NEIGHBORHOOD PEER EFFECTS IN SECONDARY SCHOOL ENROLLMENT DECISIONS

Gustavo J. Bobonis and Frederico Finan*

Abstract—This paper identifies neighborhood peer effects on children's school enrollment decisions using experimental evidence from the Mexican PROGRESA program. We use exogenous variation in the school participation of program-eligible children to identify peer effects on the schooling decisions of ineligible children residing in treatment communities. We find that peers have considerable influence on the enrollment decisions of program-ineligible children, and these effects are concentrated among children from poorer households. These findings imply that policies aimed at encouraging enrollment can produce large social multiplier effects.

I. Introduction

LOW secondary school enrollment rates remain an important concern for much of the developing world. Despite significant improvements over the past forty years, secondary school enrollment rates in 2000 were only 54% among low-income countries (Glewwe & Kremer, 2005). Given that education fosters growth and improves welfare, promoting secondary school enrollment represents an important policy issue.[1] To design appropriate and effective policies as a redress for low enrollment rates, it is necessary to understand individuals' decisions to enroll in secondary school.

Although many factors affect the decision to enroll in secondary school, recognition is increasing that individuals' neighborhoods or communities influence their educational attainment. Residents of poor neighborhoods tend to attain lower educational levels and fare substantially worse on a wide range of socioeconomic outcomes than individuals living in more affluent ones, in both developed and developing country settings (Case & Katz, 1991; Kling, Liebman, & Katz, 2007; Gray-Molina, Perez de Rada, & Jiménez, 2003; Sanchez-Peña, 2007).[2] Several existing theories attempt to explain why residential location may affect an individual's schooling outcomes. For instance, a child's decision to enroll in school may be influenced by a desire to conform with others in his or her reference group due to peer pressure or social norms (Bernheim, 1994; Akerlof, 1997; Akerlof & Kranton, 2002; Glaeser & Scheinkman, 2003). Additionally, there may be informational externalities as individuals learn about the benefits of schooling from the actions of their peers (Bikhchandani, Hirschleifer, & Welch, 1992). Finally, social interactions may generate important strategic complementarities in student learning and teachers' effort (Kremer, Miguel, & Thornton, 2009; Lazear, 2001), which may attract students to school.[3] Thus, neighborhood-level social interactions could play an important role in an individual's schooling decision process. Understanding these effects can lead to policies that encourage the internalization of these interactions, making human capital investments more efficient (Bénabou, 1993; 1996). However, to our knowledge, existing empirical research has not opened the black box of neighborhood interactions to understand how particular behaviors of neighbors influence individuals' schooling decisions.[4]

In this paper, we use evidence from a human development program in rural Mexico to examine the role of neighborhood-level behavioral social interactions on a child's decision to enroll in secondary school. The PROGRESA program, initiated by the Mexican government in 1997, provides cash transfers to marginalized households in rural areas. The transfer is paid to mothers contingent on their children's primary and secondary school attendance and family visits to health services. The 506 communities selected to participate in an experimental evaluation of the program were randomly divided into two groups, with the treatment group being phased in to the program in March–April 1998 and the control group in November–December 1999. Within these selected communities, a poverty indicator was constructed at baseline to classify eligible and

Received for publication May 18, 2007. Revision accepted for publication March 28, 2008.

* Department of Economics, University of Toronto (Bobonis); Department of Economics, University of California, Berkeley, and NBER (Finon).

We are grateful to Josh Angrist, David Card, Ken Chay, Alain de Janvry, Weili Ding, Chris Ferrall, John Hoddinott, Caroline Hoxby, Asim Khwaja, David S. Lee, Steve Lehrer, Craig McIntosh, Rob McMillan, Ted Miguel, Elisabeth Sadoulet, Aloysius Siow, T. Paul Schultz, Duncan Thomas, the editor, and two anonymous referees, whose suggestions greatly improved the paper. We also thank seminar participants at the University of California at Berkeley, Queen's University, University of Toronto, CIRPÉE, and NEUDC 2003 and 2005 Conferences for helpful comments. We thank Caridad Araujo, Paul Gertler, Sebastián Martínez, Iliana Yaschine, and the staff at Oportunidades for providing administrative data and for their general support throughout. Bobonis acknowledges financial support from the Institute of Business and Economics Research at the University of California at Berkeley and NICHD Training Grant (T32 HD07275). Finan acknowledges financial support from the Social Science Research Council.

[1] School enrollment is perhaps a necessary but not a sufficient condition for improving education attainment. Low school quality remains an important obstacle for education attainment in developing countries (Banerjee et al., 2007).

[2] On the other hand, Oreopoulos (2003) uses quasi-experimental variation in assignment to different types of public housing units in Toronto and finds no long-term neighborhood effects on individuals' labor market outcomes.

[3] Also, resources for local public goods such as schools may be limited by the resources available to community residents or the capacity of residents to attract and direct government funding toward these (Benabou, 1993).

[4] Some contributions to the literature on social learning and social interactions in technology adoption have been successful in opening this black box. Examples are Kremer and Miguel (2006), Munshi and Myaux (2006), and Duflo and Saez (2003).

The Review of Economics and Statistics, November 2009, 91(4): 695–716

ineligible households. While household eligibility was determined within all (treatment and comparison group) communities, only households below a welfare threshold and within the treatment villages became program beneficiaries during the evaluation period.

Using experimental variation in the induced school participation of the subset of eligible children in these communities, we can identify neighborhood peer effects in secondary school enrollment decisions among children who were ineligible for the program within the program communities. The use of this experimental design enables us to overcome many of the identification problems that plagued previous literature on social interactions (Manski, 1993).

Our first set of results suggests that children from ineligible households residing in the PROGRESA villages increased their secondary school enrollment rate by 5.0 percentage points relative to ineligible households in control villages. Moreover, there were significant differential effects on school enrollment by household's welfare index level and grade level. For instance, among ineligible households with a value of the welfare index below the median for ineligible households, PROGRESA increased secondary school enrollment by 5.5 percentage points but had no effect for children among the upper welfare index group. Overall, these findings indicate a significant spillover effect on the secondary school enrollment rates of noneligible households residing in the treatment villages.

In the second stage of the study, we exploited the fact that PROGRESA created an exogenous shock to secondary school participation of children residing in the same villages and examined the extent to which social interactions affect children's decisions to enroll in secondary school. We find that children have an increased likelihood of attending secondary school of approximately 5 percentage points as a result of a 10 percentage point increase in the village network enrollment rate. Substantially larger effects of approximately 6.5 percentage points are also found for ineligible children of relatively poorer households—a subgroup of children more likely to interact with treated children in these villages. These estimates indicate that the policy intervention benefited from important social multipliers as behavioral social interactions in effect doubled the direct effects of the school enrollment subsidy.

A potential concern with our identification strategy is that the program may have affected ineligible children through other mechanisms. The focus of PROGRESA was not limited strictly to education but also encouraged investments in health and nutrition while providing eligible households with substantial monthly payments. With the program inducing behavioral changes among eligible households along several dimensions, the increase in enrollment among ineligible households was not necessarily due to peer effects but rather a response to some other change in the behavior of eligible households.

Our results are consistent with three alternative explanations. First, although we do not find any evidence that the program had a direct effect on school quality, we cannot definitively reject the hypothesis that PROGRESA did not improve teacher quality or effort indirectly, as teachers could have responded to children becoming more interested in school, leading to an increased school enrollment of ineligible children (Kremer et al., 2009; Duflo, Dupas, & Kremer, 2007). Second, we cannot reject the hypothesis that noneligible children enrolled in secondary school with the expectation that this would affect their future program participation. A final alternative interpretation for our findings is that ineligible households may have simply responded to information regarding the benefits of schooling and attaining an education (Jensen, 2007). If PROGRESA led parents and students to update their priors on the value of enrollment, the program may have affected the enrollment decisions of noneligible households directly.

The data are, however, inconsistent with several other hypotheses. We do not find any evidence that the program affected either the consumption of ineligible households or children's health, which may have led to greater school enrollment rates. Also, we condition on a large number of predetermined mean village-level contextual and environmental characteristics that may be correlated with the impacts of the intervention and show that the effects are robust to these specifications. Finally, we present evidence inconsistent with a relative reduction in transportation costs faced by program village children and with potential contamination bias concerns. This sensitivity analysis confirms the validity of the identifying assumptions of the model.

Our study contributes to the growing literature on neighborhood-based peer effects in schooling decisions. The seminal paper by Case and Katz (1991) identifies neighborhood-based peer effects in idleness among youth in high-poverty areas in Boston using an instrumental variables strategy to address the reflection problem. Two recent contributions also use instrumental variable strategies to estimate behavioral peer effects in schooling decisions in various contexts. Cipollone and Rosolia (2007) use plausibly exogenous variation in the school attainment of men as a result of a policy following an earthquake in southern Italy to identify the effect on the school attainment of women in these regions. Lalive and Cattaneo (2009) extend our analysis to test whether social interactions affected the schooling decisions of primary and secondary school children. They also use subjective information on parents' perception of children's ability and school efforts to understand the reasons for endogenous social interactions in schooling decisions. Their results are complementary and confirm many of ours.

The paper is structured as follows. Section II provides a brief discussion of the PROGRESA program and its evaluation component, as well as the data used in the analysis. In section III, we present an empirical model of social interaction effects and discuss its identification problems. We then describe our research design and how it avoids these identification pitfalls. The main estimates are reported in section IV, followed by sensitivity tests of the identifying assumption in section V, a discussion of alternative interpretations in section VI, and a conclusion in section VII.

II. PROGRESA Program, Evaluation, and Data

A. *Background on the PROGRESA Program Evaluation*

In 1997, the Mexican government initiated a large-scale education, health, and nutrition program, the PROGRESA program, aimed at improving human development among children in rural Mexico. The program targets the poor in marginal communities, where 40% of the children from poor households drop out of school after the primary level. The program provides cash transfers to the mothers of over 2.6 million children conditional on school attendance, health checks, and health clinics participation, at an annual cost of approximately $1 billion, or 0.2% of Mexico's GDP in 2000. The education component of PROGRESA consists of providing subsidies, ranging from $70 to $255 pesos per month (depending on the child's gender and grade level), to children attending school in grades 3 to 9 of primary and lower secondary school. Overall, the program transfers are sizable, representing 10% of the average expenditures of beneficiary families in the sample.

A distinguishing characteristic of PROGRESA is that it included a program evaluation component from its inception. PROGRESA was implemented following an experimental design in a subset of 506 communities located across seven states: Guerrero, Hidalgo, Michoacán, Puebla, Querétaro, San Luis Potosí, and Veracruz. Among these communities, 320 were randomly assigned into a treatment group, with the remaining 186 communities serving as a control group, thus providing an opportunity to apply experimental design methods to measure its impact on various outcomes. In addition, within these selected communities, a poverty indicator was constructed using the household income data collected from the baseline survey in 1997. A discriminant analysis was then separately applied in each of the seven regions in order to identify the household characteristics that best classified poor and nonpoor households. These characteristics, which were unknown to the households, were then used to develop an equation for computing a welfare index that determined eligibility into the program (see Skoufias, Davis, & de la Vega, 2001, for a more detailed description of the targeting process).[5] While household eligibility was determined within all (treatment and comparison group) communities, only households classified as eligible and within the treatment villages became program beneficiaries during the evaluation period. That the eligibility classification exists for both treatment and control communities and treatment was randomly assigned are critical design aspects for the identification of the neighborhood peer effects, as will be discussed in section III.

An issue in the initial implementation (during the first year) of the program involved an increase (by the program administrators) in the number of eligible households after it was discovered that households with certain characteristics—the elderly poor who no longer lived with their children—were excluded from the initial eligibility criteria. Because of this oversight, a new discriminant analysis was conducted, and households were reclassified as either eligible (poor) or noneligible (nonpoor) households. Households that were originally classified as nonpoor but included in this second set of eligible households—the *densificado* group—became program beneficiaries approximately eight months after the start of the program (Skoufias, Davis, & de la Vega, 1999). As a result of this change in program implementation, there are eligible households above and below the initial region-specific eligibility thresholds. For our analysis we classify these *densificado* households as eligible, since these are eligible for treatment at some point during the evaluation period.

B. *Data and Measurement*

Since the baseline census in October 1997, extensive biannual interviews were conducted during October 1998, May and June 1999, and November 1999 on approximately 24,000 households of the 506 communities.[6] Each survey is a community-wide census containing detailed information on household demographics, income, expenditures and consumption, and individual socioeconomic status, health, and school behavior. More specifically, the surveys in October 1997, October 1998, May and June 1999, and November 1999 collected information on the school enrollment and grade completed of each child in the household between 6 and 16 years old. We thus have information on enrollment during three consecutive school years (1997–98, 1998–99, and 1999–2000) and grade promotion during two consecutive school years. Since primary school enrollment is almost universal in rural Mexico, we restrict our interest to the enrollment and promotion decisions of children who have attained at least a primary education but have not completed secondary school at baseline. Secondary school enrollment is the most problematic decision for school attainment,[7] and also the grade levels where PROGRESA has had its greatest impact among eligible households (Schultz, 2004). In our

[5] In addition to capturing the multidimensionality of poverty, another advantage of a welfare index is that it permits the classification of new households according to their socioeconomic characteristics other than income.

[6] There was a round of data collection in March 1998 just prior to the start of the intervention.

