IEG WORLD BANK | IFC | MIGA
Independent Evaluation Group

Do Conditional Cash Transfers Lead to Medium-Term Impacts?

EVIDENCE FROM A FEMALE SCHOOL STIPEND PROGRAM IN PAKISTAN

IEG WORLD BANK | IFC | MIGA
INDEPENDENT EVALUATION GROUP

Do Conditional Cash Transfers Lead to Medium-Term Impacts?

Evidence from a Female School Stipend Program in Pakistan

2011
The World Bank
Washington, D.C.

1818 H Street NW
Washington DC 20433
Internet: http://ieg.worldbankgroup.org
E-mail: ieg@worldbank.org

1 2 3 4 5 14 13 12 11

How to cite this report:
IEG (Independent Evaluation Group). 2011. *Do Conditional Cash Transfers Lead to Medium-Term Impacts?: Evidence from a Female School Stipend Program in Pakistan.* Washington, DC: World Bank.

Cover: Young girl reading; Pakistan. Photo by Curt Carnemark, courtesy of the World Bank Photo Library.

ISBN-13: 978-1-60244-175-0
ISBN-10: 1-60244-175-8

Independent Evaluation Group
Strategy, Communication, and Learning
E-mail: ieg@worldbank.org
Telephone: 202-458-4497
Facsimile: 202-522-3125

Contents

Tables

Figures

Abbreviations

CCT	Conditional Cash Transfer
DD	Difference-in-differences
EDO	Executive District Officer for Education
FSSP	Female School Stipend Program
GSP	Girls' Stipend Program
MICS	Multiple Indicator Cluster Survey
PESRP	Punjab Education Sector Reform Program
PIHS	Pakistan Integrated Household Survey
PMIU	Program Monitoring and Implementation Unit
PSLM	Pakistan Social and Living Standard Measurement
RDD	Regression Discontinuity Design

Acknowledgments

This impact evaluation was prepared by a core team led by Javier Baez, and consisting of Ximena del Carpio, Andaleeb Alam, and Tu Chi Nguyen. The team was assisted by Diana Hakobyan and Jesse Torrence. Helen Chin edited the report for publication. We are grateful to World Bank operational staff who provided assistance, especially Mariam Qaiser, Dhushyanth Raju, and Sofia Shakil. We also thank government agencies and research centers in Islamabad and Lahore, Pakistan, especially Muhammad Asif and Sohail Malik (Program Monitoring and Implementation Unit), Feyza Bhatti (Mahbub Ul Haq Human Development Center), Minhaj ul Haque (Population Council), and Shamim Rafique (Bureau of Statistics – Planning and Development Department). Many thanks go to others inside and outside of IEG and the World Bank who provided helpful comments, especially Andrew Warner and Martha Ainsworth.

The work was conducted under the general guidance of Cheryl Gray (Director) and Mark Sundberg (Manager). The team is grateful for the extensive and excellent advice provided by peer reviewers John Hoddinott (Senior Research Fellow, International Food Policy Research Institute) and Emmanuel Skoufias (Lead Economist, PRMPR, World Bank

).

Executive Summary

The Punjab Female School Stipend Program (FSSP), a targeted conditional cash transfer program in Pakistan, was implemented within the context of a larger education sector reform and in response to gender gaps in education. Difference-in-differences and regression discontinuity analyses show that, four years into program implementation, adolescent girls in stipend districts are more likely to progress through and complete middle school and work less. Although less significant in a statistical sense, there is also some suggestive evidence that participant girls delay marriage by more than a year, and have fewer births by the time they are 19. In addition, girls who are exposed to the program later on and are eligible for the benefits given in high school also increase their rates of matriculation into and completion of high school grade levels. While these are only early outcomes, they have the potential to translate into significant gains in future productivity, consumption, fertility, and intergenerational human capital accumulation. There is no evidence that the FSSP has adverse indirect effects on similar educational outcomes of boys living in the same household with eligible girls.

Conditional cash transfer (CCT) programs have become more prevalent in recent years with the expectations that they would reduce short-term poverty and encourage children of poor and vulnerable households to accumulate human capital critical for their future growth. Impact evaluations have investigated these programs, and the majority appears to show positive short-term effects on the behaviors on which they are conditioned. Despite differences in design and implementation context, most CCTs increase immediate consumption, encourage the use of educational and health services, and reduce child labor. However, little is known about whether these impacts actually translate into longer-term effects on children's accumulation of human capital and their future employment and income.

