1

TRANSFERS TO THE POOR INCREASE THE SCHOOLING OF THE NON-POOR: THE CASE OF MEXICO'S PROGRESA PROGRAM

Gustavo J. Bobonis and Frederico S. Finan

University of California at Berkeley

September 25, 2002

Introduction

Recent literature has argued that peer effects are an important determinant of educational outcomes. This paper examines the existence of peer effects in secondary school enrollment in poor rural villages in Mexico. We exploit a uniquely designed educational subsidy program, which allows us to plausibly identify a spillover effect among the non-recipients of the program. After investigating several potential mechanisms, we find that the role of peer effects provides the most convincing explanation for this program externality.

Much of the empirical literature that rely on non-experimental data to identify schooling peer effects, estimate the association between own outcomes and group behavior. However, Manski (1995) points out three methodological problems with this identification strategy, namely: (i) individuals may have self-selected into their reference groups (the correlated effect), (ii) exogenous characteristics of the reference group (observed or unobserved) affect individuals' schooling decisions (the contextual effect), and (iii) simultaneity between the peer group effect on own outcomes and own outcome effects on peer group outcomes (the endogenous effect).

Several approaches have been used to address this identification problem, such as instrumental variables techniques (Case and Katz (1991), Gaviria and Raphael (1999), Borjas (1992), and structural estimation (Oates and Schwab (1992), Kremer (1997)). However, the usual caveats of the validity of instruments or the misspecification of the structural model have cast doubts on many of these results. In a recent innovative study, Sacerdote (2000) uses a random assignment of college freshmen roommates to provide strong evidence for peer effects on grade point averages and the decision to join fraternities. In a different context, Miguel and Kremer (2002), using an experimental design, suggests that peer effects influence the adoption of deworming medicine for school children in Kenya.

Our paper utilizes the research design of an incentive-based cash transfer program (PROGRESA) aimed at improving education, health, and nutrition among the rural poor population of Mexico. PROGRESA included a program evaluation component from its inception. In the evaluation design, it randomized communities into treatment and control groups, thus providing a unique opportunity to apply experimental design methods to measure spillover effects in educational outcomes. The random assignment, along with an objective selection of beneficiaries and the panel structure of the data will permit us to convincingly identify these effects and avoid several of the identification pitfalls that have plagued much of the social interaction literature.

In particular, we examine whether PROGRESA increased secondary school participation among the non-beneficiaries residing in the treatment villages relative to control villages. Our research shows that the enrollment rates of children from non-beneficiary households near the welfare threshold for selection into the program increased by 10 percentage points. Moreover, we find differential spillover effects by gender, age, and other household characteristics. We use the variation in this impact along with other data to try to understand the mechanism underlying this spillover effect. Exploring various hypotheses, we can reject various contextual factors and argue that peer effects are the most plausible explanation for this increased enrollment.

Our findings contribute to a controversy put forth by two other papers that have analyzed the educational spillover effects of PROGRESA. Both Behrman, Sengupta, and Todd (2001) and Handa et al (2001) attempt to measure spillover effects in school enrollment and in the end report inconsistent results. Behrman, Sengupta, and Todd (2001) use a Markov schooling transition model to assess the impact of the program. Estimating this model for non-eligible children of the treatment and control villages, they cannot reject that the conditional probability of enrollment is the same, which rejects the spillover hypothesis. However, they fail to consider that the program may have only impacted those non-beneficiary households that were just above the cut-off criteria.

Conversely, Handa et al (2001) find that school continuation rates aggregated at the village level are higher for PROGRESA communities than non-PROGRESA communities among the 10-12 years old non-eligible cohort. They find the continuation rates among 11-12 non-eligible girls is 9.5 percentage points higher in PROGRESA communities relative to non-PROGRESA communities. While these findings do present preliminary evidence of a spillover effect, their identification suffers from potential omitted variables bias and endogeneity problems. Instead, we propose a model of individual school enrollment that controls for unobservable characteristics in a fixed-effects framework. We also extend the analysis by exploring the determinants of the program's externality.

In a broader context, our results provide important policy implications. To design policies that increase levels of educational attainment, it is necessary to understand what motivates children to enroll. Our findings suggest that peer effects are an important determinant of school enrollment. By not considering these social interactions, we ignore important social multipliers that further justify program costs and other potential targeting methods.

The paper is structured as follows. Section 2 explains PROGRESA and its research design. Section 3 describes our empirical strategy, and a description of the data follows in Section 4. The results are reported in Section 5 and reasons for the spillover are explored in Section 6. Section 7 concludes the paper.

Background on PROGRESA and its Research Design

In this section, we provided a brief overview of the program and its experimental design. Several key features of the PROGRESA data will help identify the spillover effect.

To break the inheritance of poverty among the rural poor the Mexican government initiated, in 1997, a large-scale program (PROGRESA) aimed at improving education, health, and nutrition among this population. The program targets the poor in marginal rural communities, where 40% of the children from poor households discontinue school after primary level. The program provides cash transfers to the mothers of over 2.5 million children conditional on school attendance, at an annual cost of approximately a billion dollars.

A distinguishing characteristic of PROGRESA is that it included a program evaluation component from its inception. PROGRESA was implemented following an experimental design on a subset of 506 communities (a diagram which is often used to explain the design of the program is depicted in Figure 1). Among these communities, 320 were randomly assigned into a treatment group, with the remaining 186 communities serving as a control group, thus providing a unique opportunity to apply experimental design methods to measure its impact on education outcomes. Within these selected communities, eligible households were identified on the basis of a welfare index, which was constructed from the resulting score of a discriminant analysis of income and various household assets and characteristics. While household eligibility was determined within all communities, only households below a welfare threshold and within the treatment villages could become PROGRESA beneficiaries.[1]

Since the baseline census in 1997, the program has conducted extensive interviews biannually, on over 24,000 households located across seven states (Guerrero, Hidalgo, Michoacán, Puebla, Querétaro, San Luis Potosí, and Veracruz). Each survey is a community level census containing detailed information on schooling, health, and nutrition. These data are available at the individual, household, and village level.

Empirical Methodology

We are interested in testing whether PROGRESA increased secondary school enrollment among the non-beneficiaries residing in the treatment villages. Our analysis compares changes in secondary school enrollment decisions of non-beneficiary children between treatment and control villages during a 3-year time period. Here, we discuss the econometric models and identification strategy used to test the possibility of a spillover effect.

Our identification strategy exploits two important features of the experimental design of the program. First, the random assignment of treatment and control villages should ensure that the observable and unobservable characteristics of households in the treatment group are similar to the control group. Randomization also helps avoid problems of selection bias that frequently arise in non-experimental data. Although randomization is not a necessary condition for our identification strategy, it does provide some level of comfort for presupposing our counterfactual assumption. The second important feature of the experiment design is the panel structure of the data, which include a pre-treatment round and two post-treatment rounds. These features will allow us to estimate a difference-in-difference model to compare changes in enrollment rates between treatment and control villages among non-beneficiaries.

