1998

Two case studies on impact evaluation of education projects

Jooseop Kim
Iowa State University

Two case studies on impact evaluation of education projects

by

Jooseop Kim

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Major Professor: Peter F. Orazem

Iowa State University

Ames, Iowa

1998

UMI Number: 9826546

Graduate College

Iowa State University

This is to certify that the Doctoral dissertation of

Jooseop Kim

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College

TABLE OF CONTENTS

LIST OF TABLES

CAN PRIVATE SCHOOL SUBSIDIES INCREASE SCHOOLING FOR THE POOR?: THE QUETTA URBAN FELLOWSHIP PROGRAM

CAN CULTURAL BARRIERS BE OVERCOME IN GIRLS' SCHOOLING?: THE COMMUNITY SUPPORT PROGRAM IN RURAL BALOCHISTAN

APPENDIX

GENERAL INTRODUCTION

The development of new methodologies and technologies for data collection and analysis has led to an increased emphasis on project evaluation. However, most evaluation studies concentrate on projects in developed countries. Because of the lack of financial and professional resources, project evaluation in developing countries is often done poorly or not done at all. The resulting lack of information on project success may lead to ineffective implementation or expansion of projects, which can result in wasteful use of scarce public resources.

The main purpose of this study is to apply evaluation methods widely adopted in developed countries to two projects in a developing country. While it is widely recognized that randomized experimental design is the most accurate method for project evaluation, it is not feasible in many cases due to financial and administrative difficulties in developing countries. A recent study (John Newman et al. 1994) reports that analyses using randomized experimental design are "not only feasible but also yield the most robust results." They discuss seven case studies with simple randomized experimental designs to support their conclusion. This study applies the method to educational projects in Pakistan. The first project evaluation case study is urban girls' fellowship school pilot project implemented in Quetta, the largest city in Balochistan province of Pakistan. The second study analyzes the impact of the community support girls' school project in rural districts of Balochistan. To test the robustness of alternative evaluation methods, different evaluation methods are used. These include reflexive comparison, matched comparison, randomized control, and cross-sectional and first-differenced econometric analyses. Both of the studies suggest that expansions of the pilot projects are likely to be successful.

Measured program effects in the two case studies are not sensitive to the differences in evaluation methods. In the urban girls' fellowship case study, six different methods are applied, and all of the results show that the fellowship program led to sharp increases of enrollment for both boys and girls. The positive effect of the program is verified in the cohort-specific sample as well as the age-specific sample, implying that initial gains in enrollment persist over time.

The program effects in the community support girls' school project are the same as those in the urban case study. The community support program led to increases of enrollment for both boys and girls. The enrollment growth for rural girls was similar to that of urban girls, but the increase for boys was smaller than that in the urban fellowship program. The results are not substantially altered when alternative evaluation methods are applied.

Dissertation Organization

This dissertation has been organized so that a general introduction and literature review precede two manuscripts with a general conclusions section following. The general conclusions section summarizes the results and discussions, and provides suggestions for further research. The appendix includes additional tables which provide supplementary analyses of the two impact evaluation studies. All references cited are listed after the appendix.

The two manuscripts have been written according to specifications for submission to the World Bank Working Paper Series on Impact Evaluation Reforms. The first manuscript is entitled "Can Private School Subsidies Increase Schooling for the Poor?: The Quetta Urban Fellowship Program." The second manuscript is entitled "Can Cultural Barriers Be Overcome in Girls' Schooling?: The Community Support Program in Rural Balochistan.

LITERATURE REVIEW

Even though some literature on evaluation research dates back to the seventeenth century, the literature on systematic, evaluation based on data analysis is relatively new. The evolution of evaluation methodology has made rapid progress, due not only to the need of accurate measure of effectiveness of socio-economic programs corresponding to the rapid pace of socio-economic changes, but also to technological progress. In particular, dramatic improvements in computer technology have made it possible for evaluation research to be performed at lower cost in both time and money.

Modern data-based evaluation became widespread in the mid-sixties. Before then, evaluations seldom went beyond a verbal description of how the programs were enacted and how they performed. Nevertheless, there were isolated examples of methods typically used in modem evaluations. For example, the so-called "before-and-after" study, which is widely employed currently has its roots in the 1930's. Roethlisberger and Dickson(1939) found that a worker's productivity rises with any change in the intensity of illumination.

Qualitative vs. Quantitative Approaches

Nowadays, the need for evaluation as a part of the policy process is widely recognized. However, there is little consensus on appropriate evaluation methodology. Several methodologies have been developed, each with its own advantages and disadvantages. No single method is universally preferred. One issue is whether to use a qualitative or quantitative methodology. The advantage of qualitative methods is that they offer more detail than quantitative methods. Since qualitative methods are usually based upon interviews with project participants, they may capture more detailed information. The most serious problem with qualitative evaluation is that it is inherently subjective. Different researchers could fail to reach the same conclusion even if they used the same methodologies on the same individuals. Another drawback with qualitative approaches is that it is too costly to gather data sets large enough to get significant results. Generally, qualitative approaches are used in the design and monitoring of projects, while quantitative approaches are appropriate for estimating net impacts.

Mixed approaches are often used in practice. A Brookings field evaluation of a job creation program in 1981 is a good example of linking qualitative and quantitative approaches. The U.S. Department of Labor used a quantitative method with aggregate data to compare government employment before and after the introduction of public service employment. They concluded that the program may have created about 50% more jobs. However, Brookings researchers who performed a qualitative evaluation found that public employment would have been lower without the program, so the evaluation by the U.S. Department of Labor may have underestimated the effect of the program.[1] From this example, we can learn that a quantitative approach can provide a more accurate measure of program effects when it is supplemented by detailed information from a qualitative research.

Microsimulation

Another methodology often used for project evaluation is a microsimulation approach. The advantage of this approach is that it can provide estimated costs and benefits for a project, which can guide the decision on whether implementation is warranted. Since this approach is easier and cheaper, it has been widely adopted in various fields. For example, in the 1975-1976 food stamp debate, the 1977 aborted Defter Jobs and Income proposal, the 1984 - 1986 tax reform debate, and most recently in the welfare reform debate, the microsimulation approach has been applied.

The following steps used for the microsimulation of policy changes on food stamp program costs illustrate the method. First, Current Population Survey(CPS) data was modified to make it compatible with the simulation model. Additional data on public assistance programs which may affect food stamp eligibility on simulated after-tax income which determines eligibility were added. Then, food stamp usage was simulated under existing rules, and under proposed changes to the current rules - so that the results could be compared to measure the expected impact of the policy change.[2]

In spite of its wide spread use, outcomes from microsimulations are often criticized. Since in many cases, microsimulation models depend on strong behavioral assumptions, the estimates from those models may be biased by the researchers' priors on outcomes. In addition, experience has shown that microsimulation has limited validity when compared to actual outcomes. Researchers in project evaluation now accept that experimental methodology offers a better approach to conduct impact analysis. While microsimulation may provide an estimate of a projects' cost, the experimental approach focuses on the benefits of the project. Thus it is recommended that each method should be applied at different stages in a project.

Experimental Design

The true impact of a project cannot be assessed correctly unless we can establish a relevant counterfactual to the project. However, it is impossible to do this exactly since it would require observing the impact of the project's presence and absence simultaneously on a certain group. Instead, experimentalists must establish a control group of non-participants for the project who have the same attributes as the group on whom the project is imposed. The participant group is called the treatment (or experimental) group, and the excluded group is called the control (or comparison) group. The project's net effect is measured as the gross outcome for an experimental group minus the gross outcome for a control group provided that all other factors which might affect behavioral outcomes are common across the treatment and control groups.

It is critical to select comparable experimental and control groups in order to estimate a project's true impact. Statistically, the best way to create comparable experimental and control groups is through randomized assignment. Randomization does not necessarily mean a random sampling, but rather a random allocation to experimental and control groups from a target population. Ideally, randomized assignment will result in experimental and control groups which are identical except for the presence or absence of the treatment. This implies that baseline data on the experimental and control groups (i.e., data on the groups collected before the treatment was put in place) would have the same statistical moments. Any differences in the statistical moments between the groups after the intervention would be interpreted as the net effect of the intervention. However, in practice, it may be impossible to obtain this degree of randomization. In small samples9 the groups may prove to be different, even if the assignment is randomized. For that reason, it may still be necessary to condition the comparison of outcomes on covariates which may differ between treatment and control groups.

An example of randomized design applied to a social program is the evaluation of the Job Training and Partnership Act(JTPA) program. This evaluation illustrates both the advantages of and the possible hurdles encountered in the experimental approach.

During the Reagan administration, JTPA was passed by Congress, replacing the previous Comprehensive Employment and Training Act (CETA). This legislation provided funding for services and job training opportunities for economically disadvantaged workers. Currently, the federal government provides JTPA services to about 900,000 persons each year at a cost about $3,000 per participant.[3]

The evaluation of job training programs including CETA and JTPA has been performed since the mid-1970s. Evaluators have depended on both non-experimental and experimental methods since these methods are believed to be appropriate to estimate the difference between the earnings with and without the programs. Several authors have found that experimental design is superior to non-experimental design because non-experimentally-based statistical analyses cannot avoid the selection bias problem. In the extreme case, some authors such as Barnow(1987) argue that"... experiments appear to be the only method available at this time to overcome the limitation of non-experimental evaluation" (p.190). The study for JTPA also shows the practical difficulties in evaluating a program. The most difficult problem was the randomization of sites in which the study was to be conducted. Ideally sites should be selected randomly, but the JTPA study selected virtually any site which would agree to participate. Moreover, performance standards and resource allocations had to be modified from the original plan in order to gain cooperation of local programs. As a result, the sites included in the model were neither representative of JTPA sites generally, nor were they functioning under normal resources or conditions for a JTPA site.

There have been very few evaluation studies in developing countries. Developing countries may be prone to rapid political change and turnover of governmental personnel which make it difficult to implement and monitor a project experiment. Also, there are relatively few professionals or consulting firms with the training necessary to carry out an impact evaluation compared to developed countries.[4]

The lack of evaluation studies in developing countries may lead to implementation or expansion of ineffective programs. In countries with limited resources, opportunity costs associated with poorly designed or implemented programs in terms of foregone uses of resources are particularly large. The need to make efficient use of resources, together with budgetary constraints which makes it impossible to reach all potential beneficiaries at once, create both the need and the means for conducting pilot interventions in developing countries. The advantages of using randomized experimentation for pilot program are that first, if it is designed at the program's first implementation phases, it can help policymakers choose among alternative program options. Second, it is less expensive than that for a fully implemented program. Third, in practice it is easier to randomize assignment to experimental and control groups when there is only partial coverage of potential beneficiaries than under a fully covered program.

Quasi-experimental Design

Program evaluators often find that randomization is difficult or impossible. In that case, evaluators must use quasi-experimental (or non-randomized experimental) methods. The basic concept of this approach is exactly the same as the randomized experimental approach except that randomization is not available. Therefore, the key issue when using this method is how to construct a control group which mimics as closely as possible the characteristics of the experimental group.

There are two types of quasi-experiments, classified by time points at which the design of evaluations begin and their implementation is undertaken. One is Ex-Ante Quasi-Experiments, in which control groups are constructed before the program intervention. This type of quasi-experiment is preferred since evaluators can get information about both control and treatment groups before intervention of a program so that comparability can be increased. In contrast, Ex-Post Quasi-Experiments, which is less preferred than Ex-Ante Quasi-Experiments, are those for which control groups are constructed after the program intervention. A common reason why this method is widely used in spite of increased problems of incomparability or selection bias is that the decision to carry out an evaluation is made after the program has already been implemented.

The most widely used quasi-experimental method is the matched comparison approach, in which control groups are selected on the basis of characteristics which are comparable to those of the treatment group rather than random assignment. Therefore, in constructing matched control groups, it is desirable for evaluators to fully understand the factors that would affect the outcome.

While the matched comparison approach is still the most widely used quasi-experiment, the flaws of this approach have been increasingly recognized. Evaluators can not with certainty avoid the selection bias problem in using this method. That is, even if observed attributes are identical, evaluators cannot ensure that participants and non-participants also had identical unobserved attributes which might affect outcomes of the treatment. A study evaluating the Salk(polio) vaccine illustrates the possible flaws of the matched comparison method than when the experimental method was used. In that study, both experimental and matched comparison methods were used for evaluation. The vaccine was evaluated to be 14 percent more effective when the matched comparison method was used. Therefore, there were significant differences in measured outcomes using the two different evaluation methods.

Reflexive Evaluation

Another quasi-experimental method is reflexive evaluation. This method doesn't require a constructed control group. Instead, the participants serve as their own control group. The outcome of the program is measured as the change in observed behavior from before implementation to after the program is put in place. Researchers need panel data for this approach, or at least recall data on behavior before implementation. Even though this approach is less robust than experimental and matched comparison methods, most evaluation studies in developing countries use this approach simply because it is less expensive and easier to carry out.[5]

A possible flaw of this approach is that it is hard to disentangle programmatic effects from naturally occurring changes without the program. Evaluations using this approach tend to be inaccurate when natural changes in the absence of a program are large.[6]

Several attempts have been made to increase the credibility of quasi-experimental design. One attempt was made by some statisticians and econometricians (Heckman and Robb, 1985) in the way to use econometric methods to adjust for the differences between constructed control groups and the experimental groups. This work was extended by James Heckman and V. Joseph Hotz(Heckman and Hotz, 1989). They focus on how evaluators can eliminate the selection bias problem which often arises when using non-random control groups approach. They argue that"... simple specification tests eliminate the most unreliable and misleading estimators that give rise to the sensitivity problem recently discussed in the evaluation literature." (p.874)

Nevertheless, the goal of any quasi-experimental evaluation is to mimic the outcome of a properly conducted evaluation using randomized design. That is why the randomized experimental approach is considered by many to be the best method to conduct an impact evaluation if it is feasible.