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

TABLE 1.—INDIVIDUAL AND HOUSEHOLD CHARACTERISTICS ACROSS PROGRAM AND COMPARISON VILLAGES

	Ineligible Households				Eligible Households			
	Mean (SD)	Program	Comparison	Difference	Mean (SD)	Program	Comparison	Difference
A: Child characteristics								
School enrollment in 1997	0.699	0.712	0.680	0.032	0.663	0.664	0.662	0.002
	[0.459]			(0.029)	[0.473]			(0.020)
School enrollment in 1998	0.655	0.679	0.618	0.061*	0.635	0.661	0.592	0.069***
	[0.475]			(0.033)	[0.481]			(0.024)
School enrollment in 1999	0.515	0.532	0.489	0.042	0.516	0.540	0.479	0.061***
	[0.500]			(0.034)	[0.500]			(0.023)
Child's age in 1997	13.43	13.41	13.46	−0.05	13.36	13.36	13.35	0.02
	[1.72]			(0.07)	[1.67]			(0.04)
Grade completed in 1997	6.25	6.27	6.23	0.05	6.03	6.03	6.04	−0.01
	[1.01]			(0.05)	[0.93]			(0.03)
Gender (boy)	0.495	0.497	0.494	0.003	0.504	0.511	0.492	0.019*
	[0.500]			(0.020)	[0.500]			(0.010)
Indigenous	0.115	0.129	0.093	0.036	0.306	0.305	0.308	−0.003
	[0.319]			(0.040)	[0.461]			(0.052)
B: Household characteristics								
Head of household's schooling	3.19	3.25	3.10	0.15	2.57	2.58	2.57	0.01
	[2.97]			(0.20)	[2.39]			(0.11)
Head of household's gender (male)	0.926	0.932	0.918	0.014	0.921	0.921	0.922	−0.001
	[0.261]			(0.013)	[0.269]			(0.007)
Head of household's age	48.78	48.82	48.73	0.08	45.88	45.62	46.30	−0.68**
	[10.65]			(0.62)	[10.84]			(0.33)
Household size	6.85	6.78	6.97	−0.19	7.34	7.33	7.38	−0.05
	[2.32]			(0.17)	[2.36]			(0.09)
Total household-level PROGRESA Transfers (posttreatment)	—	—	—	—	111.48	170.27	14.93	155.34***
					[131.44]			(5.84)

Note: Standard deviations of variables are reported in brackets. Differences estimated in OLS regression models. Robust standard errors in parentheses; disturbances are allowed to be correlated within village; significantly different from zero at *90%, *95%, and *99% confidence. The numbers of ineligible and eligible children are 2,738 and 11,147, respectively.

sample, this concerns approximately 2,120 children who are eligible at baseline to enter any of three lower secondary school grade levels. By selecting the sample based on grade completed at baseline rather than including children who start completing their primary schooling during the post-treatment evaluation period, we avoid issues of dynamic selection into secondary school (Cameron & Heckman, 1998). Also, with village-level censuses, we can reliably construct village-level means of household and individual characteristics, including schooling decisions and contextual variables that may affect it.

Table 1 presents the mean of various individual and household-level characteristics for both eligible and noneligible children and their differences between treatment and control villages. The first row in the table demonstrates the hurdle that secondary school represents for children in rural Mexico and highlights a clear objective of the program (table 1, panel A). In 1997, the enrollment rate of eligible children in secondary school was 66% on average. Although enrollment rates were on average 4 percentage points higher among ineligible children, only 70% of these were enrolled in secondary school. As one would expect from the random assignment, the preprogram difference in enrollment rates between treatment and control villages among both eligible and ineligible households was small and statistically insignificant. In addition, the simple difference in 1998 and 1999 enrollment rates between treatment and control communities provides a straightforward measure of the program's impact on school participation. In both years, enrollment rates in treatment villages were roughly 6 percentage points higher than in control villages among the beneficiary households. Table 1 also shows the first indication of a possible spillover effect. Although the difference is statistically insignificant (in the second year), secondary school enrollment in the treatment villages is approximately 6 and 4 percentage points higher than in control villages among children of ineligible families in 1998 and 1999, respectively. Given these low enrollment rates, it is perhaps not too surprising that the mean educational level of heads of households is also quite low, as heads of eligible and ineligible households have completed only 2.6 and 3.2 years of schooling, respectively (panel B). These children also tend to come from large households; the average household size in these villages is 7.3 for eligible households and 6.8 for ineligible ones.

We also compare mean attributes at baseline (October 1997) across treatment and control villages to evaluate the randomization of our sample (table 1, columns 2–4, 6–8). As one would hope from the random assignment, there are no statistically significant differences in the observed characteristics of these individuals on most dimensions.[8]

[8] Behrman and Todd (1998) conduct an exhaustive analysis of the degree of success of the random assignment of villages in the PROGRESA program and conclude that the randomization was successful.

In addition to the village census data, we use administrative data on the number of PROGRESA transfers received by the households per survey round. As expected, the administrative data on transfers show that eligible households in treatment villages received 170 pesos per month (on average) during the April 1998–December 1999 period (table 1, panel B). Average transfers for control households are nonzero because they begin to receive program transfers by December 1999. The difference in transfers between the two groups is large and substantial. More importantly, the administrative data show no evidence of program leakage (i.e., ineligible households receiving cash transfers).[9]

Finally, we make use of administrative data on secondary schools in the evaluation regions (which contain information on number of pupils by grade, teachers, number of classrooms, and other infrastructure characteristics of the schools). Without information on which school each child attends, we match, using GPS data, children from the same village to the secondary school closest in distance to the village.[10] These administrative data allow us to rule out alternative hypotheses and to test our identifying assumptions (see the discussion in section V). Means of baseline characteristics of schools attended by the children in the sample are reported in table 2; there are no systematic differences between treatment and control villages, as expected.

Given our panel data structure, an important issue in the empirical analysis is the extent of sample attrition. If being out of sample is correlated with the likelihood of being in the program (treatment) group, then this could bias the coefficient estimates. Sample attrition rates through the two posttreatment survey rounds are approximately 20% for the sample of children in secondary school, in both eligible and ineligible households (table A1, columns 1 and 4), and the likelihood of attrition is highly correlated with individuals' observable characteristics (columns 2 and 5). Fortunately, across program and comparison groups, attrition rates are balanced, and the observables correlates of attrition are not significantly different (columns 3 and 6). We use baseline individual, household, and community characteristics to control for any potential attrition bias in all our estimations.

Table 2.—School Characteristics across Program and Comparison Villages

	All Villages		
	Program	Comparison	Difference
Tele-secondary school	0.85	0.88	−0.03 (0.03)
General secondary school	0.05	0.04	0.01 (0.02)
Technical secondary school	0.09	0.07	0.02 (0.02)
Rural	0.93	0.95	−0.02 (0.02)
Semiurban	0.06	0.04	0.02 (0.02)
Classrooms in grade 7	1.14	1.15	−0.01 (0.08)
Classrooms in grade 8	1.05	1.02	0.03 (0.08)
Classrooms in grade 9	0.98	0.94	0.04 (0.08)
Number of teachers	3.10	3.06	0.04 (0.37)
Pupil-teacher ratio	22.14	21.59	0.55 (0.92)

Note: Differences estimated in OLS regression models. Robust standard errors in parentheses; disturbances are allowed to be correlated within village; significantly different from zero at *90%, **95%, and ***99% confidence. The number of secondary schools is 506.

III. Identification of Neighborhood Peer Effects

In this section, we discuss the econometric model used to estimate neighborhood peer effects and the assumptions needed for identification. We base our empirical model on a simple decision problem of school enrollment in the presence of social interactions. This will allow us to postulate various mechanisms through which peers can play a role in school enrollment decisions.

An individual's secondary school enrollment decision (y_{ic}) can be modeled as a function of (1) the child's expected learning (which is determined by the child's learning, i.e., cognitive, ability, the school organization and environment, common to all children, and the ability distribution of peers attending school); (2) a desire to conform with the reference group's (i.e., the neighbors') school enrollment and participation behaviors (($y_{1c}, \ldots, y_{-i,c}, y_{i+1,c}, \ldots, y_{I,c}$) denoted by $y_{-i,c}$) due to either peer pressure or social norms (Bernheim, 1994); (3) individual opportunity costs of attending school, which may vary as a result of a government subsidy for school participation, as well as the perceived safety of commuting to school; and (4) variation in the tastes for schooling. In this theoretical framework, peer effects enter the child's utility through three main mechanisms. First, it captures the idea of strategic complementarities in peer participation and effort in the education production function: if the child enrolls in school, the time that peers spend in class, as well as their effort levels inside and outside the classroom, can enhance the child's learning, in addition to his or her own ability and the school environment. Also, the preference-based mechanisms for social interactions incorporate the role that the desire to attend school may be increasing in the school enrollment of peers in the reference group (i.e., the proportional complementarity utility function), influenced by a desire to conform with others due to peer pressure or social

[9] Although this does not prove that leakage was not an issue in the program's implementation, there is no evidence of it at the central level.

[10] Although there may be some measurement error associated with matching children to their geographically closest school, there are at least two reasons that the misclassification should be minimal. First, households in these villages have a very limited choice of schools due to the scarce number of secondary schools in these marginal areas (only 10% of households have access to a secondary schools in their village). Second, based on fieldwork we conducted in 2003, we were able to perfectly match the villages visited to the secondary schools reported as attended in informal interviews with village members.

norms, resulting in children not wanting to deviate from choices made by others in her reference group (Akerlof's 1997 quadratic conformist utility function), or due to changes in the expected costs of commuting to school due to their peers' school going (e.g., safety in numbers).

Under the assumption that children make school participation decisions taking other individuals' choices as given, maximizing utility yields an equation for the child's optimal school participation level, which results in the standard linear-in-means empirical model used to estimate neighborhood peer effects:

$$y_{ic} = \alpha + \beta X_{ic} + \gamma \bar{X}_c + \lambda Z_c + \theta \bar{y}_c + u_{ic}, \tag{1}$$

where y_{ic} is an indicator variable for the school enrollment behavior of child i in village c; X_{ic} are exogenous characteristics of the individual; $\bar{X}_c$ are the mean exogenous characteristics of the reference group; Z_c are characteristics of the environment (village or school) that may influence individuals' school enrollment decisions; and $\bar{y}_c$ is the enrollment rate of the reference group.[11] This linear-in-means model provides a formal expression to various hypotheses often advanced to explain the common observation that individuals belonging to the same group or neighborhood tend to behave similarly. The first, *correlated effects,* proposes that individuals in the same group tend to behave similarly because they have similar characteristics or face similar environments; these are represented in the model by the vector of parameters β and λ. The second, *contextual peer effects,* proposes that exogenous characteristics of the reference group (e.g., parental involvement in children's education in the village) influence individual behavior; the vector of parameters γ captures these contextual effects. Finally, the hypothesis of *endogenous peer effects* proposes that the school enrollment behavior of the group influences individual behavior; the parameter θ in the model captures this effect. In the empirical analysis, we cannot and do not distinguish from production or tastes-based motivations for the social interaction effects; these are captured in the θ reduced-form parameter.[12]

As Manski (1993) shows, OLS estimation of the linear-in-means model cannot separately identify the two types of social interaction effects as a result of the simultaneity of individuals' actions.[13] Equation (1) represents individual i's school enrollment best-response function given peers' potential school enrollment decisions and exogenous characteristics. However, the data consist of equilibrium behavioral choices of all individuals in a reference group, and therefore the individuals' school enrollment decisions are jointly determined, leading to simultaneity bias (Moffitt, 2001).

Identification of parameter θ is possible, however, under a partial-population experiment setting, whereby the outcome variable of some randomly chosen members of the group is exogenously altered (Moffitt, 2001). Formally, we can assume that individuals' school enrollment decisions follow model (1), augmented for the existence of an exogenous treatment T_{ic} that equals unity for a subset of individuals in the reference group c and zero otherwise. The individual characteristics of this subgroup are denoted by the superscript E:

$$y^E_{ic} = \alpha + \beta X^E_{ic} + \gamma \bar{X}_c + \lambda Z_c + \theta \bar{y}_c + \delta T^E_{ic} + u^E_{ic}. \tag{1$'$}$$

In addition, there are individuals within the same reference group c (denoted with superscript NE) who do not receive treatment:

$$y^{NE}_{ic} = \alpha + \beta X^{NE}_{ic} + \gamma \bar{X}_c + \lambda Z_c + \theta \bar{y}_c + u^{NE}_{ic}. \tag{1$''$}$$

Using equations (1′) and (1″), and recalling that group averages are related to within-village treated (E) and untreated (NE) group averages by

$$\begin{aligned} \bar{y}_c &= m^E_c \bar{y}^E_c + (1 - m^E_c)\,\bar{y}^{NE}_c \\ \bar{X}_c &= m^E_c \bar{X}^E_c + (1 - m^E_c)\bar{X}^{NE}_c, \end{aligned} \tag{2}$$

where m^E_c is the share of treated individuals in the reference group c, we can show, based on Moffitt (2001), that the mean equilibrium outcome in the reference group satisfies the following condition:

$$\begin{aligned} \bar{y}_c = {} & \frac{\alpha}{1-\theta} + \frac{\beta+\gamma}{1-\theta}\bar{X}_c \\ & + \frac{\lambda}{1-\theta} Z_c + \frac{\delta}{1-\theta} m^E_c T_c. \end{aligned} \tag{3}$$

Substituting equation (3) in equation (1″), we can solve for the reduced-form relationship of the school enrollment outcomes of untreated individuals as a function of the partial-population treatment in the reference group and exogenous individual, reference group, and environmental characteristics:

[11] In this specification, we are assuming that the reference group and the environment are one in the same. This clearly need not be the case.

[12] We refer interested readers to Becker and Murphy (2000), Durlauf and Young (2001), and Glaeser and Scheinkman (2003) for a thorough discussion of the literature on choice in the presence of social interactions. Duflo and Saez (2003) examine reduced-form endogenous interaction effects with respect to retirement savings decisions in the United States using an analogous experimental design.

[13] To see this, taking the expectation of equation (1) conditional on X and Z, integrating over Z, and solving for $\bar{y}_c$ results in the mean equilibrium outcome in group c, which, substituted in equation (1), yields the reduced form for individual outcomes: $y_{ic} = (\alpha/1 - \theta) + \beta X_{ic} + (\gamma + \beta\theta/1 - \theta)\bar{X}_c + (\lambda/1 - \theta)Z_c + u_{ic}$. Manski (1993) shows that conditional on $\theta \neq 1$, this equation has a unique solution, parameters γ and θ are unidentified, but composite parameters $(\alpha/1 - \theta)$, $(\gamma + \beta\theta/1 - \theta)$, and $(\lambda/1 - \theta)$ are identified. Although the identification of the composite parameters does not allow distinguishing between endogenous and contextual social interaction effects, it permits determining whether some social effect is present.

$$y_{ic}^{NE} = \frac{\alpha}{1-\theta} + \beta \bar{X}_{ic}^{NE} + \frac{\theta\beta + \gamma}{1-\theta} \bar{X}_c + \frac{\lambda}{1-\theta} Z_c + \frac{\theta\delta}{1-\theta} m_c^E T_c + u_{ic}^{NE}. \tag{4}$$

The partial-population treatment terms in the two reduced-form equations have intuitive interpretations. In equation (3), the $(\delta/(1-\theta))m_c^E$ term can be decomposed into two additive terms: (1) the direct effect of the treatment on the mean enrollment of the reference group, which is assumed to affect a subsample of the reference group (δm_c^E), and (2) the indirect effect as a result of behavioral social interactions ($(\theta/(1-\theta))\delta m_c^E$). For the untreated group, equation (4), the partial-population treatment term accounts for the fact that the untreated group is not directly affected by the treatment (by definition) and includes only the indirect effect: the social interaction effect.