[1] See Skoufias et al (2001) for an evaluation of the targeting methods used to identify the program's beneficiary household.

Following the simple difference-in-difference framework, suppose that school enrollment, S_{it}, among the children from non-eligible households prior to the start of the PROGRESA program can be expressed as the sum of a common year effect, β_t, and a village fixed-effect, δ_V:

(1) $$E[S_{io} \mid T, V] = \beta_0 + \delta_V$$

Also suppose that after the introduction of the program, the effect of PROGRESA on the non-beneficiaries, denoted γ, is constant, i.e.

(2) $$E[S_{i1} \mid T, V] = \beta_1 + \delta_V + \gamma$$

Then the difference in secondary school enrollment rates across treatment and control villages, and time yields:

(3) $$\begin{aligned}\gamma = \{&E[S_{i1} \mid T=1, V=treatment] - E[S_{i0} \mid T=0, V=treatment]\} \\ &- \{E[S_{i1} \mid T=1, V=control] - E[S_{i0} \mid T=0, V=control]\}\end{aligned}.$$

Econometrically, this difference-in-difference (DID) model can be expressed as follows,

(4) $$S_{i,t} = \alpha + \delta T_i + \sum_t \beta_t Y_t + \gamma_t (Y_t \times T_i) + \sum_j \varphi_j X_{i,t} + \varepsilon_{i,t}.$$

Where $S_{i,t}$ is the enrollment decision of individual *i* at time *t*, Y_i is an indicator variable for the enrollment year (year dummies), T_i is an indicator for if the village is a treatment or control village, $X_{i,t}$ is a set of observable characteristics, and $\varepsilon_{i,t}$ are the unobservable determinants of school enrollment.

With repeated observations on an individual, we can augment the model (4) above to incorporate individual fixed-effects, η_i,

(5) $$S_{i,t} = \alpha + \sum_t \beta_t Y_t + \gamma_t (Y_t \times T_i) + \sum_j \varphi_j X_{i,t} + \eta_i + \varepsilon_{i,t}.$$

This will help control for any omitted characteristics, such as school ability and other preference characteristics that remain fixed over time.[2]

[2] With non-experimental data, fixed-effects are a useful way of controlling for unobserved characteristics that may bias the treatment effect. In principle, the randomization should remove any correlation between the treatment effect and unobserved characteristics, in which case fixed-effects serve only to increase precision.

An advantage of this model is that it controls for any time-invariant determinant of school enrollment, as well as any time varying factor that equally impacts both treatment and control villages. Thus, the key assumption underlying this difference-in-difference model is that in the absence of PROGRESA, school enrollment rates in the treatment villages grow at the same rate as school enrollment rates in the control villages. While the random assignment between villages does provide some justification for the counterfactual assumption, it is by no means a sufficient condition. An important part of the identification procedure will be to find indirect tests for the validity of this assumption.

Data Source

Our empirical strategy uses individual level data on a child's decision to enroll into secondary school from 1997 to 1999. Here, we describe the data used for this study and our motivation for exploring these particular grade levels. We also compare mean attributes across treatment and control villages to evaluate the randomization of our sample.

Our empirical analysis is based primarily on data collected for the education component of PROGRESA. Educational subsidies are provided to mothers, contingent on their children's regular attendance to school. These cash transfers are available for each child attending school in any of the three upper grades of primary school or the first three grades of secondary school (lower-secondary school). The transfers increase with grade level and are higher for girls than for boys. These cash transfers range from $200 to $255 pesos for a child in secondary school, which is roughly half of what a child would earn if working full time,[3] with a maximum of $625 per month for the family in 1998. Overall, PROGRESA transfers are important, representing 22% of the income of beneficiary families.

The data used for this research include the 1997 pre-treatment census, and the follow-up surveys in October 1998 and November 1999. We thus have information on enrollment during three consecutive school years 1997-98, 1998-99, and 1999-2000. For the econometric analysis, we restrict our interest to the decisions to enroll into secondary

[3] The average daily wage of 16-18 years old in the sample is 25 pesos in 1997. A full time work of 20 days per month would generate an income of 500 pesos or $59 per month.

school since this is the most problematic decision for school attainment, and the grade levels that PROGRESA has its greatest impact (see Figure 2).[4] In our sample, this concerns approximately 19,500 children who are eligible to enter any of three lower secondary school grade levels.

Two important sample restrictions come from our use of the program's welfare index in the estimation strategy. First, the original welfare index, that classified 52% of the population as eligible was re-estimated in July of 1999 to account for a bias against the elderly poor who no longer lived with their children. As a result, some of the households that were originally considered as non-poor at the start of 1999 were later classified as poor. Second, because the welfare index was estimated by region, two regions had different threshold levels for classifying eligible households. While one of these regions only represented 1.2% of the sample, the other represented nearly 12%. We address these two issues by simply dropping these observations, and recognize that our inferences cannot be extended to these excluded regions.

For our final estimation sample, Table 1 reports differences in individual and household average characteristics between treatment and control villages for both beneficiary and non-beneficiary households. As Table 1 indicates, even among non-poor households, only 65% of the children were enrolled in secondary school during 1997. Fortunately, in this pre-program year the difference in enrollment rates between treatment and control villages is both small and statistically insignificant. Although it is statistically insignificant, we also see our first indication of a possible spillover effect. If the randomization were perfect and done at the individual level, then we could simply take the difference in the 1998 enrollment rates between treatment and control as our measure of the spillover effect. As reported, secondary school enrollment is 3 percentage points higher in the treatment villages than in the control villages among children of non-beneficiary families.

While many of the important determinants of school enrollment (such as head of household education and gender) are not statistically different between treatment and control villages, there are a few categories were the randomization did not eliminate the difference. There is a statistically significant difference in the proportion of indigenous

[4] In 1997, primary school enrollment was close to 96.5% compared to 65% enrollment into secondary school.

households between treatment and control localities, as well as a few welfare characteristics. These results are consistent with Behrman and Todd (1999)'s evaluation of the randomization of PROGRESA. Their study concluded that the random assignment was successful at the locality level, but cannot reject some significant differences in household characteristics between the treatment and control groups.

Empirical Results

In this section, we present evidence consistent with the possibility of a spillover effect in education. After providing a few tests of our counterfactual assumption, we explore the possible mechanisms underlying the spillover effect.

Figure 3 presents suggestive evidence for the possibility of a spillover effect. This figure depicts enrollment rates in secondary school by the welfare index used to classify beneficiary households. The top two panels plot enrollment rates among the children of beneficiary households for the pre-treatment year 1997 and the post-treatment year 1998[5]. The top panels demonstrate the well-documented impact of PROGRESA on the enrollment of the poor. In the pre-intervention year, there is very little difference in enrollment rates between the treatment and control villages, as one would expect given the random assignment. Also, school enrollment clearly increases by welfare level. In comparison, post-treatment enrollment in the treatments villages appears approximately 8 percentage points higher than in the control villages. In the difference-in-difference framework, the corresponding impact of the program is thus slightly less than 8 percentage points given the very slight advantage of the treatment villages in the pre-intervention year. This estimate is consistent with those found in Schultz (2002) and Behrman, Sengupta, and Todd (2001). Also note that the impact of the program is fairly uniform across welfare levels among this beneficiary population.

