally how identification works in this case.

Regarding the correlated unobservables problem, in the Progresa localities, schools were generally poorly equipped, so in coordination with the program, schools were to receive additional resources to match the increased demand. However, the general view of the school principals in the semi-structured interviews was that the expected increase in resources did not happen, at least it didn't happen in the first years of the program, mitigating potential group level unobservables from the supply side (Adato *et al.*, 2002). Nonetheless, to address any other potential problem, the panel feature of the Progresa data is used to also estimate a first difference version of the model, which removes the time invariant unobservable components. This approach removes the contextual effects because they are all measured at baseline and are time invariant.

Finally, a recent understudied source of downward bias in the estimation of peer effects is exclusion bias (Caeyers and Fafchamps, 2020). This negative bias arises because the exclusion of i from the pool of i's peers creates a small sample negative relationship between i's characteristics and that of his peers, when fixed effects are included at the level of selection pool. In the present context, exclusion bias is not a concern because the use of instrumental variables can eliminate it. Specifically, to remove the bias it is necessary to control for i's own value of the instrument, that is, in this case x_{ipc} in equation (3).[35]

4.3 Treatment Effects and Social Interactions

One particular feature of the Progresa program is that it benefits only a subset of the population, those identified as poor. This kind of interventions is known as a partial population intervention (Moffitt, 2001), which alter the private incentives for a subset of a group, with or without the explicit intention of affecting the entire group. Exploiting this feature of the data and using the behavioral model to provide a framework for interpretation, it is possible to decompose the overall program treatment effect and show the potential effect of social disruptions.

exogenous (contextual) effects is related to properties of the network (sociomatrices). Blume *et al.* (2015) provide a generalization of their results by allowing for distinct social structures for contextual and endogenous effects.

[35]For further details, see appendix A.3.2 in Caeyers and Fafchamps (2020).

First, using the law of total probability and the experimental design of the program, the average treatment effect can be decomposed as follows (steps are in the appendix),

$$ATE \equiv \lambda = \lambda^B \Pr(P=1) + \lambda^{NB} \Pr(P=0) \tag{6}$$

The average treatment effect is a weighted average of the treatment effect on the beneficiary (poor), λ^B, and the treatment effect on the non-beneficiary (non-poor), λ^{NB}; and the weight is given by the fraction of beneficiaries (poor), $\Pr(P=1)$. Angelucci and De Giorgi (2009) define an average indirect effect as the effect of the treatment in the subset of the population that were not meant to be treated, that is, λ^{NB} here. In the case of Progresa, this subpopulation is identified as the non-beneficiary (non-poor) within the treated localities. In a subsequent paper, Angelucci and Di Maro (2016) label this effect as a spillover effect.

Second, considering the reduced-form in equation (5), the treatment effect on beneficiaries, non-beneficiaries, and overall can be expressed as,

Treatment effect on the beneficiaries (poor)

$$\begin{aligned}\lambda^B &= \mathbb{E}(\omega_1|T=1,P=1) - \mathbb{E}(\omega_0|T=0,P=1)\\ &= \beta + \frac{\gamma_0}{1+\phi}\mathbb{E}(x|T=1,P=1) + \frac{\phi\gamma_0}{1+\phi}\mathbb{E}(\bar{x}_p|T=1,P=1) + \delta_0\mathbb{E}(\bar{x}_c|T=1,P=1)\end{aligned} \tag{7}$$

Treatment effect on the non-beneficiaries (non-poor)

$$\begin{aligned}\lambda^{NB} &= \mathbb{E}(\omega_1|T=1,P=0) - \mathbb{E}(\omega_0|T=0,P=0)\\ &= \beta + \frac{\phi\gamma_0}{1+\phi}\mathbb{E}(\bar{x}_p|T=1,P=0) + \delta_0\mathbb{E}(\bar{x}_c|T=1,P=0)\end{aligned} \tag{8}$$

Overall program treatment effect

$$\lambda = \beta + \underbrace{\frac{\gamma_0}{1+\phi}\mathbb{E}(x|T=1,P=1)\Pr(P=1)}_{\text{Direct effect}} + \underbrace{\frac{\phi\gamma_0}{1+\phi}\mathbb{E}(\bar{x}_p|T=1) + \delta_0\mathbb{E}(\bar{x}_c|T=1)}_{\text{Spillover program effect}} \tag{9}$$

The equation shows that the treatment effect on the beneficiaries is composed of a direct effect from the receipt of the transfer, an endogenous effect that comes from peers' behavior, and a contextual effect from the transfer at the locality level. Notably, the treatment effect on the non-beneficiaries comes exclusively from the endogenous and contextual effects, since they did not receive any transfer. So, without this peer effect there would be no spillover effect. This implies that the overall program treatment effect is composed of a direct effect on the beneficiaries (poor) and a spillover effect from both beneficiaries and non-beneficiaries; and in particular, this spillover effect is produced by both the endogenous and contextual effects of the transfer. This shows how social interactions can amplify program effects through the complementarities that arise from social interactions.[36] The decomposition shows that the program is not only spilling over the non-beneficiaries, but also on the beneficiary population through the endogenous effect. When the distinction introduced by the program targeting between beneficiaries and non-beneficiaries results in social division that disrupt the structure of the endogenous-effects network, the endogenous spillover effect on the non-beneficiaries not longer exists because beneficiaries now belong to a separate reference group and in that case $\mathbb{E}(\bar{x}_p|T = 1, P = 0) = 0$. In this case, the endogenous spillover effect would be the corresponding one in equation (7). Finally, β captures pre-program differences between treated and control localities. In the results, this coefficient is not significant, reflecting the quality of the randomization.

Equation (5) shows that not only the treatment received by an agent affects his own outcome, but his response also depends on the treatment assigned to his contextual- and endogenous-effects networks. In this sense, the above decomposition relates to the literature on treatment response in a setting with interference.[37]

5 Empirical Results

5.1 Benchmark Model

As described in Section 4.2, the benchmark specification of the model considers contextual effects at the locality level and endogenous effects at the school grade level, more concretely, the latter network is comprised by children who have reached the same grade level at baseline and live in the same locality.[38] Children in this network have reached the same schooling level and are at the

[36]In contrast to Arduini, Patacchini and Rainone (2019), this paper considers and makes explicit the potential effect of the transfer that comes from the contextual-effects network in the treatment decomposition. In the social context of Progresa localities, where parents make contributions for school resources and take school and educational decisions, this contextual effect is relevant.

[37]Manski (2013) studied identification of treatment response in a setting with social interactions by defining the treatment response to be a function of the entire vector of treatments received by the population. Sobel (2006) proposed causal estimands for assessing housing voucher effects when the household's decision to move or not to move may be affected by whether their neighbors receive a housing voucher to move.

[38]The elements of the contextual-effects network, C, are assumed to be $c_{ij} = \frac{1}{n_l - 1}$ if the pair of children (i, j) live in the same locality, and zero otherwise; where n_l is the number of children in the locality and $c_{ii} = 0$. The elements of the peer-effects network, A, are assumed to be $a_{ij} = \frac{1}{n_g - 1}$ if the pair of children (i, j) reached the same grade and live in the same locality, and is equal to zero otherwise; n_g is the number of children in the corresponding grade. The entry $a_{ii} = 0$ in the sociomatrix A because social utility captures the effect of peers' behavior and not own behavior. This unweighted scheme reflects the lack of available information with respect to specific ties among the children in

same stage in their career. Figure 3 in the appendix shows an example of this networks.

Table 4 contains the estimates of equation (3) considering this benchmark specification. Column 1 in table 4 reports the OLS estimate, which is presumably upward biased because it does not address the endogeneity of peers' behavior and the reflection bias magnifies the peer effect. In this case, the estimated endogenous effect is $\phi/(1+\phi) = 0.3961$.

Following the logic of the model, the average transfer and the average individual characteristics at the grade level (endogenous-effects network) can be used to instrument peers' behavior and overcome the endogeneity problem, equation (4). Column 2 and 3 in table 4 report the corresponding IV estimates considering the whole sample of children (treated and control) and the sample of children in treated localities, respectively. Looking at the treated sample, the magnitude of the estimated endogenous effect is 0.2913, which means that a 10 percent increase in peers' attendance increases the likelihood of attending by 2.9 percent.[39] Considering the utility function, equation (1), this estimate translates to a marginal rate of substitution (*MRS*) between social and private utility of order $\phi = 0.4110$. That is, private utility should increase by 0.41 to account for the marginal cost incurred from deviating from the social norm.

It is important to recognize the presence of potential group level unobservables in the error term of equation (3). To account for this unobserved heterogeneity, column 4 exploits the panel feature of the data and considers the difference between pre- and post-program school attendance. In this case, the average difference in peers' attendance replaces the average peers' behavior in equation (3). The estimated endogenous effect is 0.1828, but it has a larger standard error. It is worth noting that the contextual effects are no longer significant as expected because they are not affected by the program in the short-run and are measured at baseline, that is, these effects were removed by the first difference operator. For this reason, differencing is not entirely advisable, but provides a robustness check on the estimated endogenous effect. Considering the sample size, the preferred specification is that in column 2. The evidence is consistent with a positive endogenous effect on school attendance of 0.2638 and *MRS* between private and social utility of 0.3583.

In all regressions, the transfer of the program has a positive effect on school attendance. For example, considering the whole sample, the effect of the transfer is 0.0282, which means that for every \$100 pesos ($\approx$ \$10 dollars in 1998), school attendance increases by 2.8 percentage point. The average transfer in the sample is around \$146, thus the estimated program treatment effect is around 4.2 percentage points. The locality treatment dummy is not significant and thus consistent with no important pre-program differences, reflecting the quality of the randomization. Being identified as poor according to the program selection has a negative effect, supporting the hypothesis that

the locality.

[39]This estimate is consistent with other results in the literature. Lalive and Cattaneo (2009) found a peer effect of 0.505 for primary school, and Bobonis and Finan (2009) found an effect of 0.649 for secondary school.

lack of resources is associated with low school attendance. Furthermore, the average education of the household's heads and the average preschoolers in the locality are relevant contextual-effects variables. Higher years of schooling among the heads in the locality have a positive influence on school attendance, while the presence of preschoolers reduces school attendance. Girls, for example, tend to take care of the small children, the elderly, and the sick.