Also note that one could use coefficient estimates from equations (3) and (4) to identify the direct treatment and peer effects parameters. Specifically, note that the ratio of the $m_c^E T_c$ reduced-form coefficients from equations (3) and (4) is equal to θ, the peer effects parameter.

The specifications that we adopt in this paper are based on equation (1″) and a slight variant of equation (3):

$$y_{ic}^{NE} = \alpha + \beta X_{ic}^{NE} + \gamma \bar{X}_c + \lambda Z_c + \theta \bar{y}_{-i,c} + u_{ic}^{NE} \tag{1″}$$

$$\bar{y}_{-i,c} = \tilde{\alpha} + \tilde{\beta}_1 X_{ic}^{NE} + \tilde{\beta}_2 \bar{X}_c + \tilde{\lambda} Z_c + \tilde{\delta} T_c + \bar{\varepsilon}_c, \tag{3′}$$

where T_c is the PROGRESA treatment village indicator variable and composite coefficients $\tilde{\alpha} = (\alpha/1-\theta)$, $\tilde{\beta}_2 = (\theta\beta + \gamma/1-\theta)$, $\tilde{\lambda} = (\lambda/1-\theta)$, and $\tilde{\delta} = (\delta m_c^E/1-\theta)$. Note that equation (3′) uses T_c rather than the interaction term $m_c^E T_c$ as the instrumental variable. We allow for this discrepancy in the model because the share of treated individuals in the reference group (m_c^E)—in this case, the share of PROGRESA-eligible children in the village—may not be exogenous if there is any sorting of individuals into and out of the village based on unobservable characteristics of the households or villages. However, estimates that use $m_c^E T_c$ as the instrumental variable (IV) provide quantitatively similar estimates to those reported in the results sections below.

Under the conditions of (a) robust partial correlation between the instrumental variable and the endogenous regressor ($\tilde{\delta} \neq 0$), and (b) lack of correlation between the excluded IV and the disturbance term in equation (1″) ($E[T_c u_{ic}^{NE}] = 0$), IV estimation is a consistent estimator of parameter θ. Condition a can be tested in the data, and results will be discussed in section IV. Condition b the exclusion restriction, is not directly testable and is a maintained assumption of the model; the random assignment of the program across villages is not sufficient to ensure that this condition holds.

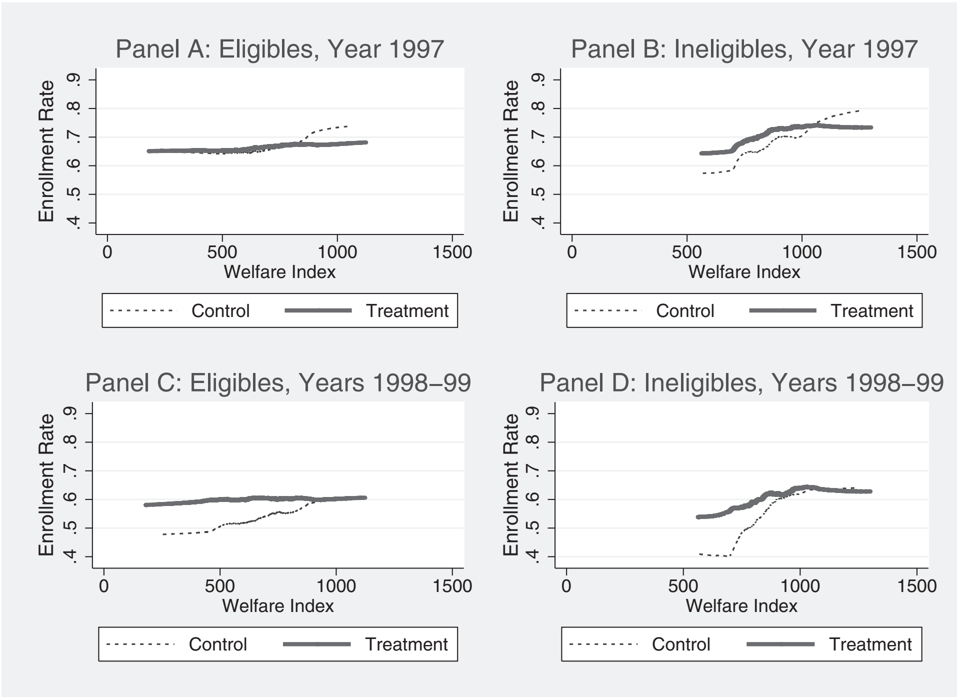
The IV exclusion restriction relies on the assumption that an increase in school participation among ineligible children in treatment villages is the effect of the exogenous increase in school participation among the eligible secondary school children within the village, not the result of changes in contextual variables affected by the program. Since it is possible, however, that the program affected ineligible children through other channels, we follow various strategies to provide evidence that this is not the case. First, using rich microdata for both eligible and ineligible households, we directly test whether other potential externalities from program impacts or particular intricacies of the program had an effect on ineligible households. We do not find any evidence of changes in the consumption patterns or health status of ineligible households or in measures of school quality, for instance. Second, we condition on a large number of predetermined mean village-level contextual ($\bar{X}_c$) and environmental (Z_c) characteristics that may be correlated with the impacts of the intervention and show that the effects are robust to these specifications. We do not find any evidence of alternative mechanisms and defer discussion of these results to section V.[14]

Finally, we also assume neighborhood peer effects to be at the village level. Although we lack information on the specific individuals who belong to a child's reference group, we believe that the assumption of village-level effects may not be problematic for the following reasons. As is common in village economies in less-developed countries, substantial ethnographic evidence documents social interactions at the village level in rural communities in Mexico (e.g., Foster, 1967). Furthermore, rural villages in this sample are quite small, with 47 households per village and only 20 children of secondary school age per village, on average. Thus, in the context of Mexico, village peer effects may be a more credible assumption than studies that use city blocks (Case & Katz, 1991), census tracks (Topa, 2001; O'Regan & Quigley, 1996), schools (Evans, Oates, & Schwab, 1992; Gaviria & Raphael, 2001), or classrooms (Hoxby, 2000).[15]

[14] We present in the appendix a more general linear-in-means model of social interactions that allows for direct treatment effects on children's contextual characteristics. To identify endogenous peer effects in this model, we need to assume that the other variables affected have neither direct nor contextual social interaction effects on children's school enrollment decisions. If the condition fails to hold, we can still identify the presence of peer effects, but we cannot distinguish between endogenous and contextual peer effects. We estimated reduced-form equations consistent with this more flexible model in which we directly explore the relationship between school enrollment and $m_c T_c$. Our results, while less precisely estimated, are consistent with the estimates reported in section IV. These results are available on request.

[15] Although social interactions are assumed to occur strictly at the village level, many of the children in the sample are attending schools located in neighboring villages. If the children are interacting strongly with children in these other villages, a social network that comprises only own-village children may be ill defined. However, our instrumental variables strategy allows us to avoid this issue, since the random assignment of the village to the experimental groups should be uncorrelated with

FIGURE 1.—NONPARAMETRIC ESTIMATES OF ENROLLMENT RATES BY HOUSEHOLD ELIGIBILITY INDEX, YEARS 1997–1999



Note: Locally weighted smoothing of the proportion of individuals enrolled in secondary school by the welfare index of program eligibility; bandwidth = 0.8. The numbers of ineligible and eligible children are 2,738 and 11,147, respectively. Vertical lines are drawn at welfare index levels 550 and 822.

IV. Estimates of Spillovers and Neighborhood Peer Effects

A. *Estimates of Reduced-Form Spillover Effects*

In this section, we present evidence on the reduced-form spillover effects of the program on school enrollment and grade promotion. We start the discussion with a graphical analysis to shed light on the patterns in the data. Figure 1 presents a series of graphs, based on nonparametric estimates, depicting enrollment rates in secondary school by the welfare index used to classify eligible and ineligible households.[16] Enrollment rates do not differ at baseline among eligible children in program and comparison villages (figure 1, panel A), and the difference is positive but small and insignificant among ineligible children (panel B). However, for 1998 and 1999, enrollment rates in program villages among both eligible and ineligible children increase substantially relative to the comparison group (panels C and D). Within the ineligible group, we observe a striking difference in enrollment rates between treatment and control villages among relatively poorer households. This enrollment difference remains until a household welfare index of approximately 900 units (the median welfare index of ineligible households), at which point the enrollment rates tend to converge. This figure suggests that any spillovers of the program may have been concentrated among ineligible households with welfare characteristics relatively similar to the eligible households but classified above the welfare qualification.

Parametric linear probability estimates of the reduced-form relationship between program and comparison villages enrollment rates mirror the results depicted in figure 1. Consistent with Schultz (2004) and Behrman, Sengupta, and Todd (2005), we find that children in eligible households increased their school enrollment by 6.3 percentage points relative to eligible children in control villages (table 3, panel A, regression 1). The point estimate with household and village-level controls implies an effect of 7.0 percentage points, or 12% (panel B, regression 1).[17] The point estimate indicates that the program had a slightly greater impact during its first year (although we cannot detect any statistically significant differences by year, p-value = 0.55) (panel C, regression 1) and among children who were to be enrolled in either sixth or seventh grade in 1998, the last year of primary school and the first of secondary school (regression 2).

The results presented in columns 3 to 5 suggest that PROGRESA may have also benefited ineligible children. On aver-

the assignment of neighboring villages to the program and the schooling decisions of children in these villages.

[16] The conditional means are estimated by taking the mean enrollment within a bandwidth of 0.8. The figure is robust to perturbations to the bandwidth size.

[17] When we exclude the *densificados* from the sample, a large share of which did not receive any benefits during the evaluation period, the effects on eligible children are even stronger (table 3, column 2). Excluding these individuals in the IV specifications does not affect our results.

TABLE 3.—SCHOOL ENROLLMENT TREATMENT AND SPILLOVER EFFECTS ESTIMATES AMONG ELIGIBLE AND INELIGIBLE CHILDREN (DEPENDENT VARIABLE: SCHOOL ENROLLMENT INDICATOR)

Sample	Eligible Children		Ineligible Children		
	All Grades 6–9 (1) OLS	All Grades 6–7 (2) OLS	All Grades 6–9 (3) OLS	Welfare < Median Grades 6–9 (4) OLS	All Grades 6–7 (5) OLS
A: No controls					
Treatment indicator	0.063***	0.083***	0.050†	0.085**	0.072*
	(0.022)	(0.026)	(0.031)	(0.035)	(0.040)
B: Controls					
Treatment indicator	0.070***	0.085***	0.036	0.057*	0.056*
	(0.016)	(0.019)	(0.027)	(0.031)	(0.033)
C: Year-specific effects, controls					
Treatment indicator, year 1998	0.075***	0.098***	0.040†	0.070**	0.057*
	(0.019)	(0.021)	(0.028)	(0.034)	(0.033)
Treatment indicator, year 1999	0.064***	0.074***	0.032	0.046	0.056†
	(0.018)	(0.021)	(0.031)	(0.034)	(0.036)
Mean of dependent variable	0.577	0.566	0.587	0.559	0.611
N observations	17,494	13,371	4,211	2,757	2,846
N individuals	8,828	6,744	2,116	1,382	1,423

Note: Coefficient estimates from OLS regressions are reported. Robust standard errors are in parentheses; disturbance terms are allowed to be correlated within but not across villages; significantly different from 0 at †85%, *90%, **95%, ***99% confidence. Individual and household level controls are the child's gender, indigenous status, the household's welfare index, education, age, and gender of the head of household, family size, and distance to secondary school.

age, children from ineligible households residing in the PROGRESA villages increased their secondary school enrollment rate by 5.0 percentage points relative to ineligible households in control villages (panel A, regression 3); however, the effect is imprecisely measured (significant at 89% confidence) and not robust to individual-, household-, and village-level controls (panel B, regression 3).[18] The differential effects on school enrollment by household's welfare index level (regression 4) are significant. Among ineligible households with a below-median welfare index, PROGRESA increased secondary school enrollment by 5.7 percentage points (statistically significant at 90% confidence), but had no effect for children among the upper welfare index group (−0.9 percentage points and not statistically significant; not reported in the tables).[19] Similar to the differential effects exhibited by treated households, the point estimates indicate that the spillover effects were also larger during the first year of the program (4.0 percentage points). Finally, the spillover effects for ineligible children just entering secondary school are large and sustained during the two academic years at approximately 5.6 percentage points (panel C, regression 5).[20]

In table 4, we investigate the effects of the program on promotion rates of ineligible children. Although the program has small and marginally significant grade promotion effects among all eligible secondary school–ready children (panel B, regressions 1 and 3), the effects are more pronounced among children just entering secondary school in both eligible and ineligible households (4.4 percentage points and 5.6 percentage points; panel B, regressions 2 and 5), and those residing in ineligible households below the median in terms of the welfare index. For instance, among ineligible households with a below-median welfare index, PROGRESA increased secondary school promotion rates by 6.1 percentage points (statistically significant at 95% confidence), which implies an increase of roughly 12%. Both the direct and indirect grade promotion effects are sustained during the 1999–2000 academic year around the range of 4.2 to 7.1 percentage points (8–14%), providing us confidence that the program promoted the school enrollment and study effort of children continuing in or reentering secondary school onto the second year of the program.