Notes

[1] V.Joseph Hotz. "Evaluation of Federal Social Programs: An Uncertain Impact". June 1993. pp. 38-39
[2] Ibid. pp. 6-7
[3] See LaLonde (1995) for a detailed discussion.
[4] See John Newman. Laura Rawlings. and Paul Gertler (1994)
[5] Joseph Valadez and Michael Bamberger. "Social Program Evaluation in Developing Countries". Oct. 1993. pp.26
[6] Ibid. pp. 27

CAN PRIVATE SCHOOL SUBSIDIES INCREASE SCHOOLING FOR THE POOR?: THE QUETTA URBAN FELLOWSHIP PROGRAM

A paper submitted to the World Bank Working Paper Series on Impact Evaluation of Education Reforms

Jooseop Kim[1], Harold Alderman[2], and Peter F. Orazem[1]

Introduction

Primary school enrollment rates in Pakistan are lower than in other countries at the same level of economic development. The proportion of children in school is about half that in India and three quarters that in Bangladesh and Nepal. Nationally, the enrollment rate is 58 percent, 69 percent for boys but only 42 percent for girls. The province of Balochistan has the lowest enrollments with only 45 percent of children aged 5-10 enrolled in school. The enrollment gender gap is even wider in Balochistan with 62 percent of boys but only 29 percent of girls enrolled.[3]

The government of Pakistan has established a goal of universal primary enrollment by the year 2006. This would require more than doubling girls' enrollment nationally and more than tripling girls' enrollment in Balochistan. However, increasing government school capacity is constrained by inadequate public budgets. The potential impact of these resources on increased enrollments has been limited by government school location decisions based more on political patronage than on local needs.

The children least served by exiting public schools are girls, children in rural areas, and children in the poor neighborhoods of cities.[4] There is evidence that enrollment rates are constrained by insufficient school supply in urban areas.[5] However, there are additional restrictions on adding new government schools in these poor neighborhoods. The government requires that a neighborhood provide land for a new government school, but many of these neighborhoods have developed as squatters' communities with no established property rights on the land. Conversations with education officials confirmed that lack of defined property rights had greatly restricted the construction of new government schools in the poorest neighborhoods.

The primary schooling opportunities for poor girls are much worse than for poor boys. There are relatively few girls' schools in poor neighborhoods, so that many girls have to enroll in boys' school if they want to get an education. Cultural prohibitions against exposing girls to the public have meant that the absence of girls' school meant a lack of educational opportunities for girls. If universal primary enrollment is to be achieved for girls, more separate girls' schools will need to be established. Given the limitations on increased government provision, an alternative is to try to increase the availability of private girls' schools in poor neighborhoods.

This study measures the success of such an effort to induce the creation of new private girls' schools in Quetta, the capital city of Balochistan. This study is unique in that it represents one of the few attempts to use experimental design methods to evaluate an educational policy innovation.[6] By randomizing the implementation of the pilot program, we are able to generate robust estimates of the impact of the program on enrollments. Random assignment avoids the bias in impact assessments inherent when the program is applied to individuals or groups believed to benefit atypically from the program.

This study shows that regardless of how the impact is measured, the program increased girls' enrollments by an average of 33 percentage points. At the same time, boy enrollments rose an average of 27.5 percentage points, partly because boys were also allowed to attend the new schools, and partly because parents would not send their girls to school and not also educate their boys. While neighborhoods differed in the success of the program, success was not clearly related to the relative wealth of a neighborhood or the education levels of the parents. As a consequence, it appears that the program offers tremendous promise for increasing enrollment rates in other poor urban areas.

The Urban Girls' Fellowship Program

In February 1995, the Balochistan Education Foundation launched the Urban Girls' Fellowship (UGF) Program in Quetta, the capital and largest city of Balochistan. The purpose of this pilot project was to determine whether establishing private schools in poor neighborhoods was a possible and cost effective means of expanding the delivery of primary educational services to girls in lower income neighborhoods of Quetta. Recent evidence from the Pakistan Integrated Household Survey suggests that about 77 percent of girls who start school finish the primary cycle. It was thought that if these poor girls started school, it was probable that many would persist in school long enough to attain literacy.

School establishment was to be encouraged through subsidies paid directly to schools. The subsidy was Rs. 100 (about $3) per month per girl. This subsidy was sufficient to cover typical tuition at the lowest priced private schools. The subsidy was scheduled to last three years, ending in February of 1998, after which the school would be on its own. The subsidy amount was reduced in the second year and reduced again in the third year. Consequently, the schools needed to generate increased revenues from other sources over time if they were to meet their recurrent expenses. Schools were required to meet certain quality standards. These included a requirement that there be no more than 50 boys and girls per classroom, that there must be one teacher for every classroom, and that there could not be more boys than girls in the school. Boys were ineligible for scholarships, so schools had an incentive to attract as girls to an initial upper scholarship limit of Rs. 10,000 (100 girls × Rs. 100 per girl) per month.

The pilot project was limited to ten initial sites. The only restrictions on choice of neighborhood were that the neighborhood had to be composed of poor households and that there be no existing government girls' school in the area. A non-governmental organization (NGO) was contracted to conduct an initial census of each site to establish the number of girls in the target age range (4-8) and to inform parents of the program.[7] The emphasis was to create a partnership between parents in a neighborhood and the school operator. This was to be accomplished by first conducting a meeting of parents in a neighborhood to see if they were interested in attracting a private school to their area. The parents were asked to form a committee, which would represent the neighborhood in negotiations with potential school operators. With the assistance of the NGO, the parents' committee developed a proposal regarding the neighborhood's requirements for a school, resources the neighborhood was willing to provide the school (i.e. land, buildings, equipment) and any other requirements an operator was expected to satisfy. Experienced school operators were provided these specifications and were allowed to make proposals in response. Each parent committee was allowed to select among the proposals. In the end, proposals were received and accepted in all ten sites, and ten schools were ultimately opened.[8]

Survey Design and Data Strategies

Because government resources are in limited supply and the need to expand enrollments is so great, the government of Balochistan needed an accurate measure of the program's success and its prognosis for expansion. It was decided to use randomized assignment into treatment and control groups to accomplish this task. However, there were several factors which constrained the experimental design.[9] With only ten possible sites, the government opted to place one school in each of ten urban slum areas of Quetta. This was considered politically expedient because it assured that all major ethnic groups would get at least one school. Ethnic groups tended to segregate into one or two of these slum areas, so the government could not be accused of favoritism.

A second problem was that there was no recent census of the population from which one could define treatment and control populations. The most recent census was fourteen years old, and the population of Quetta was estimated to have grown at about seven percent per year since then. Consequently, an area frame sampling strategy was chosen to define the treatment and control neighborhoods.

The area frame was designed as follows. A map of Quetta was produced with each of the ten slum areas outlined. In each area, three sites, literally points on a map, were selected. One of these areas was chosen randomly to be the treatment neighborhood for the creation of a private school. The other neighborhoods were reserved to be controls. The only criterion for the treatment neighborhood was that it could not already have a government girls' school. While it was possible that the control sites would contain a government girls' school, it turned out that none of the control sites had girls' schools either.

By randomizing site selection, it was hoped that there would be no systematic differences in characteristics and behavioral patterns between the control and the treatment neighborhoods. However, the lack of information on population characteristics and the small number of pilot sites led to the possibility that the two groups would differ in important ways. Therefore, it was important to collect information on population attributes in all sites to enable us to test for statistically important differences in treatment and control populations which might also affect differences in enrollment outcomes between the two groups. We also have an interest in determining if relative success of a school depends upon observable neighborhood characteristics.

The baseline data collected in the treatment and control sites included information on socioeconomic characteristics of the households, parents' education, and educational attainment and current enrollment status of all children in the household. All households in the treatment neighborhoods were surveyed at the time of the promotion of the scholarship program in the summer of 1994 before any fellowship schools were opened. The baseline survey of households in the control group neighborhoods was conducted in July 1995. Because most of the data on socioeconomic status of the household does not change over time, the difference in the timing of the surveys should not be problematic. Information on the enrollment status of control neighborhood children was obtained for the current year (1995) and retrospectively for the previous year. This does raise the possibility of recall bias, although parents should be able to remember whether their children were in school a year earlier.[10] Subsequently, enrollment data was collected in

1996 in both treatment and control neighborhoods. All data collection and training of surveyors was supervised by the Balochistan Education Management Information System (BEMIS) to insure data comparability.

Theory of Enrollment Response to the Girls' Fellowship Program

Before conducting the statistical comparison of the treatment and control neighborhoods, it is important to identify the possible endogenous responses to the program. It is also important to identify the exogenous variables that might condition the magnitude of those responses. Households are assumed to have parents, a daughter and a son. Parents are assumed to derive utility from their own consumption of goods (Z_h) and from the human capital of their daughter (H_f) and their son (H_m). The utility function has the form $U=U(Z_h, H_f, H_m, T)$, where T is a vector of taste indicators that are not subject to choice. Parents maximize utility subject to a budget constraint. Sending their daughters to school requires that the household sacrifice current consumption and human capital investment for their sons.

Let Y be household income, P_z be the price of consumption goods, and P_f and P_m are the prices of schooling for their daughter and son, respectively. The schooling price includes school fees, the cost of transportation, and the cost of materials. The income constraint on parental utility maximization is $P_z Z_h + P_f H_f + P_m H_m = Y$.

For cultural reasons, parents may face some disutility from sending their daughters to school. Social prohibitions against exposing their daughters to the outside world will cause them to discount the utility they get from their daughter's education by some factor

$d_f<1$. Then the parents utility will have the form $U(Z_h, d_fH_f, H_m, T)$, with $U_{Hf} = d_fU_H(Z_h, H_f, H_m, T)$ and $U_{Hm} = U_H(Z_h, H_f, H_m, T)$.

The first order conditions yield the following relation:

$$(2) \qquad \frac{U_{Hf}}{U_{Hm}} = \frac{d_fU_H(Z_h, H_f, H_m, T)}{U_H(Z_h, H_f, H_m, T)} = \frac{P_f}{P_m}$$

where U_{Hf} and U_{Hm} represent the marginal utility of girls schooling and boys schooling, respectively. To get parents to equalize schooling for their boys and girls so that $H_f = H_m = H$, the cost of girls schooling must be discounted by $P_f = d_fP_m < P_m$. Alternatively, if the pecuniary costs of schooling are the same for boys and girls so that $P_f = P_m$, then the right-hand-side of (2) will equal one. Then, $d_fU_H(H_f) = U_H(H_m)$, which implies that $U_H(H_f) > U_H(H_m)$. Diminishing marginal utility would then imply that $H_m > H_f$ at the optimum.

Reduced form equations for boy's and girl's schooling have the following functional forms:

$$(3) \qquad H_f = H_f\,(P_f, P_m, d_f, Y, P_z, T)$$

$$(4) \qquad H_m = H_m\,(P_m, P_f, d_f, Y, P_z, T)$$

The reduced form equations suggest that enrollment will depend on school fees, the rate at which parents discount girls' education relative to boys, income, the price level, and tastes. Numerous studies suggest that education is a normal good so that $\partial H_m/\partial Y > 0$ and $\partial H_f/\partial Y > 0$. Those conditions are sufficient to insure that $\partial H_m/\partial P_m < 0$ and $\partial H_f/\partial P_f < 0$. The discount factor d_f acts as an additional price on girls' schooling, so $\partial H_f/\partial d_f < 0$. The girls' fellowship program will lower P_f, so girls' schooling will increase. The impact of the fellowship program on boys' enrollment is ambiguous. However, there are two reasons to believe that the girls' fellowship program will have a positive impact on boys' schooling. First, the program creates a new low priced private school that can accept boys, lowering P_m, although it lowers P_f even more. Second, boys' education may increase as their sisters go to school for a very practical reason - parents may want their boys to escort their sisters to and from school. This implies that boys' education and girls' education may be complementary goods so that $\partial H_m/\partial P_f < 0$. In any event, it will be important to monitor both boys' and girls' enrollments in response to the program.

Equations 3) and 4) suggest that income, the cost of schooling, and the disamenity of sending girls to school may condition the enrollment response to the fellowship program. Schooling costs are measured by fees charged in existing neighborhood schools, average distance to schools (a proxy for transport costs) and the opportunity cost of child time (measured by the child's age and its square). The parents' disamenity for sending girls to school is assumed to be inversely related to fathers' and mothers' education. Parents' taste for education are also assumed to depend on the child's birth order (there may be a preference for educating the eldest child, particularly the eldest boy) and on citizenship (refugees may value education less or may feel the return from education is less). These variables comprise the vector of exogenous variables we will use in the analysis below.

Differences and Similarities Between Treatment and Control Neighborhoods

Statistical properties of the baseline data are described in Table 1 for both treatment and control neighborhoods. Sample statistics are reported separately for boys and girls. The treatment sample included 1,310 children, 781 girls and 529 boys. The control sample included 1,358 children, 697 girls and 661 boys. Enrollment rates for boys and girls in the treatment group were higher than those in the control group: 6.6% higher for girls and 8.8% higher for boys. The other variables in Table 1 represent the exogenous variables believed to affect parental enrollment choices for their children. Most of the variables come directly from the questionnaire. However, distance to school, annual fees, and household income were generated from information on the survey. Distance to school and annual fees were measured as the neighborhood average distance and annual fees of the children in school. Averages are the appropriate measure since people all live in the same neighborhood. Household income was estimated using information on the number of adults in the household with various educational attainments and various household assets. Details on the estimated measure of household income are contained in the Appendix.