Although there is no test for instrument validity, the test statistics in the bottom part of table 4 provide confidence to the results. Considering the Hansen-J statistic, the null hypothesis that the instruments are uncorrelated with the error term is not rejected in all cases (valid instruments). Using the Lagrange Multiplier Kleibergen and Paap rk statistic, the null hypothesis that the endogenous variable and the excluded instruments are uncorrelated is rejected in column 2 and 3 (relevant instruments). Finally, weak instruments are not a main concern. Under homoskedastic errors, the Cragg-Donald Wald F statistics reported are over 28.1 in column 2 and 3. Considering a robust measure to *non-i.i.d.* errors, the Kleibergen and Paap rk Wald statistic is also reported.[40]

[40] As a rule of thumb, weak instruments are a problem if the F-statistic is less than 10. There are no critical values for the *non-i.i.d.* case.

Table 4: Benchmark Model. OLS and IV Estimates of Contextual- and Endogenous-Effects

	(1) OLS		Instrumental Variables (2) All sample		(3) Treated sample		(4) DiD	
First stage								
Transfer ($)			0.0304***	(0.0069)	0.0378***	(0.0073)	0.0148**	(0.0059)
Poor			-0.0345**	(0.0175)	-0.0716***	(0.0232)	-0.0017	(0.0153)
Male			0.0144	(0.0119)	0.0202	(0.0143)	-0.0097	(0.0111)
Other			Yes		Yes		Yes	
Second stage								
Treatment	0.0043	(0.0073)	0.0037	(0.0086)	-	-	0.0046	(0.0075)
Transfer ($)	0.0282***	(0.0021)	0.0282***	(0.0021)	0.0313***	(0.0024)	0.0143***	(0.002)
Poor	-0.0223***	(0.0055)	-0.0225***	(0.0055)	-0.0341***	(0.0078)	0.0010	(0.0055)
Male	0.0166***	(0.0036)	0.0166***	(0.0036)	0.0169***	(0.0046)	0.0017	(0.0034)
Other	Yes		Yes		Yes		Yes	
Endogenous effect	0.3961***	(0.0198)	0.2638***	(0.0976)	0.2913***	(0.1056)	0.1828	(0.2323)
MRS (ϕ)	0.6558		0.3583		0.4110		0.2237	
Contextual effects (means and proportions at the locality level)								
Transfer ($)	-0.0115**	(0.0055)	-0.0070	(0.0068)	0.0051	(0.0086)	-0.0001	(0.0067)
Poor	0.0016	(0.0170)	-0.0091	(0.0210)	0.0271	(0.0276)	-0.0087	(0.0189)
Male	0.0022	(0.0308)	0.0094	(0.0366)	-0.0412	(0.0472)	-0.0059	(0.0336)
Indigenous	-0.0056	(0.0226)	-0.0113	(0.0262)	0.0208	(0.0326)	-0.0497*	(0.0266)
Disable	0.0787**	(0.0377)	0.0913**	(0.0450)	0.0684	(0.0607)	0.0025	(0.0403)
Head indigenous	0.0198	(0.0226)	0.0248	(0.0264)	-0.0043	(0.0310)	0.0539**	(0.0263)
Head literate	-0.0535**	(0.0213)	-0.0631**	(0.0259)	-0.0734**	(0.0361)	-0.0263	(0.0225)
Head's schooling	0.0065*	(0.0036)	0.0093*	(0.0049)	0.0146**	(0.0067)	0.0005	(0.0037)
Head agr. worker	0.0071	(0.0155)	0.0089	(0.0182)	-0.0655***	(0.0228)	-0.0062	(0.0183)
Head ejidatario	-0.0015	(0.0185)	-0.0020	(0.0215)	-0.0783**	(0.0334)	-0.0037	(0.0207)
Head married	-0.0329**	(0.0154)	-0.0379**	(0.0183)	-0.0405*	(0.0224)	0.0171	(0.0174)
Preschoolers (0-5)	-0.0157**	(0.007)	-0.0187**	(0.0084)	-0.0215**	(0.0108)	0.0048	(0.0076)
School-agers (6-17)	0.0095*	(0.0056)	0.0108	(0.0067)	-0.0038	(0.0081)	0.0022	(0.0055)
Young adults (18-44)	-0.0026	(0.0086)	-0.0035	(0.0101)	-0.0182	(0.0129)	-0.0037	(0.0097)
Adults (45-64)	0.0224**	(0.0114)	0.0273**	(0.0135)	0.0118	(0.0176)	0.0265*	(0.0136)
Seniors (65+)	0.0277	(0.0222)	0.0319	(0.0267)	0.0098	(0.0368)	0.0176	(0.0232)
Hansen-J $\chi^2_{(11)}$			14.91	[0.187]	16.90	[0.111]	16.63	[0.119]
KP rk LM $\chi^2_{(12)}$			69.43	[0.000]	59.64	[0.000]	15.70	[0.206]
KP Wald rk $F_{(12)}$			6.03		5.26		1.31	
Cragg-D Wald F			34.16		28.12		7.45	
Observations	31,182		31,177		19,300		30,969	

Source: ENCASEH 1997 and ENCEL 1998. Coefficient estimates from IV regressions using GMM estimator. Standard errors in parentheses are clustered at the locality-grade level. State and municipalities effects are included but not reported. Other characteristics are: age, school level, ethnicity, disabilities, age, gender and years of education of the household's head, household demographic structure, dwelling characteristics, presence of secondary school facility, community activities, and school parent association. There are 12 excluded instruments in the first stage, which correspond to peers' average of individual characteristics (see equation 4). The contextual-effects network corresponds to all children in the locality. The peer effects network considers children who have reached the same grade level and live in the same locality. Ejidatario means his a shareholder of common land. Hansen's J statistic is used for testing overidentifying restrictions, allowing observations to be correlated within groups. The null hypothesis is that instruments are uncorrelated with the error term (valid instruments), and that excluded instruments are correctly excluded from the estimated equation. Kleibergen-Paap rk LM is used for testing underidentification, whether the excluded instruments are correlated with the endogenous regressor (relevant instruments) when errors are non-iid. The null hypothesis is that the equation is underidentified. Kleibergen-Paap Wald rk F is used for testing weak identification, whether the excluded instruments are weakly correlated with the endogenous regressor when errors are non-iid. Cragg-Donald Wald F test for the iid errors weak identification case is included. P-values in brackets. Significantly different than zero at 99 (***), 95 (**), and 90 (*) percent confidence.

5.2 Beneficiary Status as Endogenous- and Contextual-Effects Networks

Given both contextual- and endogenous-effects networks, the utility function in equation (1) describes how private and social incentives affect school attendance decisions. In particular, the social utility component captures the cost of deviating from peers' behavior and the parameter ϕ is the marginal rate of substitution (*MRS*) between social and private utility, while $\phi/(1+\phi)$ is the endogenous effect. Relative to another network structure, a higher *MRS* means that the preference to conform is stronger under that network, and thus deviating from peers' behavior (social norm) is costlier, other things equal. This behavioral interpretation of the parameters and the flexibility of the sociomatrices permit the assessment of the relative importance of different networks structures in the data, such as the one implied by social divisions due to program targeting.

As mentioned in Section 2 the qualitative evidence showed that the distinction introduced by the program between beneficiaries and non-beneficiaries led to social tensions and social divisions that affected their interactions. To study empirically whether social divisions appeared in the localities and learn about their effect on the complementarities arising from social interactions and on the program impact, the contextual- and endogenous-effects networks are modified to account for the social disruptions.

Table 5 contains the main contribution of the empirical analysis using the sample of children in treated localities.[41] To assess whether social divisions manifested in the endogenous-effects network, this network is partitioned by beneficiary status. That is, it considers the beneficiaries (poor) and non-beneficiaries (non-poor) as two separate groups that make up their own endogenous-effects network. To maintain the comparability with the benchmark model, the contextual-effects network is unchanged at the locality level. Figure 4 in the appendix shows an example of such networks and table 5 reports the corresponding estimates. Compared to the benchmark model (column 3 in table 4), the magnitude of the endogenous effect decreases from a significant 0.2913 to an insignificant 0.0845 (column 1 in table 5), and the corresponding *MRS* changes from 0.4110 to 0.0923. This means that the beneficiary-based network configuration does not entail any social pressure to conform to peers' behavior and thus the social utility becomes irrelevant under this situation. If this network configuration was relevant, a positive and significant peer effect would have been found as evidence of the complementarities that arise from such network structure. However, the results do not support the case that beneficiaries and non-beneficiaries children each conform to their peers' behavior. In other words, the data do not support the case where social divisions manifested within the classroom between the children of beneficiary and non-beneficiary families.

An important thing to notice is that even if the program effectively created social divisions

[41] It is in the treated localities where the hypothesized distinction between beneficiaries and non-beneficiaries is supposed to be meaningful. In the control localities, although the program identified beneficiaries and non-beneficiaries, no treatment was offered, and thus there were no distinctions between these two groups.

between beneficiaries and non-beneficiaries, the distinction introduced by the program cannot alter the relationships among the children if only a small fraction of them were beneficiaries. For instance, in the extreme case, if only one child is a beneficiary of the program in the peer group, then there should be no important disruption of the interactions within this group. Similarly, in those peer groups where most of the children were beneficiaries. If the program treats everyone in the peer group, then there are no social distinctions introduced by the program. In these cases, the estimated endogenous effects and *MRS* should be similar to those found under the benchmark model. This strategy is used to further investigate the previous result and it provides a robustness check. To operationalize it, the sample is partitioned into four quartiles according to the share of beneficiaries (poor) children in the peer group.