B. *Estimates of Neighborhood Peer Effects*

Table 5 reports neighborhood peer effects (θ) estimates form OLS and IV estimation of equations ($1''$) and ($3'$). The IV estimate of the overall neighborhood peer effect implies that a 1 percentage point increase in the reference

[18] This result is consistent with Behrman, Sengupta, and Todd's (2005) lack of an overall effect among ineligible children. That said, we find positive spillover effects among children in the 10–13 years age group, consistent with their finding of a spillover effect for 12 year olds. Our effects are more precisely estimated due to the fact that we concentrate on individuals of secondary school age and pool observations across age-specific groups.

[19] The difference in effects is statistically significant at 90% confidence.

[20] One exception is the lack of a spillover effect on girls. Despite the fact that PROGRESA had a larger impact on eligible girls (Schultz, 2004; Behrman, Sengupta, & Todd, 2005), we do not find a similar differential spillover effects between boys (the point estimate is 0.033, standard error 0.030, not statistically significant) and girls (point estimate of 0.027, standard error 0.031, not statistically significant) once we include household and village-level controls (not reported in the tables).

To further check robustness, we estimate program spillover effects using a specification with village contextual characteristics and find largely similar results: overall effect estimates of 2.8 percentage points (standard error 2.5) and larger effects for the subgroup of children in households below the welfare index median (5.4 percentage points, standard error 2.9) and those just entering secondary school (5.0 percentage points, standard error 3.1). These estimates are also robust to the inclusion of municipality fixed effects and to employing probit specifications. Available from the authors on request.

TABLE 4.—GRADE PROMOTION TREATMENT AND SPILLOVER EFFECTS ESTIMATES AMONG ELIGIBLE AND INELIGIBLE CHILDREN (DEPENDENT VARIABLE: GRADE PROMOTION INDICATOR)

	Eligible Children		Ineligible Children		
Sample	All Grades 6–9 (1) OLS	All Grades 6–7 (2) OLS	All Grades 6–9 (3) OLS	Welfare < Median Grades 6–9 (4) OLS	All Grades 6–7 (5) OLS
A: No controls					
Treatment indicator	0.029	0.043*	0.045	0.075**	0.067*
	(0.021)	(0.023)	(0.031)	(0.034)	(0.038)
B: Controls					
Treatment indicator	0.032*	0.044**	0.040†	0.061**	0.056*
	(0.017)	(0.019)	(0.027)	(0.029)	(0.031)
C: Year-specific effects, controls					
Treatment indicator, year 1998	0.022	0.033*	0.036	0.058*	0.041
	(0.017)	(0.019)	(0.028)	(0.033)	(0.031)
Treatment indicator, year 1999	0.042**	0.055***	0.044†	0.065**	0.071**
	(0.018)	(0.021)	(0.030)	(0.032)	(0.035)
Mean of dependent variable	0.515	0.481	0.549	0.515	0.505
N observations	17,327	13,245	4,179	2,738	2,822
N individuals	8,828	6,749	2,114	1,381	1,421

Note: Coefficient estimates from OLS regressions are reported. Robust standard errors in parentheses; disturbance terms are allowed to be correlated within but not across villages; significantly different from zero at †85%, *90%, **95%, ***99% confidence. Individual and household-level controls are the child's gender, indigenous status, the household's welfare index, education, age, and gender of the head of household, family size, and distance to secondary school.

group's enrollment rate leads to a 0.65 percentage point increase in a child's probability of enrollment (significant at 99% confidence; table 5, panel A, regression 1). The magnitude of the peer effect estimate decreases to 0.54 percentage points once individual and household-level controls, as well as state fixed effects, are included (significant at 95% confidence; panel A, regression 2), and reduces further to 0.49 percentage points once the following village-level predetermined contextual variables are included: the proportion of secondary-school-age girls and the proportion of indigenous children in the village, mean village-level family size and educational level, age, and gender proportions of heads of households (significant at 89% confidence; panel A, regression 3).[21] In contrast, the OLS estimate of the overall peer effect for the control villages, which does not take into account the problems of self-selection into reference groups, the reflection problem, and unobserved heterogeneity in the

[21] A specification that uses $m_c T_c$ as the excluded instrument gives an estimate of the endogenous peer effects (θ) of 0.481 (standard error = 0.276, significant at 92% confidence).

TABLE 5.—OLS AND IV ESTIMATES OF ENDOGENOUS PEER EFFECTS AMONG INELIGIBLE CHILDREN

Sample	All Children, Grades 6–9 (1)	All Children, Grades 6–9 (2)	All Children, Grades 6–9 (3)	Welfare < Median (4)	All Children, Grades 6–7 (5)	All Children Grades 6–7, Year 1999 (6)	All Children Grades 6–7, Year 1999 (7)
	Dependent Variable: School enrollment indicator						
A: IV estimates							
Social network enrollment rate	0.649***	0.541**	0.492†	0.671***	0.675***	0.606**	0.574**
	(0.239)	(0.263)	(0.310)	(0.246)	(0.242)	(0.261)	(0.275)
Individual and household controls	No	Yes	Yes	Yes	Yes	Yes	Yes
State indicators	No	Yes	Yes	Yes	Yes	Yes	Yes
Village contextual controls	No	No	Yes	Yes	No	No	Yes
First-stage F-statistic	[8.7]	[8.9]	[7.6]	[13.9]	[9.4]	[8.4]	[9.7]
Observations	4,211	4,211	4,211	2,757	2,846	1,423	1,423
Mean of dependent variable	0.587	0.587	0.587	0.559	0.567	0.524	0.524
	Dependent Variable: Social network enrollment rate						
B: First-stage regressions							
Treatment indicator	0.077***	0.067***	0.057***	0.082***	0.083***	0.086***	0.084***
	(0.026)	(0.022)	(0.021)	(0.022)	(0.027)	(0.030)	(0.027)
Observations	4,211	4,211	4,211	2,757	2,846	1,423	1,423
	Dependent Variable: School enrollment indicator						
C: OLS estimates, control group							
Social network enrollment rate	0.884***	0.708***	0.716***	0.903***	0.768***	0.812***	0.880***
	(0.058)	(0.073)	(0.076)	(0.070)	(0.072)	(0.086)	(0.083)
Observations	1,678	1,678	1,678	1,075	1,151	578	578

Note: Coefficient estimates from OLS and IV regressions are reported. Robust standard errors in parentheses; disturbance terms are allowed to be correlated within villages, but not across villages; significantly different from zero at †85%, *90%, **95%, ***99% confidence. Individual and household-level controls are the child's gender, indigenous status, the household's welfare index, education, age, and gender of the head of household, family size, and distance to secondary school. Village contextual controls are the proportion of secondary school–age girls and the proportion of indigenous children in the village, mean village-level family size and educational level, age, and gender proportion of heads of households.

population, implies peer effects in the 0.71 to 0.88 percentage point range as a result of a 1 percentage point increase in the reference group's enrollment rate (significant at 99% confidence, table 4, panel C, regressions 1–3). The IV estimates suggest that peer effects are quite large for this population. And although we cannot necessarily reject that the OLS and 2SLS estimates are significantly different from each other, the results do suggest that the OLS estimates are biased upward.

Substantially larger peer effects are found among the relatively poorer children within the ineligible group and among those in the lower secondary school grades. The point estimate on the effect for children in the below-median welfare index group is 0.671 (panel A, regression 4) and that on the children just entering secondary school is 0.675 (regression 5). The estimates with and without contextual controls for the 1999–2000 academic year, with point estimates of 0.574 and 0.606 percentage points, indicate that these effects are sustained into the second year of the program (regressions 6–7). The OLS estimates of social network enrollment rate effects for these subgroups in control villages imply effects of 0.903 and 0.768, respectively (panel C, regressions 4–7). Again, the experimental evidence suggests that the OLS estimates are biased upward, although we cannot reject that the coefficients are equal.[22]

That there exists a differential effect by the household welfare index is consistent with various explanations. First, this differential effect may simply suggest that households that are relatively poor and more credit constrained are more responsive to a positive inducement of attending school. Alternatively, these differential effects may reflect differences in social ties between ineligible households that are just above the welfare cutoff and those that are better off. In particular, if children from ineligible households that are slightly above the cutoff are more likely to interact with eligible children in the village, then the induced school participation of eligible children should have a more pronounced effect on this subgroup of children. However, the differential effects may be strictly due to the fact that the instrument is stronger for the subsample of children residing in low welfare index households.

To test this hypothesis, without information on the exact peer network of each student, we construct a measure of the number of extended family members who live in different households and can enroll in secondary school for each child in the village. This measure serves as a proxy for a child's number of family-related peers in the village (a potential subset of a child's peer group).[23] Comparing ineligible children from households below the median of the welfare index to those above the median, we find that the number of eligible extended family links at baseline is significantly greater for ineligible children in the first group (0.97 children) relative to the latter group (0.65 children), among children with some extended family link in the village. This difference of approximately 0.31 children (standard error, 0.09, significant at 99% confidence; not reported in the tables) implies that the number of eligible links is 48% higher among households classified in the below-median welfare index group.[24] While we do not expect all interactions to occur in these villages solely at the extended family level, this evidence is consistent with poorer ineligible children tending to interact more with eligible children.

As noted by other researchers (Graham, 2008; Hoxby & Weingarth, 2006), the linear-in-means model is unable to provide answers to the equity-efficiency trade-offs that pervade in theoretical discussions of peer effects. Kling, Liebman, and Katz (2007), using experimental variation in the poverty rates of neighborhoods in which individuals reside in the United States, find no evidence of nonlinear poverty effects. For comparability reasons, we assess whether there are nonlinearities in peer effects by allowing the parameter estimates to vary according to children's baseline enrollment decision and baseline village-level enrollment rates. Although point estimates suggest that effects are greater among children in communities with low baseline enrollment (results not shown), we cannot reject the linearity assumption.[25]

Weak instruments are not a main concern in the estimation. There is a robust partial correlation between the program village treatment indicator and the potentially endogenous regressor, the village-level secondary school

[22] A specification that uses $m_c T_c$ as the excluded instrument gives an estimate of endogenous peer effects (θ) of 0.573 (standard error = 0.258, significant at 97% confidence). In specifications that include baseline enrollment as an additional regressor (to take into account potential pretreatment differences), the estimated effects vary between 0.370 (standard error = 0.236; significant at 89% confidence) and 0.595 (standard error = 0.279; significant at 97% confidence) given small perturbations in the welfare cutoff. Moreover, none of these specifications suffers from weak IV problems (results available from the authors on request).

We cannot identify effects on children with a high household welfare index, since the average enrollment rate effect is small and indistinguishable from 0 in these villages (point estimate: 0.009, standard error 0.044) and thus the first-stage correlation is weak for this subgroup (point estimate 0.025, standard error 0.026). Therefore, no inferences can be made on the peer effects for children in the wealthier households.

[23] We construct identifiers for extended families in the villages by grouping children according to unique identifiers of their parents' last names. In Latin America, each individual has two last names, the first being the father's first last name and the second the mother's first last name. Therefore, we can construct the households where individuals are related (within reasonable errors) by using unique numerical identifiers of each combination of last names.

[24] Assuming that other children who are not matched to an extended family network actually have no extended family eligible links (therefore, we can impute a 0 number of extended family links for all these children), we can construct measures for all ineligible children in the village. We also find a greater number of links for children in the below-median welfare index group (0.58 children) relative to other ineligible children (0.41 children), for a difference of 0.16 children (standard error 0.06, significant at 99% confidence).

[25] Estimates are available from the authors on request.

enrollment rate. The F-test statistics reflecting the significance of the IV in the first-stage equations excluding and including controls are 8.74 and 7.60 in the overall effect model (panel B, regressions 1–3), and the first-stage F-statistics for the poorer ineligible and the lower-grade groups are respectively 13.9 and 9.4 (regressions 4–5).[26]

In summary, this evidence is consistent with the hypothesis that changes in reference groups' school enrollment behavior affect children's own enrollment behavior and that these effects differ depending on children and their family's inherent opportunity costs, as well as by the types of peers they interact with. As will be shown in section V, these results are robust to specifications and identifying assumptions.

V. Sensitivity Analyses and Tests of Identifying Assumptions

It has been well documented that the impact of PROGRESA was not restricted to schooling. That the program may have affected ineligible children in ways other than an increase in the enrollment rates of their reference groups remains a potential concern for our identification strategy. Such a situation would invalidate our exclusion restriction, and we would be mistakenly attributing the effects of other mechanisms to peer effects. In this section, we present a series of robustness checks and tests of our underlying counterfactual assumption to show that we are in fact providing consistent estimates of neighborhood peer effects.

A. *Reduced-Form Tests of Alternative Mechanisms*

In order for the treatment village indicator to serve as a valid instrument, the program cannot have indirectly affected other determinants of an ineligible child's enrollment decisions. This is a substantive assumption in the case of PROGRESA, where the program's multidimensionality affected the livelihoods of beneficiary households through a series of mechanisms. Apart from the increases in secondary school enrollment rates among eligible children (Schultz, 2004), researchers have found significant increases in household consumption levels, food consumption, and food quality (Hoddinott & Skoufias, 2004); improvements in health status; and increases in health care utilization (Gertler, 2004; Gertler & Boyce, 2001).[27] If any of these program impacts create externalities—in the form of, for example, interhousehold resource transfers, correlated positive shocks to income, or positive health externalities—that increase school enrollment rates for ineligible children, then we would be confounding behavioral peer effects with the positive externalities from these other mechanisms.

In additional to other program externalities, changes in environmental or institutional factors affecting children's school enrollment decisions may also pose concerns. A set of particularly important changes affecting school enrollment decisions were school supply-side interventions that accompanied the implementation of the program. Although this was done to mitigate potential congestion effects due to the expected increase in schooling demand, the improvement in schooling facilities may have attracted children from ineligible households.

To verify whether any of these factors play a role in explaining the enrollment spillover effect, we test for the existence of any posttreatment differences in household per capita consumption and expenditures, and the health status of children that may have been affected by the program (table 6).