The purpose of the control group is to get information on the counterfactual state.[11] A reasonable approximation of the change in outcome due to the program intervention is to measure the difference in outcomes between the treatment and control groups before and after the intervention. However. it is important to check whether there are important differences between the treatment and the control groups which might also result in different outcomes.[12]

Tests for statistical significance of the differences between the treatment and the control groups were performed in two ways. First, in order to check if the randomization yielded observationally equivalent treatment and control populations, we conducted tests of the equality of means of the endogenous and exogenous variables. A second analysis was based on estimated enrollment equations in the baseline data. These tested the null hypothesis of the equality of behavioral coefficients in the enrollment choice models for the treatment and control neighborhoods.

Tests of Equality of Means in the Baseline Data Sets

The third and sixth columns of Table 1 report the corrected t-values and the degrees of freedom for tests of the hypotheses that the means of variables are equal across the treatment and control groups.[13] The results show that baseline enrollment rates for both sexes were significantly higher in the treatment group. In addition, there were significant differences in average mother's education and birth order between the treatment and control girls. Nevertheless, the differences were small numerically. For boys, father's highest grade and citizenship were significantly higher in the treatment group. Once again, the differences in means were small numerically. The joint test that the means were equal across all variables was easily rejected. Based upon the results, we can reach a statistical conclusion that the treatment and control samples are not identical.

Tests of Equality Coefficients in the Baseline Enrollment Model

A second way that the treatment and control neighborhoods may differ is in the decision-making process of parents. To check this, we estimated the following binary model of parental decision regarding their children's schooling:

(5) $$R_{it}^* = \beta_t' X_{it} + U_{it}$$

where $R_{it} = 1 \quad \text{if } R_{it}^* > 0$

$R_{it} = 0 \quad \text{if } R_{it}^* \leq 0$

In equation (5), an unobserved variable R_{it}^* depends on the index function, $\beta_t' X_{it}$, where X_{it} is the vector of characteristics in equations (3) and (4) which affect parental choices regarding their children's enrollment. When R_{it}^{*} is positive, we observe the child in school and $R_{it}=1$. Otherwise, the child will not enroll.

Table 2A presents the baseline probit estimates of enrollment choice for boys and girls. Separate estimates are presented for the treatment and control neighborhoods. The estimated parameters exhibit the same sign patterns in the treatment and control groups and are qualitatively similar to results obtained in other studies of enrollment. The coefficient on household income is positive in both samples. Parental education positively influences their children's enrollment, and mother's education is a more important factor influencing girl's education than father's education. Enrollment increases with age, but at a diminishing rate. First-born children have a higher probability of enrollment than their younger siblings, but the coefficient is not significant. Native Pakistanis have a higher probability of enrollment. After pooling the treatment and the control data, we can also estimate the effects of neighborhood average distance to school and average annual fees. They have negative coefficients except for a positive but insignificant effect of annual fees on boy's schooling.

Table 2B shows the result of the tests of equality of coefficients between the treatment and the control groups. The coefficients for the two groups are not statistically different, except for father's educational level in the girls enrollment equation. This result suggests that parental decision making on education is similar in the treatment and control neighborhoods. Despite significant differences in characteristics as reported in table 1, we can still measure the change in enrollment due to the program by measuring the difference between treatment and control group enrollment rates, holding constant the differences in the exogenous variables.

Evaluation Strategy

The evaluation problem is essentially a missing data problem. A child i cannot be simultaneously in both the treatment state (R_{1it}) and the control state (R_{0it}). Letting $d_i = 1$ if child i was eligible for the fellowship, and $d_i = 0$ otherwise, the observed outcome(R_{it}) can be expressed as $R_{it} = d_i R_{1it} + (1-d_i)R_{0it}$. Given the impossibility of observing the true impact of the fellowship program ($\alpha_t = R_{1it} - R_{0it}$), the goal is to get an unbiased estimator of α_t.

One way to get an unbiased estimator of α_t is to use a control group to derive estimates of the counterfactual state. The difference in outcomes between the treatment and control groups is used as an estimate of α_t. Different estimators of α_t depend on different assumptions about the control group. Three estimators depend only on comparisons of endogenous outcomes without controlling for the exogenous variables. Mathematically, these are defined as

(6) Reflexive: $E^R(\alpha_{it}|d_i=1) = E(R_{Tt}) - E(R_{iB}|d_i=1)$

(7) Quasi-experimental: $E^Q(\alpha_{it}|d_i=1) = [E(R_{Tt}) - E(R_{Nt})] - [E(R_{iB}|d_i=1) - E(R_{iB}|d_i=0)]$

(8) Experimental: $E^E(\alpha_{it}|d_i=1) = [E(R_{Tt}) - E(R_{Nt})]$

where subscripts T, B, and N represent treatment neighborhoods, baseline data, and control neighborhoods, respectively. In equation (6), the expected program effect, $E(\alpha_{it}|d_i=1)$, is measured by the gap between the expected enrollment rate after the program, $E(R_{Tt})$, and the expected enrollment rate before the program was implemented, $E(R_{iB}|d_i=1)$. This only requires information on enrollments for the treatment group. The method based on the equation (6) is called the reflexive method. The underlying assumption of this method is that the period t outcome without the program would have been identical to the observed pre-program outcome.

In equation (7), the expected program effect is measured by the gap between the post program outcome in the treatment group, $E(R_{Tt})$, and that in the control group, $E(R_{Nt})$, adjusted by the initial difference between the two groups, $E(R_{iB}|d_i=1) - E(R_{iB}|d_i=0)$. This method is called the quasi-experimental, nonequivalent control method. In this method, it is assumed that the difference in outcomes between the two groups before the program intervention would be constant over time if it were not for the program, so the difference in outcome between the two groups after the program intervention reflects the difference due to the program as well as to the initial difference. Differencing the differences yields an estimate of the program effect.

In equation (8), the expected program effect is measured by the observed gap in outcomes between the treatment group and the control group after the program intervention. This is called the experimental method, in which it is assumed that the control group mimics perfectly the treatment group.

The methodological differences follow from different assumptions about the unobserved counterfactual state, which is the state when the program was not established. The methods based on equations (6) through (8) assume the counterfactual state is non-stochastic. If we relax that assumption so that the counterfactual state, R_{0it}, follows a stochastic process, it is possible to set up the following model:

(9) $$R_{0it} = X_{it}\,\beta_t + U_{it}$$

In equation (9), X_{it} is the vector of observed characteristics as in equations (3) and (4), U_{it} is an error term, and β_t is a vector of parameters to be estimated. Modifying equation

(9) using the definition of the program effect, α_{it}, and assuming that the program effect is invariant across individuals but not time so that $\alpha_{it}=\alpha_t$, we have

(10) Covariate post-test: $R_{it} = X_{it}\,\beta_t + d_i\alpha_t + U_{it}$, for $t = 0, 1, \ldots\ldots\ldots, T$

In equation (10), R_{it} is the observed enrollment rate, and d_i is a dummy variable indicating residence in a fellowship school neighborhood. Assuming X_{it} is independent of the unobserved variables U_{it}, so that $E(U_{it}|X_{it})=0$ for all i,t, we can estimate equation (10) using a cross-sectional data set. An alternative way to estimate the program effect using econometric analysis is to use a first-difference model. Assuming intervention occurs between time t and t-1, and since all the variables except the age variable are time invariant, we can modify equation (10) to be

(11) First difference with covariates: $R_{it} - R_{it-1} = X_i\,(\beta_t - \beta_{t-1}) + d_i\alpha_t + U_{it} - U_{it-1}$.

A further assumption that β_t is also time invariant simplifies equation (9) to

(12) First difference without covariates: $R_{it} - R_{it-1} = d_i\alpha_t + U_{it} - U_{it-1}$.[14]

Results

There are two ways of measuring the effect of the program on enrollments. One is to measure the change in enrollment for children in the target age of 5 to 8. The other is to measure enrollment rates longitudinally for children aged 5 to 8 in the initial year of the fellowship program.

Results of Age-specific Analysis

Because the program was aimed to improve access to school for a target age group, this section estimates the effect on enrollments of 5-8 year old children of access to the fellowship program.

The first four columns of table 3 report age-specific enrollment rates for boys and girls before and after the program intervention. The enrollment rate decreased 7.6 % for boys and rose 1.3 % for girls in the control neighborhoods. At the same time, enrollment in the fellowship school neighborhoods rose 19.8 % for boys and 26.0 % for girls. From this information, we can apply three different methods based on equations (6) through (8). The results using those methods are reported in the first three rows of table 8. The results are similar across all three methods. All imply that the fellowship program had a positive effect on girls in the target age group, and that parents sent their boys to school in increasing numbers as well. Applying the same methods to two years of data yield even larger estimates of the enrollment effects of the fellowship schools.

An alternative method based on equation (10) can also be applied to the same sample. The first two and the last two columns of table 4 report the results of the covariate post-test probit analysis of the probability of enrollment using cross-sectional data. The enrollment rate in fellowship neighborhoods rose 33.4 % for girls, and 22.4 % for boys in the first year of the fellowship program. After two years, enrollment in the fellowship neighborhoods had risen 42.7 % for girls and 38.4 % for boys. These results are consistent with the results in table 3. First, this result shows that parents made responses very quickly to the fellowship program. Considering the fellowship schools were established in February in 1995, and that survey data were collected in July of that year, the response of the parents in the target area was nearly instantaneous. This supports the view that there was excess demand for primary education in these poor areas. Second, the result suggests that the fellowship program was continuously successful year by year.

For girls, the estimated program effect increased by almost 10 % in 1996 over the estimated program effect in 1995. Boys' enrollment rates grew 16 % in the year after implementation.

Results of Cohort-specific Analysis

The cohort-specific analysis follows the enrollments of a fixed group of children over time. This sample has two distinct advantages over the age-specific analysis. First, it allows us to see if initial gains in enrollment persist over time. Because it is assumed that five years of schooling are needed to attain permanent literacy, this program will be truly successful only if children remain in school for several years. The other advantage of the cohort-specific analysis is that we can control for individual specific unobservable effects which might also be correlated with program outcomes.

The last four columns of table 3 report the enrollment rates before and after the program intervention for fixed cohorts of boys and girls in the treatment and the control groups. We begin with the cohort aged 4-7 in 1994 to capture the children aged five to eight in 1995.[15] Because enrollments increase with age at least initially, some of the enrollment growth in the cohort-specific analysis will reflect a maturity effect. Nevertheless, the comparison between the fellowship and control neighborhoods should difference away this maturity effect, leaving an unbiased estimate of the program effect.

Estimates of the fellowship effect for the cohort-specific sample are summarized in the last three rows in table 3. The reflexive method will yield upward biased estimates because of the maturity effect, as evidenced by the 46.8 % increase in boys' enrollment, and 44.3 % increase in girls' enrollment. These estimated effects are much bigger than the reflexive estimates in the age-specific analysis.

The estimates from the quasi-experimental, nonequivalent control method and the experimental method remove the maturity effect under the assumption of common maturity effects across neighborhoods. Consequently, the measured program effects using quasi-experimental and experimental methods are smaller than the reflexive estimate, and are more comparable to the estimates using the age-specific sample. All the results show large gains in both boy and girl enrollments following the opening of the fellowship schools. Most estimates show slightly higher enrollment gains for girls than for boys. Looking across the age-specific and cohort-specific estimates, we can conclude that girl enrollments rose by 25-35 percent as a result of the program, and that boy enrollments rose by a few percentage points less.

The first four columns of table 4 report the cohort-specific post-test probit analysis of the probability of enrollment. The inclusion of quadratic terms in age control for maturation, so the coefficients on the treatment dummy can be interpreted as an estimate of the program effect controlling for the maturity effect. The program effects for the children aged five to eight in 1995 were measured as 33.4 % and 22.4 % increase in enrollment for girls and boys, respectively. One year later, the measured effects grew an additional 6.5 % for girls, and 4.4 % for boys. Rising effects over time indicate that the large initial enrollment gains persisted over time. The persistence of the effect is a promising sign for the continued survival of these schools, particularly since fees rose in the second year in many schools.

Results of First-Difference Analysis

Another possible source of bias in the estimate of the program effect is unobserved heterogeneity in children that is correlated with the program effect. If cross-sectional differences in individual fixed effects are responsible for measured program effects, then we can remove the fixed effect by differencing the dependent variable.

Table 5 presents the results of the first difference analysis under the maintained assumption that the coefficients of the regressors are time invariant. The dependent variable is the change in enrollment status from before to after the implementation of the program. The coefficient on the treatment dummy measures the effect of the program on enrollment choice. The last two specifications of the first difference analysis allow the coefficients on the individual and neighborhood effects to vary over time. The results corroborate results presented above in the sense that the coefficient representing the treatment effect, is significantly positive, and the program effect was larger for girls' enrollment than boys' enrollment. However, now the estimated program effect is larger after one year than after two years, in contrast to the cross-sectional results. The cause of the discrepancy is unclear, but must be related to the control for fixed effects. Note that the enrollment rates were initially higher in the fellowship neighborhoods, and children in school before the fellowship schools opened will not contribute to the measured fellowship school effect in the first difference analysis. Note also that it is possible that the opening of the fellowship schools encouraged parents to send their children to school at younger ages, and the smaller effect over time reflects the first-time enrollment of older children in the control neighborhoods. In fact, some of the later enrollment growth in control neighborhoods may be related to the fellowship program if the promotion of children's education in fellowship neighborhoods spilled over to the control neighborhoods. Nevertheless, the estimated two years enrollment growth effects in specification four are large. Controlling for fixed effects lowers the estimated effect by 12 to 30 percent, leaving the estimated enrollment impact of the fellowship school to be 24.2 percent for boys and 28.1 percent for girls.

Results of Neighborhood-Specific Analysis

Given the strong average estimated enrollment growth due to the creation of the fellowship schools, an important issue is whether there is any significant variation in program effects across the neighborhoods. If so, are there any identifiable neighborhood attributes which increase the likelihood of program success? This analysis is necessarily speculative since there are only 10 neighborhoods and therefore 10 degrees of freedom. Neighborhoods were divided into two groups, neighborhoods with over 30 % increase in girls' enrollment, and those below 30 %.

