

# Getting Girls into School: Evidence from a Scholarship Program in Cambodia

DEON FILMER and NORBERT SCHADY  
World Bank

## I. Introduction

Raising the schooling levels of girls is often regarded as an important priority for developing countries (World Bank 2001; Schultz 2002). Frequently, the Mincerian rate of return to schooling for women is at least as large as that for men (for examples, see Deolalikar [1993] for Indonesia; Schultz [1993] for Thailand, Côte d'Ivoire, and Ghana). There is also a large literature documenting associations between female education and a variety of social outcomes, including lower fertility, decreases in child mortality, improvements in child health and nutrition, and better education and child cognitive development (see reviews in Strauss and Thomas 1995; World Bank 2001).

In a large number of countries, especially in the Middle East and parts of Asia and Africa, overall education for girls is low and lags significantly behind that of boys. Over 180 governments have adopted universal primary education and gender parity in schooling as Millennium Development Goals. Despite these commitments, increasing the schooling attainment of girls continues to be a challenge in much of the developing world.

There is surprisingly little evidence on policies and programs that effectively raise school attainment, including for girls (see the reviews by Glewwe 2002; Case 2004). Most of the existing evidence on the impact of demand-side incentives on schooling outcomes is based on evaluations from middle-income countries, especially in Latin America. Conditional cash transfer (CCT) programs have been found to have positive effects on enrollment in Mexico

We thank Luis Benveniste, Pedro Carneiro, Eric Edmonds, T. Paul Schultz, John Strauss, Miguel Urquiola, and two anonymous referees for many useful comments. We also thank the World Bank's Research Support Budget and the Bank Netherlands Partnership Program for financial support; and staff of the Cambodian Ministry of Education Youth and Sports, Asian Development Bank, and UNICEF, whose collaboration was instrumental in aiding us to complete this work. All errors are, of course, our own. The findings, interpretations, and conclusions expressed in this paper are entirely ours. They do not necessarily represent the views of the World Bank and its affiliated organizations or those of the executive directors of the World Bank or the governments they represent.

(Schultz 2004; Behrman, Sengupta, and Todd 2005), Colombia (Attanasio, Fitzsimmons, and Gomez 2005), Nicaragua (Maluccio and Flores 2004), Honduras (Glewwe and Olinto 2004), and Brazil (Cardoso and Portela Souza 2004). An important question is the extent to which the observed changes are a result of the income effects associated with the transfer, or the conditions—see Das, Do, and Özler (2005) for a theoretical discussion, and Schady and Araujo (2008) for an empirical application to Ecuador. Finally, Angrist et al. (2002) evaluate a school voucher program in Colombia; they find that the program increased high school graduation rates and test scores.

Although the findings from Latin America are informative, relatively little is known about the impact of demand-side programs in lower-income countries, with much weaker institutional contexts and lower quality education. Exceptions include Kim, Alderman, and Orazem (1999) on the Quetta Fellowship program in Pakistan, Ravallion and Wodon (2000) on the Food-for-Education program in Bangladesh, and Kremer, Miguel, and Thornton (2007) on a merit-based scholarship program in Kenya. All of these report large effects on school enrollment.

In this study we evaluate the impact of a program designed to increase the enrollment of girls in secondary school in Cambodia. The program we evaluate is the Japan Fund for Poverty Reduction (JFPR) scholarship program. This program, which began in the 2004 school year, awarded scholarships to poor girls who were completing sixth grade, the last grade of primary school. We show that the scholarship program increased the enrollment and attendance of recipients at program schools by about 30 percentage points. Larger impacts are found among girls with the lowest socioeconomic status at baseline. The results are robust to a variety of controls for observable differences between scholarship recipients and nonrecipients, to unobserved heterogeneity across girls, and to selective transfers between program schools and other schools. We conclude that there is substantial potential for demand-side interventions in lower-income countries like Cambodia.

The rest of the paper proceeds as follows. In Section II we briefly describe the setting, the JFPR scholarship program, and the data used in this study. Section III discusses the empirical strategy. In Section IV, we present the results. Section V concludes.

## **II. Data, Variables, and Program Selection Rules**

### ***A. Education in Cambodia and the JFPR Scholarship Program***

Cambodia is recovering from many years of internal and external strife. The Khmer Rouge regime of the 1970s and Vietnamese occupation in the 1980s had severe repercussions for all aspects of the economy and society, including

the education sector (De Walque 2004). Most Cambodian children attend school, but a large share complete only very few grades. According to the 2000 Demographic and Health Survey (DHS), 85% of 15–19-year-olds had completed grade 1 while only 27% had completed grade 7, the first year of lower secondary. These percentages are lower for rural areas, 83% and 21%, respectively, and lower yet for rural girls, 78% and 17%, respectively (Filmer 2007). To address these problems, the Cambodian government has initiated a series of reforms in the education sector, including scholarship programs for students from disadvantaged backgrounds.

The program we evaluate in this study is the Japan Fund for Poverty Reduction (JFPR) scholarship program. This program selected 93 lower secondary schools, and, within each of these schools, approximately 45 girls who were beginning seventh grade were awarded scholarships of \$45 each. (No more than 45 scholarships could be awarded per school; in a handful of schools, the number of scholarships awarded was smaller.) The value of the scholarship is large—in 2002, mean per capita gross domestic product in Cambodia was approximately \$300 (World Bank 2005). Another way of benchmarking the magnitude of the transfer is by comparing it with expenditures by households on education. On average, the \$45 transfer is almost exactly equivalent to average household spending per student in lower secondary school, as reported in a recent household survey. However, average expenditures on schooling by poor families, such as those that received JFPR scholarships, are generally lower.<sup>1</sup>

Once a girl is selected for a JFPR scholarship, she is automatically eligible to continue receiving a scholarship for the 3 years of the lower secondary cycle. The JFPR program therefore attempts to increase the fraction of girls who make the transition from primary school to lower secondary school and to encourage girls to complete the lower secondary school cycle. In the 2004 school year, there were 698 lower secondary schools in Cambodia, so the JFPR scholarship program covered approximately 15% of lower secondary schools in the country.

Although the JFPR program is known as a “scholarship” program, it does not directly subsidize the fees paid by parents for the education of their daughters; rather, families receive cash transfers provided their daughter is enrolled in school, maintains a passing grade, and is absent without “good

<sup>1</sup> We use the 2004 Cambodia Socio-Economic Survey (CSES) for these calculations. Average spending per female student in seventh through ninth grades is \$44. These expenses cover school fees, the cost of textbooks, school supplies, school allowance, transport costs, and gifts. In households with annual consumption expenditures below \$200, which more closely resemble the sample of JFPR scholarship recipients, expenditures on schooling per female student are only \$13 per year.

reason” fewer than 10 days in a year. (Scholarship recipients agreed to use funds toward education, but no attempt was made to enforce this agreement.) The JFPR program therefore functions much like a CCT program of the sort that has been implemented and evaluated in many Latin American countries.

### **B. Data**

The main sources of data for the analysis in this study are two: application forms to the scholarship program and data on school enrollment and attendance from an unannounced school visit. The application form contains 28 questions about parental education, demographic composition of the household, ownership of various assets, housing materials, and distance to the nearest secondary school; applicants were also asked to specify the name of the secondary school they would like to attend, with the understanding that this should be one of the 93 schools that were eligible for the scholarship program. These forms were completed by a girl in class or at home. Some provisions were made for teachers to verify that the information that was provided was accurate, but teachers were not encouraged to make additional markings on the form to indicate the poverty or academic potential of a given applicant.<sup>2</sup>

We use the information on the application forms to correct for observable differences across girls. Two things are worth noting here. First, application forms were handed out at the primary “feeder schools” to the 93 secondary schools included in the scholarship program. In practice, selection of the feeder schools and selection of the girls who were encouraged to fill out application forms by the primary school teachers appear to have been somewhat ad hoc.<sup>3</sup> Second, almost 30% of the application forms were not filled out completely. Because we use the characteristics on the application form to correct for differences across girls, most of the results we report are based on a sample of girls with completed applications. We also report the results from a specification that includes all girls who applied for scholarships, regardless whether the application forms were complete. In this specification, the missing control

<sup>2</sup> We inspected a large number of application forms to verify this and found no evidence that additional information was recorded.

<sup>3</sup> We were able to get the application forms from 91 of the 93 schools that received JFPR scholarships. In one school, the records were destroyed by a fire, and in another the Local Management Committee did not send the forms despite numerous requests. The fraction of girls in primary feeder schools who applied for a scholarship can be approximated by merging the application forms with administrative data on female enrollment kept by the Ministry of Education. In the median primary feeder school, approximately 31% of girls applied for a scholarship. However, this figure should be treated with caution: because the application forms included the school name but not the school code, merges had to be done phonetically, which likely introduced noise into the process.

variable is replaced with a zero, and a set of dummy variables indicating missing data (one for each covariate) is included in the regression. Appendix table A1 shows that the differences between the full sample of girls and the sample of girls with completed applications are generally small.

Once application forms had been filled out, they were forwarded from the primary school to the Local Management Committee (LMC) of the relevant JFPR secondary school. The LMCs were then tasked with identifying the 45 girls who were most needy and awarding them scholarships. To assist in this process, LMCs were given a set of weights for each question and a formula to aggregate responses into a final score. According to program administrators, these weights were developed somewhat arbitrarily. For example, applicants who had between three and five brothers received 1 extra point, while those with more than five brothers received 2 extra points; applicants with one to three sisters received 1 extra point, while those with four or five sisters received 2 extra points, and those with six or more sisters received 3 extra points. (More points were meant to increase the probability of receiving a scholarship.) A number of characteristics, including having at least one parent who had completed at least primary school were meant to disqualify applicants from scholarships. However, the LMCs were given considerable discretion over the specifics of how to apply the selection rules—the *Programme Implementation Manual* states that “LMCs will have flexibility to adapt this process as they see fit” (Royal Government of Cambodia 2003, 18). As we show, the evidence suggests that LMCs used the information on the form to assess an applicant’s socioeconomic status without strictly following the proposed weights.

When LMCs had chosen the girls who would be awarded scholarships, all applicants were notified whether they had been selected. The list of girls selected for scholarships was posted at the feeder primary schools as well as at the JFPR secondary schools. Girls then received scholarship payments at a public school ceremony held three times a year. These provisions were meant to increase transparency and minimize the potential for corruption. In addition, the program included a complaints mechanism whereby recipients or other community members could complain to district, provincial, or national Ministry of Education officials if they felt there was fraud.

The second source of data for this evaluation is based on an unannounced visit to each program school. These visits were carried out by an independent firm, hired specifically for this purpose. Enumerators were given a list of applicants to each JFPR school without information on their scholarship status. Note that girls applied for the scholarship in sixth grade during the 2003 school year, scholarships were first awarded during the 2004 school year, and school visits took place during the course of the 2005 school year. During

the school visit, enumerators checked eighth-grade enrollment rosters and physically verified whether or not a girl was attending an eighth-grade class on the day of the visit.

On the basis of the school visits, we construct measures of enrollment and attendance at a JFPR school. The measure of enrollment is a dummy variable that takes the value of one if the girl in question was enrolled in eighth grade at the JFPR school she had applied to. The measure of attendance is a dummy variable that takes the value of one if a girl was present in an eighth-grade class at the time of the visit by the enumerators. Finally, we correct for the possibility of selective school transfers using data provided by students in JFPR schools on the enrollment status of other applicants, as well as data from the Education Management Information System (EMIS) on district-level enrollment figures, by year, separately for boys and girls. Both of these are discussed in greater detail below.