Consumption Externalities. We do not find any evidence that monthly household expenditures increased in the three posttreatment survey rounds, among ineligible households with children entering or who have completed some secondary school in program relative to comparison villages (panel A, rows 1–3).[28] Since expenditures do not take into account consumption from household production, we also estimate household consumption in the first two posttreatment periods (the periods for which we have complete consumption data) and, again, find no significant difference in total consumption among these households (rows 4–5). Moreover, differential estimates for the subgroups of below-median welfare index households and those whose children are just entering secondary school also result in insignificant differences in expenditures and consumption (columns 2–3). These expenditure and consumption patterns, as well as the evidence from the transfers data, provide evidence inconsistent with the possibility of interhousehold income transfers from beneficiary to nonbeneficiary households, correlated positive income shocks at the village level, or evidence of program leakage (where some ineligible households may have been able to receive program transfers).[29]

[26] The LIML estimates of equations (1″) and (3′), which are robust to the weak instruments problem (under certain conditions, see Hayashi, 2000) give interaction effects very similar to the IV results reported in the text. Results are available on request.

[27] There is also evidence that the program improved women's relative bargaining power within the household (see Adato, Mindek, & Quisumbing, 2000, and Bobonis, 2009, for a discussion). Evidence of program impacts on other outcomes, including children and adults' labor supply (Parker & Skoufias, 2000), ability to mitigate shocks (de Janvry et al., 2004), and interhousehold transfers (Attanasio & Rios-Rull, 2000) suggest relatively small changes in these margins.

[28] We use household expenditures and consumption as proxies for household income, since income is usually measured with substantial error in agricultural households, and these may better represent permanent incomes of households. The data on consumption from home production are available only in the October 1998 and May 1999 survey rounds, not in the last (November 1999) survey. We thus restrict the consumption analysis to the first two posttreatment survey rounds and compare expenditures per capita across all survey rounds.

[29] It is also possible that the liquidity injection from the program may have relaxed lending constraints of eligible households, enabling ineligible households to borrow when hit by negative idiosyncratic shocks and making them less likely to remove their children from secondary school in the event of a shock (Jacoby & Skoufias, 1997; Angelucci & De Giorgi, 2009). To examine this alternative channel, we estimate expenditure responses of ineligible households to natural shocks in both program and

TABLE 6.—TESTS OF ALTERNATIVE MECHANISMS FOR SPILLOVER EFFECT

Dependent Variables	Sample	Coefficient Estimate on Treatment Village Indicator (SE)			
		All Children, Grades 6–9 (1) OLS	Welfare < Median (2) OLS	All Children, Grades 6–7 (3) OLS	Mean of Dependent Variable (4)
A: Household consumption and expenditures					
Expenditures per capita, October 1998		−5.39	1.56	−3.43	163.7
		(8.15)	(8.84)	(9.00)	
Expenditures per capita, May 1999		1.65	1.78	−0.70	161.5
		(6.98)	(7.97)	(7.15)	
Expenditures per capita, November 1999		2.88	8.35	6.55	152.6
		(5.53)	(5.97)	(6.06)	
Consumption per capita, October 1998		−7.42	−0.19	−6.02	192.4
		(8.42)	(9.12)	(9.35)	
Consumption per capita, May 1999		−3.39	−0.83	−5.83	170.5
		(7.76)	(7.52)	(8.73)	
School expenditures per capita, October 1998		1.68	2.54**	3.27***	8.5
		(1.23)	(1.19)	(1.18)	
School expenditures per capita, November 1999		0.28	1.27	1.45	8.5
		(1.51)	(1.67)	(1.55)	
Food expenditures per capita, October 1998		−2.47	−0.29	−0.43	106.7
		(5.41)	(6.13)	(6.56)	
Food expenditures per capita, November 1999		5.42	9.58**	6.64*	91.7
		(3.54)	(3.83)	(3.93)	
B: Child health spillovers					
Days ill, October 1998		0.10	−0.01	0.04	0.357
		(0.11)	(0.14)	(0.12)	
Days of difficulty with daily activities due to illness, November 1999		−0.047	−0.127	−0.039	0.178
		(0.061)	(0.093)	(0.083)	
Days of no daily activities due to illness, November 1999		0.029	−0.029	0.023	0.093
		(0.046)	(0.071)	(0.059)	
Days in bed due to illness, November 1999		0.030	−0.031	0.021	0.046
		(0.044)	(0.067)	(0.057)	

Note: Each coefficient is from a separate regression. Coefficient estimates from OLS regressions are reported. Robust standard errors in parentheses; disturbance terms are allowed to be correlated within but not across villages; significantly different from zero at *90%, **95%, and ***99% confidence levels. Individual and household-level controls are the child's gender, indigenous status, the household's welfare index, education, age, and gender of the head of household, family size, and distance to secondary school. Village contextual controls are the proportion of secondary school–age girls and the proportion of indigenous children in the village, mean village-level family size and educational level, age, and gender proportion of heads of households.

Households may be changing the composition of household expenditures as a result of their children's enrollment in school. Consistent with the evidence on increased school participation, estimates suggest an increase in the resources spent on schooling per capita (e.g., school supplies, school contributions), particularly during the first posttreatment round. Although the point estimate for the overall sample is insignificantly different from 0, the estimates for the relatively poor and lower-grade subgroups imply average increases in educational expenditures per capita of 2.54 and 3.27 pesos (30% and 38%, respectively, significant at the 95% confidence; panel A, row 6). The spillovers on school expenditures are somewhat muted by the last survey round; although the point estimates suggest increases in school expenditures per capita in the order of 14% to 17% for the various subgroups, none of these are significant at conventional confidence levels. Finally, note that ineligible households also increase expenditures per capita on food items by 7% to 10% in November 1999, during the second academic year (row 9). This evidence, consistent with the evidence presented by Angelucci and De Giorgi (2009) on the spillover effects of the program on food consumption per capita, would suggest that a possible mechanism through which increased school enrollment could be affected is through improved nutrition and the health status of these children more generally.

comparison villages for our subsample and find no evidence that ineligible households in program villages that suffer natural shocks have higher expenditure levels than those in comparison villages (not reported in the tables). Furthermore, the school enrollment effects are lower among "shock" than among "no-shock" households (not reported in the tables), suggesting that these mechanisms do not drive our results.

Health Externalities. We do not find evidence of significant improvements in the health status of secondary school–aged children as a result of an increase in household food expenditures or due to other potential health externalities, such as a reduction in the transmission of communicable diseases (Miguel & Kremer, 2004) or potential improvements in access to health facilities. Unfortunately, the survey collected data from different questions across rounds regarding the self-reported health status of children. Therefore, we show evidence from the first posttreatment round (October 1998) on the number of days the child was ill in the past four weeks and on answers to questions of difficulty with activities of daily living (ADL) in the last survey round

TABLE 7.—ROBUSTNESS CHECKS OF ENDOGENOUS PEER EFFECTS

	Coefficient Estimate on Social Network Enrollment Measure (s.e.)		
Specification (dependent variable is the school enrollment indicator)	Sample: All Children, Grades 6–9 (1)	Welfare < Median (2)	All Children, Grades 6–7 (3)
OLS[a]	0.668***	0.660***	0.740***
	(0.042)	(0.053)	(0.044)
IV, no contextual controls[a]	0.546**	0.652***	0.675***
	(0.260)	(0.235)	(0.242)
	[8.98]	[14.37]	[9.36]
IV, predetermined contextual controls[a,b]	0.495†	0.671***	0.641***
	(0.308)	(0.246)	(0.273)
	[7.59]	[13.92]	[9.36]
IV, predetermined and expenditure-related household-level and contextual controls[a,b,c]	0.512*	0.660***	0.663**
	(0.302)	(0.250)	(0.303)
	[7.68]	[13.28]	[8.64]
IV, predetermined contextual and school characteristics controls[d]	0.495†	0.636**	0.602**
	(0.305)	(0.263)	(0.277)
	[7.83]	[13.31]	[9.30]
IV, predetermined, expenditure-related contextual controls and school characteristics[a,b,c,d]	0.523*	0.650**	0.610*
	(0.294)	(0.257)	(0.318)
	[7.87]	[13.19]	[8.41]
IV, predetermined contextual controls and characteristics of other children attending secondary school[a,b,c]	0.560	0.691**	0.686*
	(0.393)	(0.311)	(0.354)
	[4.76]	[9.82]	[6.47]
Mean of dependent variable	0.587	0.559	0.567
Observations	4,211	2,757	2,846

Note: Each coefficient estimate is from a separate regression. Coefficient estimates from OLS and 2SLS regressions are reported. Robust standard errors in parentheses; disturbance terms are allowed to be correlated within but not across villages; significantly different from 0 at *90%, **95%, and ***99% confidence. First-stage F-statistics of significance of partial correlation between IV (treatment indicator) and social network measure are reported in brackets.

[a] Individual and household-level controls are the child's gender, indigenous status, household's welfare index, education, age, and gender of the head of household, family size, and distance to secondary school. These are included in all specifications.

[b] Village-predetermined contextual controls are the proportion of secondary school–age girls and the proportion of indigenous children in the village, mean village-level family size and educational level, age, and gender proportion of heads of households.

[c] Expenditure-related contextual characteristics are mean village-level household expenditures, mean educational, food, boys' clothing, girls' clothing, alcohol and tobacco expenditure shares, and an indicator variable for whether the village suffered a flood shock.

[d] School characteristics are indicator variables for general, technical, secondary schools (relative to *tele-secundaria* schools), urban and semiurban school indicators (relative to rural schools), school-level pupil-to-teacher ratio, and the number of home teachers, teaching assistants, physical education teachers, and art teachers in school.

[e] School composition controls are the mean village-level contextual characteristics mentioned for children enrolled in secondary school.

(November 1999).[30] There is no significant reduction or increase in the number of days reported ill among ineligible children in October 1998 (the point estimate is 0.10, not statistically significant; panel B, row 1, column 1). Differential effects by welfare subgroups suggest no difference in the morbidity of relatively poorer and lower-school-grade children households (panel B, row 1, columns 2–3). Similar results are found using the ADL measures in November 1999 (rows 2–4). One caveat from this analysis is that morbidity and ADL measures are unlikely to capture more subtle health effects that may have occurred, such as worm infestations that lead to sluggishness or malnutrition (Strauss & Thomas, 2007). That said, the peer effects estimates are robust to the inclusion of controls for the child's health status in the two distinct survey rounds once we condition on the child's morbidity measure and the various ADL measures, respectively.[31] Therefore, to the extent that the available data provide information on the children's health status, the evidence is inconsistent with any positive health externality hypothesis.

B. *Robustness Checks to Contextual Effects and Other Correlated Unobservables*

In addition to these reduced-form tests, we report estimates of the neighborhood peer effect conditioning on a series of expenditure-related village contextual controls (in addition to the predetermined contextual controls): mean village-level household expenditures, mean educational, food, boys and girls' clothing, alcohol and tobacco expenditure shares, and an indicator variable for whether the village suffered a rainfall shock (i.e., flood) in the past six months. Table 7 reports estimates of θ from a series of regressions that gradually condition on village-level predetermined and expenditure-related contextual

[30] See Gertler and Boyce (2001) for a detailed discussion of this self-reported data in the PROGRESA evaluation surveys and a thorough analysis of the health impacts on eligible households.

[31] The estimates of θ for the overall sample, conditioning on the available health measures, are 0.64 (standard error 0.29, significant at 5% confidence) for the 1998–99 academic year and 0.44 (standard error 0.31, significant at 15% confidence) for the 1999–2000 academic year. Estimates for the relevant subgroups are also robust to these controls. Available from the authors on request.

variables and also compares these to OLS estimates of θ. Conditioning on these sets of contextual variables reduces the point estimate of the overall effect slightly, from 0.54 to 0.49 (table 7, column 1). The point estimates for the below-median welfare index and lower-grade subgroups do not vary significantly with the inclusion of additional controls (columns 2–3). Also note that the F-statistics of the first-stage regression coefficients (reported in brackets) do not vary substantially once we condition on potential exogenous interaction factors. This exercise suggests that the estimates are robust to these potential contextual effects, especially among the specific subgroups that experience significant behavioral responses.

Changes in School Reference Group Composition. One potential source of bias could stem from changes in the composition of students attending secondary school. If children (or parents) base their enrollment decisions on the cognitive ability or socioeconomic background composition of their potential classmates (i.e., changes in contextual characteristics at the school level), then our estimates could be confounded by the composition changes in the student body that PROGRESA induced. Although baseline measures of cognitive ability or school achievement are not available, we can verify whether the socioeconomic composition of children attending secondary school changed in the PROGRESA villages using the predetermined contextual characteristics defined above. As expected, children attending secondary schools in treatment villages are disproportionately selected from lower-SES households—households with larger family sizes or lower school attainment of the head of household—relative to children attending secondary school in comparison villages.[32] To the extent that the reduction in the mean "quality" or achievement of students lowers the incentives for children to enroll in secondary school, this potential mechanism would bias our estimates downward.[33] We test for this possibility by estimating models that condition on the mean contextual characteristics of children in the village attending secondary school. The estimates in these specifications increase slightly to 0.56 for the overall sample and to 0.69 for the relevant subgroups (table 7, row 7).

Transportation Costs. Another potential concern is that the program somehow reduced school transportation costs, and this induced ineligible children in the program villages to enroll. Although data on school transportation costs were, unfortunately, not collected, one possibility is to test whether there is a differential effect on children who live less than 1 kilometer from a secondary school, and presumably do not require school transportation.[34] As shown in table 8, although there appear to be effects of different magnitudes in the overall sample (statistically insignificant), there is a small positive 7 percentage point and statistically insignificant difference in the estimated effect between children living less than 1 kilometer away from a secondary school and those who live farther out among the below-median welfare index group (row 1, columns 3–4). In specifications that include predetermined contextual and school composition characteristics controls, the point estimate of the differential effect varies between 7 percentage points lower and 3 percentage points greater for children residing within 1 kilometer of secondary school (rows 2–3). The differential effects by distance to the secondary school for the subsample of children in lower school grades are consistently negative but never significantly different from 0 (column 6, rows 1–3). In the most robust specification, which includes village-contextual and school characteristics controls, the differential effect is 1 percentage point lower (row 3).[35]

Program Contamination. Another concern is the potential contamination of the experimental design, given that some children from treatment and control villages attended the same secondary school. Among the ineligible children in both treatment and control villages, 9.6% of the overall sample and 11.9% of the sample below the median welfare index were matched to the same secondary school. While these interactions could bias our results in either direction, if children from the control villages experienced a crowd-out effect as a result of PROGRESA, then our estimates may be overstated. To test for this bias, we reestimate the peer effects model for the sample of villages in treatment and control groups that are not assigned the same secondary school. Our estimates for this subsample reported in table 8, rows 3 and 4, which are similar to those presented above, suggest that the possibility of contamination is not a particular source of bias.