Table 6 reports the summary statistics for these more and less successful neighborhoods. Eight neighborhoods out of ten neighborhoods fell into the more successful group, so the less successful neighborhoods were the clear exception. Several important findings are apparent. For most variables, the sample means are similar in the two groups. One apparent difference is in household income. However, the higher average income is in the less successful neighborhoods. Clearly if the concern was that poor neighborhoods could not benefit from a subsidized private school, that fear was exaggerated.

Average parental education in neighborhoods is also not a prerequisite for success. Differences in parental education were insignificant. Taking the averages at face value, the program was more successful in the neighborhoods with more educated mothers but less educated fathers.

The variables which differed significantly between the more and less successful neighborhoods were citizenship and distance, although the numerical differences were not great. Citizenship was positively related to enrollment of both boys and girls in the baseline estimates. It is reasonable to assume that the greater success in neighborhoods with higher proportions of citizens reflects a stronger taste for schooling.[16]

Shorter distance to school reflects higher density of schools in a neighborhood. It is not clear why fellowship schools in neighborhoods with more competing schools should do better. On the other hand, the difference in commuting time between more and less successful neighborhoods was only two minutes, so the difference is probably unrelated school success.

An intriguing result was that boys' enrollments rose in neighborhoods with more success raising girls' enrollments, but that boys' enrollments fell in the neighborhoods that were relatively less successful. Why this happened is unclear. However, the result is consistent with a presumption that boys' enrollment and girls' enrollment are complementary so that successfully increasing enrollment of one gender will also increase enrollments overall.

Comparison with Alternative Policy Options

Given the apparent success of the fellowship program, is it cost effective when compared to alternative policy options? Table 7 reports the estimated changes in alternative policies needed to match enrollment increase that resulted from the fellowship program. Two policy options were considered: income transfers to poor households and construction of new schools. Our estimates are based on estimated enrollment choice elasticities with respect to income and distance.

Girls' schooling is much more sensitive to household income than is boys' enrollment. Based on our estimates, 3471 Rs. of direct subsidy to a household would be needed to raise the probability of girls' enrollment by 25 %. This is well above the initial subsidy of 1400 Rs. per year per girl in the fellowship program.[17] A similar increase in boys' enrollment probability would require an income transfer of 15030 Rs.. For comparison, Alderman, Orazem, and Paterno (1996) reported that a 10% increase of household income causes a 1.2 % increase in the enrollment rate in private school, implying that an income transfer of 14808 Rs would be required to raise enrollment 25 % for both sexes. Therefore, the fellowship option would be less expensive than income subsidies.

A policy which decrease distance to schools implies establishing more schools. Our estimates imply that 78% more schools would be required to raise enrollment by 25 %. However, the coefficient on distance was not precisely estimated, at least in part because there was little variation in distance to schools across neighborhoods. However, Alderman, Orazem, Paterno (1996) estimated the enrollment elasticity with respect to distance was 0.1. Based upon this estimation, 250% more private schools would be required to meet the target effect.

Conclusions and Extensions

A summary of all measured program effects is contained in table 8. All of the results show that the fellowship program has positively affected enrollment for both boys and girls. Most show that the effect was larger for girls' enrollment. One can conclude that the estimated program effects are robust to differences in assumptions about possible biases due to measured and unmeasured differences between treatment and control neighborhoods. Before the project was implemented, it was not clear whether low girls' enrollment was due to cultural barriers which cause parents to withhold their daughters from school or to inadequate supply of girls' schools. The results of the urban fellowship experiment provide strong evidence that subsidizing the establishment of girls' primary schools can lead to sharp increases of girls' enrollment. In addition, even though the fellowship was given only to girls, boys' enrollment in those neighborhoods also sharply increased. This suggests that there also may have been excess demand for boys' primary education in these poor areas. The measured change over two years yielded mixed evidence on whether the enrollment growth advantage in fellowship neighborhoods over control neighborhoods continued to grow over time. However, even if the initial enrollment gain decreased in subsequent years, the enrollment gains after two years are still around 25 percentage points. School success appears not to depend on neighborhood income or other observable socioeconomic variables, suggesting that expansion of the program to other poor neighborhoods is also likely to be successful.

Future work will be required to assess the long term effects of the fellowship program. In particular, the future sustainability of the schools and the enrollment effects after the subsidies expire will need to be assessed. The short term success of the fellowship program does not guarantee long term success when the financial burden of supporting the schools are fully borne by the neighborhoods. School outcomes will also need to be assessed. The ultimate success of the fellowship program depends on whether children attain literacy.

Table 1. Summary Statistics of Baseline Datasets and Tests of the Equality of Means Between the Treatment and Control Groups[1), 2)]

Variable	Girls Treatment	Girls Control	t-value [3)]	Boys Treatment	Boys Control	t-value[3)]
Enrollment rate	0.366	0.300	2.67	0.486	0.398	3.03
	(0.482)	(0.459)	[1468]	(0.500)	(0.490)	[1180]
Household Income	7.108	6.808	1.03	7.005	6.592	1.41
	(7.157)	(3.011)	[1476]	(6.815)	(2.847)	[1188]
Age	6.026	6.001	0.19	6.040	6.003	0.44
	(1.403)	(1.429)	[1476]	(1.426)	(1.444)	[1188]
Mother's highest grade	0.619	0.395	2.08	0.623	0.414	1.74
	(2.243)	(1.844)	[1466]	(2.208)	(1.918)	[1183]
Father's highest grade	3.405	3.079	1.27	3.635	2.723	3.38
	(4.745)	(4.882)	[1417]	(4.579)	(4.548)	[1162]
Birth order	2.832	3.004	2.21	3.074	2.965	1.27
	(1.474)	(1.510)	[1476]	(1.447)	(1.482)	[1188]
Citizenship	0.868	0.835	1.79	0.877	0.814	2.98
	(0.339)	(0.371)	[1476]	(0.329)	(0.389)	[1188]
Distance to school	17.77	17.81	0.05	16.93	16.42	0.62
	(9.443)	(9.991)	[491]	(9.338)	(9.394)	[515]
Annual fees	244.3	187.0	1.19	531.3	391.7	1.73
	(536.0)	(502.5)	[480]	(1036.8)	(765.1)	[505]
Joint test[4)]			121.61			82.20
Number of observations	781	697		529	661	

1) Age 4 to 8 for both groups. The baseline data was collected in 1994 for the treatment group, and collected in 1995 for the control group.

2) The numbers shown in parentheses are the standard deviations and those in the square brackets are the degrees of freedom. The degrees of freedom differ due to missing information in the surveys.

3) The null hypothesis is that the mean of the variable in the treatment group is equal to that in the control group. If the t-value is smaller than 1.96, the null hypothesis cannot be rejected at 0.05 significance level.

4) Reported numbers are F statistics with degrees of freedom (9, 1478) for girls and (9,1190) for boys. The null hypothesis was that the means of the variables between the treatment group and the control group are equal for all variables.

Table 2A. Baseline Probit Analysis of the Probability of Enrollment

Variable	Girls and Boys			Girls			Boys		
	Treatments	Controls	Pooled	Treatments	Controls	Pooled	Treatments	Controls	Pooled
Household	0.138	0.422	0.143	0.171	0.572	0.196	0.037	0.218	0.043
Income/10.000	(2.362)	(2.879)	(2.710)	(2.377)	(2.870)	(2.921)	(0.346)	(0.954)	(0.455)
Age	1.820	2.235	2.060	1.611	2.623	2.045	2.176	1.927	2.065
	(5.226)	(6.323)	(8.396)	(3.612)	(4.864)	(6.048)	(3.674)	(3.986)	(5.650)
Age square	-0.101	-0.140	-0.124	-0.089	-0.174	-0.127	-0.119	-0.119	-0.118
	(3.621)	(5.014)	(6.347)	(2.508)	(4.127)	(4.745)	(2.513)	(2.884)	(4.044)
Mother's	0.051	0.094	0.065	0.067	0.118	0.082	0.007	0.072	0.036
Highest grade	(2.443)	(3.422)	(4.091)	(2.500)	(2.649)	(3.739)	(0.197)	(1.963)	(1.516)
Father's	0.023	0.065	0.049	0.027	0.084	0.057	0.025	0.050	0.040
Highest grade	(2.369)	(6.634)	(7.337)	(2.271)	(5.997)	(6.570)	(1.498)	(3.550)	(3.794)
Birth order	-0.029	-0.036	-0.030	-0.017	-0.020	-0.021	-0.036	-0.053	-0.036
	(0.918)	(1.214)	(1.407)	(0.416)	(0.461)	(0.732)	(0.717)	(1.251)	(1.153)
Citizenship	0.693	0.335	0.569	0.628	0.214	0.482	0.762	0.538	0.716
	(5.207)	(2.556)	(6.375)	(3.590)	(1.079)	(3.839)	(3.545)	(2.888)	(5.464)
Girl	-0.419	-0.541	-0.474						
	(4.878)	(5.340)	(7.106)						
Distance						-0.001			-0.012
To School						(0.036)			(0.598)
Annual						-0.380			0.073
Fees/1.000						(1.185)			(0.271)
Number of Observations	1.231	1.324	2.555	725	677	1.402	506	647	1.153
Pseudo R^2	0.277	0.295	0.273	0.230	0.331	0.254	0.358	0.293	0.304

Table 2B. Equality Test on Coefficient Between Treatment and Control Group

Variable	Girls and Boys		Girls		Boys	
	χ^2	result	χ^2	result	χ^2	result
income	2.03	not reject	2.58	not reject	0.25	not reject
age	0.09	not reject	0.02	not reject	0.14	not reject
age square	0.89	not reject	0.47	not reject	0.99	not reject
mother's highest grade	1.50	not reject	0.73	not reject	1.95	not reject
father's highest grade	7.68	reject	6.88	reject	1.09	not reject
birth order	0.07	not reject	0.00	not reject	0.04	not reject
citizenship	1.20	not reject	1.36	not reject	0.00	not reject
girl	0.05	not reject				
Joint Test	29.90	reject	23.33	reject	13.68	reject

Significance level: $\alpha = 0.05$

Table 3. Enrollment Rate Before and After Program Intervention

	Age-specific				Cohort-specific			
	Treatment		Control		Treatment		Control	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Enrollment Rate Before Program(E_B)	56.33	45.29	51.06	34.86	38.75	34.06	36.55	29.03
Enrollment Rate in 1995(E_{95})	64.29	63.93	49.68	38.37	64.29	63.93	49.68	38.37
Enrollment Rate in 1996(E_{96})	76.15	71.30	43.50	36.20	85.50	78.36	59.87	45.97
E_{95} - E_B	7.96	18.64	-1.38	3.51	25.54	29.87	13.13	9.34
E_{96} - E_B	19.82	26.01	-7.56	1.34	46.75	44.30	23.32	16.94

Measure of Effect	Age-specific		Cohort-specific	
	Boys	Girls	Boys	Girls
Reflexive, 1994-1995	8.0	18.6	25.5	29.9
Reflexive, 1994-1996	19.8	26.0	46.8	44.3
Quasi-experimental, 1994-1995	9.3	15.1	12.4	20.5
Quasi-experimental, 1994-1996	27.4	24.8	23.4	27.4
Experimental, 1994-1995	14.6	25.6	14.6	25.6
Experimental, 1994-1996	32.7	35.1	25.6	32.4

Note: Since 1994 baseline data for the control group was not available, they were estimated from the 1995 baseline data in the way that children who enrolled in advances grades in 1995 and enrolled in recall data were considered in enrolled in 1994.

Table 4. Post-test Probit Analysis of Probability of Enrollment Using Cross-sectional Data[1)]

Variable	1995[2)]		1996, Cohort-specific[3)]		1996, Age-specific[4)]	
	Girls	Boys	Girls	Boys	Girls	Boys
Treatment dummy	0.334	0.224	0.399	0.268	0.427	0.384
	(10.148)	(5.143)	(9.679)	(5.511)	(8.488)	(5.495)
Household Income/10,000	-0.001	-0.003	0.012	0.072	0.034	0.128
	(0.022)	(0.080)	(0.333)	(1.513)	(0.724)	(1.872)
Age	0.141	0.276	0.229	0.936	0.615	1.330
	(0.652)	(1.197)	(0.797)	(3.416)	(1.970)	(3.925)
Age square	-0.008	-0.016	-0.011	-0.057	-0.036	-0.083
	(0.496)	(0.890)	(0.570)	(3.113)	(1.546)	(3.268)
Mother's highest grade	0.016	0.030	0.029	0.011	0.027	0.018
	(0.040)	(2.330)	(1.505)	(0.867)	(1.822)	(1.231)
Father's highest grade	0.013	0.003	0.030	0.011	0.035	0.020
	(3.383)	(0.707)	(6.293)	(2.433)	(6.656)	(3.523)
Birth order	-0.008	-0.026	-0.016	-0.020	-0.0002	-0.031
	(0.720)	(2.042)	(1.214)	(1.516)	(0.016)	(1.904)
Citizenship	0.152	0.225	0.143	0.201	0.187	0.173
	(3.040)	(4.362)	(2.374)	(3.501)	(2.783)	(2.465)
Distance to School	-0.008	0.003	-0.029	-0.027	-0.035	-0.036
	(1.074)	(0.358)	(3.190)	(2.347)	(3.361)	(2.511)
Annual Fees/1,000	-0.443	-0.030	-0.170	-0.362	-0.316	-0.618
	(3.640)	(0.241)	(1.088)	(2.535)	(1.719)	(2.723)
Number of observations	1,031	830	845	700	764	650
Pseudo R^2	0.141	0.100	0.312	0.215	0.350	0.380

1) The coefficients reported here are dF/dX, where F is dependent variable and X is independent variable, not actual coefficients. Since the dependent variable is a discrete variable, dF/dX is not identical to actual coefficients. The numbers shown in the parentheses are z-values corrected for cluster effect. Dummy variables for each neighborhood included.