Column 2 in table 5 reports the estimates considering the sample where no more than one-quarter of the children in the peer group received the benefits. The magnitude of the endogenous effect is 0.3743 and the corresponding *MRS* is 0.5981, slightly larger than the estimate obtained using the benchmark model. Similarly, the partition that considers between one-quarter and one-half of the children as beneficiaries, column 3, produces a similar endogenous effect and *MRS*, the magnitudes are 0.3191 and 0.4686, respectively. Considering those where the share of beneficiaries is between one-half and three-quarters, column 4, the endogenous effect is not significant.[42] The latter peer groups are the ones in which the program is thought to alter the most the social relations. The results are consistent with the previous one, that is, a configuration that accounts for social divisions by beneficiary status does not lead to any pressure to conform. Finally, in the peer groups where the program targeted more than three-quarters of the children, the magnitude of the endogenous effect is an insignificant 0.2411 with a *MRS* of 0.3178. A possible explanation for this lower endogenous effect is that the program might had brought back students who were out of school in these groups since a large fraction received the benefits, thus changing the network structure and peer's willingness to conform. Overall, the evidence is consistent with peer effects at the school-grade level rather than by beneficiary status within the school-grade.[43]

[42]Also notice that one of the requirements for the existence of the equilibrium in the game is $\phi \geq 0$.

[43]Table 10 in the appendix replicates this exercise considering five quantiles instead of four quartiles. The results do not change. The endogenous effect is significant, except for the case when 60 to 80 percent of the children are beneficiaries.

Table 5: Beneficiary Status as Endogenous-Effects Network. IV Estimates of Contextual- and Endogenous-Effects

			Partition by share of beneficiary children in the grade							
	(1) Treated sample		**(2) [0-25%]**		**(3) (25-50%]**		**(4) (50-75%]**		**(5) (75-100%]**	
First stage										
Transfer ($)	0.0247***	(0.0048)	-0.0176	(0.0420)	0.0291***	(0.0105)	0.0204***	(0.007)	0.0265***	(0.0085)
Other	Yes		Yes		Yes		Yes		Yes	
Second stage										
Transfer ($)	0.0314***	(0.0024)	0.0229	(0.0311)	0.0291***	(0.0066)	0.03***	(0.0046)	0.0324***	(0.0032)
Endog. effect	0.0845	(0.1085)	0.3743*	(0.1944)	0.3191*	(0.1855)	-0.2602	(0.2632)	0.2411	(0.1609)
MRS (ϕ)	0.0923		0.5981		0.4686		-0.2065		0.3178	
Contextual effects (locality)										
Transfer ($)	0.0157	(0.0108)	0.0504	(0.0823)	-0.0112	(0.0307)	-0.0003	(0.0292)	0.0089	(0.0139)
Other	Yes		Yes		Yes		Yes		Yes	
Hansen-J $\chi^2_{(10)}$	17.56	[0.063]	na		na		na		na	
KP rk LM $\chi^2_{(11)}$	76.96	[0.000]	23.25	[0.016]	21.34	[0.030]	23.20	[0.017]	29.35	[0.002]
KP Wald rk $F_{(11)}$	7.43		2.16		1.98		2.37		2.82	
Cragg-D Wald F	23.80		5.27		5.26		5.35		13.99	
Observations	18,505		1,276		3,333		5,071		8,825	

Source: ENCASEH 1997 and ENCEL 1998. Coefficient estimates from IV regressions using GMM estimator. Standard errors in parentheses are clustered at the locality-grade level. State and municipalities effects are included. Other characteristics are: age, school level, ethnicity, disabilities, age, gender and years of education of the household's head, household demographic structure, dwelling characteristics, presence of secondary school facility, community activities, and school parent association. There are 12 excluded instruments in the first stage, which correspond to peers' average of individual characteristics (see equation 4). The contextual-effects network corresponds to all children in the locality and the peer-effects network considers children who have reached the same grade level and live in the same locality by beneficiary status, beneficiary (poor) and non-beneficiary (non-poor). Hansen-J test is omitted due to the small number of clusters. Hansen's J statistic is used for testing overidentifying restrictions, allowing observations to be correlated within groups. The null hypothesis is that instruments are uncorrelated with the error term (valid instruments), and that excluded instruments are correctly excluded from the estimated equation. Kleibergen-Paap rk LM is used for testing underidentification, whether the excluded instruments are correlated with the endogenous regressor (relevant instruments) when errors are non-iid. The null hypothesis is that the equation is underidentified. Kleibergen-Paap Wald rk F is used for testing weak identification, whether the excluded instruments are weakly correlated with the endogenous regressor when errors are non-iid. Cragg-Donald Wald F test for the iid errors weak identification case is included. P-values in brackets. Significantly different than zero at 99 (***), 95 (**), and 90 (*) percent confidence.

The evidence from the focus group and semi-structured interviews described how social tensions and divisions occurred in the localities. Importantly, the division between beneficiary and non-beneficiary families at the locality level cannot alter the endogenous effect, because this effect works exclusively through the endogenous-effects network. This highlights the important distinction between endogenous- and contextual-effects networks. When considering a model specification where the contextual-effects network is partitioned between beneficiaries and non-beneficiaries, keeping the endogenous-effects network at the grade level (benchmark), the endogenous effect remains largely unchanged. Column 1 in table 6 contains the estimates using this network configuration, and shows an endogenous effect of 0.3489 and a corresponding *MRS* of 0.5359. This result is consistent with the previous ones, implying that it is costlier to deviate from the social norm.[44] Notably, the average education level of the household's heads is no longer an

[44]Although these endogenous effects are higher in magnitude than the benchmark model, they are not statistically different.

important contextual effect. It is worth noting that the contextual effect of the transfer is negative and significant in column 1 and 4, offsetting school attendance and reflecting potential non-linear effects of the treatment.

When considering a partition of the sample by the fraction of beneficiary (poor) children in the locality, the endogenous effect is positive and significant in almost all the cases. For example, taking into account the sample where no more than one-quarter of the children in the locality were beneficiaries (column 2 in table 6), the endogenous effect is 0.4424 and the *MRS* is 0.7933. Moreover, locality-level demographics remain as important contextual effects. Considering the partitions with bigger shares of beneficiary children in the locality, for example, between one-quarter and three-quarters, columns 3 and 4, the contextual effects tend to disappear. This suggests that the contextual effects from a configuration that accounts for beneficiary status are no longer significant. In other words, the evidence is consistent with contextual effects happening at the locality level in the Progresa localities.

In summary, although the qualitative evaluation described how social divisions manifested in Progresa localities between beneficiaries and non-beneficiaries due to the program targeting, the empirical evidence based on a behavioral model of social interactions shows that the distinction introduced by the program didn't alter the endogenous-effects network and the complementarities in schooling choices. In other words, social divisions did not alter the complementarities that lead to a positive spillover effect among non-beneficiary children.

Table 6: Beneficiary Status as Contextual-Effects Network. IV Estimates of Contextual- and Endogenous-Effects

			Partition by share of beneficiary children in the locality							
	(1) Treated sample		**(2) [0-25%]**		**(3) (25-50%]**		**(4) (50-75%]**		**(5) (75-100%]**	
First stage										
Transfer ($)	0.0367***	(0.0061)	0.0544	(0.0461)	0.0621***	(0.0217)	0.0364***	(0.0116)	0.0315***	(0.0083)
Other	Yes		Yes		Yes		Yes		Yes	
Second stage										
Transfer ($)	0.0339***	(0.0025)	0.0659***	(0.0242)	0.0296***	(0.0085)	0.0323***	(0.0043)	0.0343***	(0.0032)
Endog. effect	0.3489***	(0.0959)	0.4424*	(0.2349)	0.5131***	(0.1246)	0.4495***	(0.1393)	0.2137	(0.1598)
MRS (ϕ)	0.5359		0.7933		1.054		0.8165		0.2717	
Contextual effects (locality by beneficiary status)										
Transfer ($)	-0.0151**	(0.0062)	0.3341	(0.2213)	-0.0238	(0.0223)	-0.0336***	(0.0112)	-0.0142	(0.0101)
Male	-0.021	(0.0281)	0.8503**	(0.3444)	-0.0169	(0.0744)	-0.053	(0.0545)	0.0583	(0.056)
Indigenous	-0.0194	(0.0282)	1.8189	(3.9247)	-0.1205	(0.0967)	-0.1159*	(0.0664)	0.0209	(0.0589)
Disable	0.0119	(0.0442)	0.4433	(1.0049)	0.0431	(0.1326)	-0.0072	(0.087)	-0.0156	(0.0889)
Head indigenous	0.0128	(0.0264)	5.0695	(5.0516)	-0.0105	(0.0718)	0.0171	(0.0587)	0.0735	(0.0656)
Head literate	-0.0321	(0.0244)	-0.3148	(0.7936)	0.0692	(0.0662)	-0.0253	(0.0432)	-0.0568	(0.0413)
Head's schooling	0.0024	(0.0041)	0.0552	(0.1818)	-0.0154	(0.0109)	-0.0073	(0.0069)	0.0134*	(0.0071)
Head agr. worker	-0.0252	(0.0164)	1.0887*	(0.6014)	-0.0386	(0.048)	0.0491	(0.0342)	-0.048	(0.0347)
Head ejidatario	-0.054**	(0.0248)	0.3105	(1.5501)	-0.0239	(0.0641)	-0.0053	(0.0464)	-0.0962**	(0.0483)
Head married	-0.0362**	(0.0173)	-1.4982***	(0.4909)	-0.0381	(0.051)	-0.0783**	(0.0344)	-0.0114	(0.0357)
Preschoolers (0-5)	-0.0155*	(0.0082)	0.5516*	(0.291)	-0.0478**	(0.0217)	-0.0156	(0.0176)	-0.0214	(0.0192)
School-agers (6-17)	0.001	(0.0053)	-0.1932**	(0.0962)	-0.0173	(0.0158)	0.0106	(0.0099)	-0.0072	(0.0099)
Young adults (18-44)	-0.0108	(0.0084)	0.7217**	(0.2823)	-0.0028	(0.0218)	-0.0105	(0.0142)	-0.0411**	(0.0177)
Adults (45-64)	0.02*	(0.0119)	0.5253	(0.3645)	0.0003	(0.0319)	0.0416*	(0.0226)	-0.0049	(0.02)
Seniors (65+)	-0.001	(0.023)	2.3203*	(1.4088)	-0.1023*	(0.0572)	-0.0086	(0.0368)	0.043	(0.0456)
Hansen-J $\chi^2_{(10)}$	9.67	[0.561]	na		na		6.79	[0.816]	18.59	[0.069]
KP rk LM $\chi^2_{(11)}$	71.1	[0.000]	16.42	[0.173]	34.61	[0.001]	32.02	[0.001]	30.13	[0.003]
KP Wald rk $F_{(11)}$	6.35		1.6		3.28		2.75		2.8	
Cragg-D Wald F	31.95		4.83		14.52		12.23		16.72	
Observations	19,285		622		2,911		6,635		9,117	