Program Eligibility Expectations. Since some ineligible households were phased in to the program in the months

[32] Estimates available from the authors on request.

[33] Alternative models of peer effects hypothesize that individuals might prefer peers in schools more similar in terms of cognitive ability or other characteristics (Hoxby & Weingarth, 2006), therefore leading to increased incentives for the marginal ineligible children to enroll in school.

[34] Based on the March 1998 survey, 97% of students attending secondary school walked to school. Although we do not have this information for posttreatment, it seems unlikely that the program would have increased the demand for public or private transportation as to be able to explain the magnitude of the spillover effect. Note also that this is an admittedly fairly weak test. Even if there is a more pronounced effect among children without a secondary school in their village, it still does not discredit a possibility of peer effects. That PROGRESA had a higher impact among eligible children without a secondary school in their village could lead to a differential effect among the ineligibles.

[35] The results are similar if we distinguish between children with and without a school in their village, or alternatively, between children who live less than and more than 2 kilometers from a secondary school.

TABLE 8.—ROBUSTNESS CHECKS OF ENDOGENOUS PEER EFFECTS ESTIMATES TO TRANSPORTATION COSTS AND PROGRAM CONTAMINATION EFFECTS

Specification (dependent variable is the school enrollment indicator)	Coefficient Estimates (s.e.)					
	Social Network Enrollment Rate Sample: All Children, (1)	Social Network Enrollment Rate × School Distance < 1 km Grades 6–9 (2)	Social Network Enrollment Rate Welfare (3)	Social Network Enrollment Rate × School Distance < 1 km < Median (4)	Social Network Enrollment Rate All Children, (5)	Social Network Enrollment Rate × School Distance < 1 km Grades 6–7 (6)
IV, no contextual controls[a]	0.561**	−0.252	0.637***	0.074	0.697***	−0.225
	(0.273)	(0.779)	(0.233)	(0.738)	(0.232)	(0.936)
IV, predetermined contextual controls[a,b]	0.516†	−0.286	0.679***	−0.066	0.657**	−0.146
	(0.317)	(0.785)	(0.251)	(0.744)	(0.258)	(0.881)
IV, predetermined contextual and school characteristics controls[a,b,c]	0.508†	−0.174	0.631**	0.031	0.607**	−0.013
	(0.318)	(0.708)	(0.272)	(0.718)	(0.267)	(0.814)
IV, sample with no overlap in secondary schools[a]	0.435		0.611**		0.653**	
	(0.349)		(0.289)		(0.291)	
	[6.0]		[10.6]		[9.4]	
IV, sample with no overlap in secondary schools, school characteristics controls[a,c]	0.405		0.551*		0.604**	
	(0.366)		(0.322)		(0.304)	
	[5.7]		[9.2]		[8.7]	

Note: Each coefficient estimate is from a separate regression. Coefficient estimated from 2SLS regressions are reported. Robust standard errors in parentheses; disturbance terms are allowed to be correlated within but not across villages; significantly different from 0 at *90%, **95%, and ***99% confidence. First-stage F-statistics of significance of partial correlation between IV (treatment indicator) and social network measure are reported in brackets.

[a] Individual and household-level controls are the child's gender, indigenous status, the household's welfare index, education, age, and gender of the head of household, family size, and distance to secondary school. These are included in all specifications.

[b] Village predetermined contextual controls are the proportion of secondary school–age girls and the proportion of indigenous children in the village, mean village-level family size and educational level, age, and gender proportion of heads of households.

[c] School characteristics are indicator variables for general, technical, secondary schools (relative to *tele-secundaria* schools), urban and semiurban school indicators (relative to rural schools), school-level pupil-to-teacher ratio, and the number of home teachers, teaching assistants, physical education teachers, and art teachers in school.

following the start of the intervention (the *densificado* households), this instability in eligibility status could have led to uncertainty about the potential future eligibility of other nonbeneficiary households. In addition, a large proportion of eligible households (27% of the total eligible population and mostly *densificado* households) never received program payments during the evaluation period.[36] To the extent that this mismanagement led to uncertainty and changes in expectations about future eligibility, ineligible households could have increased their children's school participation in order to maximize their opportunity of becoming beneficiaries (although it is equally plausible that they would have reduced their children's school participation as well).

Although expectations of program eligibility are unfortunately unobserved, rendering this hypothesis nontestable, we do provide some indirect evidence to address this issue. If the extent of uncertainty surrounding the implementation of the program was more prevalent in villages where the incorporation of *densificado* households was higher, we should expect higher increases in the school participation of ineligible children in these specific villages. However, when we estimate a schooling decision reduced-form model with an interaction term of the PROGRESA treatment indicator and the proportion of *densificado* households in the village, we find that the interaction term is small and not significantly different from 0 (not reported in the tables).[37] Additionally, if uncertainty about future program eligibility during the year 1998 was the main mechanism at play, we should not observe positive school enrollment outcomes during the second year of the program, once the uncertainty had been resolved. However, we do find positive spillover effects on school enrollment and grade promotion during the 1999–2000 academic year, especially among the relatively poor ineligible children, and the subgroup of children just entering secondary school (see tables 3–5). Although these results do not disprove the eligibility expectations hypothesis, they diminish its plausibility.

VI. Discussion of Alternative Interpretations

Our results suggest that PROGRESA had a significant impact on the secondary school enrollment and grade promotion of children from ineligible households residing in the treatment villages. These findings support a simple model of social interactions where ineligible children are changing their enrollment decisions in response to their village peers. There are, however, at least two other alter-

[36] Previous researchers of the program suspect that these households were never formally incorporated into the program (Hoddinott & Skoufias, 2004).

[37] Estimates are available from the authors on request.

TABLE 9.—EVIDENCE ON SCHOOL QUALITY IMPROVEMENTS

	Year 1998				Year 1999			
	Treatment	Control	Diff.	s.e.	Treatment	Control	Difference	s.e.
A: Type of school								
Tele-secondary school	0.85	0.87	−0.02	(0.03)	0.85	0.86	−0.01	(0.03)
General secondary school	0.05	0.04	0.01	(0.02)	0.05	0.05	0.00	(0.02)
Technical secondary school	0.09	0.08	0.01	(0.03)	0.08	0.07	0.01	(0.03)
Rural	0.92	0.94	−0.02	(0.02)	0.92	0.93	−0.01	(0.03)
Semiurban	0.07	0.04	0.03	(0.02)	0.07	0.05	0.02	(0.02)
B: Groups and classrooms								
Number of groups, grade 7	1.36	1.36	0.00	(0.09)	1.39	1.37	0.02	(0.09)
Number of groups, grade 8	1.32	1.31	0.01	(0.08)	1.38	1.35	0.03	(0.08)
Number of groups, grade 9	1.25	1.24	0.01	(0.07)	1.28	1.26	0.02	(0.07)
Number of classrooms, grade 7	1.24	1.20	0.04	(0.09)	1.24	1.23	0.01	(0.09)
Number of classrooms, grade 8	1.10	1.03	0.07	(0.09)	1.18	1.12	0.06	(0.09)
Number of classrooms, grade 9	1.01	0.94	0.07	(0.09)	1.08	1.09	−0.01	(0.08)
Number of shared classrooms	0.20	0.32	−0.12	(0.08)	0.26	0.26	0.00	(0.09)
C: Number of teachers								
Number of home teachers	2.92	2.89	0.03	(0.32)	3.05	3.13	−0.08	(0.33)
Physical education teachers	0.11	0.12	−0.01	(0.04)	0.12	0.15	−0.03	(0.04)
Art teachers	0.10	0.11	−0.01	(0.04)	0.12	0.10	0.02	(0.03)
Teaching assistants	0.32	0.29	0.03	(0.10)	0.32	0.29	0.03	(0.09)
D: Home teacher qualifications								
Incomplete primary–secondary school	0.01	0.00	0.01	(0.00)	0.01	0.00	0.01	(0.00)
Technical degree	0.02	0.01	0.01	(0.02)	0.01	0.01	0.00	(0.01)
High school (*bachillerato*)	0.01	0.02	−0.01	(0.02)	0.00	0.01	−0.01	(0.01)
Teacher's college, primary	0.01	0.01	0.00	(0.01)	0.00	0.00	0.00	(0.01)
Teacher's college, superior incomplete	0.31	0.27	0.04	(0.07)	0.29	0.33	−0.04	(0.06)
Teacher's college, superior intern	0.45	0.59	−0.14	(0.13)	0.55	0.65	−0.10	(0.13)
Teacher's college, superior complete	1.15	1.04	0.11	(0.17)	1.17	1.07	0.10	(0.16)
Bachelor's degree, incomplete	0.05	0.04	0.01	(0.02)	0.05	0.08	−0.03	(0.03)
Bachelor's degree, intern	0.39	0.40	−0.01	(0.08)	0.38	0.36	0.02	(0.08)
Bachelor's degree, complete	0.31	0.38	−0.07	(0.08)	0.38	0.42	−0.04	(0.08)
MA, incomplete	0.08	0.07	0.01	(0.04)	0.08	0.06	0.02	(0.03)
MA, complete	0.04	0.04	0.00	(0.04)	0.05	0.09	−0.04	(0.04)
E: Enrollment								
All boys and girls, grades 7–9	93.79	91.02	2.77	(9.13)	100.84	96.62	4.22	(9.28)
Girls entering grades 7–9	43.96	41.19	2.77	(4.17)	48.19	45.06	3.13	(4.37)
Boys entering grades 7–9	49.26	49.36	−0.10	(4.99)	52.16	51.15	1.01	(4.96)
Pupil-to-teacher ratio	24.86	23.65	1.21	(1.03)	25.21	23.75	1.46	(0.96)

Note: Differences estimates in OLS regression models. Robust standard errors in parentheses; disturbance terms are allowed to be correlated within but not across villages; significantly different from zero at †85%, *90%, **95%, and ***99% confidence.

native interpretations that, due to data constraints, are difficult to reject.

One alternative interpretation for our findings is that ineligible children are responding to some of PROGRESA's supply-side interventions. However, as reported in table 9, there appear to be few differences in observable school characteristics between schools attended by children from treatment villages relative to those attended by children from control villages. Compared to the schools attended by the control villages, the number of classrooms and teachers is slightly higher on average in the treatment villages in 1998, but the differences, which are quite small and statistically insignificant, are reduced even further by the 1999–2000 academic year (panels B and C). There is also only minimal evidence that the qualification of home teachers was related to any PROGRESA supply-side-related improvement in school quality. There are slightly more teachers who have completed a superior education in the treatment schools, but again the difference is not statistically significant (panel D). Interestingly, there is evidence of an increase in the mean pupil-teacher ratio in program schools during both academic years (1.21 and 1.46 during the 1998–99 and 1999–2000 academic years; statistically insignificant). Moreover, the secondary schools attended by the poorer ineligible children suffered a (marginally significant) increase in pupil teacher ratios of 1.78, as expected from the increased school enrollment among eligible and ineligible children from these villages (not reported in the table).[38] If any negative congestion effect took place, we

[38] These estimated increases are within the expected ranage from the household survey estimates of increases in school enrollment. A back-of-the-envelope calculation implies an expected increase of 1.42 in the pupil-teacher ratio. Approximately 76% of children in the villages were eligible, and approximately 59% of the ineligibles belonged to the below-median welfare index group. In addition, there are approximately twenty

TABLE 10.—TEST OF SUPPLY-SIDE RESPONSES TO INCREASED SCHOOL ENROLLMENT

Specification (Dependent variable: school enrollment indicator)	Coefficient Estimate on Treatment Village Indicator (SE) — Sample: Ineligible Children in Treatment and Control Group Villages Attending the Same Secondary School		
	All Children, Grades 6–9 (1) OLS	Welfare < Median, Grades 6–9 (2) OLS	All Children, Grades 6–7 (3) OLS
No fixed effects	0.050	0.072	0.060
	(0.063)	(0.073)	(0.075)
School-period fixed effects	0.052	0.068	0.082
	(0.081)	(0.096)	(0.146)
Individual and household controls	Yes	Yes	Yes
Contextual controls	Yes	Yes	Yes
Observations	406	327	281
Mean of dependent variable	0.525	0.489	0.441

Note: Coefficient estimates from OLS regressions are reported. Robust standard errors in parentheses; disturbance terms are allowed to be correlated within but not across villages; significantly different from zero at †85%, *90%, **95%, and ***99% confidence. Individual and household-level controls are the child's gender, indigenous status, the household's welfare index, education, age, and gender of the head of household, family size, and distance to secondary school. Village contextual controls are the proportion of secondary school–age girls and the proportion of indigenous children in the village; mean village-level family size; and educational level, age, and gender proportion of heads of households.

would expect in equilibrium a reduction in school enrollment among ineligible children, biasing our peer effects estimates downward.

We also estimate neighborhood peer effects models including controls for a large set of these contemporaneous school characteristics and find that these estimates are quite robust: the point estimate for the overall sample is 0.52 (significant at 90% confidence), and those for the low-welfare and lower-grade subgroups, are respectively, 0.64 and 0.60 (both significant at 95% percent confidence) (table 7, row 6). Overall these estimates suggest that it is unlikely that ineligible children are responding to differences in observable school quality changes.

To further investigate the potential supply-side responses, Table 10 reestimates our reduced-form models, restricting the sample to the set of children likely to attend secondary schools that serve ineligible children from both treatment and control villages. If supply-side responses were driving our results, we would not expect PROGRESA to affect this sample of ineligible households. Yet the reduced-form estimates for this sample are consistent with those presented in the overall sample.[39] Among households below the median of the welfare index, PROGRESA increased secondary school enrollment by 7.2 percentage points. Even when we account for school-period fixed effects, which would capture potential supply-side responses at the school level, the point estimates change only slightly. In general, these are imprecisely measured due to the small sample size.

In sum, our analysis suggests that our reduced-form results are not simply due to improvements in school characteristics such as the number of classrooms or the quality of the teachers. But unfortunately we cannot reject that there were systematic improvements in supply-side input that are more difficult to measure, such as teacher motivation or effort (Kremer et al., 2009; Duflo, Dupas, & Kremer, 2007; Banerjee & Duflo, 2006).