Similarly, the bottom two panels depict enrollment rates of the treatment and control villages for the non-beneficiary groups. Again, we see little difference in participation rates between treatment and control villages in the pre-treatment year. There does however, appear to be a more distinct association between enrollment and welfare

[5] Enrollment in 1999 are not shown but are incorporated in the estimation of the difference-in-difference model.

status compared to the beneficiary sample. In the post-treatment year the difference in enrollment rates between treatment and control is concentrated at the threshold level and decreases by welfare. Figure 3 suggests that a spillover effect maybe encountered among the non-eligible households that are the most similar to the beneficiary households in terms of welfare status; those near the threshold.

Table 2 reports estimates of the effect of PROGRESA on the school enrollment of the beneficiary population using our difference-in-indifference model with fixed effects. To verify the relationship between the program's treatment effect and welfare levels, the estimation was done by welfare percentile groupings. The estimates are fairly consistent across intervals, largely confirming the nonparametric estimates presented in Figure 2. In 1999, the impact of the program is slightly diminished from the impact of 1998, but is once again fairly consistent across each quintile. Interestingly, when we estimate our model for those children in the decile closest to the program eligibility threshold, we find that the effect of the program is close to 10 percentage points. This implies that PROGRESA is most effective for families with a low opportunity cost for attending school. The size of the transfers may not be sufficient to induce the poorest households to send their children to school.

Estimates of the impact of PROGRESA on school participation of the *non-beneficiary* population are presented in Table 3, again using the difference-in-difference model with fixed-effects[6]. Since a spillover effect is more likely to be encountered near the eligibility threshold, the model is also re-estimated by welfare terciles above the threshold.[7] Using the entire non-beneficiary sample, there is a slight yet insignificant spillover effect, which is consistent with the results Behrman, Sengupta, and Todd (2001) found. In 1998, the non-beneficiaries in treatment villages were 1.1 percentage points more likely to enroll into secondary school than non-beneficiaries in control villages. When we decompose the sample by welfare terciles, we notice the interesting result that was demonstrated in Figure 3. For the first welfare tercile, we estimate a spillover effect of 9.9 percentage points, and

[6] A simple linear probability model with pooled observations was also estimated. The estimates are similar when a full set of controls is included.

[7] To assess the robustness of sample partitioning, we estimated the model by quartiles. We found that the results were robust to this partition: there is an estimated increase of 10 percent in enrollment rates for non-beneficiary children in the first welfare-level quartile, and no significant spillover effect for upper quartiles in 1998. We also partitioned the sample by deciles, but the sample sizes became too small and constrained the power of the estimates.

statistically significant at conventional significance levels. Thus, PROGRESA increased the secondary school participation of *non-beneficiary* children by 15.2%, which is similar to the treatment effect among the beneficiary households near the threshold. This point estimate is robust to controlling for other time varying characteristics, such as the child's age, and grade level. Beyond the first tercile, the point estimate starts to decline and is no longer statistically significant, again mirroring our non-parametric estimation. By 1999, the spillover effect has declined to 3.9 % but is no longer statistically significant.

There are several possible mechanisms to explain this possible spillover effect, but before discussing these issues, it is important to note that this estimate depends on an unverifiable identifying assumption: in the absence of PROGRESA, school enrollment rates in the treatment villages would have grown at the same rate as school enrollment rates in the control villages. Below, we show evidence that this assumption is reasonable

Test of counterfactual

A critical aspect of this exercise is to provide indirect evidence that our counterfactual assumption is plausible. A convincing test of the counterfactual would be to use a times series of data prior to the start of the program to show that there are no significant differences in secondary school enrollment rates between treatment and control villages.[8] Unfortunately, with only one year of pre-treatment data such an analysis is currently not possible. Instead, we compare differences in educational attainment between treatment and control villages by age cohort. While educational attainment does not necessarily correspond to school enrollment if there was frequent grade repetition, large differences between treatment and control villages would cast some doubt on our assumption. For this to be a valid test, we would also need to assume that a majority of the families did not migrate into or out of the community after attaining their education. With these caveats in mind, Figure 4 depicts differences in cohort grade attainment between

[8] To further investigate the underlying assumptions of our difference-in-difference model, we will try to gather information on economic shocks prior to 1997. For this we plan on using another data source, INEGI, which contains several economic and social indicators for all of the PROGRESA localities. If it is possible to show that school enrollment or other economic indicators changed in similar manners between the treatment and control villages, this will provide robustness to the reported results.

treatment and control village along with the 95 percent confidence intervals. We see that there are few significant differences in educational attainment between treatment and control villages, and most of these differences occur prior to the cohort born in 1933. Again, in the absence of a time series of pre-treatment enrollment data, Figure 4 only provides some weak justification for our assumption.

In another test of our assumption we see if there exist any major shocks that have occurred during the course of the program. Table 6 provides some evidence consistent with our assumption. It displays several types of exogenous shocks to each community during 1998 and 1999. On average, treatment localities were not subjected to more frequent shocks than control localities and the difference in the occurrence of these shocks is also not significantly different from zero. Unfortunately, the lack of 1997 data also limits the power of this test.

Heterogeneity and insights in possible spillover mechanisms

In this section, we outline several hypotheses to explain the spillover effect. To test several potential mechanisms, we investigate variation in the spillover effects across certain household characteristics.

Table 4 presents different estimates of the spillover effect by various subgroups. Exploring the possible heterogeneity can provide insights into the various mechanisms that contribute to this externality. For example, one might expect to see a greater spillover effect among girls if the increase in general enrollment provides greater safety among children who must travel greater distances to school. Similarly, if PROGRESA reduces travel costs for beneficiary children, then it is conceivable that cost may be shared among non-beneficiaries as well. In this case, the impact of the program may be greater for children without immediate access to a school.

The first three columns in Table 4 report the estimation split by subgroup and for the entire sample. In case that these divisions simply capture welfare differences, the estimation is redone for the first tercile of the welfare index, as well. The first set of rows in the table show that contrary to the result reported in Handa et al (2001), the spillover effect is much larger among boys than among girls.[9] Thus, the data rejects the hypothesis that girls are increasing their enrollment because of such possibilities as safety issues.[10] In fact, for those households within the first welfare tercile above the threshold, PROGRESA increased school enrollment among these boys by 15.5 percentage points. For comparison Table 5 presents the treatment effect of the program estimated within the same subgroups. As intended, the impact of PROGRESA was higher among girls than boys, with an increase in enrollment of 9 percentage points for households near the welfare threshold.[11]

We also explore the difference in spillover effect between those children that have access to a school in their village and those that must travel more than 1.5 km. Here, we find that both the spillover effect and the treatment effect are largest among those children who do not have immediate access to a secondary school. This result is consistent with the hypothesis that sharing travel costs induces higher enrollment. We also compare treatment and spillover effects by age cohort. Interestingly, we see that the impact of the program on both beneficiaries and non-beneficiaries households near the threshold is very similar among children in the 10-13 years old cohort. Large households with 10 members or more experience a considerable spillover effect in 1998 and these estimates are robust.