Source: ENCASEH 1997 and ENCEL 1998. Coefficient estimates from IV regressions using GMM estimator. Standard errors in parentheses are clustered at the locality-grade level. State and municipalities effects are included. Other characteristics are: age, school level, ethnicity, disabilities, age, gender and years of education of the household's head, household demographic structure, dwelling characteristics, presence of secondary school facility, community activities, and school parent association. There are 12 excluded instruments in the first stage, which correspond to peers' average of individual characteristics (see equation 4). The contextual-effects network corresponds to all children in the locality by beneficiary status, beneficiary (poor) and non-beneficiary (non-poor. The peer-effects network considers children who have reached the same grade level and live in the same locality. Hansen-J test is omitted due to the small number of clusters. Hansen's J statistic is used for testing overidentifying restrictions, allowing observations to be correlated within groups. The null hypothesis is that instruments are uncorrelated with the error term (valid instruments), and that excluded instruments are correctly excluded from the estimated equation. Kleibergen-Paap rk LM is used for testing underidentification, whether the excluded instruments are correlated with the endogenous regressor (relevant instruments) when errors are non-iid. The null hypothesis is that the equation is underidentified. Kleibergen-Paap Wald rk F is used for testing weak identification, whether the excluded instruments are weakly correlated with the endogenous regressor when errors are non-iid. Cragg-Donald Wald F test for the iid errors weak identification case is included. P-values in brackets.Significantly different than zero at 99 (***), 95 (**), and 90 (*) percent confidence.

5.3 Other Endogenous-Effects Networks

Extending the previous analysis, this section follows the same strategy to study the conforming behavior of three particular social groups typically defined in terms of exogenous characteristics, namely, gender, ethnicity (indigenous and non-indigenous), and landownership. First, the literature on peer effects in education has shown that the gender composition of peer-effects network plays an important role in academic achievement and career choices. For instance, Hoxby (2000) finds that a higher share of girls in one's cohort raises achievement levels both through the impact of having higher-scoring peers and also by changing the culture in the class. Lavy and Schlosser (2011) mention that a higher proportion of girls in the classroom lower the level of disruption and violence, and improves inter-student and teacher student relationships. Second, regarding ethnicity, Mexico has one of the largest indigenous population, over 7 million in 2015, and a large share is concentrated in the states where Progresa took place. Indigenous people are the first inhabitants who were in the continent before the arrival of the Europeans. They not only speak a different language, but also have their own set of political and social institutions, customs, folklore traditions, and religious beliefs. These distinctions make ethnicity an important reference group to explore for potential peer effects. Finally, the distribution of land is among the main determinants of the rural income distribution. Importantly, land is not only a source of income that reflect socioeconomic condition, but it also indicates social status and class in rural Mexico (Stinchcombe, 1961). In the context of peer effects and schooling decisions, considering landownership as a potential reference group is important because the social and economic differences created by land holdings can translate into meaningful reference groups. The empirical analysis considers landless households and landowners.[45]

Again, to maintain comparability with the benchmark model, the contextual-effects network is held fixed at the locality level and the three different reference groups are considered as endogenous-networks. That is, the benchmark peer-effects network is partitioned by each of the three groups mentioned above. Table 7 reports the estimates considering these three different network structures. The endogenous effects considering ethnicity and gender are 0.2019 and 0.2217 with a *MRS* of 0.253 and 0.2849, respectively, both lower than the benchmark model. The estimate considering landownership is small and not significant. These results suggest that endogenous social interactions are stronger at the grade level rather than by gender, ethnicity, or socioeconomic status within the school grades. In terms of policy, other things equal, there is nothing to be gained by distinguishing by gender or ethnicity. Given the size of the localities, the characteristics of the school facilities, and the number of students, there is typically at most one classroom per grade, and thus the results provide further evidence that peer effects in education are best studied at the classroom level.

[45]Landless people have only their labor supply, while landowners tend to produce for self-consumption and sometimes for profits.

Table 7: Other Endogenous-Effects Networks. IV Estimates of Contextual- and Endogenous-Effects

	(1) Ethnicity		(2) Landownership		(3) Gender	
First stage						
Transfer ($)	0.0315***	(0.0063)	0.0279***	(0.0046)	0.0307***	(0.0044)
Other	Yes		Yes		Yes	
Second stage						
Transfer ($)	0.0280***	(0.0021)	0.0279***	(0.0022)	0.0282***	(0.0022)
Endog. effect	0.2019**	(0.0931)	0.0932	(0.0900)	0.2217***	(0.0802)
MRS (ϕ)	0.2530		0.1027		0.2849	
Contextual effects						
Transfer ($)	-0.004	(0.0074)	-0.0033	(0.0085)	-0.0067	(0.0073)
Other	Yes		Yes		Yes	
Hansen-J $\chi^2_{(11)}$	11.56	[0.397]	12.53	[0.325]	na	
KP rk LM $\chi^2_{(12)}$	84.45	[0.000]	111.51	[0.000]	118.65	[0.000]
KP Wald rk $F_{(12)}$	7.37		9.91		11.73	
Cragg-D Wald F	36.41		28.22		34.83	
Observations	30,840		29,320		30,010	

Source: ENCASEH 1997 and ENCEL 1998. Coefficient estimates from IV regressions using GMM estimator. Standard errors in parentheses are clustered at the locality-grade level. State and municipalities effects are included. Other characteristics are: age, school level, ethnicity, disabilities, age, gender and years of education of the household's head, household demographic structure, dwelling characteristics, presence of secondary school facility, community activities, and school parent association. There are 12 excluded instruments in the first stage, which correspond to peers' average of individual characteristics (see equation 4). The contextual-effects network corresponds to all children in the locality and consider the peer-effects network as children who have reached the same grade level and live in the same locality by ethnicity, landownership, and gender respectively. Hansen's J statistic is used for testing overidentifying restrictions, allowing observations to be correlated within groups. The null hypothesis is that instruments are uncorrelated with the error term (valid instruments), and that excluded instruments are correctly excluded from the estimated equation. Kleibergen-Paap rk LM is used for testing underidentification, whether the excluded instruments are correlated with the endogenous regressor (relevant instruments) when errors are non-iid. The null hypothesis is that the equation is underidentified. Kleibergen-Paap Wald rk F is used for testing weak identification, whether the excluded instruments are weakly correlated with the endogenous regressor when errors are non-iid. Cragg-Donald Wald F test for the iid errors weak identification case is included. P-values in brackets. Significantly different than zero at 99 (***), 95 (**), and 90 (*) percent confidence.

5.4 Progresa Treatment Effect

Welfare enhancing interventions in rural areas typically target a particular population or subpopulation. These type of interventions are known as partial population interventions (Moffitt, 2001) because they alter the private incentives for a subset of a group with or without the explicit intention of affecting the entire group. For example, the change in the income should affect the beneficiaries but should not, in the absence of social interactions (and ignoring general equilibrium effects), affect the outcomes of the non-beneficiaries. Whether intentionally or not, non-participants can be affected by programs, producing important spillover program effects. This section calculates such effects.

Using the estimates from column 3 in table 4 and the decomposition in equation (9), the overall program treatment effect corresponds to 5.1 percentage points increase in school attendance. Around 61 percent of that effect comes from the direct impact of the transfer, while the remaining 39 percent is a spillover effect due to the complementarities that arise from social interactions. In particular, the endogenous effect accounts for 64 percent of the spillover effect and 25 percent of the overall effect.[46,47]

$$\lambda = \underbrace{\underbrace{\beta + \frac{\gamma_0}{1+\phi}\mathbb{E}(x|T=1, P=1)\Pr(P=1)}_{\textbf{Direct effect}: 0.0311} + \underbrace{\underbrace{\frac{\phi\gamma_0}{1+\phi}\mathbb{E}(\bar{x}_p|T=1)}_{\textbf{Endogenous}: 0.0130} + \underbrace{\delta_0\mathbb{E}(\bar{x}_c|T=1)}_{\textbf{Contextual}: 0.0072}}_{\textbf{Spillover effect}: 0.0203}}_{\textbf{Overall effect}: 0.0514} \tag{10}$$

In the hypothesized case in which social divisions manifest between beneficiaries and non-beneficiaries, the endogenous spillover effect would need to be weighted by the faction of beneficiaries, $\Pr(P = 1) = 0.6824$ (see equation 7). And thus, other things equal, this effect will become 0.0089 instead of 0.0130. However, the endogenous-effects network is different when accounting for social divisions, thus the $\mathbb{E}(\bar{x}_p|T = 1) = 1.0133$ should be replaced by $\mathbb{E}(\bar{x}_p|T = 1, P = 1) = 1.4608$. Taking into account these issues, the endogenous spillover effect decreases slightly from 0.0130 to 0.0128. Interestingly, assuming that the pressure to conform to peers' behavior (*MRS* between private and social utility) does not change among beneficiary children despite the change in group membership, the overall program impacts will not be largely affected by social divisions between beneficiaries and non-beneficiaries. When social divisions appear, the average transfer at the endogenous-effects level increases for the beneficiary group, offsetting the forgone impact on the non-beneficiaries.