Objectives of this Evaluation

In an effort to fill the gap in evidence on the medium- and long-term effects of CCTs, this report evaluates one such program at public girls' schools in Pakistan, the Punjab Female School Stipend Program (FSSP). In particular, the report attempts to answer three questions: (1) What are the program's impacts on outcomes that might affect future productivity of adolescent girls who have been participating in the intervention for up to four years, including sustained school enrollment, transition and completion of middle school and high school, early labor market outcomes, and marriage and fertility decisions?, (2) Does the program affect different beneficiary subgroups differently?, and (3) Are there positive or negative indirect effects of the program on boys who live in the same household with eligible girls?

The Punjab Educational Context and the FSSP

Punjab Province has the highest welfare indicators in Pakistan, but these are still among the lowest in South Asia and some of the lowest among countries of a similar development level. There is a shortage of middle schools and high schools, and many of those that do exist are inaccessible to many households, which prevents children in rural

villages from enrolling at these grade levels. Moreover, there are large disparities in educational opportunities and attainment between girls and boys as well as across districts.

In response to the issues above and in alignment with the national decentralization of social service delivery, the government of Punjab launched the Punjab Education Sector Reform, which aimed to address issues in the educational system on both the supply and demand sides. On the supply side, it aimed to improve financial capacity at the provincial and district levels, increase educational spending, enhance school infrastructure, distribute free textbooks, recruit new teachers, and strengthen school councils, among other things. These interventions are accounted for in the statistical analysis, but not investigated in depth in this evaluation.

As part of the demand-side interventions, the FSSP was designed to improve educational attainment among girls and decrease gender gaps at the middle school level (grades 6–8). The program targeted the 15 districts with the lowest literacy rates to address their educational disadvantages. Launched at the end of 2003, the program provides quarterly subsidies (around $10) to the families of girls enrolled in middle school, with the condition that they attend at least 80 percent of classes. In 2006, the stipend was extended to girls enrolled in high school (grades 9–10). By 2007, 245,000 girls were covered by the program in middle school.

Overview of Previous Evidence

Evidence from other country experiences show that gender-targeted CCTs help reduce the gender disparities in educational opportunities by raising school attendance, enrollment, and the school progression of girls. CCTs have also been shown to help delay marriages and childbirth for adolescent girls. Evidence from an impact evaluation undertaken two years after the program started indicates that the program increased female enrollment in middle school by around 10 percent between 2003 and 2005.

Empirical Research Design

Because the FSSP was not randomly allocated, this evaluation relies on quasi-experimental techniques to construct comparable treatment and control cohorts at the school and household levels. At the school level, the annual Punjab public school censuses from 2003 to 2009 are used to create treatment cohorts of girls who were enrolled in stipend schools during the program. These girls should therefore be exposed to the program for at least one year, and they are compared against the similar group of girls in nonstipend districts during this period to estimate whether the program impacts on enrollment are sustained.

At the household level, the analysis uses two waves of the Punjab Multiple Indicator Cluster Survey (MICS), in 2003 and 2007–08, to construct cohorts of girls in stipend districts who are likely to have at least one year of exposure to the program. These cohorts are compared with counterfactual cohorts in nonstipend districts in terms of the likelihood of completing middle school, transitioning to and completing high school, as well as their labor force participation and marriage and fertility decisions.

This empirical research design is implemented with two econometric models: difference-in-differences (comparing changes between baseline and follow-up, across treatment

and control groups) and regression discontinuity design (using the cutoff rate of literacy for district eligibility). These techniques help address possible selection bias due to time-fixed differences between stipend and nonstipend districts induced by the nonrandom placement of the program. In addition, the econometric models include covariates that account for socioeconomic differences between stipend and nonstipend districts, including differences in school facilities, access, and local infrastructure. The analysis also includes models based on samples that that do not condition on enrollment in middle school to check the sensitivity of the findings to the potential issue of selection in participation.

Findings on Program Impacts

After four years of implementation, the FSSP is found to help adolescent girls progress through and complete middle school, reduce labor participation, delay marriage, and have fewer children.