Spillover Mechanisms: Peer Effects and Alternative Hypotheses

Are households near the threshold engaged in a mutual exchange of transfers that increase enrollment among non-beneficiary children? Or, are the resources that PROGRESA provided to schools attracting more students? Or, is this simply a peer effect among the children from households near the threshold? Here, we explore various hypotheses and in the end argue that peer effects appear to be the most likely explanation for these spillover effects.

Our first hypothesis is to test whether economic and social relationships exist between beneficiary and non-beneficiary households in the PROGRESA villages. And if

[9] Handa et al (2001) found a larger spillover effect among girls, ages 10-12 years old. We also test this classification and still find contradictory evidence.
[10] We also tested to see if the spillover effect was larger for girls who lived far from a secondary school but found similar results.
[11] While it might appear strange that the treatment effect is smaller than the spillover, given the standard errors of the spillover effect we cannot reject that these point estimates are the same.

so, do beneficiary households provide large enough transfers to non-beneficiary households to induce an increase in secondary school enrollment? To test this hypothesis, we use data on the exchange of goods and labor between households within a village to construct a series of measures of the social relationships of non-beneficiary households in treatment and control villages. The October 1998 post-treatment household survey asks a series of questions on whether someone in the household *received* a transfer in cash, clothes, food, or labor, and the amount of cash transfers received. In addition, the survey collects data on whether the individual making the transfer lives in the same village as the respondent, his/her name, gender, age, and educational level. Using this transfer data for each household, we match the transfer provider identity data to his/her household's respective beneficiary status and welfare level.

Table 7 presents a decomposition of the number of links of non-beneficiary households by their welfare terciles and a percentile decomposition of their respective transfer links. Surprisingly, approximately half of these households' links are with program beneficiaries (52% among treatment villages and 42% among control villages). These informal financial bonds between non-eligible and eligible households suggest that these households interact at a very personal and social level. However, since the number of households participating in these transfer relationships is small, within-village inter-household transfers cannot be driving the large increase in boys' enrollment in the identified non-beneficiary households. [12]

Table 8 reports average cash transfer to non-beneficiary households for treatment and control village. Approximately 2.7 percent of these households report receiving an informal transfer. There does not appear to be any significant differences in the total transfers received between households in treatment and control villages, neither for all non-beneficiary households nor for non-beneficiary households close to the welfare threshold. Although we cannot test whether the change in these incoming household transfers is significantly different, the evidence tends to support the hypothesis that there is no contextual effect through these informal arrangements. Finally we calculate how large the transfer would have to be to induce our large spillover effect. Based on the income-

[12] Out of 11,558 non-beneficiary households in the October 1998 household survey, only 1.6 percent (190 households) report receiving a transfer from another individual within the same village.

enrollment elasticity reported in Demombyenes (2002), a non-beneficiary would have to receive more than twice the average household income to induce such a spillover effect.[13]

Another possible hypothesis is that the increased school participation rates make it safer for girls to travel longer distances to attend secondary school. As mention above, our data clearly rejects this hypothesis since the spillover effect is strictly among boys.

Without the support of additional data, there exist at least four important alternative explanations that we currently cannot reject. The first possibility is that the program's externality effects are merely the result of some supply-side interventions. In order to prevent the deterioration of school quality that might result from a sudden increase in school participation, the program provided additional resources to the PROGRESA schools. However, if better schools were available to the entire community, why is there so much heterogeneity in the impact and particularly near the welfare threshold? If this program externality is really due to supply-side effects, then these results are also interesting given the observed heterogeneity and the general lack of evidence supporting the effects of school quality on child performance (see Hanushek and Betts).

Positive health externalities are another possible explanation for the increase in enrollment among non-beneficiaries. However, in an impact analysis of the program on the health of children and adults, Gerlter (2000) does not find any effects on either the health care utilization or the self-reported health status measures of beneficiary children aged 6 to 17. Thus we can reject this possibility because we would expect any positive health externality to occur among children of the same age group. Nonetheless, the improved health of young siblings ages 0 to 5, who experience improvements in health as a result of PROGRESA (Gertler (2000)), might allow older sisters to enroll in secondary school. Again since the identified spillover is only for boys, this hypothesis is also rejected.

A fourth possibility is that households are forming expectations about possible future program participation and might believe that they will be rewarded with PROGRESA if they start to enroll their children. Although, this possibility seems far-fetched, it is consistent with our lack of spillover effects in 1999.

[13] Demombynes (2002) find an income-enrollment elasticity of 6.2%.

A final explanation can be that, by increasing school participation in these villages, transportation costs may have been reduced significantly, inducing more children to attend secondary school. This hypothesis could be especially relevant, since we find a significant spillover effect in villages that are distant to a secondary school.

Conclusion

In the marginal rural communities of Mexico, 40% of the children from poor households discontinue school with only a primary level education. As a result the Mexican government, in 1997, launched the ambitious Program for Education, Health, and Nutrition (PROGRESA) aimed at developing the human capital of these poor rural households. While recent studies of PROGRESA have documented a significant impact among the targeted population, they fail to account for the program's externality benefits.

This paper examines whether PROGRESA increased the secondary school participation among the children of the non-beneficiary households. Exploiting both the experimental design of the program and its panel data structure, we estimate a difference-in-difference model with fixed-effects. Our results show that the external benefits of PROGRESA were concentrated among the non-eligible households near the program's eligibility threshold. Among the school children of these households, PROGRESA increased secondary school enrollment by 15.2 percent, an impact comparable to the program's effect on the treated. Exploring the variation in this impact, we find that the spillover effect was only among boys and most pronounced among the 10-13 year-old cohort.

The heterogeneity of the impact along with other data provides some insight into the possible mechanisms underlying this spillover effect. We find that the data is inconsistent with several hypotheses. We can reject that girls are more likely to enroll due to safety concerns, a hypothesis previously put forth in the literature. And, it is also unlikely that health externalities are the cause. We can also reject that intra-villages transfers between beneficiary and non-beneficiary households are responsible for the increased enrollment.

However due to several data constraints, we cannot reject two plausible mechanisms. First, although the spillover effect is concentrated among only boys from households near the threshold, the provision of additional schooling resources might explain this increase in enrollment. Second, our data are consistent with the possibility that the travel costs of non-beneficiaries have been reduced, and thus leading to the spillover effect.

While many questions have been left unanswered, this study is merely a first step in a broader research plan. More work and additional data sources are needed to either reject our peer effects story or rule out the other possibilities. Our future work agenda consists of in-depth case studies of these PROGRESA communities to complement this current analysis.

References

Behrman, Jere R., Piyali Sengupta, and Petra E. Todd. (2001). "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment". Mimeo, University of Pennsylvania.