### III. Empirical Strategy

#### A. Main Results of Program Impact

The empirical strategy in this study is based on comparisons between scholarship recipients (the “treated” group) and nonrecipients (the “control” group). The analysis begins with OLS regressions of a dummy variable that takes the value of one if girl  $i$  is enrolled in JFPR school  $s$  at time  $t$  (in the enrollment regressions) or attending school (in the attendance regressions),  $Y_{ist}$ , on a set of school-level fixed effects,  $\alpha_s$ , a vector of baseline characteristics from the application form,  $\mathbf{X}_{ist-1}$ , and a dummy variable for scholarship recipients,  $T_i$ :

$$Y_{ist} = \alpha_s + \mathbf{X}_{ist-1}\boldsymbol{\beta} + T_i\delta + \varepsilon_{ist}. \quad (1)$$

The parameter  $\delta$  is a measure of the difference in the probability of enrollment or attendance between girls who received a scholarship and those who did not. Linear probability models are used to estimate (1); estimation by probit yielded very similar results.

Regression models rely heavily on extrapolation to impute potential (or counterfactual) outcomes. This may introduce biases if the distribution of covariates  $\mathbf{X}_{ist-1}$  is very different for scholarship recipients and nonrecipients. As a first attempt to make scholarship recipients and nonrecipients more closely comparable, we trim the sample. To do this, we regress a dummy variable for scholarship recipients on the vector of observable characteristics:

$$T_i = \theta_s + \mathbf{X}_{ist-1}\boldsymbol{\eta} + \mu_i. \quad (2)$$

This regression is used to predict the probability that each girl in the sample

was awarded a scholarship,  $\pi_i$ , also known as the propensity score. The sample is then trimmed to remove girls with “very high” or “very low” values of the propensity score. Specifically, we follow Crump et al. (2006), who propose a trimming criterion to obtain an optimal subpopulation; in our data, the optimal subpopulation has a value of the propensity score larger than 0.077 and smaller than 0.923.<sup>4</sup> We then run OLS regressions within this sample. As with other trimming procedures, ensuring greater overlap in the distribution of covariates makes the results more robust to specification choices.

We also present results based on a bias-adjusted matching estimator proposed by Abadie and Imbens (2002). This estimator uses nearest neighbor matching to match treated and control girls on the basis of their covariates and then adjusts for any remaining differences in the distribution of the covariates with regression techniques. Matching is a nonparametric alternative to OLS and is often believed to be less subject to misspecification biases. We apply the bias-corrected matching estimator to the full and trimmed samples.<sup>5</sup>

### B. Heterogeneity of Program Effects

In addition to the overall results on program impact, the study presents evidence of heterogeneity in the JFPR program effects. For this purpose, we construct three dummy variables. These correspond to girls who are below the median of a composite measure of socioeconomic status (SES), described in greater detail below; girls for whom neither parent has completed primary

<sup>4</sup> The trimming criterion in Crump et al. (2006) requires that the sample be in the interval  $(0.5 - \sqrt{(0.25 - \gamma^{-1})}, 0.5 + \sqrt{(0.25 - \gamma^{-1})})$ , where  $\gamma$  is the solution to  $\gamma = 2E\{[P(\mathbf{X})(1 - P(\mathbf{X}))]^{-1} | [P(\mathbf{X})(1 - P(\mathbf{X}))]^{-1} < \gamma\}$ . Crump et al. show that trimming by this criterion minimizes the variance of the estimated effects. They apply the estimator to the well-known LaLonde (1986) experimental data and show that the trimming procedure appears to work well in practice; for another recent application, see Chen, Mu, and Ravallion (2006). We also experimented with procedures that trimmed girls who did not receive scholarships and had propensity scores below the lowest propensity score among scholarship recipients (as suggested by Dehejia and Wahba [1999]) or dropped scholarship recipients and nonrecipients with values of the propensity score below a given cut-off (as suggested by Smith and Todd [2005]). Both of these procedures have the disadvantage that the choice of cut-off for trimming is arbitrary. However, results are very similar when these methods are used to trim the data; these results are available from the authors upon request.

<sup>5</sup> Abadie and Imbens (2002) show that their estimator works well in practice by applying it to the LaLonde (1986) experimental data; other recent applications include McKenzie, Gibson, and Stillman (2006) and Schady and Araujo (2008). See also Imbens (2004) for a thoughtful discussion. The results we report are very similar when, instead of the Abadie and Imbens bias-corrected matching estimator, we match on the propensity score,  $\pi_i$ , as first proposed by Rosenbaum and Rubin (1983), or when we run OLS regressions in which scholarship recipients are given a weight of  $1/\pi_i$ , and nonrecipients are given a weight of  $1/(1 - \pi_i)$ , as proposed by Hirano, Imbens, and Ridder (2003); these results are available from the authors upon request.

school; and girls who lived more than 4 kilometers from the JFPR secondary school at the time they completed the application form. We then run separate regressions that include one of these dummies  $D_i$ , the dummy variable for whether a girl received a scholarship  $T_i$ , and the interaction term ( $D_i \times T_i$ ):

$$Y_{ist} = \psi_s + \mathbf{X}_{ist-1}\boldsymbol{\gamma} + D_i\tau + T_i\phi_1 + (D_i \times T_i)\phi_2 + e_{ist}. \quad (3)$$

Interpretation of the coefficients on these variables is straightforward: for example, in the specification that tests for heterogeneity of treatment effects by socioeconomic status, the coefficient  $\tau$  is an estimate of the difference in enrollment (or attendance) between girls of “low” and “high” socioeconomic status; the scholarship effect for high-SES girls is given by the coefficient  $\phi_1$ ; the corresponding effect for low-SES girls is given by the sum of the coefficients  $\phi_1$  and  $\phi_2$ . If  $\phi_2$  is statistically significant, there is evidence of heterogeneity of treatment effects.

### C. Correcting for Possible Selection on Unobservables

The two most important potential concerns for our results are selection on unobservables and selective transfers of girls to non-JFPR schools. We do not believe that selection on unobservables is likely to be a serious problem. Students were enrolled in primary school at the time they applied for scholarships, while the selection of beneficiaries was done by the LMC of a JFPR secondary school. Therefore, members of the LMC would generally not know the girls in question, many of whom would live in different villages or neighborhoods, and would have had little information (if any) on the academic ability or socioeconomic status of a given applicant above and beyond the information provided on the application form.

Nevertheless, we use regression discontinuity (RD) to estimate treatment effects that are arguably robust to unobserved differences between treatment and control groups. The basic logic of RD exploits a discontinuous jump in the probability of receiving a program for observations above and below an eligibility threshold. If the control function used to capture the relationship between the covariate that determines eligibility and outcomes is correctly specified, and if there is no discrete jump in unobservables at the threshold, RD can provide local estimates of treatment effects that are as good as those derived from a randomized experiment (see Hahn, Todd, and Van der Klauw 2001; Lee 2008). In the literature on schooling outcomes, RD has been used to estimate the impact of class size on test scores in Israel (Angrist and Lavy 1999) and Bolivia (Urquiola 2006), financial aid on college enrollment in the United States (Van der Klauw 2002), financial incentives for schools on test

score outcomes in Chile (Chay, McEwan, and Urquiola 2005), and pension income on enrollment in South Africa (Edmonds 2006).

The JFPR scholarships were generally awarded to girls with low socioeconomic status. However, program rules established that no more than 45 scholarships could be awarded in any given school. As a result, when we rank girls within a school by a composite measure of SES, we observe a sharp drop in the probability of receiving a scholarship around the threshold given by the 45th girl with the lowest SES. The RD estimates exploit this discontinuity to estimate program impact:

$$Y_{it} = \chi_c + f(R_{it-1}) + T_i\pi + v_{it}, \quad (4)$$

where  $R$  is the within-school rank by the composite measure of socioeconomic status, also known as the control function, and the function  $f$  is a flexible parameterization—in practice, a quadratic or a cubic. Because the relationship between the rank of an application by this measure of SES and scholarship receipt is not deterministic, this is a case of “fuzzy” (rather than “sharp”) RD. Equation (4) is therefore estimated by two-stage least squares, with  $T_i$  instrumented with a dummy variable that takes the value of one if a girl is ranked below 45 by the composite measure of SES (following Hahn et al. 2001).

#### ***D. Correcting for Possible Selective School Transfers***

The treatment effects we estimate are measures of the impact of the JFPR scholarship program on enrollment and attendance at JFPR-eligible schools. These are of interest in and of themselves. However, these estimates may not be an accurate reflection of the impact of the scholarship program on overall enrollment and attendance if some applicants to the JFPR program enrolled elsewhere. Moreover, because JFPR scholarships were not portable—recipients had to enroll at the school they applied to in order to receive the \$45 transfer—applicants who were offered scholarships may have been less likely to transfer to a non-JFPR school than those who were turned down for scholarships.

We do not believe that selective transfers across schools are likely to be an important consideration. Anecdotal evidence suggests that the scope for school choice in Cambodia is limited. Nevertheless, we approximate the impact of the JFPR program on overall enrollment (rather than enrollment at JFPR-eligible schools) in two ways. The first approach involves estimating bounds. This is arguably the most flexible, nonparametric approach to dealing with selection—see, in particular, Manski (1989), Smith and Welch (1989), and the discussion in Heckman, LaLonde, and Smith (1999).

When an applicant to a given JFPR school did not appear on the eighth-

grade enrollment rosters and was not present on the day of the school visit, enumerators asked other eighth-grade students in that school whether they knew the missing girl. If someone knew the girl in question, they were asked whether they knew she was definitely enrolled elsewhere, definitely not enrolled, or whether respondents were not certain about the enrollment status of the missing girl. On the basis of the enrollment rosters and the questions asked of eighth-grade girls during the school visits, we were able to establish the enrollment status of 95.2% of all girls who had completed scholarship applications.

One way to think of the approximately 5% of girls whose enrollment status could not be established is as attritors out of the sample. Note that the enrollment rate among scholarship recipients is given by the following identity:

$$E_t = \phi_{ot}E_{ot} + \phi_{at}E_{at}, \quad (5a)$$

where the subscript  $t$  corresponds to “treated” girls. In this identity,  $E_t$  is the proportion of treated girls who are enrolled in any school;  $E_{ot}$  is the enrollment rate among scholarship recipients who have not attrited, and  $\phi_{ot}$  is the proportion of scholarship recipients in this group;  $E_{at}$  is the enrollment rate among scholarship recipients who have attrited, and  $\phi_{at}$  is the proportion in this group. The terms  $\phi_{ot}$ ,  $E_{ot}$ , and  $\phi_{at}$ , but not  $E_{at}$ , can be calculated from the data, and  $E_t$  can be estimated under alternative values for  $E_{at}$ . Similarly, for girls who were not awarded scholarships,

$$E_c = \phi_{oc}E_{oc} + \phi_{ac}E_{ac}, \quad (5b)$$

where the subscript  $c$  corresponds to girls in the “control” group. Here too  $\phi_{oc}$ ,  $E_{oc}$ , and  $\phi_{ac}$  can be calculated, and assumptions have to be made regarding  $E_{ac}$ . Following Smith and Welch (1989), we calculate  $(E_t - E_c)$  for values for  $E_{at}$  and  $E_{ac}$  ranging from 0 to 1. This provides estimates of the effect of the JFPR program on total enrollment under alternative assumptions about the behavior of girls whose enrollment status could not be determined.