Finally, it is important to realize that the empirical results imply that the increase in school attendance among the non-beneficiaries in treated localities is due to social interactions. Two other potential mechanisms behind this result are: 1) unobserved changes in the source of income of the non-beneficiary population through the informal risk-sharing activities within the treated commu-

[46]Table 11 in the appendix provides other calculations using data from column 2 and 3 in table 4. It also provides a decomposition that ignores the contextual effect from the transfer because the corresponding coefficient is not significant in the regressions. Considering this, the endogenous effect accounts between 24 and 30 percent of the overall effect.

[47]Lalive and Cattaneo (2009) decomposed the treatment effect of the program into a direct effect that arises from the transfer and an indirect effect arising from the social environment of the children. For children in primary school, they found that the direct effect of Progresa is roughly the same size as the indirect effect. Dieye *et al.* (2014) account for peer effect in program response using a social network model. Their results reveal that treated and untreated students interact and that ignoring peer effects would have led to overstate the program actual impact. They found a positive direct program and endogenous peer effect, but a negative contextual effect.

nities, and 2) potential general equilibrium effects through price and wage changes. In fact, Angelucci and De Giorgi (2009) provide evidence of an increase in monthly food consumption among the non-beneficiary households in treated localities.[48] However, these program effects on the non-beneficiary population are not found in the first follow-up survey, October 1998, which is the one used here for the analysis. Moreover, they also show that there is no evidence of price changes in the labor and good markets leading to increasing consumption among the non-beneficiary population. In other words, at least for the first year of the program, the evidence is consistent with the hypothesis of a spillover program effect due to social interactions.

6 Conclusions

Welfare enhancing interventions typically involve households that interact with each other, such as the conditional cash transfer (CCT) program Progresa. The Progresa program was launched in 1997 to provide immediate monetary assistance and to foster development of the human capital of poor rural households. The quantitative evaluations of the program followed an experimental approach and showed positive impacts associated with the program in education, health, nutrition, and child and adult work. However, the qualitative evidence from focus groups and semi-structured interviews showed that the distinction introduced by the program between beneficiaries and non-beneficiaries led to social tensions and divisions, which are undoubtedly and unintended effect of any social program. CCT programs like Progresa have now been implemented in over sixty countries on five continents (Parker and Todd, 2017), however, the potential effects of program intervention on social relationships and program outcomes have remained understudied. Understanding the implication of these issues is important because in developing countries, households depend on friends and extended family and kinship networks for their livelihood and sometimes their survival (Cox and Fafchamps, 2008).

This paper estimates a behavioral model of school attendance and social interactions and presents evidence against social divisions between children of beneficiary and non-beneficiary families in the Progresa program. In other words, the program did not alter the endogenous-effects network and the complementarities that arise from social interactions. Nonetheless, the failure to reject this particular test of social divisions doesn't rule out other possible forms of conflict between beneficiaries and non-beneficiaries as a result of the program. Importantly, these complementarities can amplify program effects, producing large spillover effects. The evidence is consistent with a significant endogenous effect from social interactions in schooling decisions that amplified program outcomes. A 10 percent increase in peers' attendance increases the likelihood of attending

[48] They also show that non-beneficiary households benefit from the transfer by receiving more gifts and loans and by reducing their savings.

by 2.9 percent. The program treatment effect is decomposed into a direct and a spillover effects. The latter accounts for both endogenous (peer) and contextual effects. The evidence shows that there is a positive spillover effect on school attendance and that the endogenous social interactions effect accounts for 25 percent of the overall treatment effect of the program.

While it was not the case in the Progresa program, when program targeting does result in social divisions between beneficiaries and non-beneficiaries, and thus altering the structure of the peer-effects network, program impacts can be smaller because there won't be any endogenous spillover program effect on the non-beneficiaries. Interestingly, in the case of Progresa, if the pressure to conform to peers' behavior (*MRS* between private and social utility) do not change among beneficiary children despite the change in group membership, the program impacts will not be largely affected. The reason for this is that if social divisions happen, the average transfer at the endogenous-effects level increases for the beneficiaries, offsetting the forgone impact on the non-beneficiaries. Finally, it is worth mentioning that social divisions can have other consequences unrelated to program outcomes, such as disruptions in the operation of local markets. These issues are left for future research.

References

[1] Adato, M. The impact of progresa on community social relationships. International Food Policy Research Institute, Washington, DC (2000).

[2] Adato, M., Coady, D., and Ruel, M. An operations evaluation of progresa from the perspective of beneficiaries, promotoras, school directors, and health staff. Washington, DC: International Food Policy Research Institute (2000).

[3] Angelucci, M., and De Giorgi, G. Indirect effects of an aid program: how do cash transfers affect ineligibles' consumption? The American Economic Review 99, 1 (2009), 486–508.

[4] Angelucci, M., De Giorgi, G., Rangel, M. A., and Rasul, I. Family networks and school enrollment: Evidence from a randomized social experiment. Journal of Public Economics 94, 3 (2010), 197–221.

[5] Angelucci, M., and Di Maro, V. Programme evaluation and spillover effects. Journal of Development Effectiveness 8, 1 (2016), 22–43.

[6] Arduini, T., Patacchini, E., and Rainone, E. Treatment effects with heterogeneous externalities. Journal of Business & Economic Statistics 38, 4 (2020), 826–838.

[7] Baird, S., Bohren, J. A., McIntosh, C., and Ozler, B. Designing experiments to measure spillover effects. World Bank Policy Research Working Paper, 6824 (2014).

[8] Banerjee, A., Breza, E., Chandrasekhar, A. G., Duflo, E., Jackson, M. O., and Kinnan, C. Changes in social network structure in response to exposure to formal credit markets. Tech. rep., National Bureau of Economic Research, 2021.

[9] Behrman, J. R., Parker, S. W., and Todd, P. E. Do conditional cash transfers for schooling generate lasting benefits? a five-year followup of progresa/oportunidades. Journal of Human Resources 46, 1 (2011), 93–122.

[10] Behrman, J. R., Sengupta, P., and Todd, P. Progressing through progresa: an impact assessment of a school subsidy experiment in rural mexico. Economic Development and Cultural Change 54, 1 (2005), 237–275.

[11] Behrman, J. R., and Todd, P. E. Randomness in the experimental samples of progresa (education, health, and nutrition program). International Food Policy Research Institute, Washington, DC (1999).

[12] Blume, L. E., Brock, W. A., Durlauf, S. N., and Ioannides, Y. M. Identification of social interactions. Handbook of Social Economics 1, B (2011), 853–964.

[13] Blume, L. E., Brock, W. A., Durlauf, S. N., and Jayaraman, R. Linear social interactions models. Journal of Political Economy 123, 2 (2015), 444–496.

[14] Bobonis, G. J., and Finan, F. Neighborhood peer effects in secondary school enrollment decisions. The Review of Economics and Statistics 91, 4 (2009), 695–716.

[15] Bramoullé, Y., Djebbari, H., and Fortin, B. Identification of peer effects through social networks. Journal of Econometrics 150, 1 (2009), 41–55.

[16] Brock, W. A., and Durlauf, S. N. Interactions-based models. Handbook of Econometrics 5 (2001), 3297–3380.

[17] Caeyers, B., and Fafchamps, M. Exclusion bias in the estimation of peer effects. Working paper, Freeman Spogli Institute for International Studies, November 2020.

[18] Calvó-Armengol, A., Patacchini, E., and Zenou, Y. Peer effects and social networks in education. The Review of Economic Studies 76, 4 (2009), 1239–1267.

[19] Comola, M., and Prina, S. Treatment effect accounting for network changes. Review of Economics and Statistics 103, 3 (2021), 597–604.

[20] Cox, D., and Fafchamps, M. Extended family and kinship networks: economic insights and evolutionary directions. Handbook of Development Economics 4 (2007), 3711–3784.

[21] Dieye, R., Djebbari, H., and Barrera-Osorio, F. Accounting for peer effects in treatment response. Working Paper (2014).

[22] Dupas, P., Keats, A., and Robinson, J. The effect of savings accounts on interpersonal financial relationships: Evidence from a field experiment in rural kenya. The Economic Journal 129, 617 (2019), 273–310.

[23] Durlauf, S. N. Neighborhood effects. Handbook of Regional and Urban Economics 4 (2004), 2173–2242.

[24] Durlauf, S. N., and Ioannides, Y. M. Social interactions. Annual Review of Economics 2, 1 (2010), 451–478.

[25] Durlauf, S. N., and Young, H. P. Social Dynamics, vol. 4. Mit Press, 2004.

[26] Epple, D., and Romano, R. Peer effects in education: A survey of the theory and evidence. Handbook of Social Economics 1, 11 (2011), 1053–1163.

[27] Heckman, J. J. Randomization as an instrumental variable: Notes. The Review of Economics and Statistics 78, 2 (1996), 336–41.

[28] Hoxby, C. Peer effects in the classroom: Learning from gender and race variation. Tech. rep., National Bureau of Economic Research, 2000.

[29] Lalive, R., and Cattaneo, M. A. Social interactions and schooling decisions. The Review of Economics and Statistics 91, 3 (2009), 457–477.

[30] Lavy, V., and Schlosser, A. Mechanisms and impacts of gender peer effects at school. American Economic Journal: Applied Economics 3, 2 (2011), 1–33.

[31] Levy, S. Progress against poverty: sustaining Mexico's Progresa-Oportunidades program. Brookings Institution Press, 2007.

[32] Manski, C. F. Identification of endogenous social effects: The reflection problem. The Review of Economic Studies 60, 3 (1993), 531–542.

[33] Manski, C. F. Identification problems in the social sciences. Harvard University Press, 1999.

[34] Manski, C. F. Economic analysis of social interactions. The Journal of Economic Perspectives 14, 3 (2000), 115–136.

[35] Manski, C. F. Identification of treatment response with social interactions. The Econometrics Journal 16, 1 (2013).

[36] Moffitt, R. A. Policy interventions, low-level equilibria, and social interactions in steven durlauf and h. p. young (eds.). Social Dynamics 4 (2001), 45–82.

[37] Parker, S. W., and Todd, P. E. Conditional cash transfers: The case of progresa/oportunidades. Journal of Economic Literature 55, 3 (2017), 866–915.