Behrman, Jere R. and Petra E. Todd. (1999). "Randomness in the Experimental Samples of PROGRESA (Education, Health, and Nutrition Program)". International Food Policy Research Institute.

Case, Anne C. and Lawrence F. Katz (1991). "The Company You Keep: The Effect of Family and Neighborhood on Disadvantaged Youths." National Bureau of Economic Research Working Paper No. 3705.

Duflo, Esther and Emmanuel Saez. (2000). "Participation and Investment Decisions in a Retirement Plan: The Influence of Colleague's Choices." Mimeo, MIT.

Edward, Miguel and Michael Kremer. (2001). "Social Networks and Learning about Health," mimeo, University of California at Berkeley.

Edward, Miguel and Michael Kremer. (2001). "Worms: Education and Health Externalities in Kenya". NBER working paper 8481.

Handa, Sudhanshu, Mari-Carmen Huerta, Raul Perez, and Betriz Straffron. (2001) "Poverty, Inequality, and Spillover in Mexico's Education, Health, and Nutrition Program." FCDN Discussion Paper 101. International Food Policy Research Institute.

Manski, Charles F. (2000). "Economic Analysis of Social Interactions." NBER Working Paper 7580.

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

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

Schultz, T. Paul. (2002). "School Subsidies for the Poor: Evaluating A Mexican Strategy For Reducing Poverty." (Forthcoming in *Journal of Development Economics*).

Soukfias, Emmanuel, Benjamin Davis, and Sergio de la Vega. (2001). "Targeting the Poor in Mexico: An Evaluation of the Selection of Households into PROGRESA." *World Development*. 29(10). 1769-1784.

Townsend, Robert M. (1994). "Risk and Insurance in Village India." *Econometrica*, 62.

Udry, C. (1994). "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria." *Review of Economic Studies*, 61

Table 1. Treatment and Control Groups Comparison

	Beneficiaries			*Non-beneficiaries*		
	Control	Treatment	Difference	Control	Treatment	Difference
Enrollment 1997	0.559	0.574	-0.016	0.646	0.654	-0.008
	(0.012)	(0.009)	(0.015)	(0.018)	(0.015)	(0.023)
Enrollment 1998	0.565	0.612	-0.047**	0.615	0.647	-0.031
	(0.014)	(0.010)	(0.017)	(0.021)	(0.016)	(0.027)
Grade 1997	6.486	6.479	0.007	6.654	6.676	-0.022
	(0.018)	(0.013)	(0.022)	(0.031)	(0.025)	(0.040)
Gender	0.491	0.524	-0.033**	0.486	0.507	-0.021
	(0.012)	(0.009)	(0.015)	(0.019)	(0.015)	(0.024)
Indigenous	0.388	0.372	0.016	0.114	0.154	-0.040**
	(0.012)	(0.009)	(0.015)	(0.012)	(0.011)	(0.017)
Age	13.891	13.860	0.031	13.883	13.891	-0.008
	(0.034)	(0.026)	(0.042)	(0.055)	(0.046)	(0.072)
Family Characteristics						
Family size	7.499	7.473	0.027	7.089	6.802	0.287**
	(0.056)	(0.041)	(0.068)	(0.090)	(0.072)	(0.115)
Number of kids	3.869	3.890	-0.020	3.307	3.277	0.030
	(0.034)	(0.026)	(0.043)	(0.057)	(0.044)	(0.072)
Mother’s Characteristics						
Age	39.850	39.767	0.082	42.818	42.839	-0.021
	(0.194)	(0.143)	(0.239)	(0.322)	(0.253)	(0.410)
Education	2.318	2.235	0.083	3.003	3.194	-0.191
	(0.058)	(0.041)	(0.070)	(0.113)	(0.094)	(0.148)
Indigenous	0.431	0.440	-0.009	0.131	0.214	-0.082***
	(0.012)	(0.009)	(0.015)	(0.014)	(0.013)	(0.020)
Father’s Characteristics						
Age	44.652	44.296	0.356	47.672	47.680	-0.008
	(0.233)	(0.168)	(0.283)	(0.380)	(0.300)	(0.485)
Education	2.745	2.700	0.045	3.328	3.517	-0.189
	(0.061)	(0.046)	(0.076)	(0.127)	(0.100)	(0.161)
Indigenous	0.438	0.439	-0.001	0.149	0.235	-0.086***
	(0.013)	(0.010)	(0.016)	(0.015)	(0.014)	(0.021)
Head of household characteristics						
Education	2.555	2.590	-0.035	3.156	3.390	-0.233
	(0.057)	(0.043)	(0.071)	(0.115)	(0.094)	(0.150)
Gender (Male)	0.919	0.927	-0.007	0.923	0.937	-0.015
	(0.007)	(0.005)	(0.008)	(0.010)	(0.007)	(0.012)

Age	45.908	45.021	0.886***	48.815	48.906	-0.091
	(0.267)	(0.190)	(0.322)	(0.414)	(0.321)	(0.521)
Village Characteristics						
Minimum distance to urban center	109.956	112.507	-2.550**	92.047	93.920	-1.873
	(1.020)	(0.717)	(1.220)	(1.599)	(1.203)	(1.976)
Distance to the capital	173.796	166.664	7.132***	124.320	132.846	-8.526***
	(1.96)	(1.36)	(2.33)	(2.27)	(2.04)	(3.14)
Distance to a secondary school	2.134	2.386	-0.252***	2.074	2.147	-0.073
	(0.047)	(0.036)	(0.059)	(0.086)	(0.053)	(0.095)
Dwelling Characteristics						
Dirt floor	0.721	0.692	0.029**	0.180	0.245	-0.065***
	(0.011)	(0.008)	(0.014)	(0.015)	(0.013)	(0.020)
Rooms	1.758	1.762	-0.003	2.609	2.583	0.026
	(0.023)	(0.019)	(0.030)	(0.049)	(0.040)	(0.063)
Water	0.252	0.329	-0.077***	0.485	0.562	-0.078***
	(0.010)	(0.009)	(0.014)	(0.019)	(0.015)	(0.024)
Bathroom	0.652	0.621	0.031**	0.765	0.694	0.071***
	(0.011)	(0.009)	(0.015)	(0.016)	(0.014)	(0.022)
Electricity	0.691	0.648	0.043***	0.949	0.888	0.061***
	(0.011)	(0.009)	(0.014)	(0.008)	(0.010)	(0.014)

Group means: standard deviations in parentheses. Differences: standard errors in parentheses.