2) Children in this data are aged 5 to 8 in 1995. Dependent variable is enrollment status in 1995.

3) Children in this data are aged 5 to 8 in 1995. Dependent variable is enrollment status in 1996.

4) Children in this data are aged 5 to 8 in 1996. Dependent variable is enrollment status in 1996.

Table 5. First Differnce Analysis for Change of Enrollment Decision[1)]

Variable	1994 - 1995		1994 - 1996		1994 - 1995		1994 - 1996	
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Treatment Dummy	0.367	0.292	0.264	0.088	0.469	0.428	0.281	0.242
	(5.518)	(3.591)	(3.165)	(0.909)	(5.833)	(3.755)	(2.931)	(1.723)
Δ age square	-0.077	-0.082	-0.047	-0.046	-0.071	0.032	0.079	0.073
	(4.785)	(4.447)	(5.006)	(4.502)	(0.343)	(0.137)	(0.641)	(1.323)
Age94 square					-0.001	-0.022	-0.047	-0.080
					(0.040)	(0.525)	(1.055)	(1.686)
Income/10.000					-0.151	-0.009	-0.009	-0.005
					(2.680)	(0.122)	(1.309)	(0.588)
Mother's highest grade					-0.007	0.016	-0.009	-0.030
					(0.374)	(0.652)	(0.380)	(1.020)
Father's highest grade					0.004	-0.028	0.043	0.014
					(0.458)	(2.751)	(4.561)	(1.226)
Birth order					-0.029	-0.050	-0.008	-0.021
					(1.152)	(1.667)	(0.250)	(0.616)
Citizenship					0.006	0.093	0.243	0.212
					(0.047)	(0.717)	(1.515)	(1.318)
Distance to school					-0.001	0.027	-0.054	0.029
					(0.051)	(1.407)	(1.936)	(1.272)
Annual fees/1.000					-0.755	-0.103	-0.588	0.765
					(2.424)	(0.299)	(1.623)	(1.813)
Number of observations					1.055	861	863	725
Pseudo R^2					0.04	0.04	0.09	0.05

1) The Coefficients reported here are dF/dX. not actual coefficients. Children in the sample are aged 4 to 7 in 1994.

Table 6. Statistical Summary of Successful and Unsuccessful Neighborhoods

Variable	Girls		
	more successful	less successful	t-value
income	6,819	8,060	2.05
	(6690)	(8468)	
mother's highest grade	0.68	0.42	1.37
	(2.36)	(1.81)	
father's highest grade	3.28	3.82	1.34
	(4.76)	(4.70)	
citizenship	0.90	0.76	4.94
	(0.30)	(0.43)	
distance to school	17.44	19.06	9.55
	(2.05)	(1.84)	
annual fees	247.0	251.6	0.42
	(128.1)	(134.9)	
number of observations	599	182	
girl's enrollment change	41.5 %	8.5 %	
boys' enrollment change	36.8 %	- 1.8 %	

Table 7 Estimated needs to meet target effect

Alternatives	elasticities		change required to meet target effect (25%)	
	girls	boys	girls	boys
Direct subsidy to household	0.503	0.115	3471 Rs./household (50 %)	15030Rs./household (150 %)
Decrease distance to school	0.320	0.732	13.48 min. (78 %)	5.71 min (34 %)

Note: Children in the sample were aged 4 to 7. Numbers in parenthesis reports the amount as percentage needed to meet target effect. For example, direct subsidy to household which leads 50 % increase in household income may increase 25 % increase in girls' enrollment rate.

Table 8 Comparison of the Effect of the Fellowship Program

Methods	Mathematical Expression	Age-specific Boys	Age-specific Girls	Cohort-specific Boys	Cohort-specific Girls
Measure of effect using means					
Reflexive	$E^R(\alpha_{it} \mid d_i=1)= E(R_{Tt})-E(R_{iB} \mid d_i=1)$	8.0	18.6	25.5	29.9
(1994-1995)		(0.42)	(0.44)	(0.43)	(0.44)
Reflexive	$E^R(\alpha_{it} \mid d_i=1)= E(R_{Tt})-E(R_{iB} \mid d_i=1)$	19.8	26.0	46.8	44.3
(1994-1996)		(0.51)	(0.53)	(0.52)	(0.54)
Quasi-experimental	$E^Q(\alpha_{it} \mid d_i=1)=[E(R_{Tt})-E(R_{Nt})]-[E(R_{iB} \mid d_i=1)-E(R_{iB} \mid d_i=0)]$	9.3	15.1	12.4	20.5
(1994-1995)		(0.53)	(0.54)	(0.54)	(0.54)
Quasi-experimental	$E^Q(\alpha_{it} \mid d_i=1)=[E(R_{Tt})-E(R_{Nt})]-[E(R_{iB} \mid d_i=1)-E(R_{iB} \mid d_i=0)]$	27.4	24.8	23.4	27.4
(1994-1996)		(0.73)	(0.70)	(0.74)	(0.71)
Experimental	$E^E(\alpha_{it} \mid d_i=1)=E(R_{Tt})-E(R_{Nt})$	14.6	25.6	14.6	25.6
(1994-1995)		(0.65)	(0.67)	(0.65)	(0.67)
Experimental	$E^E(\alpha_{it} \mid d_i=1)=E(R_{Tt})-E(R_{Nt})$	32.7	35.1	25.6	32.4
(1994-1996)		(0.59)	(0.65)	(0.60)	(0.66)
Measure of effect using regression					
Covariate post-test	$R_{it} = X_{it}\beta_t + d_i\alpha_t + U_{it}$	22.4	33.4	22.4	33.4
(1995 cross-sectional)		(0.04)	(0.03)	(0.04)	(0.03)
Covariage post-test	$R_{jt} = X_{jt}\beta_t + d_j\alpha_t + U_{jt}$	38.4	42.7	26.8	39.9
(1996 cross-sectional)		(0.07)	(0.05)	(0.05)	(0.04)
First-difference without X's	$R_{it}-R_{it-1}=d_i\alpha_t+U_{it}-U_{it-1}$			29.2	36.7
(1994-1995)				(0.08)	(0.07)
First-difference without X's	$R_{it}-R_{it-2}=d_i\alpha_t+U_{it}-U_{it-2}$			8.8	26.4
(1994-1996)				(0.10)	(0.08)
First-difference with X's	$R_{it}-R_{it-1}=X_i(\beta_t-\beta_{t-1})+d_i\alpha_t+U_{it}-U_{it-1}$			42.8	46.9
(1994-1995)				(0.11)	(0.08)
First-difference with X's	$R_{it}-R_{it-2}=X_i(\beta_t-\beta_{t-2})+d_i\alpha_t+U_{it}-U_{it-2}$			24.2	28.1
(1994-1996)				(0.14)	(0.10)

Note: Numbers in parenthesis report standard errors.

Appendix 1

It is our primary concern whether the fellowship program affected the distribution of the coefficient of the variables over time. Even though four different specifications of econometric models depending on the different assumptions on the stability of the coefficient over time and the length of the time lag were used in the first difference analysis, all of the results are not necessarily valid for measuring the program effect. The result of the stability test of coefficients in the enrollment probability reported in the table A-1. Even though the hypothesis of no structural change over time was rejected, the fellowship program did not significantly affect the distribution of the individual coefficients over time. This suggests that the measurement of the program effect using specification without X effect terms may yield more unbiased than using specification with X effect terms.

Table A-1. Stability Test of Coefficients in the Enrollment Probability Equations
H_0 : Coefficients in 1994 = coefficients in 1995

Variable	Boys			Girls		
	t-value	d.f.	result	t-value	d.f.	result
income	0.821	893	not reject	0.165	937	not reject
age	1.907	893	not reject	1.305	937	not reject
age square	2.213	893	not reject	1.911	937	not reject
mother's highest grade	0.287	893	not reject	0.714	937	not reject
father's highest grade	0.776	893	not reject	0.507	937	not reject
birth order	0.345	893	not reject	0.158	937	not reject
citizenship	0.209	893	not reject	0.507	937	not reject
Joint Test	12.73	893	reject	11.67	937	reject

significance level: $\alpha = 0.05$

Appendix 2

It is difficult to derive income estimates for households in Pakistan. The relative importance of production for home consumption, informal labor market arrangements, barter trade and other economic activity occurring outside formal markets complicate income measurement. The budget for this project did not include resources sufficient to conduct a careful analysis of income for each household. However, the Pakistan Integrated Household Survey (PIHS) had conducted such a detailed survey of household incomes and socioeconomic attributes in 1991. The PIHS allows us to predict household income based on a regression of income on easily observed household attributes. The current study collected information on these household attributes and then used the PIHS estimates to generate predicted incomes based on these attributes.

The PIHS income equation is reported in Table A2. The specification follows that used by Alderman and Garcia (1996). That study estimated income and expenditure equations for 217 households in a single city in Balochistan. The Alderman-Garcia estimates can serve as independent validation of the income estimates we derive from the PIHS data. The Alderman-Garcia estimates are less useful for our purpose than is the PIHS because their data include rural households and the data are from 1986. The PIHS has sufficient urban observations to estimate an income equation for urban households, and it is chosen to our 1994 base period. The variables in the income equation include the number of adult males and females, the number of males and females with primary, secondary and tertiary level schooling, and the value of household assets. Alderman and Garcia found that this income specification generated predicted values that performed well in explaining household savings, loans, and nutrition status.

In general, the PIHS income estimates are sensible. Households with more capital assets, more human capital and more adult males have higher incomes. The results corresponded reasonably well in sign with those in Alderman and Garcia. More importantly, the two estimates generate equivalent estimates of relative household income. The correlation in predicted income based on the PIHS versus the Alderman-Garcia estimates is 0.82.

Table A2 Income Equations

Variable	**Alderman and Garcia**	**PIHS**
Intercept	5,999	3,303
	(2.61)	(4.64)
Number of males	938	1,219
aged 16 or more	(0.92)	(3.73)
Number of males	1,691	a
aged 6-16	(2.09)	
Number of females	-709	-188
aged 16 or more	(-0.54)	(-0.57)
Number of females	1,009	a
aged 6-16	(0.64)	
Number of children	2,820	a
5 or below	(2.99)	
Number of males with	6,140	-1,171
primary schooling	(2.95)	(-2.55)
Number of males with	2,279	-364
secondary schooling	(1.69)	(-0.92)
Number of males with more	6,435	147
than secondary schooling	(1.41)	(0.96)
Number of females with	6,707	-406
primary schooling	(1.85)	(-0.69)
Number of females with	7,758	889
middle schooling or more	(1.35)	(3.68)
Rainfed land	110	b
	(2.34)	
Irrigated land	665	b
	(4.93)	

[a]Not available in the PIHS
[b]Not relevant for urban areas

Table A2. (continued)

Variable	**Alderman and Garcia**	**PIHS**
Acres of orchards	4,065	b
	(2.57)	
Value of livestock	0.335	b
	(1.05)	
Value of vehicles	0.171	0.012
	(8.55)	(2.48)
Value of machinery and tools	0.125	0.007
	(1.27)	(1.88)
R^2	0.747	0.03
N	217	2,112

[b]Not relevant for urban areas

Notes

[1] Iowa State University.
[2] World Bank.
[3] Statistics based on 1996 data provided by the Pakistan Education Management Information System.
[4] See Alderman, Orazem and Paterno (1996) for an analysis of enrollment demand for boys and girls by income group.
[5] Alderman, Orazem and Paterno (1996) found that enrollment rates in poor urban neighborhoods fell significantly as distance to school increased.
[6] Newman, Rawlings and Gertler (1994) found that there were very few examples of the use of randomized control in evaluations of social projects in developing countries since 1980. Boruch, McSweeny, and Soderstrom (1978) found that in the period before 1980, less than five percent of studies using randomized control in nonlaboratory settings were in developing country contexts.
[7] The NGO, the Society for Community Support of Primary Education in Balochistan (SCSPEB), had several years of experience, primarily school promotion efforts in rural communities.
[8] This eventually led to an eleventh school. Divisions in one neighborhood caused the parents to divide into two schools with half of the 100 scholarships going to each school.
[9] The areas selected were primarily areas where squatters had established residence on government land that was not served by the Quetta municipal sewer system.
[10] In the analysis that follows, we use multiple methods to evaluate the change in enrollment in the treatment neighborhoods. The conclusions are not sensitive to differences in evaluation method. Therefore, the potential recall bias does not drive any conclusions about program success.
[11] Grossman (1994) classifies a randomly assigned counterfactual group as a "control group", and a nonrandomly assigned counterfactual group as a "comparison group".
[12] Newman, Rawlings, and Gertler (1994) pointed out that tests are rarely done for statistical significance of the differences, so that probabilities of receiving the program may not be equal for individuals or communities in many of the evaluation studies in developing countries.
[13] Since the sampling method in this study follows a cluster sampling, we applied Huber's method to correct for the intercorrelation problem within groups.
[14] Note that the estimated program effect on first difference analysis is sensitive to the stability of the coefficient over time and the length of the time lag. To validate (12), it was necessary to perform a statistical test of stability of the coefficients over time. The result of the stability test shows that the distributions of the coefficients are not significantly changed over time due to the fellowship program. Further results are attached in the Appendix.
[15] The cohort-specific enrollment rates in 1994 are lower than the 1994 average for the age-specific analysis. The reason is that the age-specific groups average one year older in 1994. By 1996, the enrollment rates were higher than in the age-specific analysis because by then, the cohort was one year older on average than the age-specific sample.
[16] Lower average citizenship may also signal a neighborhood with greater ethnic diversity. Because the success of the school depended on an agreement among parents to form a committee, divisions among ethnic lines may have hindered the success of the school.
[17] In addition to the 100 Rs. per month, each school received 200 Rs. per girl to defray start up costs.