[38] Sacerdote, B. Peer effects in education: How might they work, how big are they and how much do we know thus far? Handbook of the Economics of Education 3, 3 (2011), 249–277.

[39] Schultz, T. P. Impact of progresa on school attendance rates in the sampled population. February. Report submitted to PROGRESA. International Food Policy Research Institute, Washington, DC (2000).

[40] Schultz, T. P. School subsidies for the poor: evaluating the mexican progresa poverty program. Journal of Development Economics 74, 1 (2004), 199–250.

[41] Skoufias, E. 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. International Food Policy Research Institute, Washington, DC (2001).

[42] Skoufias, E., and Parker, S. W. Conditional cash transfers and their impact on child work and schooling: Evidence from the progresa program in mexico [with comments]. Economía (2001), 45–96.

[43] Sobel, M. E. What do randomized studies of housing mobility demonstrate? causal inference in the face of interference. Journal of the American Statistical Association 101, 476 (2006), 1398–1407.

[44] Stinchcombe, A. L. Agricultural enterprise and rural class relations. American Journal of Sociology 67, 2 (1961), 165–176.

Appendix A

A.1 Identification Strategy

Without lost of generality, consider a simpler version of the model.

$$\omega_{ipc} = \alpha + \frac{\beta_1}{1+\phi} y_{ipc} + \frac{\beta_2}{1+\phi} \bar{y}_c + \frac{\phi}{1+\phi} \bar{\omega}_{pc} + \eta_{ipc} \tag{A1}$$

$$\bar{\omega}_p = \alpha(1+\phi) + \beta_1 \bar{y}_{pc} + \beta_2 \bar{y}_c + \eta_{pc} \tag{A2}$$

$$\omega_{ipc} = \alpha(1+\phi) + \frac{\beta_1}{1+\phi} y_{ipc} + \frac{\phi\beta_1}{1+\phi} \bar{y}_{pc} + \beta_2 \bar{y}_c + \tilde{\eta}_{ipc} \tag{A3}$$

where ω_{ipc} denotes outcome of agent i, with peer-effects network p and contextual-effects network c; y_{ipc} is a vector of individual characteristics, $\bar{y}_c$ denotes the corresponding average characteristics in the contextual-effects network; and $\bar{\omega}_p$ denotes the average outcome in the peer-effects network.

In this case, the model is just identified and the parameters can be recovered from the reduced-from $(A3)$. Another way to understand the identification is by considering the rank condition: equation $(A1)$ is identified if $(A2)$ contains one exogenous variable, e.g. $\bar{y}_{pc}$, with non-zero coefficient ($\beta_1 \neq 0$) which is excluded from $(A1)$. Moreover, $\bar{y}_{pc}$ is a potential instrument to assess the simultaneity bias.

Peers might self-select into groups in ways unobserved to the econometrician. For example, sorting can occur if children decidedly repeat the grade. Following Epple and Romano (2011), suppose peers get together based in part on their (un)observed characteristics, that is, self-selection in the peer-effect network.

$$\omega_{ipc} = \alpha + \frac{\beta_1}{1+\phi} y_{ipc} + \frac{\beta_2}{1+\phi} \bar{y}_c + \frac{\phi}{1+\phi} \bar{\omega}_{pc} + \underbrace{\frac{\theta}{1+\phi} \bar{y}_{pc}}_{\text{self-selection}} + \epsilon_{ipc} \tag{A4}$$

$$\bar{\omega}_p = \alpha(1+\phi) + (\beta_1 + \theta) \bar{y}_{pc} + \beta_2 \bar{y}_c + \epsilon_{pc} \tag{A5}$$

$$\omega_{ipc} = \alpha(1+\phi) + \frac{\beta_1}{1+\phi} y_{ipc} + \frac{\phi(\beta_1+\theta)}{1+\phi} \bar{y}_{pc} + \beta_2 \bar{y}_c + \tilde{\epsilon}_{ipc} \tag{A6}$$

In the above case the parameters are not identified. However, consider an exogenous variable, such as the cash transfer from a program intervention, x_{ipc}.

$$\omega_{ipc} = \alpha + \frac{\beta_1}{1+\phi} y_{ipc} + \frac{\beta_2}{1+\phi}\bar{y}_c + \underbrace{\frac{\gamma}{1+\phi} x_{ipc}}_{\text{cash transfer}} + \frac{\phi}{1+\phi}\bar{\omega}_{pc} + \underbrace{\frac{\theta}{1+\phi}\bar{y}_{pc}}_{\text{self-selection}} + \upsilon_{ipc} \tag{A7}$$

$$\bar{\omega}_p = \alpha(1+\phi) + (\beta_1+\theta)\bar{y}_{pc} + \beta_2\bar{y}_c + \gamma\bar{x}_{pc} + \upsilon_{pc} \tag{A8}$$

$$\omega_{ipc} = \alpha(1+\phi) + \frac{\beta_1}{1+\phi} y_{ipc} + \frac{\phi(\beta_1+\theta)}{1+\phi}\bar{y}_{pc} + \beta_2\bar{y}_c + \frac{\gamma}{1+\phi} x_{ipc} + \frac{\phi\gamma}{1+\phi}\bar{x}_{pc} + \tilde{\upsilon}_{ipc} \tag{A9}$$

In this case, the endogenous effect is identified, and $\bar{x}_p$ serves as an instrument for peers' behavior. If self-selection occurs at the contextual-effects network, the endogenous effect parameter is still identified because social pressure comes from the individuals in the peer-effects network. However, contextual effects won't be identified in this case.

Suppose the treatment is the reference group and peers self-select in part on their eligibility (beneficiary) status, P_{ipc}. That is, within the classroom there are two non-overlapping groups, beneficiaries (poor) and non-beneficiaries (non-poor).

$$\omega_{ipc} = \alpha + \frac{\beta_1}{1+\phi} y_{ipc} + \frac{\beta_2}{1+\phi}\bar{y}_c + \frac{\gamma}{1+\phi} x_{ipc} + \frac{\phi}{1+\phi}\bar{\omega}_{pc} + \underbrace{\frac{\delta}{1+\phi} P_{ipc}}_{\text{self-selection}} + \nu_{ipc} \tag{A10}$$

$$\bar{\omega}_p = \alpha(1+\phi) + \beta_1\bar{y}_{pc} + \beta_2\bar{y}_c + \gamma\bar{x}_{pc} + \delta\bar{P}_{pc} + \nu_{pc} \tag{A11}$$

$$\omega_{ipc} = \alpha(1+\phi) + \frac{\beta_1}{1+\phi} y_{ipc} + \frac{\phi\beta_1}{1+\phi}\bar{y}_{pc} + \beta_2\bar{y}_c + \frac{\gamma}{1+\phi} x_{ipc} + \frac{\phi\gamma}{1+\phi}\bar{x}_{pc} + \frac{\delta}{1+\phi} P_{ipc} + \frac{\phi\delta}{1+\phi}\bar{P}_{pc} + \tilde{\nu}_{ipc} \tag{A12}$$

In this case, the parameters are still identified by exogenous cash transfer from the program intervention. Moreover, in this case self-selection is not unobserved and is introduced explicitly in the model.

A.2 Decomposition of the Average Treatment Effect

Let ω_1 denote the potential outcome when an individual receives a treatment, and ω_0 the potential outcome for the no treatment case. The average treatment effect (ATE) is the expected value of the difference, $\mathbb{E}(\omega_1 - \omega_0)$. Let $T = 1$ denote the case when a locality was randomly selected into treatment, and $T = 0$ otherwise; and let $P = 1$ denote the beneficiary (poverty) status of the individual. Using the law of total probability and the experimental design of the program,[49] the average treatment effect can be decomposed as follows,

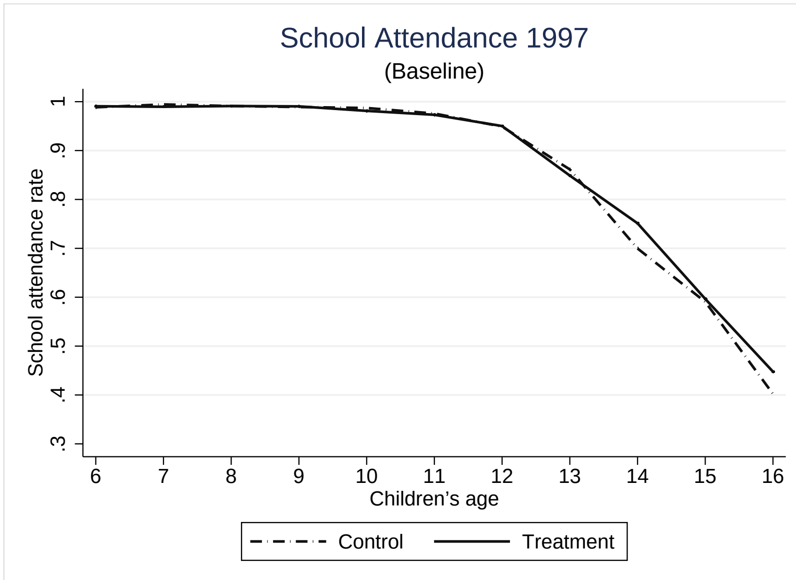
$$
\begin{aligned}
ATE \equiv \lambda &= \mathbb{E}(\omega_1 - \omega_0) \\
&= \mathbb{E}(\omega_1|T = 1) - \mathbb{E}(\omega_0|T = 0) \\
&= \mathbb{E}(\omega_1|T = 1, P = 1)\,\Pr(P = 1|T = 1) + \mathbb{E}(\omega_1|T = 1, P = 0)\,\Pr(P = 0|T = 1) \\
&\quad - \mathbb{E}(\omega_0|T = 0, P = 1)\,\Pr(P = 1|T = 0) - \mathbb{E}(\omega_0|T = 0, P = 0)\,\Pr(P = 0|T = 0) \\
&= [\,\underbrace{\mathbb{E}(\omega_1|T = 1, P = 1) - \mathbb{E}(\omega_0|T = 0, P = 1)}_{\lambda^B}\,]\,\Pr(P = 1) \\
&\quad + [\,\underbrace{\mathbb{E}(\omega_1|T = 1, P = 0) - \mathbb{E}(\omega_0|T = 0, P = 0}_{\lambda^{NB}}]]\,\Pr(P = 0) \\
\lambda &= \lambda^B \Pr(P = 1) + \lambda^{NB} \Pr(P = 0) \qquad \text{(A13)}
\end{aligned}
$$

The average treatment effect, λ, is a weighted average of the treatment effect on the beneficiary (poor), λ^B, and the treatment effect on the non-beneficiary (non-poor), λ^{NB}; and the weight is given by the fraction of beneficiaies (poor), $\Pr(P = 1)$.