- The increase in enrollment identified in the short-term impact evaluation is sustained across different grades in middle school. All cohorts that progress from grades 6 to 8 (between 2003–04 and 2009–10) show significant increases in the number of girls enrolled, ranging from 11 to 32 percent. These effects appear to be largely driven by reductions in dropout rates.
- Girls in stipend districts are also 3–6 percentage points more likely to complete middle school. This is equivalent to an increase of 4.5–10 percent relative to the baseline value of the treatment group. Younger cohorts who are exposed to the program later and are eligible for the expansion of the stipend to high school also benefit in terms of a greater likelihood of transitioning to high school (by 4–6 percentage points) and completing at least one high-school grade (by 5 percentage points). Impacts of similar magnitude have been found for comparable CCTs in Brazil and Honduras.
- Labor force participation of adolescent girls in stipend districts falls by 4–5 percentage points (nearly half of the participation at baseline), as compared with those in nonstipend areas. This reduction is largely driven by a drop in girls' participation in unpaid family work. In addition, the younger cohorts work at least 2–3 hours less per day as a result of the program.
- Girls in stipend districts age 15 to 19 tend to delay marriage by 1.4 years and have 0.3 fewer children (marginally significant in a statistical sense).

The distribution of impacts of the FSSP appears to favor girls who live in urban centers, in poorer households, and with more-educated parents.

- Girls who live in urban centers and have parents with at least a primary education benefit more in terms of middle-school completion, but less in terms of a reduction in labor force participation.
- The reduction in labor participation accrued by girls in poorer households is more than twice the average impact.
- The length of exposure seems to have an increased marginal benefit on the completion of high school grades and delayed marriage.

While there is no indication that the FSSP has indirect effects on the educational outcomes of boys living with eligible girls, the program seems to divert these boys to private schools at the primary level.

- The school enrollment and completion rates of boys with female siblings participating in the FSSP are not affected by the program.
- However, boys in participating households are 4–5 percentage points more likely to be enrolled in private schools, especially at the primary level. The share of these schools is growing, they are affordable (relative to the value of the stipend), and they are of higher quality relative to public schools. While such diversion is beneficial for the learning of boys, it might widen the learning gap between these boys and eligible girls who are increasing their enrollment in public schools of lower quality.

Robustness Analysis

In order to check the validity of the impact estimates derived from the quasi-experimental methods used, some robustness analyses were conducted.

- Based on prebaseline and baseline data, there is no indication that the outcomes evolved differently before the FSSP began. Therefore, the evidence does not suggest that the findings of the paper are driven by preprogram differences in trends.
- Measurement error on exposure to treatment could bias the outcomes toward zero. If this is the case, the results presented above may thus be interpreted as lower bounds of the true impacts.
- It is expected that there is an endogenous compositional change due to the self-selection of lower-ability students into stipend schools, but this bias is likely negative and the findings would still provide a lower bound of the actual impacts. Besides, there is no sign that there are compositional changes due to shifts in age-grade distribution, migration, or crossover effects between contiguous stipend and nonstipend districts.
- The model at the household level includes both public and private schools, so program impacts already account for a potential switch from private to public schools caused by the program.
- The findings of the paper do not appear to be driven by systematic differences in school supply between stipend and nonstipend schools or the incidence of positive or negative income shocks.

Discussion

The positive impacts of the FSSP on educational attainment suggest that the program may have important implications for future productivity and welfare of beneficiaries. For instance, women in Punjab that complete middle and high school live in households that enjoy up to 30 percent higher consumption per capita, relative to women with less than middle school education. One possible reason is that more educated women are able to marry men who have almost twice as much education as the husbands of women with less education. Another explanation is that higher schooling may enable Pakistani women to increase their own earnings by as much as 150 percent. This estimated annual income increase more than compensates for the annual stipend cost of the program.

Finally, the impacts of the FSSP may have further dynamic effects on other dimensions. Women in Punjab with middle and high school education have around 1.8 fewer children than those with less education by the end of their reproductive life. It is also estimated that the 1.4-year delay in marriage attributed to the program may lead to 0.4 fewer births by the end of the women's childbearing years. Furthermore, the evidence also shows that these women not only have fewer children but also invest more in their human capital, which may lead to positive intergenerational effects.