Calculating bounds is convincing because it imposes no assumptions on identification. However, bounds may not provide unbiased estimates of program effects on total enrollment if the information provided by students is inaccurate. Our second approach to selection uses administrative data from the Cambodia EMIS. These data are available at the school level, although school openings, closures, merges, and changes in school name make it difficult to use the school-level data. For our purposes, it is more useful to aggregate

enrollment data to the district level.<sup>6</sup> In 2004, the last year for which the EMIS data were available for this study, there were 173 districts in Cambodia. We exclude 17 districts from the analysis because in these officially designated “ethnic minority areas” a parallel scholarship scheme was available for boys; as we discuss below, boys serve as an additional control for our district-level estimates. Of the remaining 156 districts, 52 included at least one JFPR-eligible school. Most districts have more than one secondary school—the median value is 4, and the corresponding values at the 25th and 75th percentile are 2 and 6, respectively. Given these figures, it seems likely that a substantial fraction of transfers across schools occur within the same district.

We use the EMIS data for an application of the familiar double differences (or differences in differences) specification:

$$E_{gdt} - E_{gdt-1} = \alpha_p + \beta_1 J + \varepsilon_{gdt} - \varepsilon_{gdt-1}, \tag{6a}$$

where  $E_{gd}$  refers to total enrollment of girls in seventh grade in district  $d$ , the subscripts  $t$  and  $t - 1$  refer to the 2004 and 2003 academic school years,  $\alpha_p$  is a set of province fixed effects, and  $J$  is a dummy variable for districts that include a JFPR school.

The underlying assumption in the double-differences framework is that the growth in girl enrollment in districts with and without JFPR schools would have been the same in the absence of the program. We show that this assumption is suspect because there appear to be different trends in girl enrollment before the program started, as well as JFPR program “effects” on the enrollment of boys, who were ineligible for scholarships. Our preferred specification is therefore a triple-difference specification:

$$(E_{gdt} - E_{gdt-1}) - (E_{bdt} - E_{bdt-1}) = \alpha_p + \beta_2 J + (\varepsilon_{gdt} - \varepsilon_{gdt-1}) - (\varepsilon_{bdt} - \varepsilon_{bdt-1}). \tag{6b}$$

In this specification, the coefficient  $\beta_2$  measures the extent to which the growth in girl enrollment, relative to the growth in boy enrollment, was larger in districts that included a JFPR school than in other districts.

Triple differences is appropriate if the JFPR program did not affect boy enrollment. In practice, where externalities of this sort have been analyzed in similar programs, they have been found to be nonexistent (Behrman et al. [2005] on PROGRESA) or positive (Bobonis and Finan [2006] on PROGRESA; Kremer et al. [2007] on a merit scholarship program for adolescent

<sup>6</sup> Districts in Cambodia are an intermediate administrative unit—larger than communes but smaller than provinces.

TABLE 1  
CHARACTERISTICS OF SCHOLARSHIP RECIPIENTS AND NONRECIPIENTS AT BASELINE

	Recipients	Nonrecipients	Difference	SE
Parents own business	.00	.08	-.08	.02
Either parent is a government employee	.01	.14	-.13	.02
Parent lends money regularly	.00	.06	-.05	.01
Parent completed primary school	.09	.40	-.30	.03
Parent completed secondary school	.01	.06	-.05	.01
Main part of house made of cement or brick	.00	.08	-.08	.02
Roof made of tiles, metal, or fiber	.15	.56	-.41	.03
Land ownership by household > 1 hectare	.07	.33	-.26	.03
Own a large asset (>1,000,000 riels)	.01	.15	-.14	.02
Own a truck	.00	.01	-.01	.00
Own a car	.00	.01	-.01	.00
Live with both parents	.70	.83	-.13	.02
Live with one parent	.23	.19	.04	.02
Live in a hut	.53	.30	.23	.03
Earthen floor	.35	.28	.07	.03
Family has motorbike or trailer	.04	.27	-.23	.02
Family has bicycle	.48	.72	-.24	.03
Family has ox and cart	.20	.44	-.24	.03
Family has pony and trap	.02	.05	-.03	.01
Family has no means of transportation	.49	.25	.24	.03
Family has debts > 100,000 riels	.79	.54	.25	.03
Number of brothers	2.26	2.36	-.10	.07
Number of sisters	2.84	2.70	.14	.07
Applicant disabled	.04	.02	.01	.01
Other household member disabled	.15	.13	.02	.02
Applicant or other household member has disease	.32	.20	.12	.02
Distance to secondary school (km)	4.17	3.61	.56	.15
Number of observations	2,765	858	3,623	

girls in Kenya). Note that if the JFPR program had a positive effect on the enrollment of boys, these triple-differenced estimates of JFPR program effects on girl enrollment would be downward biased.

#### IV. Results

##### A. Comparisons of Scholarship Recipients and Nonrecipients at Baseline

We begin with simple comparisons of the enrollment, attendance, and baseline characteristics of girls who were awarded scholarships and those who were not in table 1. The table shows that scholarship recipients had significantly lower socioeconomic status than nonrecipients: recipients had parents with lower education levels; they were more likely to live in a hut or a house with an earthen floor and less likely to have houses made of high-grade materials like cement, brick, tiles, metal, or fiber; recipients were also less likely to own any one of a number of means of transportation, less likely to own more than a

hectare of land, less likely to regularly lend money, and more likely to have debts. These results confirm that the LMCs were successful at targeting scholarships to girls with lower socioeconomic status.

Covariate imbalance is a possible source of concern for regression-based estimates of JFPR program effects. Appendix table A2 summarizes baseline differences between scholarship recipients and nonrecipients, before and after trimming and matching (Abadie and Imbens 2002; Crump et al. 2006). The average raw difference in covariates between scholarship recipients and nonrecipients for the full sample is equivalent to 0.41 standard deviations; the difference in the trimmed sample is 0.27 standard deviations. The average difference in the matched pairs of scholarship recipients and nonrecipients is 0.23 standard deviations in the full sample and 0.20 in the trimmed sample.

Trimming and matching remove a substantial fraction of the imbalance in the covariates. As an additional robustness check, we apply the bias-corrected matching estimator to the trimmed sample and estimate “program effects” for each of the covariates in table 1. These results, which are reported in the last column in appendix table A2, show that the estimated “effects” are generally not significant. There are a handful of exceptions to this general pattern, but in every case the socioeconomic status of scholarship recipients is lower than that of matched nonrecipients. For example, after matching with the trimmed sample, scholarship recipients are significantly less likely to have parents who have completed primary school or secondary school. We would therefore expect that, if anything, any remaining difference in the distribution of covariates would lead to downward biases in our estimates of JFPR program effects.

### **B. Main Results of Program Impact**

The main results of program impact on attendance (top row) and enrollment (bottom row) are presented in table 2. Column 1 presents the raw difference between scholarship recipients and nonrecipients, without any adjustments for covariates. Column 2 is based on OLS regressions that include all of the characteristics on the application form. Column 3 supplements these controls with school fixed effects. Column 4 corresponds to a specification that includes all of the girls that applied for scholarships, regardless whether the application forms were complete; in this specification, the missing control variable is replaced with a zero, and a set of dummy variables indicating missing data (one for each covariate) is included in the regression. Column 5 is based on OLS regressions with the trimmed sample. Columns 6 and 7 report the results from the bias-corrected matching estimator, for the full and trimmed samples, respectively. Applicants from a given area may have error terms that are

TABLE 2  
BASIC RESULTS ON PROGRAM IMPACT

	Raw Difference (1)	OLS (2)	OLS + School FE (3)	OLS + School FE, Dummies for Missing (4)	OLS + School FE, Trimmed (5)	Nearest Neighbor Matching (6)	Nearest Neighbor Matching, Trimmed (7)
Attending on the day of school visit:							
Scholarship	.223 (.024)***	.299 (.025)***	.313 (.023)***	.305 (.020)***	.319 (.024)***	.299 (.040)***	.313 (.045)***
Observations	3,623	3,623	3,623	5,138	3,065	3,623	3,065
R <sub>2</sub>	.05	.09	.16	.15	.18		
Enrolled:							
Scholarship	.222 (.023)***	.292 (.025)***	.303 (.024)***	.294 (.021)***	.314 (.025)***	.300 (.037)***	.310 (.042)***
Observations	3,623	3,623	3,623	5,138	3,065	3,623	3,065
R <sub>2</sub>	.06	.10	.18	.17	.19		

**Note.** In cols. 5 and 7 trimming is determined as per Crump et al. (2006), which in this case is the 7.7th and 92.3rd percentile of the propensity score. Robust standard errors are in parentheses.

\* Significant at 10%.

\*\* Significant at 5%.

\*\*\* Significant at 1%.

correlated; to address this concern, all specifications correct for clustering at the level of an applicant's primary school.<sup>7</sup>

The first column in table 2 shows that the raw difference in enrollment and attendance between scholarship recipients and nonrecipients is about 22 percentage points.<sup>8</sup> This is a large difference, considering that scholarship recipients are generally poorer. One would therefore expect that the difference in enrollment and attendance would increase as adjustments are made for the distribution of covariates. The remaining columns in the table show that this is indeed the case. Introducing applicant characteristics as controls in a regression framework (col. 2) results in an estimated program effect of 0.299 on attendance (with a standard error of 0.025), and 0.292 on school enrollment (with a standard error of 0.025). Including school fixed effects (col. 3), including girls with incomplete application forms (col. 4), and running regressions on the trimmed sample (col. 5) all make little difference to the estimates in column 2. Columns 6 and 7, finally, show that the results from the bias-corrected matching estimator are also very close to those from the OLS regressions. For example, in the trimmed sample, the estimated effect on attendance is 0.313 (with a standard error of 0.045), and the estimated effect

<sup>7</sup> Results are similar when we cluster at the level of the LMC or when no correction is made for clustering.

<sup>8</sup> On the day of the visit, 80.2% of recipients, but only 57.9% of nonrecipients, were attending; 86.7% of recipients, but only 64.6% of nonrecipients, were enrolled in a JFPR program school as per the school enrollment rosters.

on enrollment is 0.310 (with a standard error of 0.042).<sup>9</sup> In sum, table 2 suggests that the effect of the JFPR scholarship program on enrollment and attendance is approximately 30 percentage points. The results are remarkably robust to trimming of the sample and to different ways of adjusting for differences in the distribution of covariates.