CAN CULTURAL BARRIERS BE OVERCOME IN GIRLS' SCHOOLING? : THE COMMUNITY SUPPORT PROGRAM IN RURAL BALOCHISTAN

A paper submitted to the World Bank Working Paper Series on Impact Evaluation of Education Reforms

Jooseop Kim[1], Harold Alderman[2], and Peter F. Orazem[1]

Introduction

Balochistan is the largest but least populous province of Pakistan. As a consequence, it has a higher proportion of its children living in rural areas. Increasing educational opportunities for rural children presents many challenges to the government. Parents are less educated and may see less value from schooling than do their educated urban counterparts. In addition, there are many uses of child time other than schooling, such as care and feeding livestock or chores on family plots, implying that opportunity costs for child time may be higher in rural areas.[3]

Even if demand for schooling were equal between rural and urban areas, there are serious impediments to addressing rural schooling needs. First, villages are quite remote with few educated adults who might meet the qualifications to teach. While urban residents can be assigned to teach in rural areas, teacher absenteeism increases with commuting distance between home and school.[4] In addition, rural teachers often apply for transfer to urban schools, leading to high turnover among rural teachers.

These problems are more daunting for rural girls' education. Parents prefer to have female teachers for their girls, but there are even fewer educated women than men who could serve as teachers in rural areas. Furthermore, social taboos on female travel make it difficult for women teachers to commute daily from urban to rural areas.

Constraints on the supply of rural schooling as well as weaker demand for schooling by rural parents have resulted in much poorer educational outcomes for rural children. Only 63 % of males and 25 % of females aged 10 years and older in rural areas have ever attended school, compared to 80 % for urban males and 57 % for urban females. The situation is even worse in Balochistan. Only 15 % of females aged 10 and older in Balochistan have ever attended school. As a consequence, the female literacy rate in Balochistan is only 8 %, while the literacy rate of females in urban areas is 49 %.[5]

The government of Pakistan has set a goal of universal primary education by the year 2006. Attainment of this goal will require rapid expansion of primary enrollments in rural areas, with particular emphasis on expansion of female primary enrollments. This study measures the success of such an effort to encourage female enrollments through the creation of community public girls' schools in rural areas of Balochistan. Ex-post matched comparison methods were applied to evaluate the success of the Community Support Process (CSP) program.

Under the assumption that girls' education is a rationed good in the absence of a girls' school in the community, we treat the creation of a CSP school as a relaxation of this rationing constraint. We find that regardless of how the impact is measured, the CSP program increased girls' enrollment by an average of 22 percentage points. It also increased boys' enrollment by an average of 9 percentage points, suggesting that boys' and girls' enrollment are complementary goods.

The Community Support Process (CSP) Program

From January 1992 to March 1993 an experiment to create community support for promotion of female primary schools was initiated in Zhob, Mekran and Naseerabad divisions of Balochistan. The purpose of the pilot program was to increase girls' primary enrollments by establishing segregated girls' primary schools taught by local female teachers. The program was based on a partnership between the government and the community. The government provided funding for the community school provided that the community supplied a temporary school facility, and female teacher from the community. A village education committee, composed of parents of daughters, was responsible for identifying the teacher, motivating parents to send daughters to the school, and monitoring the progress of the school, the children and the teacher. If the school was successfully operated for a probationary period, it was made a permanent government girls' school.

From the start, the program design was to accommodate parental preferences by using female teachers from the community. Because educated females are in short-supply in rural areas, the educational qualifications were relaxed relative to the standard requirements for a government teacher. Women qualified as potential teachers if they had, at minimum, eight years of schooling and were residents of the same village or lived within walking distance of the village. To make up for lack of educational background and teacher training, women were given a short course in teaching methods before the school opened. Teachers were also given in service training afterward. Those with educational deficiencies were required to make them up over time.

Theory of Enrollment Response to the Creation of CSP Girls' Schools

Before conducting the statistical comparison of the CSP and comparison groups, it is important to identify the possible endogenous responses to the program. It is also important to identify the exogenous variables that might condition the magnitude of those responses. Creating girls' school in a community will alter how parents allocate resources between boys' schooling, girls' schooling, and other consumer goods. To illustrate the parents' choice, we develop a model in which households are assumed to have parents, daughters and sons. Parents in this model are altruistic in the sense that they are willing to sacrifice their own consumption to invest in their children's schooling. Parents are assumed to derive utility from their own consumption of goods (Z) and from the human capital of their daughters (H_f) and their sons (H_m). The utility function has the form

(1) $U = U(Z, H_f, H_m, T)$,

where T is a vector of taste indicators that are not subject to choice.

Let Y be household income, P_z be the price of consumption goods, and P_f and P_m are the prices of schooling for their daughters and sons, respectively. The schooling price includes school fees, the cost of transportation, and the cost of materials. The income constraint on parental utility maximization is

(2) $P_zZ + P_fH_f + P_mH_m = Y.$

Now assume that schooling for girls is rationed to the amount of S_f, so that parents would invest more on girls' schooling if it were in sufficient supply. The supply constraint on parental utility maximization is

(3) $H_f \leq S_f.$

The parents allocate their household income to Z, H_m, and H_f in order to maximize utility given by (1), subject to $P_zZ + P_fH_f + P_mH_m = Y$ and $H_f \leq S_f$. The first order conditions are

(4) $U_m - \lambda Pm \leq 0$ or $H_m=0$

(5) $U_f - \lambda P_f - \mu \leq 0$ or $H_f=0$

(6) $U_z - \lambda P_z = 0$

(7) $Y - P_mH_m - P_fH_f - P_zZ = 0$

(8) $S_f - H_f \geq 0$,

where λ and μ are the Lagrangean multipliers associated with household budget and girls' school supply, respectively. If the constraints bind, the multipliers will be positive. When the constraint is binding (i.e., $\mu > 0$ and $H_f > 0$, and equation (5) holds with equality) and the prices for boys' and girls' schooling are the same, equations (4) and (5) imply $U_f > U_m$. Diminishing marginal utility would then imply that $H_m > H_f$ at the optimum.

The CSP program relaxes the restrictions on girls schooling.[6] The effect of the program on the girls schooling comes directly from total differentiation of equation (8) as follows:

(8)' $dH_f / dS_f = 1$.

This implies that a marginal increase on the rationed amount of girls' schooling will raise girls' enrollment by the same amount. Totally differentiating equations (4) through (7) and inserting $dH_f = dS_f$, we get the following comparative static results:

(9) $$\frac{dH_m}{dS_f} = \frac{U_{mz}P_fP_z + U_{fz}P_mP_z - U_{zz}P_mP_f - U_{mf}P_z^2}{|H|}$$

(10) $$\frac{dH_m}{dY} = \frac{-U_{mz}P_z + U_{zz}P_m}{|H|},$$

where $|H| = -2U_{mz}P_m + U_{zz}P_m^2 + U_{mm}P_z^2 < 0$. From equations (9) and (10), we get

(11) $$\frac{dH_m}{dS_f} = -P_f\frac{dH_m}{dY} + \frac{P_z(U_{fz}P_m - U_{mf}P_z)}{|H|}.$$

Assuming diminishing marginal utility, $U_{ii} < 0$ for i = m, f, and z. Hence, from equation (10), we can see the income effect depends on the sign and magnitudes of U_{mz} and U_{zz}. However, numerous studies suggest that education is a normal good so that equation (10) is positive. The effect of the program on boys' schooling is ambiguous. Equation (11) shows that the effect can be decomposed into an income effect and another term. It is apparent that the effect of relaxing rationing of girls' schooling imposes a cost on the household which can lower boys' schooling through the income effect. The program lowers net income available for other purposes including boys' schooling. However, we can see that under some circumstances, the creation of a girls' school can lead to an increase in boys' schooling as well. If this happens, we can conclude that boys' and girls' education are complements. Boys' and girls' education are more likely to be complements when the price of girls' schooling is low so that $P_f = 0$. In addition, they are more likely to be complements when $U_{mf} > 0$ and/or when $U_{fz} < 0$. The first condition implies that the marginal utility from boys' schooling increases at higher levels of girls' schooling. The second implies that the marginal utility of consumer goods decreases as girls' schooling rises.

Reduced form equations for boys' and girls' schooling have the following functional forms:

(12) $H_f = H_f(P_f, P_m, Y, P_z, T, S_f)$

(13) $H_m = H_m(P_m, P_f, Y, P_z, T, S_f)$

The reduced form equations suggest that enrollment will depend on school fees, income, the price level, tastes, and the available supply of girls' schooling.

Survey Design and Data Strategy

To evaluate the enrollment effect of the CSP schools, we required a comparison group of villages without CSP schools. A sample of villages was drawn from the 1990 Balochistan Human Resources Survey. This survey contained information on useful village attributes including the number and type of schools in the village, and the population of girls and boys. Villages of size comparable to the CSP village without a girls' school were taken to be potential comparison villages. A total of 30 villages in three divisions were selected by the Balochistan Education Management Information System (BEMIS). Three villages with CSP schools in each division were also drawn from lists of CSP schools.

For each CSP and comparison village, household information was collected by BEMIS in the same manner to insure comparability. The household survey information included socioeconomic characteristics of each household including parental education and occupation, age, gender, educational attainment and current enrollment status for all children in the household.

Differences and Similarities Between Treatment and Comparison Group

To begin the analysis, we investigate whether the comparison and CSP villages are statistically similar. Statistical properties are summarized in Table 1 for both treatment and comparison groups. Sample statistics are reported separately for boys and girls. The treatment sample included 355 children, 175 girls and 180 boys. The comparison sample included 1,023 children, 595 girls and 428 boys. Enrollment rates for both boys and girls are higher in the treatment than in the comparison group, a first-pass estimate of the impact of the CSP school effect. Nevertheless, there may be other factors responsible for the higher CSP village enrollments. The other variables in Table 1 are those believed to affect parental enrollment choices for their children. Children's age, father's education attainment, and birth order of the child in a family were taken directly from the survey data. There was information on mother's education attainment in the survey data, but almost all the mothers had never attended school. Lack of variation in mother's education led to exclusion from the analysis. Household income was generated from information on the number and educational attainment of adults in the household, land holdings, and other productive household assets. Details on the estimated measure of household income are contained in the Appendix.

We performed two different tests of the equivalence between the CSP and comparison villages. One is an equality test of means in characteristics between the treatment and the comparison groups. This test provides information on how closely the comparison villages match the treatment villages. The other test examines whether the behavioral coefficients in the enrollment choice model (based on equation (14)) are equal between the treatment and the comparison villages.

Tests of Equality of Means

The third and sixth columns of Table 1 report corrected t-values and degrees of freedom for hypothesis tests that the means of the variables are equal across the treatment and the comparison groups. The sample statistics for boys were statistically equivalent. For girls, fathers were more educated in the CSP villages. Birth order was also significantly larger in the CSP villages, implying somewhat larger numbers of children per household. Nevertheless, the joint test that the means were equal across all variables other than the enrollment rate was not rejected at the 0.05 level of significance. Based upon the results, we can reach a statistical conclusion that the treatment and comparison samples are drawn from the same universe of villages.

Test of Equality of Behavioral Coefficients in the Enrollment Choice Model

Parental decision-making regarding their children's education may differ between the treatment and comparison group. To check this, we estimate the following model of parental choice regarding their children's schooling:

(14) $R_i^* = \beta' X_i + U_i$

where $R_i = 1 \quad \text{if } R_i^* > 0$

$R_i = 0 \quad \text{if } R_i^* \leq 0$

In equation (14), an unobserved variable R_i^* depends on the index function, $\beta' X_i$, where X_i is the vector of characteristics in equations (12) and (13) which affect children's enrollment. When R_i^* is positive, we observe the child in school and $R_i=1$. Otherwise, the child will not enroll.

Table 2A reports the coefficients and z values of the probit analysis of enrollment choices for boys and girls. Separate estimates are presented for the treatment and the comparison groups. All of the estimates exhibit the same sign pattern of coefficients with the exception of the generally small and imprecisely estimated coefficients on birth order for boys and girls. The coefficient on household income is positive in both samples.

Parental education also positively influences children's enrollment. Enrollment increases with age, but at a diminishing rate.

Table 2B shows the result of the tests of equality of coefficients between the treatment and the comparison groups. The coefficients for the two groups are not statistically different, except for father's educational level in the enrollment equation for girls. In addition, the joint test of equality of behavioral coefficients across the CSP and comparison villages does not reject the null hypothesis of equality for the girls sample, while it was rejected for the boys' sample and the pooled sample. This result suggests that the patterns of the parental decision-making regarding their girl's education are identical in the treatment and the comparison groups. It should be emphasized, however, that in this ex post design, the tests of equality is not definitive. If the implementation of the CSP schools changed parental behavior, then ex ante identical groups will look different ex post. Therefore, the previous test of equality of time invariant household attributes is more relevant.

Evaluation Strategy

Accuracy of a quantified impact evaluation depends mainly on how well the comparison group was constructed. Since a child i cannot be simultaneously in both the treatment state (R_{1i}) and the non-treatment state (R_{0i}) of the CSP program, we cannot observe the true impact of the program $\alpha_i = R_{1i} - R_{0i}$. Instead, the observed outcome (R_i), can be expressed as $R_i = d_i R_{1i} + (1-d_i)R_{0i}$, where $d_i = 1$ if child i lives in a CSP village. Given the impossibility of observing the true impact of the fellowship program, the goal is to get an unbiased estimator of α_i.

One way to get an unbiased estimator of α_i is to use a comparison group to derive estimates of the counterfactual state. If we have a comparison group that mimics the treatment group very well, the comparison group is a good proxy for the counterfactual state of the CSP program. In this case, the expected program effect can be measured by the gap between the post program outcome in the treatment group and that in the comparison group. Mathematically, this is defined as

(15) Matched comparison: $E^M(\alpha_i|d_i=1) \ = E(R_T) - E(R_C)$

where subscripts T and C represent treatment group and comparison group, respectively. This method is called the ex-post matched comparison method.