Another alternative interpretation for our findings is that ineligible children may have simply responded to information regarding the benefits of schooling and attaining an education. For instance, a village-wide information campaign by PROGRESA organizers (*promotoras*), or perhaps simply knowing that the government was willing to provide large transfers to raise secondary school enrollment rates, may have induced parents and students to update their priors on the value of enrollment, affecting the decisions of non-eligible households directly (Jensen, 2007). While we cannot refute this possibility, de Brauw and Hoddinott (2007) show that among PROGRESA households for which the conditionality constraints were not enforced, but where information about the program's requirements was well known, secondary school enrollment program impacts among eligible households were close to zero. This would suggest that information about the program, in combination with unconditional cash transfers, did not increase the secondary school enrollment of children in eligible households.

Finally, to the extent that the information transmission effect is homogeneous across children having completed grades 5 and 6 relative to those having completed grades 7 and 8 at baseline, specifications comparing children within the same household (household fixed effects models), for which we estimate the differential effect for children in lower grades relative to those in upper grades, should be

children of secondary school age per village. Using the estimate program impacts among eligible children of 8.3 percentage points increase in the secondary school enrollment rate and the 5.5 percentage point increase among the below-median welfare index group of ineligibles, we can estimate the mean increase in the number of pupils as $20 \times [(0.76) \times 0.083 + (0.14) \times 0.055] = 1.42$ pupils.

[39] Due to the small sample size, we lose precision in the first-stage regressions and cannot get consistent IV estimates of the neighborhood peer effects for this subgroup of children. Also, note that any crowding-out effect that is systematically more likely to affect control group children could lead to overestimation of the spillover effects, although, based on our discussion in section 5.2, this does not seem to be taking place.

purged of any household-level information effect. Estimates from these reduced-form specifications, although imprecisely estimated, suggest that the effects are robust to the household-level information effects: the point estimate for the differential spillover effect excluding household fixed effects is 0.074 percentage points (standard error 0.048, significant at 13% confidence; not reported in the table) and increases to 0.076 percentage points (standard error 0.083, insignificantly different from zero; not reported in the table) once household fixed effects are included. In sum, although these pieces of evidence suggest that it is unlikely that ineligible households would respond in such a way to information regarding the program and the value of secondary schooling, we cannot reject that possible systematic changes in households' beliefs and expectations about the value of a secondary-level education may have promoted an increase in enrollment.

VII. Conclusion

In 1997, the Mexican government introduced a randomly phased-in human development program designed to increase human capital among the rural poor. This study uses experimental variation in the school enrollment rates among program-eligible households to estimate how peers' school enrollments influence the school decisions of children ineligible to receive these program benefits. Our findings suggest that the enrollment behavior of one's peers has an important role on a child's decision to enroll in school. A 10 percentage point increase in the enrollment rate of a child's reference group increases his likelihood of attending secondary school by approximately 5 percentage points. These peer effects are more pronounced among children of relatively poorer households within the group of those who were ineligible. Furthermore, we are able to reject hypotheses on other potential contextual interaction effects using rich microdata on household consumption and expenditures, health of individual members, and administrative data on program transfers and school characteristics. These sensitivity analyses confirm the validity of the identifying assumptions of the empirical social interactions model.

Our estimates lie in the upper range of existing social multiplier estimates of school enrollment and dropout behavior in both neighborhood-based and school-based contexts. The point estimates imply social multiplier effects in the range of 2.0 and 3.0, with a preferred estimate of approximately 2.5 (peer effects of 0.595). This indicates that behavioral social interaction effects approximately doubled the direct effects of school enrollment subsidies among secondary-school-aged children in these marginalized areas. In contrast, Case and Katz's (1991) estimates of peer effects in idleness among youth in high-poverty neighborhoods in Boston imply a social multiplier effect of 1.33 (peer effects estimate of 0.25). However, Ginther, Haveman, and Wolfe (2000) and Aaronson (1998) report small estimates of social multipliers in peers' dropout behavior—in the range of 1.02 to 1.06 (often statistically insignificant)—from a sample of youth in the PSID. Estimates based on school-based reference groups suggest that high school dropout behavior of students in the United States follows social multiplier of approximately 1.20 (Gaviria & Raphael, 2001). Our estimates for children in marginal villages in rural Mexico, although not directly comparable to these, are more in line with those of Case and Katz (1991), who report estimates for a sample of marginalized youth. Notwithstanding the differences in sample and methodology, our results suggest that peer effects may be much more prevalent for marginal populations in less developed countries and consequently have important implications for the design of education policy especially in these contexts.

Future research should empirically differentiate the specific mechanisms for which we observe these reduced-form interactions. Theoretical models within economics incorporate behavioral peer effects as a result of identity formation behavior (Akerlof & Kranton, 2002), conformity behavior (Bernheim, 1994; Akerlof, 1997), and informational externalities (Bikhchandani, Hirschleifer, & Welch 1992), among others. Current work attempting to distinguish these effects, such as Akerlof and Kranton (2002), Kremer and Miguel (2007), and Munshi and Myaux (2006), could serve researchers as guides for these types of studies.

REFERENCES

Aaronson, Daniel, "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes," *Journal of Human Resources* 33 (1998), 915–46.

Adato, M., B. de la Brière, D. Mindek, & A. Quisumbing, *Final Report: The Impact of PROGRESA on Women's Status and Intrahousehold Relations* (Washington, DC: International Food Policy Research Institute, 2000).

Akerlof, George, "Social Distance and Social Decisions," *Econometrica* 65:5 (1997), 1005–1027.

Akerlof, George, and Rachel Kranton, "Identity and Schooling: Some Lessons for the Economics of Education," *Journal of Economic Literature* 40:4 (2002), 1167–1201.

Angelucci, Manuela, and Giacomo De Giorgi, "Indirect Effects of an Aid Program: How Do Cash Injections Affect Ineligibles' Consumption?" *American Economic Review* 99 (2009), 486–508.

Attanasio, Orazio, and José Víctor Ríos-Rull, "Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare?" *European Economic Review* 44:7 (2000), 1225–1258.

Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden, "Remedying Education: Evidence from Two Randomized Experiments in India," *Quarterly Journal of Economics* 122 (2007), 1235–1264.

Banerjee, Abhijit, and Esther Duflo, "Addressing Absence," *Journal of Economic Perspectives* 20:1 (2006), 117–132.

Becker, Gary S., and Kevin M. Murphy, *Social Economics: Market Behavior in a Social Environment* (Cambridge, MA: Harvard University Press, 2000).

Behrman, Jere R., Piyali Sengupta, and Petra E. Todd, "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment," *Economic Development and Cultural Change* 54 (2005), 237–275.

Behrman, Jere R., and Petra E. Todd, *Randomness in the Experimental Samples of PROGRESA (Education, Health, and Nutrition Program)* (Washington, DC: International Food Policy Research Institute, 1998).

Bénabou, Roland, "Equity and Efficiency in Human Capital Investment: The Local Connection," *Review of Economic Studies* 63 (1993), 237–264.
——— "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review* 86 (1996), 584–609.
Bernheim, B. Douglas, "A Theory of Conformity," *Journal of Political Economy* 102 (1994), 841–877.
Bikhchandani, Sushil, David Hirschleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, Cultural Changes as Informational Cascades," *Journal of Political Economy* 100 (1992), 992–1026.
Bobonis, Gustavo J., "Is the Allocation of Resources within the Household Efficient? New Evidence from a Randomized Experiment," *Journal of Political Economy* 117 (2009), 453–503.
Cameron, Stephen, and James Heckman, "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy* 106 (1998), 262–333.
Case, Anne C., and Lawrence F. Katz, "The Company You Keep: The Effect of Family and Neighborhood on Disadvantaged Youths," NBER working paper no. 3705 (1991).
Cipollone, Piero, and Alfonso Rosolia, "Social Interactions in High School: Lessons from an Earthquake," *American Economic Review* 97 (2007) 948–965.
de Brauw, Alan, and John Hoddinott, "Must Conditional Cash Transfer Programs Be Conditioned to Be Effective? The Impact of Conditioning Transfers on School Enrollment in Mexico," mimeograph, International Food Policy Research Institute (2007).
de Janvry, Alain, Frederico Finan, Elizabeth Sadoulet, and Renos Vakis, "Can Conditional Cash Transfers Serve as Safety Nets at School and from Working When Exposed to Shocks?" *Journal of Development Economics* 79 (2006), 348–373.
Duflo, Esther, Pascaline Dupas, and Michael Kremer, "Peer Effects, Pupil-Teacher Ratios, and Teacher Incentives: Evidence from a Randomized Evaluation in Kenya," Mimeograph, Massachusetts Institute of Technology (2007).
Duflo, Esther, and Emmanuel Saez, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics* 118 (2003), 815–842.
Durlauf, Steven N., and H. Peyton Young (Eds.), *Social Dynamics* (Washington, DC: Brookings Institution Press, 2001).
Evans, William N., Wallace E. Oates, and Robert M. Schwab, "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy* 100 (1992), 966–991.
Foster, George M., *Tzintzuntzan: los campesinos mexicanos en un mundo en cambio* (Colo Bosqua del Pedregal, México: Fondo de Cultura Económica, 1967).
Gaviria, Alejandro, and Steven Raphael, "School-based Peer Effects and Juvenile Behavior," this REVIEW 83 (2001), 257–268.
Gertler, Paul, "Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment," *American Economic Review* 94 (2004), 336–341.
Gertler, Paul, and Simone Boyce, "An Experiment in Incentive-Based Welfare: The Impact of PROGRESA on Health in Mexico," Mimeograph, University of California, Berkeley (2001).
Ginther, Donna, Robert Haveman, and Barbara Wolfe, "Neighborhood Attributes as Determinants of Children's Outcomes: How Robust Are the Relationships?" *Journal of Human Resources* 35 (2000), 603–642.
Glaeser, Edward and José A. Scheinkman. "Nonmarket Interactions" (vol. 1, pp. 339–369), in Mathias Dewatripont, Lars Peter Hansen, and Stephen Turnovsky (Eds.), *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress* (Cambridge: Cambridge University Press, 2003).
Glewwe, Paul, and Michael Kremer, "Schools, Teachers, and Education Outcomes in Developing Countries" (vol. 2, pp. 946–1143), in Erik Hanushek and Finis Welch (Eds.), *Handbook on the Economics of Education* (New York: Elsevier, 2005).
Graham, Bryan S., "Identifying Social Interactions through Conditional Variance Restrictions," *Econometrica* 76 (2008), 643–660.
Gray-Molina, George, Ernesto Pérez de Rada, and Wilson Jiménez, "Residential Segregation in Bolivian Cities" (pp. 25–43), in Behrman, Jere, Alejandro Gaviria, and Miguel Székely (eds.), *Who's In and Who's Out: Social Exclusion in Latin America* (Washington, DC: Inter-American Development Bank, 2003).
Hayashi, Fumio, *Econometrics* (Princeton, NJ: Princeton University Press, 2000).
Hoddinott, John, and Emmanuel Skoufias, "The Impact of PROGRESA on Food Consumption," *Economic Development and Cultural Change* 53 (2004), 37–61.
Hoxby, Caroline M., "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER working paper no. 7867 (2000).
Hoxby, Caroline M., and Gretchen Weingarth, "Taking Race Out of the Question: School Reassignment and the Structure of Peer Effects," Mimeograph, Harvard University (2006).
Jacoby, Hanan G., and Emmanuel Skoufias, "Risk, Financial Markets, and Human Capital in a Developing Country," *Review of Economic Studies* 64 (1997), 311–335.
Jensen, Robert. "The Perceived Returns to Education and the Demand for Schooling," Mimeograph, Brown University (2007).
Kling, Jeffrey R., Jeffrey B., Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica* 75 (2007), 83–119.
Kremer, Michael, and Edward Miguel, "The Illusion of Sustainability," *Quarterly Journal of Economics* 112 (2007), 1007–1065.
Kremer, Michael, Edward Miguel, and Rebecca Thornton, "Incentives to Learn," this REVIEW 91 (2009).
Lalive, Rafael, and Alejandra Cattaneo, "Social Interactions and Schooling Decisions," this REVIEW 91 (2009).
Lazear, Edward, "Educational Production," *Quarterly Journal of Economics* 116 (2001), 777–803.
Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies* 60 (1993), 531–542.
Miguel, Edward, and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica* 72 (2004), 159–217.
Moffitt, Robert, "Policy Interventions, Low-Level Equilibria, and Social Interactions" (pp. 45–82), in Steven N. Durlauf and H. Peyton Young (eds.), *Social Dynamics* (Washington, DC: Brookings Institution Press, 2001).
Munshi, Kaivan, and Jacques Myaux, "Social Norms and the Fertility Transition," *Journal of Development Economics* 80 (2006), 1–38.
O'Regan, K., and J. Quigley, "Spatial Effects upon Employment Outcomes: The Case of New Jersey Teenagers," *New England Economic Review: Federal Reserve Bank of Boston* (1996), 41–57.
Oreopoulos, Philip, "The Long-Run Consequences of Living in a Poor Neighborhood," *Quarterly Journal of Economics* 118 (2003), 1533–1575.
Parker, Susan, and Emmanuel Skoufias, *Final Report: The Impact of PROGRESA on Work, Leisure, and Time Allocation* (Washington, DC: International Food Policy Research Institute, 2000).
Sánchez-Peña, Landy L., "Gender and Socio-economic Residential Segregation in Mexico City," paper presented at the Population Association of America 2007 annual meeting (2007).
Schultz, T. Paul, "School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program," *Journal of Development Economics* 74 (2004), 199–250.
——— "An Addendum to the Final Report: An Evaluation of the Selection of Beneficiary Households in the Education, Health, and Nutrition Program (PROGRESA) of Mexico Dated June 4, 1999," (Washington, DC: International Food Policy Research Institute, 1999).
Skoufias, Emmanuel, Benjamin Davis, and Sergio de la Vega, "Targeting the Poor in Mexico: An Evaluation of the Selection of Households into PROGRESA," *World Development* 29 (2001), 1769–1984.
Strauss, John, and Duncan Thomas, "Health over the Life Course," in T. P. Schultz and J. Strauss (Eds.), *Handbook of Development Economics,* Vol. 4 (Amsterdam: North-Holland Press, 2007).
Topa, Giorgio, "Social Interactions, Local Spillovers and Unemployment," *Review of Economic Studies* 68 (2001), 261–296.