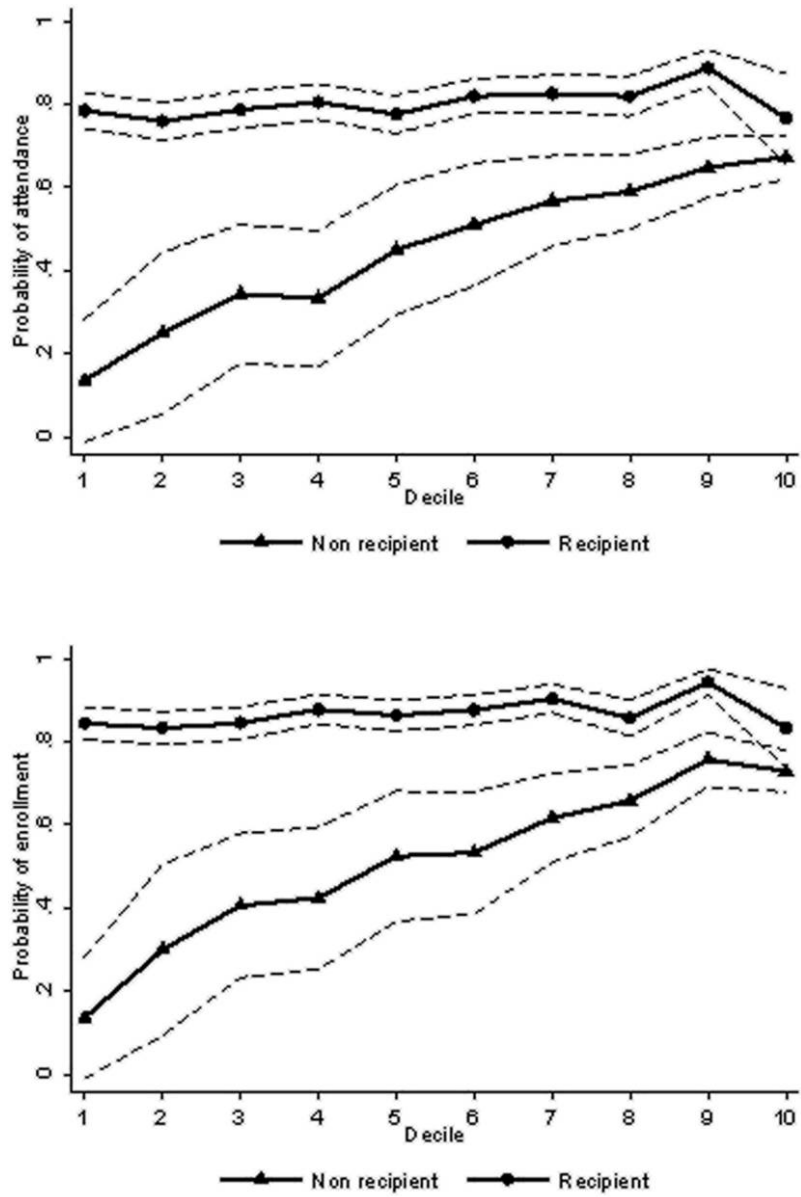
### C. Heterogeneity of Program Effects

We next turn to evidence of heterogeneity in the estimated program effects, concentrating on comparisons by a composite measure of socioeconomic status, by parental education, and by distance to school. To generate the composite measure of SES, we aggregated all of the characteristics on the application form by principal components (as described in Filmer and Pritchett 2001). We also considered two alternative ways of aggregating these characteristics. First, we defined responses on the application form as “goods” or “bads”; we then generated a variable that is a simple sum of these recoded responses (see Case, Paxson, and Ableidinger [2004]; Paxson and Schady [2007] for similar approaches).<sup>10</sup> Second, we aggregated the responses on the application form using the weights suggested in the *Programme Implementation Manual*. These alternative measures of SES are both worse predictors of scholarship receipt than the measure calculated by principal components. This is a disadvantage for RD estimates of program impact, so the results we report are based on aggregation by principal components. However, our results are not sensitive to this choice.<sup>11</sup>

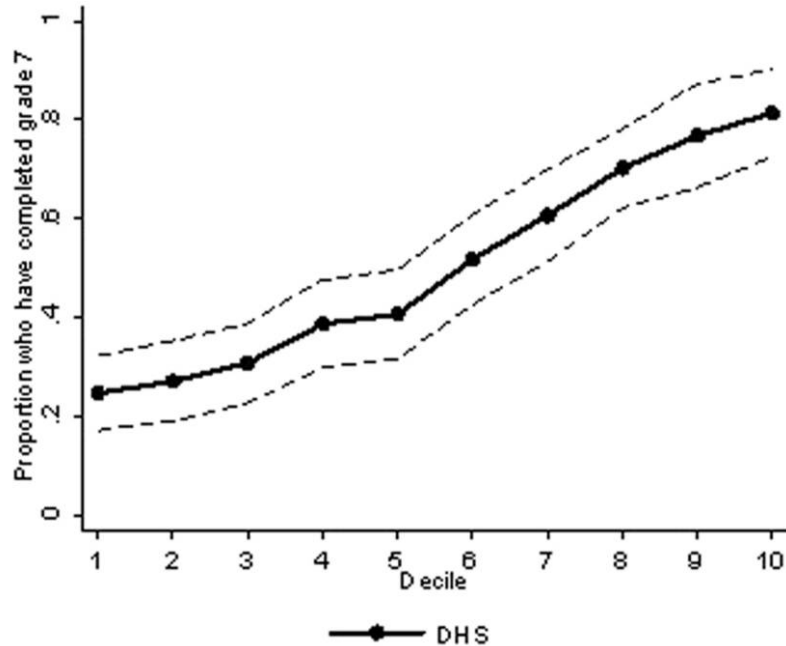
<sup>9</sup> The estimates in table 2 are based on matching on the four nearest neighbors. However, our results are not sensitive to this choice. For example, the estimated program effect on attendance for the full sample of girls with completed applications reported in table 2 is 0.299 (with a standard error of 0.040); when matching is performed using only one neighbor the value is 0.336 (with a standard error of 0.038); and when matching is performed using 16 neighbors, the estimate is 0.302 (with a standard error of 0.042).

<sup>10</sup> Specifically, we generated a score that increases by one unit if an applicant has any characteristic that should reasonably be correlated with higher SES—for example, an applicant whose parents own a business or one who has at least one parent who has completed primary school—and decreases by one unit for every characteristic on the application form that should reasonably be correlated with lower SES—for example, living in a hut or having an earthen floor. Distance to school is first recoded as a dummy variable for children living more or less than 4 kilometers from school, roughly corresponding to median distance, and applicants who live closer to a school receive one more point. The variables for the number of brothers and sisters are not used as it is not clear how to sign them. The resulting measure of SES has a maximum value of 8, corresponding to applicants with the highest SES, and a minimum value of  $-6$ , corresponding to applicants with the lowest SES; the median value is  $-1$ .

<sup>11</sup> In a simple regression of a dummy variable for receiving a scholarship on the measure of SES that follows the program guidelines the  $R$  squared is 0.18; in a comparable regression in which the measure of SES is based on a simple sum the  $R$  squared is 0.25; and that in a comparable



**Figure 1.** Attendance and enrollment status by decile. *Top*, Attendance by decile. *Bottom*, Enrollment by decile. Dashed lines show 95% confidence interval. Deciles derived from aggregate measure of SES.



**Figure 2.** Proportion of girls who complete grade 7 among those who completed grade 5, by economic status decile, Demographic and Health Survey data. Dashed lines show 95% confidence interval. Deciles derived from aggregate measure of SES.

We begin our analysis of heterogeneity in figure 1. The figure graphs the probability of attendance (upper panel) and enrollment (lower panel) by deciles of the composite measure of SES, separately for scholarship recipients and nonrecipients. Among scholarship recipients, there are essentially no differences in enrollment or attendance by socioeconomic status. Figure 1 shows, however, that there are steep socioeconomic “gradients” in enrollment and attendance among girls who were turned down for scholarships. Figure 2 presents similar calculations based on the 2000 Demographic and Health Survey; here too

---

regression with the measure of SES generated by principal components is 0.27. We also ran regressions that included a cubic in the within-school rank, with the rank generated by principal components, a simple sum, or the program guidelines, as well as dummies for girls above or below the threshold given by the applicant with the 45th lowest rank within a school. In these regressions, the *R* squared is 0.23 when the rank is generated with the program guidelines, 0.29 when it is generated with a simple sum, and 0.32 when it is generated by principal components. Estimates of treatment heterogeneity as well as RD measures of program impact that use these alternative measures of SES are similar but less precise; these are available from the authors upon request.

there is a clear socioeconomic gradient.<sup>12</sup> The gap between the two lines in figure 1 is a reduced-form estimate of the impact of the JFPR scholarship program on school enrollment and attendance for girls of different SES. The figure is therefore consistent with larger JFPR program effects among girls with the lowest SES, a point we examine more carefully below.

The main results on treatment heterogeneity are presented in table 3. The first two columns correspond to estimates of treatment effects for girls above and below the median of the composite measure of SES; the next two columns correspond to treatment effects for girls whose parents have both not finished primary school, compared to girls with at least one parent who has completed primary school; the last two columns correspond to girls who live more than 4 kilometers away from the JFPR school they applied to, compared to those who live closer.

Table 3 shows clear evidence of treatment heterogeneity. Focusing first on the composite measure of socioeconomic status, the results suggest program effects of 26.4 to 26.7 percentage points on attendance among high-SES girls, and 42.8 to 43.1 among low-SES girls. A similar pattern of larger program effects among low-SES girls is found for enrollment. Another way of looking at these results is to analyze the effect of the JFPR program on differences in school enrollment and attendance between high-SES and low-SES girls; this is closely related to the “gradient” discussed above. Table 3 suggests that low-SES girls are 18.7 to 19.1 percentage points less likely to be attending school in the absence of the program, and 19.5 to 20.2 percentage points less likely to be enrolled in school. The scholarship program removes approximately three-quarters of the difference in school attendance between high- and low-SES girls, and more than 90% of the difference in school enrollment. These are very large effects by any standard.

Table 3 shows that the JFPR program also had larger effects on girls whose parents had little formal schooling, although the differences are only significant when the dependent variable is school enrollment. Among daughters of high education parents, the impact of the program on enrollment is between 22.1 and 24.2 percentage points, compared to an impact of 32.3 to 34.1 percentage points among daughters of low-education parents. Finally, the impact of the JFPR program is larger among girls living further away from a secondary school. The interaction term is always significant, regardless whether the de-

<sup>12</sup> The exact list of socioeconomic variables in the DHS is not identical to that on the JFPR application form. We derive the index of socioeconomic status in the DHS from variables describing the ownership of a bicycle, cart, boat, motorbike, car, truck, radio, television; the conditions of the dwelling, such as hard roofing and finished flooring; the availability of electric lighting; the main source of drinking water; the type of toilet facilities; and the main type of cooking fuel used.

**TABLE 3**  
HETEROGENEITY OF TREATMENT EFFECTS

	Low SES		Parents without Education		School More than 4 km Away	
	OLS + School FE (1)	OLS + School FE, Trimmed (2)	OLS + School FE (3)	OLS + School FE, Trimmed (4)	OLS + School FE (5)	OLS + School FE, Trimmed (6)
Attending on the day of school visit:						
Scholarship	.264 (.028)***	.267 (.031)***	.297 (.032)***	.286 (.040)***	.274 (.026)***	.281 (.028)***
Dummy variable	-.187 (.059)***	-.191 (.063)***	-.072 (.032)**	-.096 (.040)**	-.120 (.042)***	-.112 (.046)**
Scholarship × Dummy variable	.164 (.054)***	.164 (.058)***	.021 (.038)	.043 (.045)	.119 (.044)***	.112 (.048)**
Enrolled:						
Scholarship	.243 (.026)***	.250 (.029)***	.242 (.028)***	.221 (.035)***	.267 (.026)***	.278 (.028)***
Dummy variable	-.195 (.057)***	-.202 (.061)***	-.126 (.032)***	-.166 (.041)***	-.105 (.038)***	-.098 (.045)**
Scholarship × Dummy variable	.196 (.054)***	.198 (.057)***	.081 (.036)**	.120 (.043)***	.112 (.040)***	.108 (.047)**

**Note.** Full sample N = 3,623; trimmed sample N = 3,065. In cols. 2, 4, and 6, trimming is determined as per Crump et al. (2006), which in this case is the 7.7th and 92.3rd percentile of the propensity score. Robust standard errors are in parentheses.