However, the findings of this analysis are still a partial picture, and further research is needed to investigate other relevant aspects of the FSSP impacts, such as:

- Impacts on the complete outcomes, such as educational attainment, labor force participation, marriage, and fertility by the end of their working and childbearing years.
- Whether the impacts on schooling are translated into improvements in cognitive development (that is, test scores).
- How different amounts of stipend could affect the outcomes differently, and more broadly, whether the program benefits offset the costs.
- The effects of various supply-side educational interventions and how they may complement the impacts of the FSSP.
- Factors within the implementation process and context that may also influence the program impacts.

1. Introduction

Motivation and Objectives

Conditional cash transfer (CCT) programs are becoming a popular tool for alleviating short-term poverty and reducing the inter-generational transmission of poverty. More than 30 developing and transition countries have implemented these programs, providing incentives to poor households to make investments in the human capital of their children. Programs vary in scale, transfer size, conditionalities, eligibility, and implementation features. Despite these differences, evidence arising from impact evaluations indicates that most of these programs fulfill their short-term objectives of increasing immediate consumption and use of educational and health services. Indeed, a recent review of impact evaluations of safety net programs shows that all of the CCTs evaluated increased the consumption of participant households, and 87 percent raised the school enrollment and attendance of their children (World Bank 2010).

However, much less is known about the medium- and long-term impacts of CCTs on human development outcomes. Most evaluations to date only assess the changes in school attendance and enrollment and visits to health centers—the conditions with which the beneficiaries must comply in order to receive a transfer. While these investments are critical for the accumulation of human capital in children from poor households, such changes indicate little about whether the accumulation is indeed taking place and if there is improvement in the future prospects of these children (that is, children progressing through school, completing school, learning more, becoming more productive and employable, and achieving higher earnings in the future). In fact, few studies examine the impacts along some of these subsequent parts of the causal chain.[1] Likewise, the evidence is thin regarding indirect effects on outcomes, such as marriage and fertility decisions, which have plausible implications in the long term on the welfare of the beneficiaries.

This report contributes to the understanding of the medium-term impacts of CCTs through an impact evaluation of a female school CCT program in Pakistan, focusing on the identification of program impacts, their distribution across different groups of individuals, and potential spillover effects. The program investigated is the Female School Stipend Program (FSSP) —a gender-targeted CCT implemented in the province of Punjab in northeast Pakistan. This study examines the program's impacts on school progression, middle- and high-school completion, early labor market outcomes, and marriage and fertility decisions of adolescents who have been participating in the intervention for up to four years.[2] In addition, the report looks at the distribution of impacts across various participant groups and who could benefit differently from the program. Finally, the evaluation explores the spillover effects of giving money to girls (conditioned upon their enrollment and attendance in middle/secondary school) on the enrollment of boys.

After four years of implementation, the FSSP is found to help girls in beneficiary districts progress through and complete middle school, which could help reduce gender gaps in schooling. Furthermore, the younger cohorts of participants, in particular, are also more likely to enroll in and complete at least one high school grade. This matches with the expectation that the younger cohorts join the program when it is more mature and stay in the program longer, indicating some marginal benefits of the length of exposure. Another hypothesis is that these girls also benefit from the expansion of the program (started in 2006) to include stipends for high school enrollment. In terms of labor force participation and marriage and fertility decisions, the program encourages girls — particularly those who live in poorer households — to work less, delay marriage, and have fewer children.

This report is structured around five chapters as follows: The first chapter gives an overview of the program, the context in which it was implemented, and available evidence on the impacts of the FSSP as well as other CCTs. Chapter two focuses on the questions this evaluation sets out to answer and the methods and information used to answer them. The third chapter presents the results from the analysis and is structured around three evaluation questions regarding average impact, heterogeneity of impacts, and spillover effects. Chapter four performs the robustness checks of the findings, examining whether they are sensitive to preprogram trends, measurement error, endogenous compositional changes, and crowding-out effects. Finally, the conclusion discusses the implications of the results, some limitations of this evaluation, and areas that require further work.