Table A1.—Relationship between Attrition and Characteristics of Children at Baseline
(Dependent variable: Attrition Indicator)

	Ineligible Children				Eligible Children			
	Treatment	Correlates	Main Effect of Correlates	Interaction of Correlates with Treatment	Treatment	Correlates	Main Effect of Correlates	Interaction of Correlates with Treatment
	(1) OLS	(2) OLS	(3) OLS		(4) OLS	(5) OLS		(6) OLS
Treatment village	−0.006	0.008	0.102		0.000	0.005	0.016	
	(0.005)	(0.010)	(0.122)		(0.003)	(0.006)	(0.075)	
Treatment × Year 1998	0.008		−0.005		0.002		0.004	
	(0.020)		(0.021)		(0.011)		(0.010)	
Treatment × Year 1999	0.034		0.017		0.022		0.029**	
	(0.025)		(0.026)		(0.014)		(0.014)	
Year 1998	0.208***	0.150***	0.154***		0.196***	0.140***	0.138***	
	(0.016)	(0.010)	(0.016)		(0.009)	(0.005)	(0.008)	
Year 1999	0.201***	0.095***	0.087***		0.199***	0.096***	0.078***	
	(0.018)	(0.013)	(0.018)		(0.011)	(0.007)	(0.010)	
Child's age		0.062***	0.058***	0.005		0.059***	0.061***	−0.005
		(0.003)	(0.004)	(0.005)		(0.002)	(0.003)	(0.003)
Grade completed in 1997		−0.029***	−0.027***	−0.004		−0.019***	−0.026***	0.012**
		(0.005)	(0.008)	(0.010)		(0.003)	(0.004)	(0.005)
Gender (boy)		−0.013*	−0.022*	0.016		−0.021***	−0.017***	−0.008
		(0.008)	(0.012)	(0.015)		(0.004)	(0.006)	(0.008)
Indigenous		0.034	0.019	0.029		0.003	0.011	−0.014
		(0.021)	(0.031)	(0.041)		(0.007)	(0.012)	(0.015)
Family size		−0.004**	−0.001	−0.005		−0.002**	−0.002	0.000
		(0.002)	(0.003)	(0.004)		(0.001)	(0.002)	(0.002)
Head of household education		0.002	0.003	−0.002		0.000	−0.001	0.001
		(0.002)	(0.003)	(0.004)		(0.001)	(0.002)	(0.002)
Head of household gender (male)		−0.010	−0.012	0.006		−0.029***	−0.027*	−0.004
		(0.017)	(0.023)	(0.034)		(0.010)	(0.016)	(0.020)
Head of household age		0.001*	0.001	0.000		0.000	0.000	−0.001
		(0.000)	(0.001)	(0.001)		0.000	0.000	0.000
Distance to secondary school		0.003	0.004	0.000		0.003*	0.003	0.000
		(0.003)	(0.004)	(0.005)		(0.002)	(0.003)	(0.003)
Distance to urban center		0.000	0.001**	−0.001*		0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)		0.000	0.000	0.000
General secondary school		−0.053	−0.012	−0.056		0.016	0.002	0.012
		(0.034)	(0.072)	(0.083)		(0.019)	(0.028)	(0.038)
Technical secondary school		−0.067**	−0.073	0.026		0.001	−0.014	0.030
		(0.027)	(0.051)	(0.066)		(0.015)	(0.025)	(0.033)
Urban school		0.011	0.149	−0.156		−0.044**	0.078	−0.145**
		(0.029)	(0.174)	(0.176)		(0.022)	(0.059)	(0.064)
Semiurban school		−0.003	−0.033	0.043		0.005	0.014	−0.009
		(0.023)	(0.025)	(0.041)		(0.013)	(0.018)	(0.026)
Number of home teachers		0.001	0.009	−0.013		0.007	−0.003	0.013
		(0.012)	(0.029)	(0.032)		(0.006)	(0.014)	(0.015)
Physical education teachers		0.029	0.023	−0.001		0.003	−0.025	0.042
		(0.038)	(0.069)	(0.086)		(0.019)	(0.025)	(0.035)
Art teachers		−0.064*	−0.077	0.043		−0.007	−0.017	0.014
		(0.034)	(0.054)	(0.067)		(0.018)	(0.021)	(0.035)
Teaching teachers		0.021	−0.008	0.039		−0.001	0.004	−0.006
		(0.013)	(0.021)	(0.026)		(0.007)	(0.008)	(0.013)
Number of teachers		0.007	0.003	0.008		−0.007	0.002	−0.012
		(0.013)	(0.031)	(0.034)		(0.006)	(0.014)	(0.016)
Pupil-teacher ratio		0.000	0.000	−0.001		0.000	0.000	−0.001
		(0.001)	(0.001)	(0.001)		0.000	(0.001)	(0.001)
Constant	0.011**	−0.681***	−0.755***		0.006***	−0.596***	−0.605***	
	(0.005)	(0.056)	(0.104)		(0.002)	(0.031)	(0.066)	
State indicators	Yes	Yes	Yes		Yes	Yes	Yes	
State × Treatment indicators	No	No	Yes		No	No	Yes	
Observations	8,214	8,184	8,184		33,441	33,351	33,351	
R^2	0.08	0.16	0.17		0.08	0.15	0.15	
Interactions F-statistic		—	1.29			—	1.36	
P-value			[0.149]				[0.105]	

Note: Robust standard errors in parentheses; disturbance terms are allowed to be correlated within but not across villages. Significantly different than zero at *90%, **95%, and ***99% confidence. F-statistic of joint significance of interaction terms ($F(28,379)$ reported at bottom of the table.

APPENDIX

Identification of Endogenous Peer Effects in a Partial-Population Experiment Model with Potential Contextual Characteristics Effects

We present a more flexible linear-in-means model of social interactions that allows direct treatment effects on children's contextual characteristics. We derive equilibrium reduced-form equations relating the partial-population treatment to individuals' school enrollment behavior given the potential effect of a second mechanism through a contextual variable. This model allows us to formally determine the conditions under which we can identify endogenous peer effects.

We start with our linear-in-means partial-population experiment model, where PROGRESA treatment essentially works as a subsidy for secondary school enrollment of eligible children. The school enrollment best response functions are:

$$y^E_{ic} = \alpha + \beta X^E_{ic} + \gamma \bar{X}_{-i,c} + \lambda Z_c + \theta \bar{y}_{-i,c} + \delta T^E_{ic} + u^E_{ic} \tag{A1$'$}$$

$$y^{NE}_{ic} = \alpha + \beta X^{NE}_{ic} + \gamma \bar{X}_{-i,c} + \lambda Z_c + \theta \bar{y}_{-i,c} + u^{NE}_{ic}. \tag{A1$''$}$$

However, we now allow the PROGRESA treatment to also affect other contextual characteristics of eligible children. We assume for simplicity that the program affects only one contextual characteristic, X_{1c}, and the vector of contextual determinants of enrollment can be decomposed into $X_c = [X_{1c}, X_{2c}]$. Moreover, the choice of X_{1c} is itself a linear function of individuals' predetermined contextual characteristics (X_{2c}), environmental factors affecting all children in the village (Z_c), and the schooling price subsidy (T^E_{ic}):

$$\begin{aligned} X^E_{1,ic} = {} & \pi_0 + \pi_1 X^E_{2,ic} + \pi_2 \bar{X}_{1,-i,c} + \pi_3 \bar{X}_{2,-i,c} + \pi_4 T^E_{ic} \\ & + \pi_5 Z_c + \nu^E_{ic} \end{aligned} \tag{A2$'$}$$

$$\begin{aligned} X^{NE}_{1,ic} = {} & \pi_0 + \pi_1 X^{NE}_{2,ic} + \pi_2 \bar{X}_{1,-i,c} + \pi_3 \bar{X}_{2,-i,c} \\ & + \pi_5 Z_c + \nu^{NE}_{ic}. \end{aligned} \tag{A2$''$}$$

Based on equations (A1′)–(A2″) and equations (3′) and (3″), we can solve for the reduced-form equilibrium school enrollment choices of children based on these best-response functions. These reduced-form equations are a complex function of the potential direct impacts of the subsidy on y and X_1, endogenous social interactions in y, and potential contextual and endogenous social interactions in X_1.

The solution involves simple algebra on the simultaneous equation model. First, averaging equations (A1′)–(A2″) at the village level, we get:

$$\begin{aligned} \bar{y}^E_c = {} & \alpha + \beta_1 \bar{X}^E_{1c} + \beta_2 \bar{X}^E_{2c} + \gamma_1 \bar{X}_{1,c} + \gamma_2 \bar{X}_{2,c} \\ & + \lambda Z_c + \theta \bar{y}_c + \delta T^E_c \end{aligned} \tag{A4$'$}$$

$$\bar{y}^{NE}_c = \alpha + \beta_1 \bar{X}^{NE}_{1c} + \beta_2 \bar{X}^{NE}_{2c} + \gamma_1 \bar{X}_{1,c} + \gamma_2 \bar{X}_{2,c} + \lambda Z_c + \theta \bar{y}_c \tag{A4$''$}$$

$$\bar{X}^E_{1c} = \pi_0 + \pi_1 \bar{X}^E_{2c} + \pi_2 \bar{X}_{1,c} + \pi_3 \bar{X}_{2,c} + \pi_4 T^E_c + \pi_5 Z_c \tag{A4$'''$}$$

$$\bar{X}^{NE}_{1c} = \pi_0 + \pi_1 \bar{X}^{NE}_{2c} + \pi_2 \bar{X}_{1,c} + \pi_3 \bar{X}_{2,c} + \pi_5 Z_c. \tag{A4$''''$}$$

Substituting these conditions into equations (3′) and (3″) yields

$$\begin{aligned} \bar{y}_c = {} & \frac{\alpha}{1-\theta} + \left(\frac{\beta_1 + \gamma_1}{1-\theta}\right)\bar{X}_{1,c} + \left(\frac{\beta_2 + \gamma_2}{1-\theta}\right)\bar{X}_{2,c} + \frac{\lambda}{1-\theta} Z_c \\ & + \frac{\delta}{1-\theta} m_c T_c \end{aligned} \tag{A5}$$

$$\bar{X}_{1,c} = \frac{\pi_0}{1-\pi_2} + \left(\frac{\pi_1 + \pi_3}{1-\pi_2}\right)\bar{X}_{2,c} + \frac{\pi_5}{1-\pi_2} Z_c + \frac{\pi_4}{1-\pi_2} m_c T_c. \tag{A6}$$

Finally, substituting equation (A6) in equation (A5) gives the reduced-form equilibrium school enrollment equation at the village level:

$$\begin{aligned} \bar{y}_c = {} & \left[\frac{\alpha}{1-\theta} + \frac{\pi_0}{1-\pi_2}\right] + \left[\frac{\beta_2 + \gamma_2}{1-\theta} + \left(\frac{\beta_1 + \gamma_1}{1-\theta}\right) \right. \\ & \times \left.\left(\frac{\pi_1 + \pi_3}{1-\pi_2}\right)\right]\bar{X}_{2,c} + \left[\frac{\lambda}{1-\theta} + \left(\frac{\beta_1 + \gamma_1}{1-\theta}\right)\right. \\ & \times \left.\left(\frac{\pi_5}{1-\pi_2}\right)\right] Z_c + \left[\frac{\delta}{1-\theta} + \left(\frac{\beta_1 + \gamma_1}{1-\theta}\right)\left(\frac{\pi_4}{1-\pi_2}\right)\right] m_c T^E_c. \end{aligned} \tag{A7}$$

The reduced-form equilibrium condition has a very intuitive explanation, since it shows how the partial-population price subsidy affects equilibrium enrollment decisions through different channels. $[\delta m_c/(1 - \theta)]$ represents the direct effect of the subsidy on school enrollment, augmented by the school enrollment social multiplier effect; $[\pi_4 m_c/(1 - \pi_2)]$ represents the direct effect of the subsidy on the contextual variable X_1, compounded by the endogenous peer effects or externalities in the choice of X_1 of individuals in the reference group. Finally, $[m_c\ (\beta_1 + \gamma_1)/(1 - \theta)]$ represents how these effects on X_1 are channeled to affect school enrollment, through direct effects (β_1), contextual peer effects (γ_1), and endogenous peer effects $[1/(1 - \theta)]$.

We can also solve for the equilibrium school enrollment choice of ineligible children as a function of the PROGRESA subsidy and other exogenous determinants of school enrollment.

Substituting equations (A4″″), (A6), and (A7) in equation (A1″) and rearranging the structural coefficients (in order to assess the potential biases in the IV estimator), we can see that the reduced-form relationship is quite complex:

$$\begin{aligned} \bar{y}^{NE}_c = G + {} & \left[\frac{(\gamma_1 + \beta_1\pi_2)\pi_4}{1-\pi_2} + \theta\left(\frac{\delta}{1-\theta} + \left(\frac{\beta_1+\gamma_1}{1-\theta}\right)\right.\right. \\ & \times \left.\left.\left(\frac{\pi_4}{1-\pi_2}\right)\right)\right] m_c T^E_c, \end{aligned} \tag{A8}$$

where G is a linear function of $X^{NE}_{2,ic}$, $\bar{X}_{2,c}$, and Z_c. Therefore, the IV estimator leads to the estimate of the following composite parameter:

$$\text{plim } \hat{\theta}_{IV} = \left[\theta + \frac{(1-\theta)(\gamma_1 + \beta_1\pi_2)\pi_4}{(\gamma_1 + \beta_1)\pi_4 + \delta(1-\pi_2)}\right].$$

As can be seen from equation (A8), in order to get a consistent estimate of θ, one of the following two conditions needs to hold: (a) $\pi_4 = 0$, that is, that the PROGRESA subsidies cannot have any impact on other contextual variables, or (b) $(\gamma_1 + \beta_1\pi_2) = 0$, that the other variables affected have neither direct nor contextual social interaction effects on children's school enrollment decisions.