\* Significant at 10%.

\*\* Significant at 5%.

\*\*\* Significant at 1%.

pendent variable is attendance or enrollment; the magnitude of the coefficient suggests that the scholarship program fully offsets differences in school enrollment and attendance between girls living closer to and further away from school. In sum, table 3 presents compelling evidence of heterogeneity of program effects, with significantly larger effects among low-SES girls, girls from households with low education levels, and girls living far away from a secondary school.<sup>13</sup>

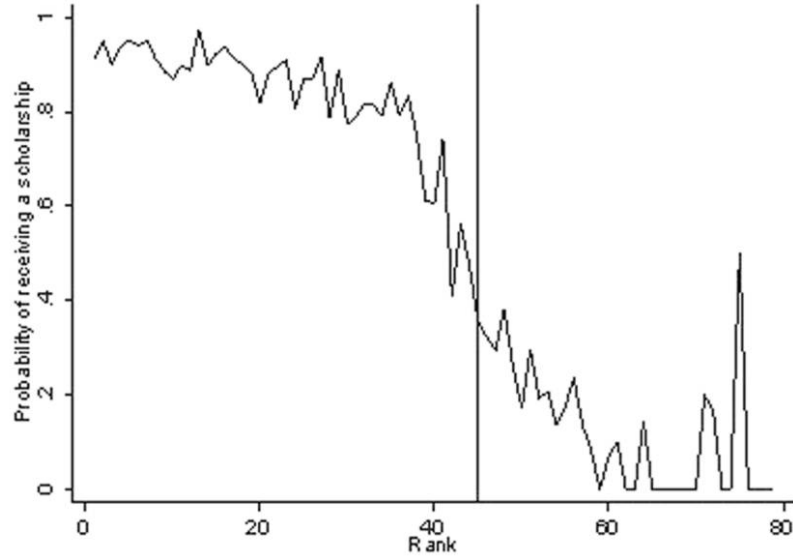
#### **D. Correcting for Possible Selection on Unobservables**

Selection on unobservables could bias estimates of program effects—upward, if LMCs selected girls who they believed had more academic potential, or downward, if LMCs selected girls who were poorer in ways that were not captured on the application form. The main approach we use to correct for selection on unobservables is RD. Before describing these results in detail, however, we note one result in table 3 that suggests that selection on unobservables was limited. LMC members of a JFPR secondary school should be less likely to know girls who lived far away from the school they were applying to. The fact that the last two columns in table 3 show significant program effects among girls who lived more than 4 kilometers from the school they applied to provides some reassurance that our estimates are not driven by selection on unobservables—at least for this group.

To assess the extent to which RD is appropriate, we begin by ranking all applicants to a given JFPR school by the composite measure of socioeconomic status. Recall that no more than 45 scholarships could be handed out per school and that LMCs were encouraged to give scholarships to the “poorest” 45 girls. We would therefore expect a sharp drop-off in the probability of receiving a scholarship around the girl with the 45th SES rank. However, since we do not know exactly how each LMC weighted the responses on the application form, we would also expect to see a fair amount of “fuzziness” around this threshold.

Figure 3 graphs the probability that a girl received a scholarship as a function of her within-school SES rank. The figure shows that, for girls with low ranks (corresponding to low SES relative to other girls applying to that JFPR school) the probability of receiving a scholarship was high; the probability that a girl ranked below 40 in a school received a scholarship is 0.876. Conversely, for girls with high ranks (corresponding to high SES relative to other girls ap-

<sup>13</sup> All of the results on heterogeneity are similar if we split the sample—for example, into high- and low-SES girls—and use the bias-corrected matching estimator within each subsample; these results are available from the authors upon request.



**Figure 3.** Probability of receiving a scholarship, within-school ranking. The ranking is derived from the aggregate measure of SES. The applicants are ranked within schools.

plying to that JFPR school) the probability of receiving a scholarship was low; the probability that a girl ranked above 50 in a school received a scholarship is 0.128. There is a sharp drop around the threshold given by the girl with the 45th rank; the probability that a girl who was ranked between 40 and 50 received a scholarship is 0.445.

Figure 3 suggests that an application of fuzzy RD to the applicant data is appropriate, and a quadratic or cubic may be a reasonable parameterization for the control function. To further investigate the extent to which the samples of applicants above and below the threshold are similar, we present the results from two sets of comparisons in appendix table A3. First, the table shows that differences in the covariates  $\mathbf{X}_{ist-1}$  between girls above and below the threshold gradually decrease as comparisons are limited to bands that are closer to the threshold. For example, the mean difference in covariates between girls who fall below the threshold (ranked 45 and lower) and those who fall above the threshold (ranked 46 and higher) is 0.54 standard deviations; the comparable difference between girls ranked 44 or 45, and those ranked 46 and 47 is only 0.12 standard deviations. Second, we regress every covariate in the vector  $\mathbf{X}_{ist-1}$  on a quadratic or cubic formulation of the control function as well as the dummy variable for girls below the threshold. Small and generally insignificant coefficients on the dummy variable for girls ranked below 45

would suggest that the RD specification effectively accounts for smooth differences in the distribution of covariates.<sup>14</sup> The results in the last columns of appendix table A3 show that RD does a reasonably good job equating the samples of individuals above and below the threshold. A handful of differences remain, so the RD results presented below should be interpreted with some caution. Note, however, that when differences are significant, they suggest that girls below the eligibility threshold have characteristics associated with lower SES that are not fully captured by the control function. If anything, estimates that compare the enrollment and attendance of low-rank and high-rank girls, such as those calculated by RD, may therefore underestimate the true underlying program effects.

The main results on the JFPR program effects estimated by RD are presented in table 4. The first four columns present the reduced-form specifications; in these specifications, the dummy variable for girls with within-school SES rank of 45 or lower enters directly in the regression. These results provide further evidence that the JFPR program had an impact on school enrollment and attendance. Girls with a lower within-school ranking have higher enrollment and attendance levels. The next four columns present the structural specifications; in these specifications, the dummy variable for scholarship receipt is instrumented with the dummy variable for girls with a within-school SES rank of 45 or lower. These results suggest program effects of 34–45 percentage points on attendance and 24–37 percentage points on enrollment.

Comparisons of program effects estimated by RD and those estimated by OLS or matching are not straightforward. RD identifies program effects around the threshold—in this case, close to the cut-off determined by the girl with the 45th SES rank within a school. The results in table 3 show that there is significant heterogeneity of treatment effects by SES, with larger effects for girls with low SES. A comparison of tables 2 and 4 shows that the program effects estimated by RD are larger than those from “comparable” specifications estimated by OLS or matching. One might therefore conclude that, if anything, LMCs favored girls with a low probability of enrollment and attendance given their observables, and that this selection on unobservables introduces a downward bias to program effects estimated by OLS or matching. However, this is a case of fuzzy RD, so the structural estimates in table 4 are based on instrumental variables. As is well known, these estimate local average treatment effects (LATE)—program effects for individuals whose likelihood of receiving the treatment (in this case, the scholarship) was affected by the instrument

<sup>14</sup> We thank an anonymous referee for this suggestion. See also DiNardo and Lee (2004) and Lee, Moretti, and Butler (2004) for a similar approach.

**TABLE 4**  
REGRESSION DISCONTINUITY ANALYSIS

	Reduced Form				Instrumental Variables			
	Quadratic Control Function (1)	Quadratic Control Function School FE (2)	Cubic Control Function (3)	Cubic Control Function School FE (4)	Quadratic Control Function (5)	Quadratic Control Function School FE (6)	Cubic Control Function (7)	Cubic Control Function School FE (8)
Attending on the day of school visit:								
Scholarship	.105 (.039)***	.094 (.037)**	.103 (.041)**	.094 (.039)**	.338 (.130)***	.331 (.134)**	.454 (.199)**	.434 (.196)**
Observations	3,623	3,623	3,623	3,623	3,623	3,623	3,623	3,623
R <sup>2</sup>	.00	.09	.00	.09	.07	.14	.05	.13
Enrolled:								
Scholarship	.087 (.037)**	.068 (.036)*	.085 (.040)**	.063 (.038)*	.282 (.122)**	.238 (.129)*	.373 (.185)**	.294 (.185)
Observations	3,623	3,623	3,623	3,623	3,623	3,623	3,623	3,623
R <sup>2</sup>	.00	.10	.00	.10	.08	.16	.08	.16

**Note.** Robust standard errors are in parentheses.

\* Significant at 10%.

\*\* Significant at 5%.

\*\*\* Significant at 1%.

(in this case, whether they fell below the cut-off given by the within-school rank; Imbens and Angrist 1994). These “compliers,” in the language of Angrist, Imbens, and Rubin (1996), cannot be identified from the data without additional assumptions, and it is not clear whether one would expect program effects to be larger or smaller among compliers. Moreover, the confidence bands around the RD estimates in table 4 are large and always include the coefficients estimated from comparable OLS specifications. With all of this in mind, we conclude that it is unclear whether LMCs selected girls for scholarships on the basis of unobservables. However, the results in table 4 indicate that selection on unobservables, if there was any, cannot on its own account for the program effects we find.

#### *E. Correcting for Possible Selective School Transfers*

The treatment effects we estimate are measures of the impact of the JFPR scholarship program on enrollment and attendance at eligible schools. In principle, if there were nonrandom transfers of students across schools—in particular, if the probability of enrolling at a different school was negatively correlated with scholarship receipt—the impact of the program on overall enrollment could be smaller than the impact on enrollment at eligible schools. Externalities could also introduce a wedge between program effects on the enrollment and attendance of scholarship applicants, reported above, and the overall effect on enrollment and attendance.

Our first attempt to correct for the possibility of nonrandom transfers is based on an application of bounds. Recall that enumerators were able to establish the enrollment status of approximately 95% of scholarship applicants on the basis of questions asked of their peers during the school visit. Note that the implied attrition rate, 4.8%, is quite low: as a point of comparison, loss to follow-up in a recent evaluation of school voucher scheme in Colombia was almost 50% (Angrist et al. 2002).

We compare the observable characteristics of attritors and other girls in appendix table A4. The table shows that overall differences are generally small, as are differences between scholarship recipients and nonrecipients who were lost to follow-up. Nevertheless, missing data are a source of concern in our analysis because attrition is correlated with treatment status: in a simple regression of a dummy variable for attrited applicants on a dummy for scholarship recipients, the coefficient is  $-.100$  (with a standard error of 0.008). Moreover, as is discussed above, girls may have been more likely to enroll in a non-JFPR school if they were turned down for a scholarship.