Overview of Previous Evidence

The last decade has seen a substantial increase in the prevalence of CCTs and, with it, a parallel increase in impact evaluations that assess their effectiveness. The wave of CCTs and evaluations started with the Mexican *Oportunidades* program, which has been operating since 1997. The program is a CCT that transfers cash to poor rural families, conditional on their children's regular attendance at school and visits to health centers for periodic checkups and growth monitoring. Evaluations of the program, based on a randomized design and multiple rounds of panel data, show significant positive impacts on short-term outcomes: school enrollment and attainment increased, child labor decreased, immediate consumption grew, and some health outcomes improved (Gertler 2000; Hoddinott and others 2000; Behrman and others 2001; Schultz 2004). These assessments were corroborated by subsequent evaluations of similar interventions in Brazil, Cambodia, Colombia, El Salvador, Honduras, Malawi, Nicaragua, Pakistan, and Turkey. For the most part, the findings are consistent with the results of the *Oportunidades* program regarding positive impacts on the use of educational and health services and short-term consumption (Olinto and others 2003; Glewwe and Olinto 2004; Attanasio and Mesnard 2005; Attanasio, Fitzsimmons, and Gomez 2005; Attanasio, Gomez, Heredia, and Vera-Hernandez 2005; Filmer and Schady 2006; Maluccio and Flores 2005; Ahmed and others 2007; Chaudhury and Parajuli 2008; Baird, McIntosh and Ozler 2009).

In addition to increases in school attendance and enrollment, grade progression also appears to improve as a result of CCTs. Often using data that span two years, evaluations of various programs identify positive effects on progression rates across grades. For instance, de Janvry and others (2006) shows that the *Bolsa Escola* program in Brazil increases grade advancement by 6 percentage points; an experimental analysis of the Honduras *Programa de Asignacion Familiar* – Phase II (PRAF II) finds that the probability of matriculating to subsequent grades increases by 2–4 percentage points among beneficiary children. Comparable effects have been documented for CCT programs in Mexico, Nicaragua, and Paraguay (Behrman and Hoddinott 2001; Maluccio and Flores 2005; Veras-Soares and others 2008).

The literature also provides evidence on the indirect effects of conditional transfers on labor market outcomes and on marriage and fertility decisions of adolescents. The influence that CCTs may have on these outcomes is expected to operate mostly through the conditionalities that require children to attend school regularly. In fact, rigorous evaluations of programs in Brazil, Cambodia, Colombia, Mexico, and Nicaragua confirm that children in stipend districts not only spend more time in school but also have relatively lower participation in income-generating activities and domestic work (de Janvry and others 2006; Ferreira and others 2009; IFS, Econometria, and SEI 2006; Behrman and others 2005a; Maluccio and Flores 2005). This effect is also observed among adolescent females. For example, the *Oportunidades* program in Mexico lowers the participation of young girls in labor markets and child care by around 10 percent and 40 percent, respectively (Behrman and others 2005a; Dubois and Rubio-Codina 2009). Furthermore, it has been shown that CCTs have led to delayed marriage and onset of sexual activity in girls in Malawi and Mexico, reduction in the number of sexual partners in Malawi, and reductions in pregnancy rates in Malawi and Turkey (Baird, Chirwa, McIntosh and Ozler, 2009; Behrman and others 2005a; Gulemetova-Swan 2009; Ahmed and others 2007).

Evaluations of the CCTs designed to explicitly address gender disparities in investments in human capital show that they have succeeded in increasing the school enrollment and attendance of girls. The gender gap in educational opportunities often plays out to the disadvantage of girls, due to cultural and labor market preferences. At least four interventions that were set up to address these issues show positive impacts in the short term. The Bangladesh Female Secondary Stipend Program raises the enrollment of girls in grades 6–10 by 8 to 12 percent (Khandker and others 2003). Similarly, the Zomba Cash Transfer experiment in Malawi more than doubles the reenrollment of girls between 13 and 22 years old who had dropped out of school (Baird, McIntosh and Ozler 2009). Another example is the Japan Fund for Poverty Reduction scholarship program in Cambodia, which has a large positive effect (around 30 percent) on the school enrollment and attendance of girls (Filmer and Schady 2006). A short-term impact evaluation of the FSSP in Pakistan —the program assessed in this report — also shows a 9 percentage point increase in girls' enrollment in stipend districts relative to nonstipend districts, between 2003 and 2005 (Chaudhury and Parajuli 2008). Although the impacts differ by magnitude, and are sometimes measured differently, this result appears to be consistent across programs and implementation context.