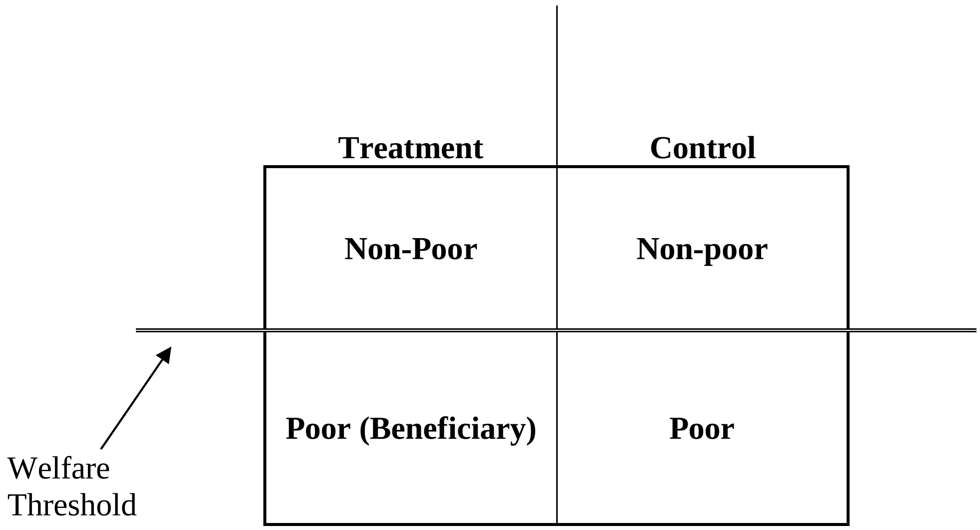
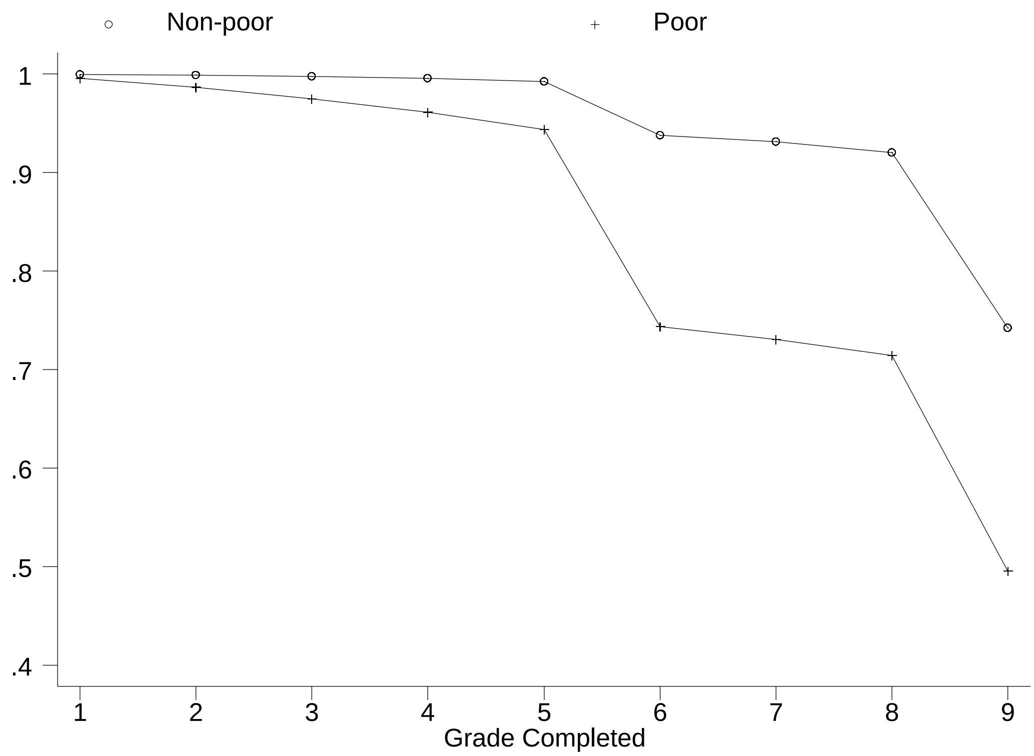


Figure 1: Program Evaluation Design


Non-poor
Poor
1
.9
.8
.7
.6
.5
.4
1
2
3
4
5
6
7
8
9
Grade Completed

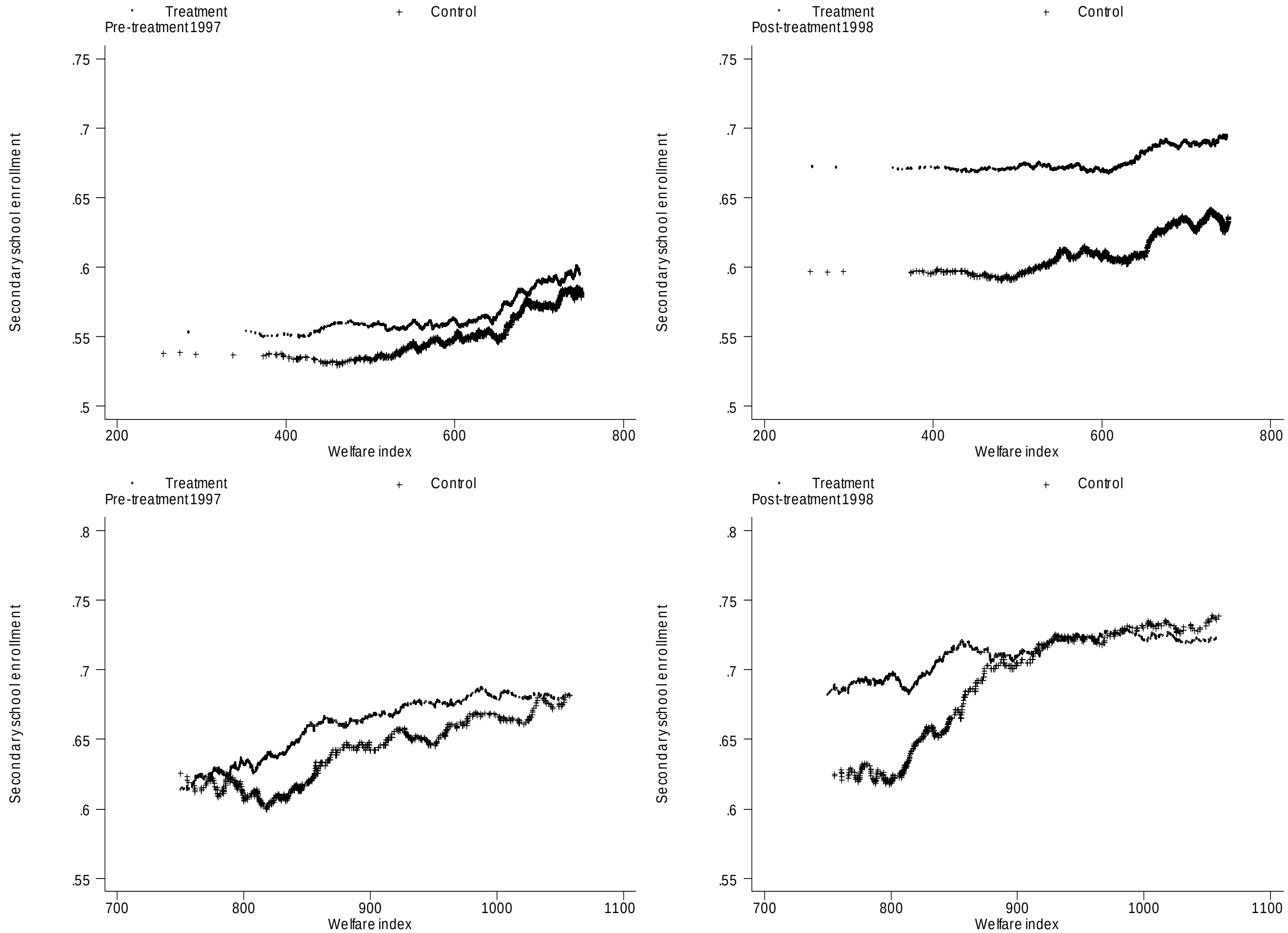


Figure 3: Nonparametric Estimates of Enrollment Levels by Program Classification, 1997 & 1998