Another way to get an unbiased estimator of α_i is to fit the following econometric model assuming the effect is invariant across individuals:

(16) Covariate post-test: $R_i = X_i\beta + d_i\alpha + U_i$

Equation (16) is the same as equation (14) except for the addition of a dummy variable indicating residence in a CSP village. Assuming X_i and d_i are independent of the unobserved variables U_i, so that $E(U_i \mid d_i, X_i) = 0$ for all i, we can estimate equation (16) using a cross-sectional data set.

Results

The first row of the Table 1 presents the enrollment rates of the treatment and the comparison group three years after the program intervention. Applying the ex-post matched comparison method given by equation (15), the measured CSP program effect was to increase girls' primary enrollment by 20.8 % and to increase boys' primary enrollment by 9.5 %.

The covariate post-test results reported in Table 3 are consistent with the results from the ex-post matched comparison method. The measured effect of the CSP program, based on equation (16), is a 21.8 % increase in girls' enrollment and a 12.9 % increase of boys' enrollment. The other estimated parameters exhibit the same sign patterns in the CSP and comparison enrollment choice equations for both girls and boys. The magnitudes are also comparable. Moreover, the coefficients are consistent with the results obtained in other studies of enrollment. The coefficient on household income is positive in both samples, so education of children is a normal good. Father's education level positively influences children's enrollment. Child age also positively affects enrollment choice, but at a decreasing rate. Peak enrollment for girls occurs 1.5 years below boys and drops off faster, so girls have smaller projected time in school. First-born children have a slightly higher probability of enrollment than their younger siblings, but the coefficient is not significant.

Income elasticities based on the coefficient of income in Table 3 and means of household income in Table 1 were 0.044 for boys and 0.049 for girls, which are highly inelastic.[7] This implies that income growth in rural areas does not guarantee an increase in primary enrollments. This also predicts that an alternative program, which aims to increase attendance through an income subsidy to households, will not be effective.

Conclusions and Extensions

There has been concern that the provision of a girls' school can boost girls schooling in the rural areas only if there is excess demand for girls' schooling. A pessimistic view is that cultural barriers or lack of parental interest about girls' schooling prevail in rural areas. Parents would not send their girls to school, even if more girls' schools are provided. However, this study shows that the creation of the CSP schools led to a substantial increase in attendance for rural girls. Although the reason for the success cannot be identified from available data, the use of parental participation and local female teachers are apparently critical to breaking cultural barriers to female schooling. An interesting side benefit of the program is that boys' enrollment also increased. The program effect on boys' enrollment is apparently due to an underlying complementarity between boys' and girls' schooling so that relaxing the constraint on girls' schools also raises incentives to send boys to school. The results were not sensitive to the methods used to measure the program effect based on different assumptions. All of the results suggest that expanding the CSP program to other villages is a promising strategy to raise rural schooling.

Future work of impact evaluation on CSP program will be required to assess the performance of the children attending CSP schools. The ultimate goal of the CSP program is to make children attain literacy. It would also be useful to compare educational outcomes in CSP schools to educational outcomes in traditional government schools. In particular, the role of local parental control may be an important factor in making all schools more effective and not just girls' schools.

Table 1 Summary Statistics of Datasets and Tests of the Equality of Means Between the CSP and Comparison Groups

Variable	Girls			Boys		
	CSP	**Comparison**	**t-value**	**CSP**	**Comparison**	**t-value**
Enrollment rate	0.623	0.415	5.546	0.761	0.666	2.844
	(0.486)	(0.493)	[768]	(0.428)	(0.472)	[606]
Household income	3197	5143	1.878	3675	6871	0.767
	(6476)	(8323)	[768]	(6552)	(14178)	[606]
Age	7.411	7.401	0.717	7.437	7.439	0.087
	(1.762)	(1.725)	[768]	(1.804)	(1.717)	[606]
Father's education	2.280	1.404	3.679	1.781	1.595	1.917
	(4.116)	(3.235)	[735]	(3.537)	(3.502)	[586]
Birth order	3.040	2.417	2.149	2.972	2.626	1.017
	(1.468)	(1.273)	[768]	(1.508)	(1.245)	[605]
Joint test			12.99			4.89
Number of observations	175	595		180	428	

Note: Critical value for joint test is 13.5.

Table 2A Probit Analysis of the Probability of Enrollment by Gender and Village Type

Variable	Girls and Boys		Girls		Boys	
	CSP	Comparison	CSP	Comparison	CSP	Comparison
Income/10000	0.269	0.111	0.212	0.115	0.324	0.146
	(1.04)	(3.43)	(0.87)	(1.11)	(1.08)	(4.46)
Age	0.782	1.471	0.501	1.296	1.171	1.724
	(13.5)	(7.79)	(3.29)	(6.13)	(3.80)	(31.0)
Age square	-0.043	-0.087	-0.029	-0.082	-0.064	-0.091
	(31.3)	(9.24)	(2.14)	(7.48)	(2.80)	(22.0)
Father's education	0.078	0.058	0.087	0.045	0.061	0.091
	(3.37)	(1.63)	(2.43)	(1.34)	(1.79)	(2.15)
Birth order	-0.053	-0.026	-0.084	0.007	0.003	-0.080
	(0.76)	(0.82)	(2.56)	(0.11)	(0.03)	(2.68)
Girl	-0.456	-0.663	–	–	–	–
	(1.64)	(14.3)				
Constant	-2.577	-5.536	-1.709	-5.204	-4.389	-7.018
	(3.88)	(4.96)	(3.35)	(3.88)	(3.64)	(14.0)
Number of Observations	353	972	175	562	178	410
Pseudo R^2	0.11	0.12	0.07	0.04	0.10	0.21

Table 2B Test of Equality of Coefficients Between CSP and Comparison Groups

Variable	Girls and Boys		Girls		Boys	
	χ^2	result	χ^2	result	χ^2	result
Income	0.38	not reject	0.15	not reject	0.51	not reject
Age	0.29	not reject	0.49	not reject	1.04	not reject
Age square	0.29	not reject	0.12	not reject	5.68	not reject
Father's education	5.72	not reject	14.36	reject	0.58	not reject
Birth order	0.01	not reject	0.86	not reject	1.49	not reject
Girl	0.47	not reject				
Joint test	7.83	reject	3.39	not reject	8.24	reject

Significance level: $\alpha = 0.05$

Note: The critical value for the joint test is 6.88.

Table 3 Post-test Probit Analysis of Probability of Enrollment Using Cross-sectional Data

Variable	Girls and Boys	Girls	Boys
CSP dummy	0.182	0.218	0.129
	(5.55)	(4.77)	(3.06)
Income/10000	0.050	0.048	0.052
	(2.84)	(1.95)	(2.15)
Age	0.495	0.428	0.498
	(5.88)	(3.92)	(4.23)
Age square	-0.029	-0.027	-0.026
	(5.20)	(3.74)	(3.38)
Father's education	0.025	0.023	0.026
	(5.58)	(3.97)	(3.75)
Birth order	-0.015	-0.010	-0.015
	(1.26)	(0.67)	(0.93)
Girl	-0.234	-	-
	(8.16)		
Number of Observations	1325	737	588
Pseudo R^2	0.13	0.06	0.18

Note: The coefficients reported here are dF/dX, where F is dependent variable and X is independent variable, not actual coefficients. Since the dependent variable is a discrete variable, dF/dX is not identical to actual coefficients.

Appendix

Estimation of household income requires information on household production, informal labor market arrangements, barter trade and other economic activities occurring outside formal markets. That type of information is difficult and costly to obtain. This project did not include sufficient resources to measure household income accurately. Instead, we utilized the Pakistan Integrated Household Survey (PIHS), which contained a detailed survey of household income and socioeconomic attributes in 1991. The PIHS allows us to predict household income based on a regression of income on easily observed household attributes. The current study collected information on these household attributes and then used the PIHS estimates to generate predicted incomes based on those attributes.

The PIHS income equation is reported in Table A. The specification follows that used by Alderman and Garcia (1996). That study estimated income and expenditure equations for 217 households in a single city in Balochistan. The predicted household income using Alderman-Garcia estimates performed well in explaining household savings, loans, and nutrition status in their study. However, the Alderman-Garcia estimates are less useful for our purpose because their data are from 1986 and some of the variables in their data do not match very well with the survey data for our current study. The variables in the income equation include the number of adult males and females, the number of males and females with primary, secondary and tertiary level schooling, and the value of household assets. As can be seen in Table A, households with more adult males, more human capital, and more capital assets have higher income in both the PIHS data and the data used in Alderman-Garcia's study.

Table A Income Equations

Variable	Alderman and Garcia	PIHS, rural
Intercept	5,999	- 777
	(2.61)	(-6.39)
Number of males	938	541
aged 16 or more	(0.92)	(4.37)
Number of males	1,691	a
aged 6-16	(2.09)	
Number of females	-709	315
aged 16 or more	(-0.54)	(2.59)
Number of females	1,009	a
aged 6-16	(0.64)	
Number of children	2,820	a
5 or below	(2.99)	
Number of males with	6,140	121
primary schooling	(2.95)	(0.69)
Number of males with	2,279	960
secondary schooling	(1.69)	(5.60)
Number of males with more	6,435	449
than secondary schooling	(1.41)	(4.65)
Number of females with	6,707	-427
primary schooling	(1.85)	(-1.88)
Number of females with	7,758	747
middle schooling or more	(1.35)	(5.87)
Value of land/1,000		1.24
		(10.97)
Rainfed land	110	a
	(2.34)	
Irrigated land	665	a
	(4.93)	
Orchards[b]	4,065	196
	(2.57)	(1.55)
Value of livestock	0.335	a
	(1.05)	
Cow[c]		62.1
		(1.88)
Camel[c]		955
		(2.19)
Donkey/horse[c]		9.45
		(0.04)
Goat/sheep[c]		11.8
		(0.65)
Value of vehicles	0.171	0.03
	(8.55)	(2.95)

Table A Income Equations (continued)

Variable	**Alderman and Garcia**	**PIHS, rural**
Value of machinery and tools	0.125	0.02
	(1.27)	(1.31)
R^2	0.747	0.430
N	217	894

[a]Not available in the current data

[b]Unit is acres for Alderman-Garcia, and value for PIHS

[c]Unit is the total number owned

Notes

[1] Iowa State University.
[2] World Bank.
[3] According to the Pakistan Integrated Household Survey (PIHS), 39 % of rural Balochistan boys who never attended school were withheld because they had to help at home. The corresponding figure was 23 % in urban Balochistan.
[4] Orazem, Paterno and Guttierez (1995) found that female teacher absenteeism decreases if teacher was single, did not have young childre, lived closer to the school, or was paid better.
[5] Statistics based on 1995-1996 Pakistan Integrated Household Survey.
[6] This is a special case of rationing in the sense that only one good is rationed. Tobin and Houthakker (1951) analyzed more general case of effects of rationing.
[7] According to Kim, Alderman and Orazem (1998), the corresponding elasticities were 0.115 for boys and 0.503 for girls in urban Balochistan.

GENERAL CONCLUSIONS

Two educational pilot projects for improving girls' enrollment were implemented in the urban and rural areas of Balochistan in Pakistan. This study evaluates those pilot projects, and provides quantitative evidence for judging whether or not to expand those projects. The most important finding is that low girls' enrollment was not due entirely to cultural barriers which caused parents to withhold their daughters from school. Both the urban and rural studies show that increasing girls' access to schooling can dramatically increase girls' enrollment.

In the urban fellowship study, it is shown that the fellowship program has positively affected enrollment for both boys and girls. The estimated impact was robust to different estimation methods. Most show that the effect was larger for girls' enrollment. The results of the urban fellowship experiment provide strong evidence that subsidizing the establishment of girls' private primary schools can lead to sharp increases in girls' enrollment. In addition, even though the fellowship was given only to girls, boys' enrollment in those neighborhoods also sharply increased. This suggests that the program may have created a low price for boys' schooling, or there also may have been excess demand for boys' primary education in these poor areas, or boys' and girls' education are complements.

The measured change over two years yielded mixed evidence on whether the enrollment growth advantage in fellowship neighborhoods over control neighborhoods continued to grow over time. However, even if the initial enrollment gain relative to control neighborhoods decreased in subsequent years, the enrollment gains after two years are still around 25 percentage points. School success appears not to depend on neighborhood income or other observable socioeconomic variables, suggesting that expansion of the program to other poor neighborhoods is also likely to be successful.

The rural CSP study also shows that the creation of a girls' school in a rural community leads to substantial increase in attendance for both boys and girls. The large growth in girls' enrollment supports the view that access to girls' schooling is rationed in rural areas. The program effect on boys' enrollment was smaller than the effect on girls, but the positive impact suggest that boys' and girls' education are complements. By exploiting this complementarity, the government can raise literacy for both boys and girls by expanding educational opportunities for girls. The results were not sensitive to the methods used to measure the program effect based on different assumptions. All of the results suggest that expanding the CSP program promise to be successful.

Future work will be required to assess each project's effects on school outcomes. The ultimate success of the primary educational projects depends on whether children attain literacy. It may be useful to assess children's performance before and after the program intervention between the treatment and control (or comparison) groups. Another way of doing this may be to compare the program outcomes among different types of educational projects, and to investigate the reason why they yield different outcomes.

Cost effectiveness of projects must also be assessed in future work. Since the resources for education projects are limited, it is important to find the most efficient way to raise enrollments and literacy for both boys and girls.