Nonetheless, the evidence remains thin regarding the connection between additional use of school inputs and the human development outcomes of adolescent girls in the medium and long term. This study assesses whether the girls who are more likely to enroll in school due to the FSSP in Pakistan also show improved school progression and completion, reduced participation in labor activities, and differences in the timing of marriage and child bearing up to four years after their first exposure to the program.

Educational Context of Punjab

Punjab is the wealthiest and most populous province in Pakistan, with average welfare indicators, such as literacy, sanitation, health, and income, higher than other provinces in Pakistan but lower than other countries in South Asia and countries at similar development levels. Divided into 36 districts, Punjab is home to over 55 percent of the total Pakistani population. For example, the literacy rate in Punjab in 2003 was approximately 54 percent, compared with 45 percent for Pakistan as a whole, 56 percent for South Asia and 63 percent for all low-income countries (Lloyd 2004; estimates from other sources).[3]

School enrollment, particularly in secondary education, is largely constrained by the scarcity of schools. The primary level is better supplied by both the public and private sectors and is becoming more affordable and accessible. The secondary level, in contrast, is dominated by the public sector, and not every village has access to middle schools (grades 6 to 8) or high schools (grades 9 and 10).[4] For instance, whereas 84 percent of the households in Punjab report that their children are within 15 minutes' walk to a primary school, only 55 percent report access to a middle or high school within the same travel time (Pakistan Social and Living Standard Measurement, or PSLM, 2004). Access is even lower in rural areas, making it less likely that children in rural villages enroll in schools at these levels (Andrabi and others 2006; Sathar and others 2003). In 2003, before the program started, while enrollment in primary school for children ages 6–10 was 58 percent, enrollment at the middle school level for children ages 11–14 was 29 percentage points lower (MICS 2003).

Education levels of females are low, both in absolute numbers and as compared with boys, especially in rural areas. At baseline, girls in program areas were less likely than boys to be enrolled in school. While enrollments in primary and middle school were very similar for boys and girls in nonstipend districts, there were large differences in stipend districts — up to 8 percentage points (MICS 2003). Gaps in enrollment, by gender, are even wider when the samples are restricted to the rural areas of Punjab. In terms of school attainment, in 2001 (before the program) urban women in Pakistan achieved, on average, 5.7 years of schooling — about 1.4 years less than men, and rural women only 1.8 years of schooling, 3.2 fewer years than men. Disparities in education, by gender, are also evident for Punjab. As shown in table 1.1, the levels of enrollment and educational attainment are systematically lower for women in rural and urban Punjab.

Table 1.1. Educational Outcomes at Baseline

	School enrollment (percent)						*Years of schooling*[a]					
	Primary (age 6–10)			*Middle (age 11–14)*			*Rural*		*Urban*		*Overall*	
	Boys	*Girls*	*Overall*	*Boys*	*Girls*	*Overall*	*Men*[b]	*Women*[b]	*Men*[b]	*Women*[b]	*Men*[b]	*Women*[b]
Stipend	50.8	42.0	46.7	24.2	15.9	20.3	4.5	1.8	6.8	5.4	5.0	2.6
Nonstipend	67.0	66.0	66.5	34.6	34.2	34.4	6.3	4.0	7.8	7.2	6.9	5.2
Punjab	59.5	55.1	57.4	30.2	26.9	28.6	5.5	3.0	7.5	6.8	6.1	4.2
Pakistan[c]	56.3	44.4	50.5	30.6	21.9	26.4	5.0	1.8	7.1	5.9	5.7	3.1

Source: MICS 2003.
Note: Those who have never attended school are assumed to have zero years of schooling. Professional degrees (law, medicine, engineering, business) are assumed to be equivalent to 16 years of education. Those whose highest grade completed was madrassa level or other are coded as missing and are therefore excluded.
a. Years of schooling is proxied by highest grade completed.
b. For men and women 15–40 years old.
c. Data source: PIHS 2001

Within Punjab, there are also marked differences in school enrollment and access between districts covered and not covered by the FSSP. Fifteen districts in Punjab with literacy rates below 40 percent were targeted by the stipend program. Enrollment levels in these districts were well below those of the 19 nonstipend districts. In 2003, for instance, enrollment in primary school for children ages 6 to 10 was 47 percent in districts selected for the stipend, as compared with 67 percent in nonstipend districts; similarly, only 20 percent of children 11–14 years old were in middle school, compared with 34 percent in nonstipend districts (table 1.1). As shown in table 1.2, access to school was also more difficult in stipend districts before the program began, where higher proportions of households live in villages without a school and, therefore, their children had to travel longer distances to go to school (MICS 2003; PSLM 2004) .