[49] Of the 506 localities, 320 were randomly assigned to the treatment group, and the rest were randomized out to form the control group. Although only the poor in treatment localities received the benefits of the program, the poor were also identified in the control localities, that is, poverty status is available for both treatment and control localities in the data. Because of the random assignment, $\Pr(P = 1|T = 1) = \Pr(P = 1|T = 0) = \Pr(P = 1)$, that is, the fraction of beneficiaries (poor) is the same in both treatment and control localities.

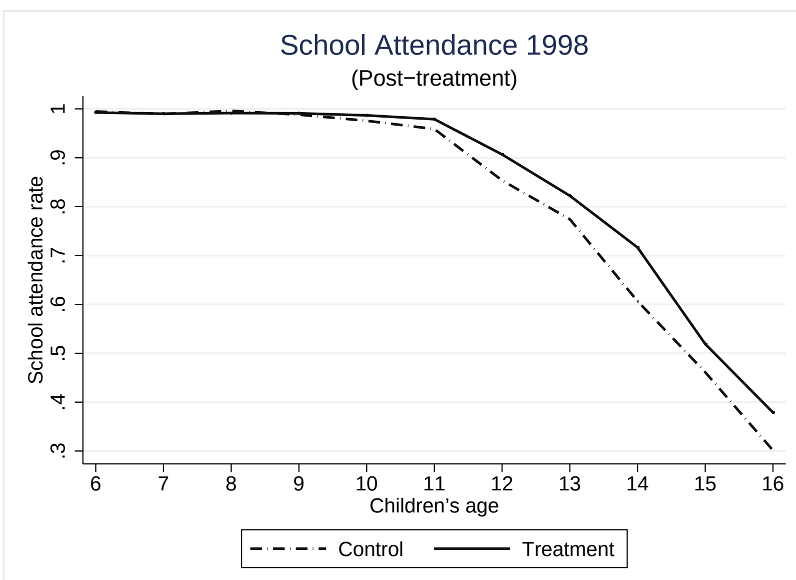
Appendix B: Supplementary Tables & Figures

Figure 1: School Attendance by Age, 1997



Source: ENCASEH 1997. The sample corresponds to children who have completed any primary or secondary school grade at baseline (October, 1997). Control and treatment refer to the locality where the children live. School attendance 1997 comes from the following item in the baseline survey, "Does your child attend to school?"

Figure 2: School Attendance by Age, 1998



Source: ENCEL 1997. The sample corresponds to children who have completed any primary or secondary school grade at baseline (October, 1997). Control and treatment refer to the locality where the children live. School attendance in 1998 comes from the following item in the first post-intervention round, "Does your child attend to school?"

Figure 3: Benchmark Model. Example of Contextual- and Peer-Effects Networks

$$
C_{8,8} = \begin{array}{cc}
 & \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} &
\begin{pmatrix}
0 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\
1/7 & 0 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\
1/7 & 1/7 & 0 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\
1/7 & 1/7 & 1/7 & 0 & 1/7 & 1/7 & 1/7 & 1/7 \\
1/7 & 1/7 & 1/7 & 1/7 & 0 & 1/7 & 1/7 & 1/7 \\
1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 0 & 1/7 & 1/7 \\
1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 0 & 1/7 \\
1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 0
\end{pmatrix}
\end{array}
\quad
A_{8,8} = \begin{array}{cc}
 & \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} &
\begin{pmatrix}
0 & 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\
1/3 & 0 & 1/3 & 1/3 & 0 & 0 & 0 & 0 \\
1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 0 \\
1/3 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \\
0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 \\
0 & 0 & 0 & 0 & 1/3 & 1/3 & 0 & 1/3 \\
0 & 0 & 0 & 0 & 1/3 & 1/3 & 1/3 & 0
\end{pmatrix}
\end{array}
$$

Example of contextual- and peer-effects networks, sociomatrices C and A, respectively. Each entry in the sociomatrices describes the strength of the relationship between the pair of children (i, j). In the benchmark specification, the contextual-effects network corresponds to all children in the locality, and the endogenous-effects network is composed by children who reached the same grade (at baseline) and live in the same locality. The example considers one locality with eight children. The weights for the contextual effects are then $1/7$ for each children. To simplify, consider two school grades in the locality. The children labeled 1 to 4 are in one grade, conforming one peer-effect network, and the children 5 to 6 comprise another network. In this case the corresponding weights are $1/3$.

Figure 4: Beneficiary Status as an Endogenous-Effects Network. Example of Contextual- and Peer-Effects Networks

$$
C_{8,8} = \begin{array}{cc}
 & \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} &
\begin{pmatrix}
0 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\
1/7 & 0 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\
1/7 & 1/7 & 0 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\
1/7 & 1/7 & 1/7 & 0 & 1/7 & 1/7 & 1/7 & 1/7 \\
1/7 & 1/7 & 1/7 & 1/7 & 0 & 1/7 & 1/7 & 1/7 \\
1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 0 & 1/7 & 1/7 \\
1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 0 & 1/7 \\
1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 0
\end{pmatrix}
\end{array}
\quad
A_{8,8} = \begin{array}{cc}
 & \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} &
\begin{pmatrix}
0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}
\end{array}
$$

Example of contextual- and peer-effects networks, sociomatrices C and A, respectively. Each entry in the sociomatrices describes the strength of the relationship between the pair of children (i, j). The beneficiary status as an endogenous-effects network corresponds to a partition of the peer-effects networks in the benchmark model by beneficiary status only. The contextual-effects network corresponds to all children in the locality, similar to the benchmark specification. The figure depicts the weights corresponding to one locality with eight children and two school grades. Children labeled 1 to 4 are in one grade, and the rest in another one; and the children labeled 1,2,3,5, and 6 are beneficiaries (in red), while 4,7, and 8 are non-beneficiaries.

Table 8: Descriptive Statistics by Treatment Status

Individual characteristics	Control		Treatment			(t-test)
	mean	s.d.	mean	s.d.	diff.	p-value
Transfer ($)			1.4580	1.2785	-1.4580	0.0000
Poor	0.6668	0.4714	0.6920	0.4617	-0.0252	0.0000
Grade 3	0.1208	0.3259	0.1236	0.3292	-0.0029	0.4336
Grade 4	0.1173	0.3218	0.1114	0.3146	0.0059	0.0936
Grade 5	0.1089	0.3116	0.1079	0.3102	0.0010	0.7629
Grade 6	0.1438	0.3509	0.1467	0.3539	-0.0029	0.4537
Grade 7	0.0476	0.2129	0.0454	0.2083	0.0021	0.3631
Grade 8	0.0371	0.1890	0.0345	0.1825	0.0026	0.2084
Grade 9	0.0200	0.1401	0.0194	0.1378	0.0007	0.6723
Male	0.5026	0.5000	0.5159	0.4998	-0.0133	0.0169
Indigenous	0.2910	0.4542	0.2948	0.456	-0.0038	0.4511
Monolingual	0.0493	0.2165	0.0323	0.1767	0.0171	0.0000
Bilingual	0.2397	0.4269	0.2610	0.4392	-0.0213	0.0000
Disabled	0.0386	0.1927	0.0364	0.1874	0.0022	0.2989
# Siblings	3.6163	2.0658	3.6279	2.0655	-0.0115	0.6175
First born (dummy)	0.1889	0.3914	0.1936	0.3952	-0.0047	0.2795
Head age	44.3321	11.8087	43.6986	11.3022	0.6335	0.0000
Head male	0.9275	0.2592	0.9317	0.2523	-0.0041	0.1462
Head indigenous	0.3413	0.4742	0.3627	0.4808	-0.0214	0.0001
Head literate	0.7360	0.4408	0.7384	0.4395	-0.0025	0.6175
Head's years of schooling	2.8591	2.5930	2.9211	2.6576	-0.0620	0.0353
Head agricultural worker	0.5208	0.4996	0.4952	0.5000	0.0256	0.0000
Head (ejidatario)	0.1270	0.3330	0.1320	0.3385	-0.0049	0.1870
Head married	0.7030	0.4570	0.7315	0.4432	-0.0286	0.0000
Household size	7.2481	2.3578	7.2108	2.3200	0.0373	0.1522
Babies (0-2)	0.4524	0.6454	0.4516	0.6388	0.0008	0.9146
Preschoolers (3-5)	0.6477	0.7267	0.6601	0.7198	-0.0124	0.1235
School-agers (6-17)	3.3385	1.4991	3.3430	1.5211	-0.0046	0.7863
Young adults (18-44)	2.0310	1.0478	2.0023	1.0443	0.0287	0.0138
Adults (45-64)	0.6163	0.7880	0.6094	0.7919	0.0069	0.4329
Seniors (65+)	0.1580	0.4310	0.1402	0.4139	0.0178	0.0001
Dependency ratio	1.4580	0.9436	1.4691	0.9196	-0.0111	0.2844
Floor sand	0.5875	0.4923	0.5943	0.4910	-0.0069	0.2118
Durable roof	0.2529	0.4347	0.2666	0.4422	-0.0137	0.0052
Durable walls	0.8079	0.3940	0.8490	0.3581	-0.0411	0.0000
# of rooms	1.9161	1.1221	1.9193	1.1055	-0.0032	0.7974
Refrigerator	0.1593	0.3660	0.1384	0.3453	0.0209	0.0000
Gas stove	0.3193	0.4662	0.2886	0.4531	0.0307	0.0000
Television	0.5440	0.4981	0.4825	0.4997	0.0616	0.0000
# Livestock (work)	0.9416	1.6824	1.0102	2.0207	-0.0687	0.0012
# Livestock (consumption)	11.9796	12.8285	12.0828	13.693	-0.1032	0.4887
Observations	12,991		21,126		Total: 34,117	