To bound the possible effects of selective transfers, table 5 reports estimates of school enrollment under alternative scenarios about the behavior of girls

**TABLE 5**  
**BOUNDING THE EFFECTS OF SELECTIVE ATTRITION ON ESTIMATES OF PROGRAM IMPACT**

Nonrecipients <sup>a</sup>			Recipients <sup>b</sup>		
Enrollment: Attritors	Ratio of Enrollment: (Attritors / Nonattritors)	Overall Enrollment	Enrollment: Attritors	Ratio of Enrollment: (Attritors / Nonattritors)	Overall Enrollment
.00	.00	.68	.00	.00	.88
.10	.13	.69	.10	.11	.88
.20	.26	.70	.20	.22	.88
.30	.39	.71	.30	.34	.88
.40	.52	.72	.40	.45	.89
.50	.65	.73	.50	.56	.89
.60	.78	.75	.60	.67	.89
.70	.92	.76	.70	.78	.89
.80	1.05	.77	.80	.89	.89
.90	1.18	.78	.90	1.01	.90
1.00	1.31	.79	1.00	1.12	.90
.12	.16	.69	.10	.11	.88

<sup>a</sup> Enrollment among nonattritors, .765. Proportion who attrited, .118.

<sup>b</sup> Enrollment among nonattritors, .895. Proportion who attrited, .018.

whose enrollment status could not be determined (Smith and Welch 1989; Heckman et al. 1999).<sup>15</sup> Consider first a scenario in which none of the missing girls are enrolled; under this assumption, total enrollment among scholarship recipients is 0.68, total enrollment among nonrecipients is 0.88, and the estimated JFPR program effect on enrollment is 20 percentage points. The corresponding scenario in which all of the missing girls are enrolled suggests program effects of 11 percentage points. Following Manski (1989), it is also possible to calculate upper and lower bounds on the treatment effects. The upper bound corresponds to a scenario in which all the missing girls who are scholarship recipients are enrolled, while all missing girls who did not receive scholarships are unenrolled (program effect:  $0.90 - 0.68 = 22$  percentage points); the lower bound corresponds to a scenario in which all the missing scholarship recipients are unenrolled, while all missing girls who did not receive scholarships are enrolled (program effect:  $0.88 - 0.79 = 9$  percentage points). These bounds do not include zero, so selective transfers on their own cannot account for the fact that the estimated program effects on enrollment are positive. Note also that these calculations of bounds do not adjust for observable or unobservable differences between girls who received JFPR scholarships and other girls—they correspond to the raw difference in enrollment

<sup>15</sup> We present results for enrollment, rather than attendance, as enumerators could obviously not ask girls in a JFPR school whether other girls not present on the day of the school visit were attending another school on that day; rather, they could only ask girls in JFPR schools whether they knew whether the missing girls were enrolled elsewhere.

in table 2. Given that scholarship recipients were poorer than nonrecipients, these are downward-biased estimates of the true program effects.

Is it possible to make a plausible guess about the enrollment of girls who were lost to follow-up? Recall that, by definition, these girls are not enrolled in a JFPR school. Arguably, the probability of enrollment for attritors who did not receive a scholarship could be approximated by the enrollment probability of other girls who were turned down for scholarships and did not enroll in a JFPR school, but whose enrollment status could be established on the basis of the answers provided by their peers during the school visits; this probability is 0.12. Similarly, the probability of enrollment for attritors who received a scholarship could be approximated by the probability of enrollment of other girls who received scholarships and did not enroll in a JFPR school, but whose enrollment status could be established; this probability is 0.10. The difference in total enrollment rates ( $E_t - E_c$ ) under these values for enrollment of attrited recipients and nonrecipients, reported in the last row of the table, is 19 percentage points ( $0.88 - 0.69$ ).

Calculating bounds is a flexible way of dealing with possible nonrandom transfers. However, bounds could still present a misleading picture if the information given by students at JFPR schools about the enrollment status of other applicants is noisy and, in particular, if measurement error in the responses is correlated with treatment status. For example, if peers know that a friend received a scholarship, they might erroneously report that the girl is enrolled in school based on the presumption that they would attend school because they received the scholarship.<sup>16</sup> As a second attempt to address concerns about selective school transfers we therefore turn to administrative data on total enrollment by district.

The results in table 6 report the coefficient from regressions with alternative measures of enrollment on a dummy variable for districts that included a JFPR-eligible school and a set of province fixed effects.<sup>17</sup> The first three columns correspond to double-differences specifications. The dependent variable in column 1 is the change in girl enrollment between the 2003 and 2004 school years,  $E_{gdt} - E_{gdt-1}$ . The next two columns are checks on the identifying assumption in the double-differences framework: the dependent variable in column 2 corresponds to the growth in boy enrollment between the 2003 and 2004 school years,  $E_{bdt} - E_{bdt-1}$ , while the dependent variable in column 3 corresponds to the growth in girl enrollment between 2002 and 2003,

<sup>16</sup> We thank an anonymous referee for pointing this out to us.

<sup>17</sup> The EMIS collects data only on enrollment, not attendance, so program effects could only be calculated for enrollment.

**TABLE 6**  
**DISTRICT ESTIMATES OF PROGRAM IMPACT USING ADMINISTRATIVE DATA ON ENROLLMENT**

	Grade 7 Enrollment Change between 2003 and 2004 (after Program) (1)	Grade 7 Enrollment of Boys Change between 2003 and 2004 (after Program) (2)	Grade 7 Enrollment of Girls Change between 2002 and 2003 (before Program) (3)	Difference in Grade 7 Enrollment: Girls – Boys Change between 2003 and 2004 (after Program) (4)	Difference in Grade 7 Enrollment: Girls – Boys Change between 2003 and 2004 (after Program) (5)	Difference in Grade 7 Enrollment: Girls – Boys Change between 2002 and 2003 (before Program) (6)
JFPR school in district	44.537 (15.586)***	20.586 (14.009)	17.552 (12.411)	23.951 (13.253)*	25.633 (14.284)*	.828 (12.336)
Observations	156	156	156	156	137	156
R <sup>2</sup>	.24	.15	.26	.19	.20	.13

**Note.** Robust standard errors are in parentheses. Grade 7 Enrollment of Girls, Grade 7 Enrollment of Boys, and Difference in Grade 7 Enrollment: Girls – Boys are dependent variables.

\* Significant at 10%.

\*\* Significant at 5%.

\*\*\* Significant at 1%.

$E_{gdt-1} - E_{gdt-2}$ —before the JFPR program had started. The remaining three columns in the table are based on triple-difference specifications. In column 4 the dependent variable is the growth in girl enrollment, relative to boy enrollment, between 2003 and 2004,  $(E_{gdt} - E_{gdt-1}) - (E_{bdt} - E_{bdt-1})$ . Column 5 is a comparable specification that excludes districts that only had one secondary school; in these districts there would have been no scope for within-district transfers, and any transfers would therefore have taken place to schools in other districts. The last column, finally, presents the results from a triple-difference estimation in which the dependent variable is the change in girl enrollment, relative to boy enrollment, between 2002 and 2003,  $(E_{gdt-1} - E_{gdt-2}) - (E_{bdt-1} - E_{bdt-2})$ . This specification is presented as a validity check on the identifying assumption in the triple-differences framework.

The coefficient in the double-difference specification in column 1 shows that the enrollment of girls grew by 44.5 more students in districts that had a JFPR school than in other districts. The coefficients in the next two columns suggest, however, that assigning a causal interpretation to this coefficient is suspect: the enrollment of boys, who were ineligible for scholarships, also grew by more in districts that included JFPR schools, and the enrollment of girls grew by more in districts that included a JFPR school even before the program started. We therefore turn to the triple-difference specifications. The results in column 4 show that girl enrollment relative to boy enrollment grew by 23.9 more students in districts with a JFPR school; the coefficient is very similar if we exclude districts that only had one school (col. 5). Column 6, finally, shows that the trends in relative girl-boy enrollment were the same in districts with JFPR schools and in other districts before the program started. The coefficient on the “treated” districts is 0.83 (with a robust standard error of 12.3), a finding that provides considerable reassurance of the triple-difference identification strategy.

What is the magnitude of the implied effect on enrollment? Thirty-four districts had exactly one JFPR school, 13 had exactly two, and 5 had exactly three.<sup>18</sup> The mean number of JFPR schools in a district that had any JFPR schools is therefore 1.44, implying that on average 65 scholarships ( $45 \times 1.44$ ) were awarded in districts that had at least one JFPR school. The coefficient in column 4 suggests that just over a third ( $23.9/65$ ) of these scholarships resulted in an additional girl enrolled in school. It is not possible to directly compare this figure—an estimate of enrollment in seventh grade, 1

<sup>18</sup> We exclude districts that were officially designated as “ethnic minority areas,” since a comparable scholarship program for boys was offered in these districts. The total number of JFPR schools in the sample of 156 districts is 75.

year after the beginning of the program—with the coefficients in tables 2 and 4—estimates of the probability of enrollment in eighth grade, 2 years after the beginning of the program. We note, however, that the implied effects are remarkably close to one another. If the bulk of transfers from JFPR schools to non-JFPR schools happen within districts, as seems plausible, the results in table 6 are robust to selective transfers across schools. Tables 5 and 6 thus suggest that estimates of the effect of the JFPR program on enrollment in eligible schools may provide a reasonable approximation of the effect on overall enrollment.

## V. Conclusion

Raising the schooling levels of girls is generally seen as an important priority for many developing countries. In spite of this, the evidence on programs that effectively increase school enrollment in low-income countries—including enrollment of girls—continues to be sparse. In this study we estimate the impact of a scholarship program for poor girls in Cambodia. Our main finding is that enrollment and attendance rates among scholarship recipients were approximately 30 percentage points higher than they would have been in the absence of the program. The effects we estimate are remarkably insensitive to different ways of “controlling for” observable differences between scholarship recipients and nonrecipients. Moreover, the implied effects are large: as a point of comparison, Schultz (2004) estimates that the highly regarded PROGRESA program in Mexico increased the transition from sixth grade, the last grade in primary school, to seventh grade, the first year of secondary school, by 11.1 percentage points—about one-third the magnitude of the effect of the JFPR program in Cambodia. Because enrollment rates in low-income countries like Cambodia tend to be much lower than those in middle-income countries like Mexico, the potential scope for program impacts may be larger in the poorest countries.<sup>19</sup>

<sup>19</sup> To put the magnitude of the estimated JFPR program effects into context, we compare them with the cross-sectional elasticity of enrollment with respect to expenditures. For this purpose, we used the 2004 Cambodia Socio-Economic Survey (CSES) and regressed completion of grade 7 on household per capita expenditures (among girls who had completed grade 5). These results suggest that a 1% increase in expenditures is associated with a 0.24 percentage point increase in grade 7 completion. By matching the characteristics of recipients among applicants to the JFPR scholarship program with girls in the CSES we estimate that mean predicted expenditures among scholarship recipients is US\$151. An increase of \$45 is therefore a 29.7% increase, from which we would expect a 7.1 percentage point ( $29.7 \times 0.24$ ) increase in grade 7 completion. Since we estimate that the impact of the program was on the order of 30 percentage points, these back-of-the-envelope calculations suggest that the JFPR program effects are substantially larger than would be expected from the simple income effect associated with the transfer.

As in many other contexts, we find significant evidence of heterogeneous treatment effects in Cambodia. The impact of the program on enrollment and attendance is largest for the most disadvantaged girls—girls of lower socioeconomic status, girls with lower parental education levels, and girls living further away from school. As a result, the JFPR program appears to have dramatically reduced socioeconomic gradients in enrollment and attendance.