Table 1.2. Access to School at Baseline

	Access to Primary School[a]			*Access to Middle School*[a]			*Access to High School*[a]		
	Rural	Urban	Overall	Rural	Urban	Overall	Rural	Urban	Overall
Stipend	88.1%	96.6%	89.7%	45.6%	84.8%	53.0%	26.2%	77.6%	35.9%
Nonstipend	96.2%	97.1%	96.5%	65.8%	89.7%	74.7%	46.5%	83.4%	60.3%
Punjab	92.3%	97.0%	93.7%	56.2%	88.4%	65.8%	36.8%	82.0%	50.2%

Source: PSLM 2004.
a. Access is defined as household travel time less than 30 minutes.

Background of the Female School Stipend Program

MOTIVATION FOR THE PROGRAM

In response to major national reforms for decentralization of social service delivery responsibilities, as well as recommendations from assessments undertaken in the 1990s, the provincial government of Punjab spearheaded its own educational reform in 2003. The Punjab Education Sector Reform Program (PESRP)—partly financed with World Bank lending in the form of multiple development policy credits—aims to address flaws identified in the educational system, including lack of infrastructure or substandard quality of infrastructure, which limited access to schools; bad quality of

service delivery and financial management; limited capacity of staff at all levels of the system; weak community and parental involvement; and a rise in the private sector provision of schooling with weak oversight for standards and quality. The reform is articulated through three major pillars with the following goals: (1) to improve the financial capacity to increase public spending for education and ensure its sustainability; (2) to devolve resources from the provincial capital to the districts in order to strengthen their financial autonomy, capacity, and accountability, which will improve educational service delivery; and (3) to improve access to services for all while ensuring equitable provision to boys and girls, and to improve the quality and governance of the education sector.[5]

The FSSP was one of several interventions launched as part of the PESRP, with a specific objective to improve educational attainment among girls and decrease gender inequities, particularly at the middle school level. Prior to the start of the program, female enrollment in primary and secondary schools was low, both in absolute terms and relative to that of boys.[6] Additionally, low enrollment for girls is accompanied by low retention and completion rates (Sawada and Lokshin 2009). A large body of analytical work in Pakistan identifies various community characteristics (access to roads, transport, share of teachers residing in the village where the school is located, scarcity and poor quality of schools for girls) and household characteristics (poverty status of the family, school costs, parental education and occupation) that are associated with such gaps in school entry for girls (Khan 1993; Sathar and Lloyd 1994; Lloyd and others 2005; Das and others 2006; Lloyd and others 2007).[7] Based on this evidence and on recommendations made by various experts, the provincial government of Punjab mobilized resources in 2003 to finance the FSSP to target a clearly disadvantaged group (girls) in districts with lagging literacy rates. None of the ineligible districts received the FSSP, but all were still eligible for other interventions sponsored by the reform, such as the distribution of free textbooks, school rehabilitation, and new teachers, among other things.

This evaluation accounts for all other programs under the PESRP that were not province-wide and were operating in parallel with the FSSP, in order to control for their possible effects. Parts of the activities undertaken to improve educational outcomes under the reform program were either phased in throughout the province, over the years, based on need or were implemented only in places where demand was most salient.[8] In this respect, the evaluation accounts for three components that could potentially affect the outcome of the FSSP: school improvements,[9] the recruitment and allocation of new (probably better qualified) teachers, and the phased distribution of free textbooks to all primary school children.