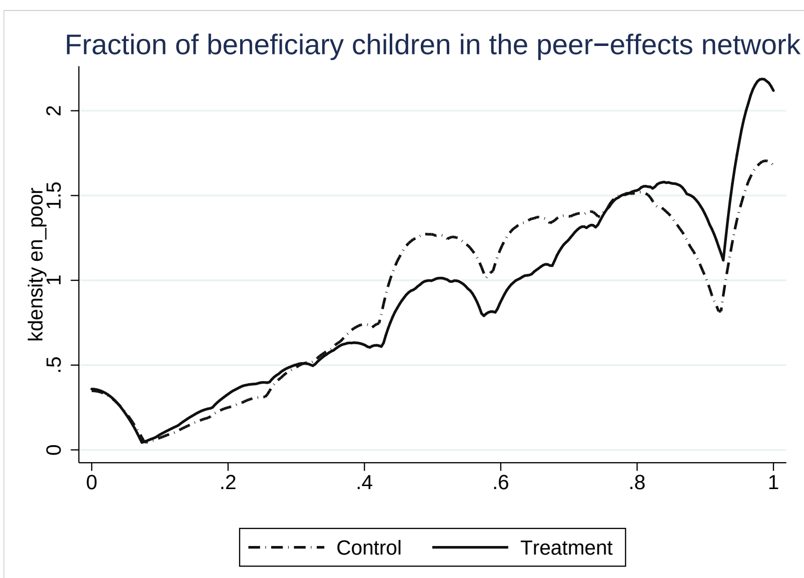
Source: ENCASEH 1997, ENCEL 1998, and administrative records. The sample corresponds to children who have completed any primary or secondary school grade at baseline (October, 1997). Control and treatment refer to the locality where the children live. The p-value comes from a t-test for the equality of means between the treatment and control samples. The 320 localities from the treatment group were randomly selected.

Table 9: Summary of Peers' School Attendance and Excluded Instruments by Treatment Status

Endogenous-effect network	**Control**		**Treatment**			(t-test)
	mean	s.d.	mean	s.d.	diff.	p-value
School attendance (1998)	0.8144	0.2756	0.8517	0.2451	-0.0373	0.0000
Transfer ($)	0.0000	0.0000	1.0089	0.7166	-1.0089	0.0000
Poor	0.6668	0.2571	0.6920	0.2680	-0.0252	0.0000
Male	0.5026	0.1982	0.5159	0.2103	-0.0133	0.0000
Disable	0.0386	0.0882	0.0364	0.0878	0.0022	0.0253
Head's years of schooling	2.8592	1.3451	2.9212	1.4608	-0.0620	0.0001
Head agricultural worker	0.5208	0.2918	0.4952	0.3031	0.0256	0.0000
Head married	0.7030	0.2507	0.7315	0.2462	-0.0286	0.0000
Preschoolers (3-5)	1.1001	0.5598	1.1117	0.5804	-0.0116	0.0683
School-agers (6-17)	3.3385	0.7315	3.3430	0.7613	-0.0046	0.5846
Young adults (18-44)	2.0310	0.4816	2.0023	0.5089	0.0287	0.0000
Adults (45-64)	0.6163	0.3697	0.6094	0.3679	0.0069	0.0926
Seniors (65+)	0.1580	0.1834	0.1402	0.1790	0.0178	0.0000
Observations	12,991		21,126		Total: 34,117	

Source: ENCASEH 1997, ENCEL 1998, and administrative records. The sample corresponds to children who have completed any primary or secondary school grade at baseline (October, 1997). Control and treatment refer to the locality where the children live. The p-value comes from a t-test for the equality of means between the treatment and control samples. The 320 localities from the treatment group were randomly selected. All averages correspond to the peer-effects network, which considers children who have reached the same grade level and live in the same locality. School attendance 1997 comes from the following item in the baseline survey, "Does your child attend to school?"

Figure 5: Fraction of Beneficiary Children in the Endogenous-Effects Network



Source: ENCEL 1997. The sample corresponds to children who have completed any primary or secondary school grade at baseline (October, 1997). Control and treatment refer to the locality where the children live. The fraction corresponds to the peer-effects network, which considers children who have reached the same grade level and live in the same locality.

Table 10: Beneficiary Status as Endogenous-Effects Network. IV Estimates of Contextual- and Endogenous-Effects

	Partition by share of beneficiary children in the grade									
	(1) [0-20%]		**(2) (20-40%]**		**(3) (40-60%]**		**(4) (60-80%]**		**(5) (80-100%]**	
First stage										
Transfer ($)	0.1567**	(0.0754)	-0.0109	(0.0157)	0.0347***	(0.0099)	0.0188**	(0.0075)	0.0216**	(0.0088)
Other	Yes		Yes		Yes		Yes		Yes	
Second stage										
Transfer ($)	0.0485	(0.0360)	0.0116	(0.0139)	0.0390***	(0.0069)	0.0274***	(0.0045)	0.0324***	(0.0033)
Endog. effect	0.5025**	(0.2415)	0.4111*	(0.2199)	0.4169***	(0.1503)	-0.1258	(0.2348)	0.3089*	(0.1716)
MRS (ϕ)	1.0100		0.6981		0.7151		-0.1117		0.4470	
Contextual effects (locality)										
Transfer ($)	-0.0723	(0.1732)	-0.0936	(0.0767)	-0.0260	(0.0297)	0.0383	(0.0333)	0.0045	(0.0137)
Other	Yes		Yes		Yes		Yes		Yes	
Hansen-J $\chi^2_{(10)}$	na		na		na		na		na	
KP rk LM $\chi^2_{(11)}$	18.86	[0.064]	20.5	[0.039]	29.87	[0.002]	23.06	[0.017]	23.95	[0.013]
KP Wald rk $F_{(11)}$	2.35		1.92		3.08		2.22		2.27	
Cragg-D Wald F	6.36		3.94		8.74		5.26		11.03	
Observations	654		1,782		3,289		4,819		7,961	

Source: ENCASEH 1997 and ENCEL 1998. Coefficient estimates from IV regressions using GMM estimator. Standard errors in parentheses are clustered at the locality-grade level. State and municipalities effects are included. Other characteristics are: age, school level, ethnicity, disabilities, age, gender and years of education of the household's head, household demographic structure, dwelling characteristics, presence of secondary school facility, community activities, and school parent association. There are 12 excluded instruments in the first stage, which correspond to peers' average of individual characteristics (see equation 4). The contextual-effects network corresponds to all children in the locality and the peer-effects network considers children who have reached the same grade level and live in the same locality by beneficiary status, beneficiary (poor) and non-beneficiary (non-poor). Hansen-J test is omitted due to the small number of clusters. Hansen's J statistic is used for testing overidentifying restrictions, allowing observations to be correlated within groups. The null hypothesis is that instruments are uncorrelated with the error term (valid instruments), and that excluded instruments are correctly excluded from the estimated equation. Kleibergen-Paap rk LM is used for testing underidentification, whether the excluded instruments are correlated with the endogenous regressor (relevant instruments) when errors are non-iid. The null hypothesis is that the equation is underidentified. Kleibergen-Paap Wald rk F is used for testing weak identification, whether the excluded instruments are weakly correlated with the endogenous regressor when errors are non-iid. Cragg-Donald Wald F test for the iid errors weak identification case is included. P-values in brackets. Significantly different than zero at 99 (***), 95 (**), and 90 (*) percent confidence.

Table 11: Progresa Treatment Effect

	All sample				**Treated sample**			
	w/Contextual		**wo/Contextual**		**w/Contextual**		**wo/Contextual**	
$\beta/(1+\phi)$	0.0037	(0.0086)	0.0037	(0.0086)				
β	0.0050		0.0050					
$\gamma_0/(1+\phi)$	0.0282	(0.0021)	0.0282	(0.0021)	0.0313	(0.0024)	0.0313	(0.0024)
$\mathbb{E}(x \mid T=1, P=1)$	1.4580	(0.0106)	1.4580	(0.0106)	1.4580	(0.0106)	1.4580	(0.0106)
$\Pr(P=1)$	0.6824	(0.0025)	0.6824	(0.0025)	0.6824	(0.0025)	0.6824	(0.0025)
Direct effect	0.0330		0.0330		0.0311		0.0311	
$\phi/(1+\phi)$	0.2638	(0.0976)	0.2638	(0.0976)	0.2913	(0.1056)	0.2913	(0.1056)
ϕ	0.3583		0.3583		0.4110		0.4110	
$\phi\gamma_0/(1+\phi)$	0.0101		0.0101		0.0129		0.0129	
$\mathbb{E}(\bar{x}_{pc} \mid T=1)$	1.0133	(0.0049)	1.0133	(0.0049)	1.0133	(0.0049)	1.0133	(0.0049)
Endogenous effect	0.0102		0.0102		0.0130		0.0130	
$\delta_0/(1+\phi)$	-0.0070	(0.0068)			0.0051	(0.0086)		
δ_0	-0.0095				0.0072			
$\mathbb{E}(\bar{x}_c \mid T=1)$	1.0089	(0.0039)			1.0089	(0.0039)		
Contextual effect	-0.0096				0.0072			
Spillover effect	0.0006		0.0102		0.0203		0.013	
Overall effect	0.0337		0.0433		0.0514		0.0442	
Edogenous / Overall	0.3040		0.2366		0.2535		0.2951	

Source: ENCASEH 1997 and ENCEL 1998. Coefficient estimates and standard errors in parentheses come from Table 4 column 2 (All sample) and column 3 (Treated sample), respectively. The columns labeled 'wo/Contextual' ignore the contextual effects because this coefficient estimate is not significant in the regressions.