Table 2: PROGRESA Treatment Effects on Beneficiary Households, decomposed by welfare levels
Dependent Variable: Enrollment in secondary school (yes=1/no=0)

	All Beneficiary Hhs Welfare index: 254-750		**1st-40th percentile** Welfare index: 254-632		**41st-80th percentile** Welfare index: 632.4-718		**81st-100th percentile** Welfare index: 718.14-750		**90th-100th percentile** Welfare index: 718.14-750	
Treatment 1998	0.051***	0.048***	0.059***	0.057**	0.041*	0.038*	0.056**	0.052*	0.096***	0.094***
	(0.013)	(0.013)	(0.023)	(0.022)	(0.021)	(0.021)	(0.027)	(0.027)	(0.035)	(0.034)
Treatment 1999	0.013	0.007	0.020	0.009	-0.019	-0.022	0.058**	0.051*	0.078**	0.074**
	(0.014)	(0.014)	(0.024)	(0.024)	(0.023)	(0.023)	(0.029)	(0.029)	(0.037)	(0.037)
Controls	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Num obs	15449	15449	6185	6185	6117	6117	3147	3147	1759	1759
R-sq: within	0.0041	0.0151	0.0054	0.0199	0.0041	0.0131	0.0087	0.0191	0.018	0.0261
Between	0.0016	0.3888	0.0068	0.3433	0.0019	0.387	0.008	0.4005	0.002	0.3203
Overall	0.0028	0.3465	0.0061	0.3021	0.0001	0.3435	0.0083	0.3644	0.0032	0.2948

* Standard errors reported in parentheses. Controls include age and grade level of the child, in addition to the year dummies.

Table 3: PROGRESA Heterogeneous Spillover Effects on Non-beneficiaries, by terciles (excluding top 15 percent of households)
Dependent Variable: Enrollment in secondary school (yes=1/no=0)

	All Non-Beneficiary Hhs Welfare index: 750-1025.6		**1st-33rd percentile** Welfare index: 750-822		**34th-66th percentile** Welfare index: 822-891.33		**67th-100th percentile** Welfare index: 891.33-1025.6	
Spillover 1998	0.014	0.011	0.107**	0.099**	-0.036	-0.037	-0.011	-0.013
	(0.024)	(0.024)	(0.043)	(0.043)	(0.043)	(0.043)	(0.040)	(0.040)
Spillover 1999	-0.020	-0.025	0.039	0.030	-0.052	-0.053	-0.030	-0.034
	(0.026)	(0.026)	(0.047)	(0.047)	(0.046)	(0.046)	(0.045)	(0.045)
Controls	NO	YES	NO	YES	NO	YES	NO	YES
Num obs	4009	4009	1325	1325	1365	1365	1319	1319
R-sq: within	0.0083	0.0195	0.0256	0.0358	0.0117	0.0239	0.004	0.0132
Between	0.0017	0.5163	0.0012	0.3874	0	0.4872	0.007	0.5076
Overall	0.0025	0.448	0.0039	0.3544	0.0008	0.4167	0.0052	0.4414

* Standard errors reported in parentheses. Controls include age and grade level of the child, in addition to the year dummies.

Table 4: Decomposition of Spillover Effects by Various Individual and Village Characteristics

	Entire Sample			1^{st}-33^{rd} percentiles		
	Sample size	Spillover in 1998	Spillover in 1999	Sample size	Spillover in 1998	Spillover in 1999
Male	1983	0.035	-0.057	681	0.155***	-0.035
		(0.033)	(0.036)		(0.058)	(0.063)
Female	2029	-0.013	.007	646	0.013	0.074
		(0.034)	(0.038)		(0.062)	(0.068)
School in village	1122	0.018	0.030	279	0.007	0.029
		(0.039)	(0.044)		(0.105)	(0.108)
No school within 1.5km	2455	0.025	-0.056	882	0.151***	0.065
		(0.033)	(0.035)		(0.049)	(0.054)
Age cohort						
10-13	1590	0.044	0.024	515	0.163**	0.062
		(0.037)	(0.052)		(0.078)	(0.109)
14-17	2420	0.013	-0.031	812	0.066	-0.048
		(0.040)	(0.042)		(0.061)	(0.066)
Family size						
≤ 6	1940	-0.033	-0.045	648	0.046	0.017
		(0.033)	(0.362)		(0.061)	(0.067)
$10 \geq$	568	0.132*	0.027	210	0.216*	0.042
		(0.074)	(0.080)		(0.126)	(0.139)

* T-statistics reported in parentheses. All regressions control for age and grade levels, in addition to the time effects.

Table 5: Decomposition of Treatment Effects by Various Individual and Village Characteristics						
	Entire Sample			90^{th}-100^{th} percentiles		
	Sample size	Treatment in 1998	Treatment in 1999	Sample size	Treatment in 1998	Treatment in 1999
Male	7909	0.049***	0.007	870	0.081*	0.013
		(0.0187)	(0.020)		(0.047)	(0.050)
Female	7579	0.049***	0.010	891	0.099*	0.123***
		(0.019)	(0.021)		(0.051)	(0.054)
School in	3604	0.039	0.039	449	0.077	0.138**
village		(0.254)	(0.028)		(0.062)	(0.069)
No school	9874	0.038**	-0.020	1121	0.099**	0.066
within 1.5 km		(0.017)	(0.185)		(0.045)	(0.047)
Age cohort						
10-13	6027	0.084***	0.0861***	680	0.146***	0.256***
		(0.024)	(0.032)		(0.061)	(0.081)
14-17	9459	0.023	-0.059	1080	0.941**	0.027
		(0.020)	(0.211)		(0.048)	(0.051)
Family size						
≤ 6	5358	0.086***	0.001	775	0.081	0.053
		(0.022)	(0.023)		(0.054)	(0.058)
$10 \geq$	2586	-0.031	0.017	272	0.087	0.121*
		(0.033)	(0.036)		(0.065)	(0.069)

* T-statistics reported in parentheses. All regressions controlled for age and grade level.