Two potential concerns with cross-sectional, nonexperimental estimates like the ones reported in this study are selection on unobservables and selective transfers from program schools to other schools. Program effects estimated by regression discontinuity are somewhat larger than those estimated by OLS or matching; we conclude that selection on unobservables is unlikely to be important in our data. An application of bounds, and triple-difference estimates of program effects using administrative data both suggest that selective transfers across schools are not empirically important; we conclude that the estimates we present of the impact of the program on enrollment at eligible schools are likely to be a good approximation of the effect on overall enrollment.

Our results show that demand-side incentives can effectively increase the school enrollment and attendance of girls in one of the poorest countries in the world, with weak public sector institutions and relatively low quality schooling. Nevertheless, this study leaves some important questions unanswered. Chief among them is the effect of the scholarship on other schooling outcomes, including repetition rates and measures of performance like test scores. Moreover, in the absence of empirical evidence on the effectiveness of other interventions to improve schooling outcomes in Cambodia, we cannot draw convincing cost-effectiveness comparisons or estimate whether the scholarship amount set by the JFPR program was “too large” or “too small” to induce a given change in girl enrollment. Nevertheless, given the paucity of evidence from low income countries, these estimates of large program impacts may be cause for optimism.

## Appendix

**TABLE A1**  
DIFFERENCES IN CHARACTERISTICS AT BASELINE: COMPLETE AND INCOMPLETE APPLICATION FORMS

	Complete Application Form	Incomplete Application Form	Difference	SE
Parents own business	-.01	.03	.04	.05
Either parent is a government employee	-.03	.06	.09	.04
Parent lends money regularly	-.03	.07	.10	.05
Parent completed primary school	-.01	.03	.04	.04
Parent completed secondary school	-.03	.08	.11	.05
Main part of house: cement or brick	-.01	.03	.04	.04
Roof made of tiles, metal, or fiber	-.05	.11	.16	.05
Land owned > 1 hectare	-.03	.08	.11	.06
Own a large asset (>1,000,000 riels)	-.06	.14	.20	.05
Own a truck	-.04	.09	.13	.06
Own a car	-.04	.10	.14	.06
Live with both parents	.00	.01	.01	.04
Live with one parent	.03	-.08	-.11	.03
Live in a hut	.06	-.15	-.21	.04
Earthen floor	.05	-.12	-.17	.04
Family has motorbike or trailer	-.05	.12	.16	.05
Family has bicycle	-.03	.06	.09	.04
Family has ox and cart	-.05	.12	.17	.04
Family has pony and trap	.01	-.02	-.03	.05
Family has no transportation	.08	-.18	-.26	.04
Family has debts > 100,000 riels	.06	-.14	-.21	.05
Number of brothers	.00	.00	.00	.04
Number of sisters	.02	-.04	-.05	.04
Applicant disabled	-.03	.07	.09	.04
Other household member disabled	-.03	.06	.09	.04
Applicant/other member has disease	.02	-.04	-.05	.04
Distance to secondary school (km)	.01	-.02	-.03	.05
Average (of absolute values)			.11	

**Note.** Variables are normalized:  $(X - \text{Mean})/SD$ .

**TABLE A2**  
**DIFFERENCES IN CHARACTERISTICS AT BASELINE, OVERALL, MATCHED, AND TRIMMED**

	Recipients – Nonrecipients						Nearest Neighbor Matching Estimates of “Program Effects” on Covariates					
	Full Sample			Trimmed			Full Sample			Trimmed		
	All	Matched	All	All	Matched	All	Coefficient	SE	Coefficient	SE	Coefficient	SE
Parents own business	-.52	-.04	-.19	-.02	-.02	-.02	-.07	.09	-.26	.14	-.26	.14
Either parent is a government employee	-.68	-.10	-.31	-.06	-.06	-.06	-.59	.12	-.89	.18	-.89	.18
Parent lends money regularly	-.47	-.01	-.25	-.03	-.03	-.03	.03	.08	-.11	.16	-.11	.16
Parent completed primary school	-.81	-.35	-.48	-.31	-.31	-.31	-.35	.09	-.39	.11	-.39	.11
Parent completed secondary school	-.39	-.02	-.24	-.07	-.07	-.07	-.23	.10	-.28	.13	-.28	.13
Main part of house: cement or brick	-.53	-.02	-.19	-.01	-.01	-.01	-.12	.09	-.22	.13	-.22	.13
Roof made of tiles, metal, or fiber	-.96	-.74	-.63	-.68	-.68	-.68	-.41	.08	-.39	.10	-.39	.10
Land owned > 1 hectare	-.79	-.36	-.39	-.36	-.36	-.36	-.15	.09	-.16	.11	-.16	.11
Own a large asset (>1,000,000 riels)	-.67	-.07	-.36	-.09	-.09	-.09	-.27	.09	-.31	.12	-.31	.12
Own a truck	-.18	.00	-.07	.00	.00	.00	-.04	.08	-.02	.10	-.02	.10
Own a car	-.13	.00	-.05	.01	.01	.01	-.09	.09	-.21	.14	-.21	.14
Live with both parents	-.30	-.26	-.20	-.26	-.26	-.26	-.02	.06	-.07	.06	-.07	.06
Live with one parent	.10	.26	.09	.20	.20	.20	.05	.06	-.01	.07	-.01	.07
Live in a hut	.46	.38	.30	.35	.35	.35	-.06	.08	-.05	.09	-.05	.09
Earthen floor	.15	.39	.17	.28	.28	.28	.22	.08	.20	.08	.20	.08
Family has motorbike or trailer	-.77	-.36	-.35	-.28	-.28	-.28	-.45	.10	-.37	.13	-.37	.13
Family has bicycle	-.49	-.50	-.30	-.44	-.44	-.44	-.12	.07	-.11	.08	-.11	.08
Family has ox and cart	-.54	-.43	-.31	-.40	-.40	-.40	-.04	.08	-.01	.09	-.01	.09
Family has pony and trap	-.20	.04	.00	.02	.02	.02	-.41	.10	-.41	.13	-.41	.13
Family has no transportation	.48	.49	.34	.44	.44	.44	.05	.08	.10	.08	.10	.08
Family has debts > 100,000 riels	.57	.77	.32	.57	.57	.57	.54	.09	.42	.11	.42	.11
Number of brothers	-.07	.17	-.05	-.04	-.04	-.04	-.01	.09	-.09	.10	-.09	.10
Number of sisters	.10	.39	.08	.00	.00	.00	.34	.09	.24	.10	.24	.10
Applicant disabled	.07	.15	.04	.13	.13	.13	.12	.08	.10	.09	.10	.09
Other household member disabled	.05	.22	.05	.12	.12	.12	.13	.09	.09	.10	.09	.10
Applicant/other member has disease	.27	.37	.15	.30	.30	.30	.16	.09	.09	.09	.09	.09
Distance to secondary school (km)	.22	.43	.19	.05	.05	.05	.02	.08	.04	.09	.04	.09
Average (of absolute values)	.41	.27	.23	.20	.20	.20	.19	.08	.21	.09	.21	.09

**Note.** Variables are normalized: (X – Mean)/SD.

**TABLE A3**  
DIFFERENCES IN CHARACTERISTICS AT BASELINE ABOVE AND BELOW 45 THRESHOLD, IN SUCCESSIVELY SMALLER BANDS AROUND THRESHOLD

	RD Estimates of "Program Effects" on Covariates														
	Full Sample			Applicants Ranked 40th to 51st			Applicants Ranked 44th to 47st			Quadratic Control Function			Cubic Control Function		
	Difference	SE		Difference	SE		Difference	SE		Coefficient	SE		Coefficient	SE	
Parents own business	-.67	.16	.01	.18	-.04	.31	-.19	.17	-.08	.16	-.15	.17	-.08	.16	
Either parent is a government employee	-.70	.13	-.20	.19	-.02	.30	-.15	.16	-.11	.15	-.15	.16	-.11	.15	
Parent lends money regularly	-.66	.19	-.16	.18	-.19	.31	-.22	.22	-.22	.19	-.22	.22	-.22	.19	
Parent completed primary school	-.94	.09	-.16	.14	-.12	.24	-.33	.15	-.23	.13	-.33	.15	-.23	.13	
Parent completed secondary school	-.33	.10	.09	.14	.21	.25	.00	.07	.07	.14	.00	.07	.07	.14	
Main part of house: cement or brick	-.82	.14	-.36	.15	-.32	.24	-.31	.20	-.31	.16	-.29	.11	-.22	.11	
Roof made of tiles, metal, or fiber	-1.13	.09	-.42	.12	-.19	.19	-.29	.11	-.22	.11	-.29	.11	-.22	.11	
Land owned > 1 hectare	-.95	.12	-.22	.14	.08	.23	-.14	.12	-.07	.12	-.14	.12	-.07	.12	
Own a large asset (>1,000,000 riels)	-1.00	.15	-.47	.20	-.11	.28	-.44	.20	-.44	.18	-.44	.20	-.44	.18	
Own a truck	-.21	.10	-.03	.12	.00	.00	-.01	.14	.00	.14	-.01	.14	.00	.14	
Own a car	-.14	.10	.00	.00	.00	.00	.14	.11	.12	.08	.14	.11	.12	.08	
Live with both parents	-.43	.05	-.16	.07	-.09	.13	-.07	.08	-.29	.08	-.07	.08	-.29	.08	
Live with one parent	.26	.06	.02	.08	.05	.14	.02	.07	.17	.08	.02	.07	.17	.08	
Live in a hut	.72	.07	.21	.09	.15	.14	.06	.10	.16	.09	.06	.10	.16	.09	
Earthen floor	.21	.08	.01	.10	.07	.15	.04	.08	.11	.09	.04	.08	.11	.09	
Family has motorbike or trailer	-1.11	.12	-.38	.16	.09	.22	-.39	.15	-.32	.14	-.39	.15	-.32	.14	
Family has bicycle	-.79	.05	-.07	.08	-.18	.14	.06	.07	.02	.08	.06	.07	.02	.08	
Family has ox and cart	-.74	.08	.00	.12	.05	.17	.00	.10	.06	.10	.00	.10	.06	.10	
Family has pony and trap	-.25	.10	.04	.18	-.12	.29	.06	.11	.12	.12	.06	.11	.12	.12	
Family has no transportation	.75	.05	.05	.07	-.05	.09	-.18	.07	-.16	.08	-.18	.07	-.16	.08	
Family has debts > 100,000 riels	.71	.08	.31	.12	.03	.18	.25	.11	.27	.11	.25	.11	.27	.11	
Number of brothers	-.19	.06	-.14	.09	-.31	.16	-.07	.10	-.16	.09	-.07	.10	-.16	.09	
Number of sisters	.06	.06	-.15	.10	-.45	.16	-.13	.10	-.24	.10	-.13	.10	-.24	.10	
Applicant disabled	.13	.04	.02	.06	-.10	.13	-.01	.06	.03	.07	-.01	.06	.03	.07	
Other household member disabled	.11	.06	.04	.09	.04	.16	.03	.08	.08	.09	.03	.08	.08	.09	
Applicant/other member has disease	.33	.05	.13	.09	.11	.15	.12	.08	.22	.09	.12	.08	.22	.09	
Distance to secondary school (km)	.28	.07	.12	.09	-.07	.13	.05	.10	.02	.09	.05	.10	.02	.09	
Average (of absolute values)	.54	.15	.15	.15	.12	.14	.14	.14	.16	.16	.14	.14	.16	.16	

Note. Variables are normalized:  $(X - \text{Mean})/SD$ .