APPENDIX

ADDITIONAL TABLES

Table 1. Enrollment Rate by Year and Age (Urban)

Age	1994 Treatment	1994 Control	1995 Treatment	1995 Control	1996 Treatment	1996 Control
Boys						
3 years	3.9	0	a	3.0	a	a
4 years	11.5	2.6	12.0	2.4	a	4.5
5 years	22.6	28.1	32.9	23.9	48.7	12.5
6 years	50.6	55.3	56.5	41.2	73.3	35.0
7 years	67.4	59.3	78.9	62.1	89.3	60.3
8 years	77.7	69.7	85.6	69.4	89.0	71.8
9 years	76.4	68.3	90.7	78.8	88.7	71.9
10 years	87.1	a	90.1	74.7	87.4	72.7
Girls						
3 years	9.1	1.5	a	1.7	a	a
4 years	8.8	8.9	23.4	2.3	a	3.3
5 years	29.7	21.5	35.4	20.9	45.5	18.2
6 years	39.2	39.6	60.5	32.2	69.0	35.7
7 years	53.6	46.0	71.8	50.5	78.1	41.6
8 years	60.0	36.0	81.5	50.4	78.3	53.1
9 years	86.0	46.6	73.9	46.0	80.8	54.0
10 years	75.3	a	94.7	49.2	76.3	41.3

[a] Not available in the current data set.

Note: Table 1.shows the enrollment rate by year and age for both boys and girls. These numbers are used to calculate the program effect with either age specific and cohort specific sample in the context. We can get additional benefits in analysis from the above table. First, by applying the quasi-experimental method, we can see the change in enrollment due to the program is bigger in younger age group for both boys and girls. Second, we can calculate the program effect on enrollment for children excluded in the sample, who were aged less than four and more than eight. The effect was also bigger in younger age group, and the interpretation is substantially the same as in the context.

Table 2 Enrollment Rate Before and After the Program by Neighborhood (Urban)

Neighborhood Name	Treatment			Control		
	Before	After	Gap	Before	After	Gap
Boys						
Pashtoon Abad	46.9	100.0	53.1	46.5	36.5	-10.0
Brewery Road	44.4	85.7	41.3	77.1	59.0	-18.2
Almo/Shabo	44.4	58.3	13.9	52.2	43.1	- 9.1
Nawa Killi	43.2	72.7	29.5	33.3	17.9	-15.5
Samungali	42.0	100.0	58.0	27.8	20.0	- 7.8
Hudda	48.9	100.0	51.1	46.5	53.7	7.2
Marriabad	43.6	96.0	52.5	81.6	63.0	-18.6
Kechi Baig	29.2	64.3	35.1	55.6	49.0	- 6.6
Irrigation Colony	60.0	85.7	25.7	34.8	38.2	3.4
Killi Sheikhan	39.0	95.2	56.2	59.5	60.5	1.0
Girls						
Pashtoon Abad	55.2	94.1	38.9	33.3	27.3	- 6.0
Brewery Road	47.1	88.1	41.0	52.1	67.4	15.4
Almo/Shabo	42.4	33.3	- 9.0	28.9	41.7	12.8
Nawa Killi	55.2	62.5	7.3	4.4	6.0	1.6
Samungali	45.0	91.3	46.3	10.0	16.7	6.7
Hudda	47.4	100.0	53.6	35.3	42.9	7.6
Marriabad	42.9	100.0	57.1	64.2	69.2	5.0
Kechi Baig	30.9	64.0	33.1	29.8	28.1	- 1.7
Irrigation Colony	61.4	100.0	38.6	23.9	12.3	-11.6
Killi Sheikhan	50.0	97.3	47.3	58.3	57.5	- 0.8

Note: An attempt to calculate the neighborhood specific program effects has been made. Table 2 reports the change of average enrollment rate from 1994 to 1996 for both boys and girls. It shows that there exists considerable variation in enrollment growth of both genders across neighborhoods. For example, changes in control group enrollment rates for boys varied from 18.6 % decline in Marriabad to a 7.2 % increase in Hudda. Even larger differences in enrollment growth occur in the target neighborhoods. Girls' enrollment changed from a 9.0 % decline in Almo\Shabo to a 57.1 % increase in Marriabad.

Table 3. Correlation between Program Effect and Neighborhood Characteristics (Urban)

	Program Effect	Average Income	Average Mother's Education	Average Father's Education	Average Cost
Boys					
Pashtoon Abad	63.0 (3)	4	4	6	6
Brewery Road	59.4 (4)	3	1	7	4
Almo/Shabo	22.9 (9)	8	9	9	2
Nawa Killi	45.0 (6)	1	6	1	9
Samungali	65.8 (2)	7	9	10	1
Hudda	43.9 (7)	5	2	2	7
Marriabad	71.0 (1)	9	3	8	5
Kechi Baig	41.7 (8)	6	7	4	10
Irrigation Colony	22.3 (10)	10	5	3	8
Killi Sheikhan	55.3 (5)	2	8	5	3
Girls					
Pashtoon Abad	45.0 (5)	5	7	5	3
Brewery Road	25.7 (8)	3	6	9	6
Almo/Shabo	- 21.8 (10)	9	9	8	2
Nawa Killi	5.6 (9)	1	3	2	8
Samungali	39.6 (6)	6	9	10	1
Hudda	45.1 (4)	4	5	3	7
Marriabad	52.1 (1)	8	1	6	5
Kechi Baig	34.8 (7)	7	4	4	10
Irrigation Colony	50.3 (2)	10	2	1	9
Killi Sheikhan	48.2 (3)	2	8	7	4

Notes:
1. Numbers in the parenthesis represent ranking of the measured effect.
2. Measured program effect is subject to Quasi-experimental method (1994-1996)
3. Numbers other than program effect indicate rankings with ascending order. For example, average cost for boys and girls education is highest in the Samungali neighborhood.

Table 3 (Continued)

4. It is hard to pinpoint why the enrollment rate grew in some neighborhoods and declined in others. Table 3. report rank orderings of neighborhood averages of parent's education, income, and school cost (before the fellowship schools opened). Given there are only ten neighborhoods, this type of analysis is suggestive at best. Nevertheless, it is apparent that income is not critical - some of the poorest neighborhoods (Marriabad, Irrigation Colony) were among the most successful while some of the richest were not successful at all (Almo/ Shabo). Similarly, it is difficult to see a link between enrollment growth and average education or schooling costs in the neighborhood.

Table 4. Enrollment Rate by Age and Sex (Rural)

	Boys		**Girls**	
Age	**Treatment**	**Comparison**	**Treatment**	**Comparison**
5 years	51.4	26.9	57.9	28.7
6 years	75.0	48.9	54.6	39.4
7 years	80.0	75.7	60.0	48.5
8 years	78.6	79.1	74.4	47.5
9 years	100.0	88.9	42.9	59.7
10 years	85.4	85.9	66.7	33.9

Note: Table 4 shows the gap of enrollment rate between the treatment and comparison group is larger in younger age group for boys, but does vary not much for girls except age 9. It suggests that the program has positively affected the new entrance of boys, while it has positively affected enrollment for entire age group of girls. However, it remains as a puzzle why the effect was negative for girls aged 9.

REFERENCES

Alderman, Harold, Peter Orazem, and Elizabeth M. Paterno. 1996. "School Quality, School Cost, and the Public/Private School Choices of Low-Income Households in Pakistan." *The World Bank Working Paper Series on Impact Evaluation of Education Reform* =2.

Behrman, Jere. R., Robert A. Pollak, and Paul Taubman. 1995. *From Parent to Child.* The University of Chicago Press.

Boruch, Robert, John McSweeny, and John Soderstrom. 1978. "Randomized Field Experiments for Program Planning, Development, and Evaluation: An Illustrative Bibliography." *Evaluation Quarterly* 2(4):655-95.

Cook, Thomas D. and Donald T. Campbell. 1979. *Quasi-Experimentation.* Rand McNally College Publishing Company.

Deaton, Angus and John Muellbauer. 1980. *Economics and Consumer Behavior.* Cambridge University Press.

Hanushek, Eric A. 1995. "Interpreting Recent Research on Schooling in Developing Countries." *The World Bank Research Observer* 10(2):227-46.

Grossman, Jean Baldwin. 1994. "Evaluating Social Policies: Principles and US. Experience." *The World Bank Research Observer* 9(2):159-80.

Heckman, James J., and V. Joseph Hotz. 1989. "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training." *Journal of the American Statistical Association* 84(408):862-74.

Hoole, Francis W. 1978. *Evaluation Research and Development Activities.* Sage Publications.

Hsiao, Cheng. 1991. *Analysis of Panel Data.* Econometric Society Monographs No 11. Cambridge University Press.

Huber, Peter J. 1980. *Robust Statistics.* John Wiley & Sons.

LaLonde, Robert J. 1986. "Evaluating the Econmetric Evaluations of Training Programs with Experimental Data." *American Economic Review* 76(4):604-18.

_______. 1995. "The Promise of Public Sector-Sponsored Training Programs." *Journal of Economic Perspectives* 9(2):149-68.

Levitan, Sar A. 1992. *Evaluation of Federal Social Programs: An Uncertain Impact, Center for Social Policy Studies*. The George Washington University Press.

Manski, Charles F., and Irwin Garfinkel. 1992. *Evaluating Welfare and Training Programs*. Harvard University Press.

Newman, John, Laura Rawlings, and Paul Gertler. 1994. "Using Randomized Control Designs in Evaluating Social Sector Programs in Developing Countries." *The World Bank Research Observer* 9(2):181-201.

Rossi, Peter H., and Howard Freeman. 1993. *Evaluation: a Systematic Approach*. Sage Publications.

Schultz, T. Paul. 1995. *Investment in Women's Human Capital*. the University of Chicago Press.

Silberberg, Eugene. 1990. *The Structure of Economics: A Mathematical Analysis*. McGraw-Hill Publishing Company.

Tobin, James and H.S. Houthakker. 1951. "The Effects of Rationing on Demand Elsticities." *Review of Economic Studies* 18:1-14.

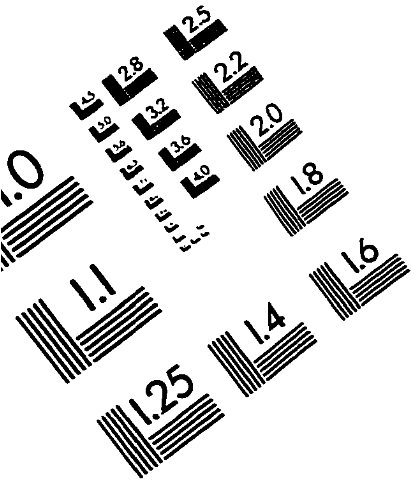
ACKNOWLEDGEMENTS

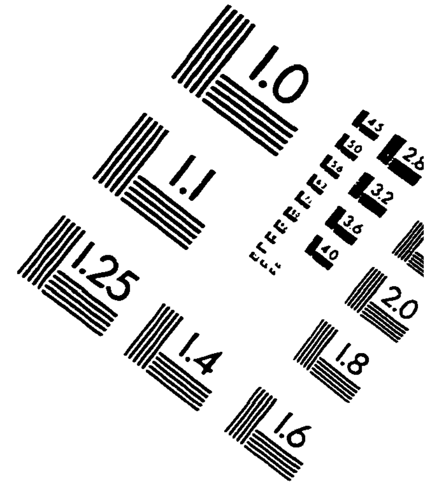
I acknowledge many who have helped in shaping this dissertation. First, I am especially indebted to my major professor, Dr. Peter F. Orazem, for his guidance, support, timely encouragement, and his positive and professional attitude throughout my graduate work. His advice and critical reactions have been vital to conduct these projects.

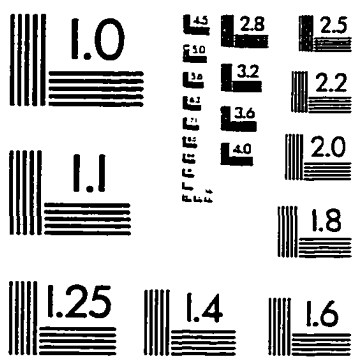
Sincere appreciation goes to my committee member, Dr. Peter J. Mattila, for his guidance providing an economist's view and support. To member of my committee I am grateful for their patience, professionalism, and availability.

I would like to dedicate this dissertation to my parents, Young-Kwan Kim and Su-Jung Chae, with my deepest thanks. Their endless love, trust, and encouragement in letting me pursue study abroad were remarkable. I must acknowledge my wife, Jung-Ah Lee, who has been willing to sacrifice for me. Without her love and support, this dissertation could not have finished.

IMAGE EVALUATION TEST TARGET (QA-3)

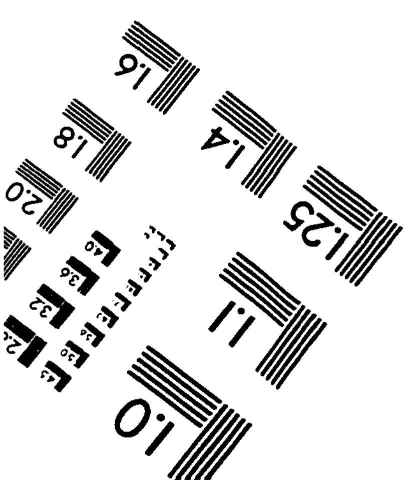


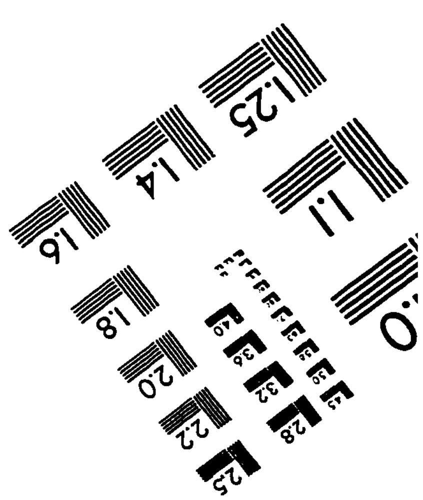




150mm

6"





APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989