Table 6: Observable Exogenous Shocks in 1998 & 1999 (Village level)

	1998			1999		
	Treatment	Control	Difference	Treatment	Control	Difference
Drought	0.703	0.726	-0.023	0.748	0.728	0.019
	(0.458)	(0.447)	(0.042)	(0.435)	(0.446)	(0.041)
Flood	0.100	0.075	0.025	0.006	0.016	-0.010
	(0.300)	(0.265)	(0.027)	(0.080)	(0.127)	(0.009)
Icestorm	0.094	0.129	-0.035	0.195	0.163	0.032
	(0.292)	(0.336)	(0.028)	(0.397)	(0.370)	(0.036)
Fire	0.159	0.134	0.025	0.048	0.065	-0.017
	(0.367)	(0.342)	(0.033)	(0.214)	(0.248)	(0.021)
Plague	0.350	0.360	-0.010	0.201	0.179	0.022
	(0.478)	(0.481)	(0.044)	(0.402)	(0.385)	(0.037)
Earthquake	0.006	0.022	-0.015	0.093	0.087	0.006
	(0.079)	(0.145)	(0.010)	(0.290)	(0.283)	(0.027)
Hurricane	0.038	0.048	-0.011	0.022	0.022	0.001
	(0.190)	(0.215)	(0.018)	(0.148)	(0.146)	(0.014)
Other	0.013	0.011	0.002	0.003	0.016	-0.013
	(0.111)	(0.103)	(0.010)	(0.057)	(0.127)	(0.008)
Number obs	320	186		313	184	

	Difference 1999-1998		
	Treatment	Control	Difference
Drought	0.038	0.000	0.038
	(0.492)	(0.490)	(0.046)
Flood	-0.093	-0.060	-0.033
	(0.312)	(0.280)	(0.028)
Icestorm	0.102	0.038	0.064
	(0.448)	(0.436)	(0.041)
Fire	-0.112	-0.071	-0.041
	(0.405)	(0.419)	(0.038)
Plague	-0.147	-0.179	0.032
	(0.575)	(0.587)	(0.054)
Earthquake	0.086	0.065	0.021
	(0.292)	(0.307)	(0.028)
Hurricane	-0.013	-0.027	0.014
	(0.211)	(0.244)	(0.021)
Other	-0.010	0.005	-0.015
	(0.126)	(0.165)	(0.013)
Number obs	320	186	

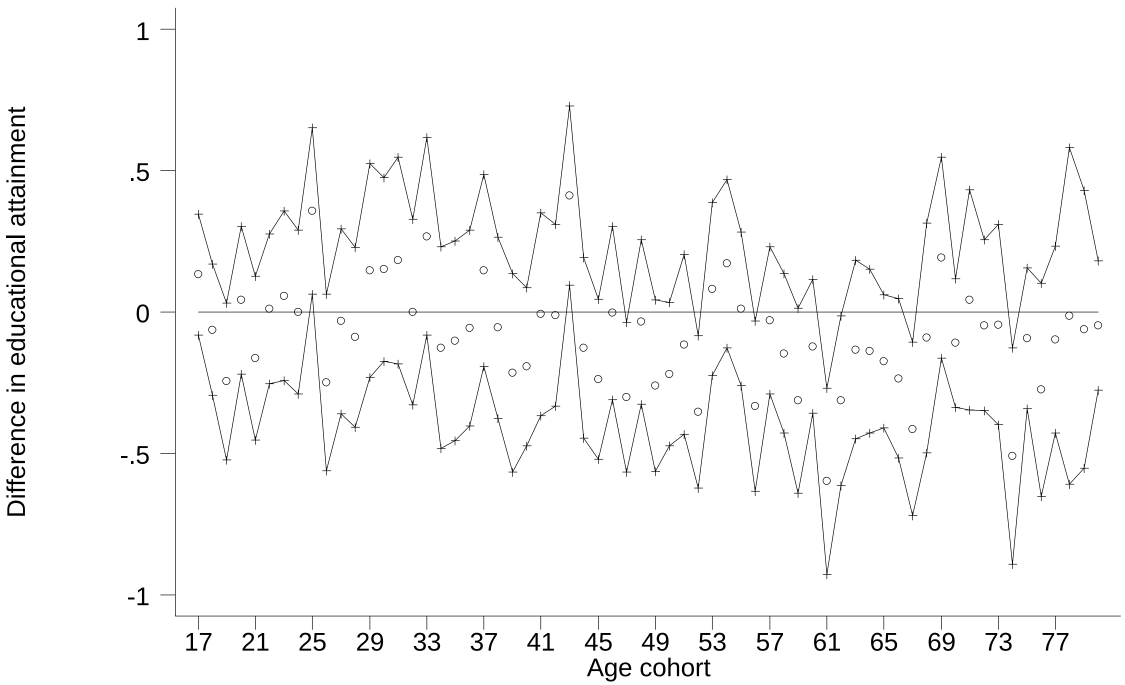


Figure 4: Differences in educational attainment between treatment and control villages

Table 7: Number of Within-Village Links for Non-Beneficiaries, by Welfare Levels

Treatment Villages

	Non-Beneficiaries		
Transfer Links	1st tercile	2nd tercile	3rd tercile
Beneficiaries			
1st - 40th percentile	20	3	3
41st-80th percentile	13	10	7
81st-100th percentile	6	3	1
Total	39	16	11
Non-Beneficiaries			
1st tercile	21	12	3
2nd tercile	11	4	0
3rd tercile	6	3	0
Total	38	19	3

Control Villages

	Non-Beneficiaries		
Transfer Links	1st tercile	2nd tercile	3rd tercile
Beneficiaries			
1st - 40th percentile	7	3	1
41st-80th percentile	8	5	2
81st-100th percentile	0	0	1
Total	15	8	4
Non-Beneficiaries			
1st tercile	10	6	2
2nd tercile	1	3	0
3rd tercile	5	5	5
Total	16	14	7

Table 8: Differences in Inter-household Transfers Received in 1998 among villages: Non-beneficiary households

Dependent Variable: Total Cash Transfers Received in the Last Month

	All non-beneficiary households				First welfare tercile of non-beneficiaries			
	Treatment	**Control**	**Difference**	**Number obs**	**Treatment**	**Control**	**Difference**	**Number obs**
All cash transfers								
All non-beneficiary households	23.09	21.64	1.45	2804	31.90	35.22	-3.33	779
			(9.09)				(22.39)	
Households with positive transfers	920.38	788.71	131.67	73	1297.2	1025.0	272.2	22
			(198.81)				(577.0)	
Within-village cash transfers								
All non-beneficiary households	0.18	0.58	-0.40	2804	-	-	-	-
			(0.36)					
Households with positive transfers	150.0	162.5	-12.5	6	-	-	-	-
			(74.80)					
Outside-village cash transfers								
All non-beneficiary households	22.91	21.06	1.85	2804	31.90	35.22	-3.33	779
			(9.06)				(22.39)	
Households with positive transfers	958.90	850	108.9	68	1297.2	1025.0	272.2	22
			(206.91)				(577.0)	

Standard errors in parentheses.