**TABLE A4**  
**DIFFERENCES IN CHARACTERISTICS AT BASELINE, ATTRITORS VERSUS NONATTRITORS**

	Regression of X on Attrition, Scholarship Receipt, and Interaction											
	Average Value of Covariate			Attritor		Scholarship Recipient		Interaction				
	Nonattritors	Attritors	Difference	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Parents own business	-.01	.25	.26	.17	.06	.25	-.52	.11	-.09	.25		
Either parent is a government employee	-.01	.17	.18	.14	-.18	.20	-.70	.09	.14	.20		
Parent lends money regularly	-.01	.28	.30	.18	.14	.26	-.45	.11	-.15	.26		
Parent completed primary school	-.01	.13	.13	.10	-.33	.14	-.85	.07	.24	.17		
Parent completed secondary school	-.01	.12	.12	.11	-.07	.17	-.40	.07	.02	.16		
Main part of house: cement or brick	-.02	.47	.49	.17	.35	.23	-.49	.10	-.24	.26		
Roof made of tiles, metal, or fiber	-.01	.13	.13	.10	-.43	.14	-1.01	.07	.33	.16		
Land owned > 1 hectare	-.01	.17	.18	.11	-.17	.17	-.80	.09	-.03	.17		
Own a large asset (>1,000,000 riels)	-.01	.29	.30	.16	-.08	.24	-.69	.10	.22	.28		
Own a truck	.00	.07	.07	.12	-.01	.17	-.18	.07	.00	.17		
Own a car	-.01	.29	.30	.23	.39	.34	-.09	.06	-.40	.34		
Live with both parents	.00	-.07	-.07	.09	-.23	.11	-.32	.04	.05	.19		
Live with one parent	-.02	.34	.35	.11	.59	.13	.17	.05	-.47	.21		
Live in a hut	.00	.04	.04	.09	.42	.11	.51	.06	-.48	.19		
Earthen floor	.00	-.07	-.07	.10	.02	.13	.16	.06	-.05	.19		
Family has motorbike or trailer	-.01	.21	.22	.13	-.20	.18	-.79	.08	.19	.21		
Family has bicycle	.00	.08	.08	.10	-.27	.12	-.52	.05	.35	.17		
Family has ox and cart	.00	.07	.08	.10	-.21	.13	-.56	.06	.11	.20		
Family has pony and trap	-.01	.31	.32	.17	.38	.23	-.16	.07	-.38	.24		
Family has no transportation	.00	-.01	-.01	.10	.32	.12	.52	.05	-.29	.18		
Family has debts > 100,000 riels	.01	-.16	-.17	.10	.08	.13	.57	.07	.05	.18		
Number of brothers	-.01	.14	.15	.08	.06	.10	-.07	.05	.19	.20		
Number of sisters	.00	.01	.01	.10	.02	.11	.10	.05	.12	.20		
Applicant disabled	.00	.08	.08	.10	.04	.10	.07	.03	.21	.24		
Other household member disabled	-.01	.13	.13	.09	.26	.12	.08	.05	-.27	.21		
Applicant/other member has disease	.00	.08	.09	.08	.28	.09	.30	.05	-.18	.18		
Distance to secondary school (km)	-.01	.12	.13	.12	.36	.14	.26	.06	-.34	.20		
Average (of absolute values)				.17								

**Note.** Variables are normalized:  $(X - \text{Mean})/\text{SD}$ .

## References

- Abadie, Alberto, and Guido Imbens. 2002. "Simple and Bias-Corrected Matching Estimators for Average Treatment Effects." NBER Technical Working Paper 283, National Bureau of Economic Research, Cambridge, MA.
- Angrist, Joshua D., Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer. 2002. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment." *American Economic Review* 92, no. 5:1535–58.
- Angrist, Joshua D., Guido Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91, no. 434:444–55.
- Angrist, Joshua D., and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* 114, no. 2:533–75.
- Attanasio, Orazio P., Emla Fitzsimmons, and Ana Gomez. 2005. "The Impact of a Conditional Education Subsidy on School Enrollment in Colombia." Unpublished manuscript, Department of Economics, University College London.
- Behrman, Jere R., Pilali Sengupta, and Petra E. Todd. 2005. "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Mexico." *Economic Development and Cultural Change* 54, no. 1:237–75.
- Bobonis, Gustavo J., and Frederico Finan. 2006. "Endogenous Peer Effects in School Participation." Unpublished manuscript, Department of Economics, University of Toronto; Department of Economics, University of California at Berkeley.
- Cardoso, Eliana, and André Portela Souza. 2004. "The Impact of Cash Transfers on Child Labor and School Attendance in Brazil." Unpublished manuscript, Department of Economics, Vanderbilt University.
- Case, Anne. 2004. "The Primacy of Education." Unpublished manuscript, Research Program in Development Studies, Princeton University.
- Case, Anne, Christina Paxson, and Joseph Ableidinger. 2004. "Orphans in Africa: Parental Death, Poverty, and School Enrollment." *Demography* 41, no. 3:483–508.
- Chay, Ken, Patrick McEwan, and Miguel Urquiola. 2005. "The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools." *American Economic Review* 95, no. 4:1237–58.
- Chen, Shaohua, Ren Mu, and Martin Ravallion. 2006. "Are There Lasting Impacts of a Poor-Area Development Program?" World Bank Policy Research Working Paper no. 4084, World Bank, Washington, DC.
- Crump, Richard K., V. Joseph Hotz, Guido Imbens, and Oscar A. Mitnik. 2006. "Moving the Goalposts: Addressing Limited Overlap in Estimation of Average Treatment Effects by Changing the Estimand." NBER Working Paper no. T0330, National Bureau of Economic Research, Cambridge, MA.
- Das, Jishnu, Quy-Toan Do, and Berk Özler. 2005. "Reassessing Conditional Cash Transfer Programs." *World Bank Research Observer* 20, no. 1:57–80.
- Dehejia, Rajeev H., and Sadek Wahba. 1999. "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs." *Journal of the American Statistical Association* 94:1053–62.
- Deolalikar, Anil. 1993. "Gender Differences in the Returns to Schooling and School Enrollment Rates in Indonesia." *Journal of Human Resources* 28, no. 4:899–932.

- De Walque, Damien. 2004. "The Long-Term Legacy of the Khmer Rouge Period in Cambodia." World Bank Policy Research Working Paper 3446, World Bank, Washington, DC.
- DiNardo, John, and David S. Lee. 2004. "Economic Impacts of New Unionization on Private Sector Employers: 1984–2001." *Quarterly Journal of Economics* 119, no. 4:1383–1441.
- Edmonds, Eric V. 2006. "Child Labor and Schooling Responses to Anticipated Income in South Africa." *Journal of Development Economics* 81, no. 2:386–414.
- Filmer, Deon. 2007. "Educational Attainment and Enrollment around the World." Development Research Group, World Bank, <http://econ.worldbank.org/projects/edattain>.
- Filmer, Deon, and Lant Pritchett. 2001. "Estimating Wealth Effects without Expenditure Data—or Tears: An Application to Enrollments in States of India." *Demography* 38, no. 1:115–32.
- Glewwe, Paul. 2002. "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes." *Journal of Economic Literature* 40, no. 2:436–82.
- Glewwe, Paul, and Pedro Olinto. 2004. "Evaluating the Impact of Conditional Cash Transfers on Schooling: An Experimental Analysis of Honduras' PRAF Program." Unpublished manuscript, Department of Economics, University of Minnesota.
- Hahn, Jinyong, Petra E. Todd, and Wilbert van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69, no. 1:201–9.
- Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs." In *Handbook of Labor Economics*, vol. 3A, ed. Orly Ashenfelter and David Card, 1865–2097. Amsterdam: North-Holland.
- Hirano, Keisuke, Guido Imbens, and Geert Ridder. 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." *Econometrica* 71, no. 4:1161–89.
- Imbens, Guido. 2004. "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review." *Review of Economics and Statistics* 86, no. 1:4–29.
- Imbens, Guido, and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62, no. 2:467–75.
- Kim, Jooseop, Harold Alderman, and Peter F. Orazem. 1999. "Can Private School Subsidies Increase Enrollment for the Poor? The Quetta Urban Fellowship Program." *World Bank Economic Review* 13, no. 3:443–65.
- Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2007. "Incentives to Learn." Unpublished manuscript, Department of Economics, Harvard University, <http://www.economics.harvard.edu/faculty/kremer/papers/IncentivesToLearn.pdf>.
- LaLonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *American Economic Review* 76, no. 4:604–20.
- Lee, David S. 2008. "Randomized Experiments from Non-random Selection in U.S. House Elections." *Journal of Econometrics* 142, no. 2:675–97.
- Lee, David S., Enrico Moretti, and Matthew J. Butler. 2004. "Do Voters Affect or Elect Policies? Evidence from the U.S. House." *Quarterly Journal of Economics* 119, no. 3:807–59.

- Maluccio, John A., and Rafael Flores. 2004. "Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan *Red de Protección Social*." Unpublished manuscript, International Food Policy Research Institute, Washington, DC.
- Manski, Charles F. 1989. "Anatomy of the Selection Problem." *Journal of Human Resources* 24, no. 3:343–60.
- McKenzie, David, John Gibson, and Steven Stillman. 2006. "How Important Is Selection? Experimental versus Non-Experimental Measures of the Income Gains from Migration." World Bank Policy Research Working Paper 3906, World Bank, Washington, DC.
- Paxson, Christina, and Norbert Schady. 2007. "Cognitive Development among Young Children in Ecuador: The Roles of Health, Wealth, and Parenting." *Journal of Human Resources* 41, no. 1:49–84.
- Ravallion, Martin, and Quentin Wodon. 2000. "Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy." *Economic Journal* 110, no. 462:158–75.
- Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41–55.
- Royal Government of Cambodia. 2003. *Targeted Assistance for Education of Poor Girls and Children in Ethnic Minority Areas: Programme Implementation Manual*. Phnom Penh: Royal Government of Cambodia.
- Schady, Norbert, and Maria Caridad Araujo. 2008. "Cash Transfers, Conditions, and School Enrollment in Ecuador." *Economía* 7, no. 2 (forthcoming).
- Schultz, T. Paul. 1993. "Investments in the Schooling and Health of Women and Men: Quantities and Returns." *Journal of Human Resources* 28, no. 4:694–734.
- . 2002. "Why Governments Should Invest More to Educate Girls." *World Development* 30, no. 2:207–25.
- . 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics* 74, no. 1, special issue (June): 199–250.
- Smith, James P., and Finis R. Welch. 1989. "Black Economic Progress after Myrdal." *Journal of Economic Literature* 27, no. 2:519–64.
- Smith, Jeffrey A., and Petra E. Todd. 2005. "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125, nos. 1–2: 305–53.
- Strauss, John, and Duncan Thomas. 1995. "Human Resources: Empirical Modeling of Household and Family Decisions." In *Handbook of Development Economics*, vol. 3A, ed. Jere Behrman and T. N. Srinivasan, 1883–2023. Amsterdam: North-Holland.
- Urquiola, Miguel. 2006. "Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia." *Review of Economics and Statistics* 88, no. 1:171–77.
- van der Klaauw, Wilbert. 2002. "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach." *International Economic Review* 43, no. 4:1249–87.
- World Bank. 2001. *Engendering Development through Gender Equality in Rights, Resources, and Voice*. New York: Oxford University Press for World Bank.
- . 2005. *World Development Indicators*. Washington, DC: World Bank.