PROGRAM STRUCTURE AND IMPLEMENTATION

The main goal of the FSSP was to promote the participation of girls in public education at middle school level (grades 6 through 8).[10] Program benefits comprise a quarterly subsidy of approximately 600 Pakistan rupees (equivalent to $10) per female student.[11] Beneficiary girls were targeted based on their district of residence (districts with the lowest literacy rates in the province, at below 40 percent), and enrollment in

eligible grades (6 through 8) in public schools. Eligibility was conditional on a minimum school attendance rate of 80 percent, as reported regularly by the school.

The FSSP started in late 2003 and covered 15 eligible districts.[12] Although planning for the FSSP began in early 2003, stipends were not distributed until the fourth quarter of 2003, with compensation for attendance in the third quarter of 2003, as reported and accounted by the monitoring unit. During the first quarter of the program in 2003, about 156,000 girls received the stipend. In 2006, the FSSP was extended to include high school (grades 9 and 10). The impact of this extension is not the main focus of this report. The FSSP is still active in Punjab; in 2007 it covered 245,000 beneficiaries in middle school, which corresponds to a participation rate of 51 percent among eligible girls.[13] According to the actual quarterly stipend data, the program spent around $7.3 million in 2007 on stipends for grades 6 through 8.[14]

Box 1.1. Implementation and Monitoring of the FSSP

The implementation of the FSSP from the provincial level to the district level, to the schools and beneficiaries, is monitored regularly. Very few noncompliance issues have been found. All public schools in stipend districts maintain a class attendance roster kept by the head teachers. Every quarter the teachers submit the list of all eligible students, based on the attendance rates, to the Executive District Office for Education (EDO), which sends the stipend request to the Program Monitoring and Implementation Unit (PMIU). Once the EDO receives the money, it transfers the funds to post offices corresponding to the respective localities, which are responsible for disbursing the stipend. Almost half (46 percent) of the stipends are disbursed directly to recipient households either by postal carrier or a postal agent (a shopkeeper), while the rest are disbursed through schools. In both cases, the recipient signs a receipt and this information is submitted, through the EDO, to the PMIU. It often takes 10–14 weeks from the time the lists of eligible students are submitted, to the time the beneficiaries receive their stipends. In 2005, a validation conducted by a third party found that bribery or withholding of stipends was not an issue. A delay in delivery by the post office to remote areas was identified but was addressed by the PMIU (by channeling the disbursement through schools). The data on the number of girls receiving stipends and funds released are reported by the PMIU on its Web site. Furthermore, the PMIU has been undertaking monthly censuses of schools since November 2006, partly to monitor the stipend delivery. For instance, in September 2007, of the 4,600 middle/high/higher secondary girls' schools surveyed, only 24 reported that stipends had either not been issued, issued but not delivered, not delivered, or not delivered fully. More details on the implementation data collection process are described in appendix B.

Source: PESRP Web site.

Previous Impact Evaluation of the Female School Stipend Program

An earlier impact evaluation, conducted within the first two years of the program, had to take into account the fact that the FSSP was targeted toward and implemented in the 15 districts with the lowest literacy rates. It therefore compared changes in the number of girls enrolled in public middle schools from 2003 (before the program started) to 2005 (two years into the program), between stipend and nonstipend districts. The difference in the trajectories of female enrollments between the two groups of

schools is interpreted as the impact of the FSSP.[15] As an additional control, the study contrasted this difference with the difference between the enrollment trajectories of boys in public middle schools in stipend and nonstipend districts (boys were not eligible for stipends in any of the districts).[16] The analysis drew primarily on the provincial school censuses of 2003 and 2005, which contain information on enrollments, by grade, in public schools (schools are the units of analysis). The evaluation used these data and the rule that determines program eligibility (the cutoff at a 40 percent literacy rate) to estimate the net growth in enrollments of girls in grades 6–8 (middle school), in stipend public schools relative to nonstipend public schools, and that difference relative to enrollments of boys in the same districts (Chaudhury and Parajuli 2008).[17]

The short-term evaluation found that the FSSP increased the number of girls enrolled in grades 6–8, on average, by 9–23 percentage points, depending on the specifications.[18] Even though the magnitude of program impact estimates varies across different models, impacts appear to be consistently positive and statistically significant. Their triangulation results—based on individual-level data from two rounds of household surveys that are representative at the provincial level and collected before and after the program—yielded similar results.

2. Evaluation Focus and Empirical Approach

Evaluation Questions

The main focus of this evaluation is to estimate the medium-term impacts of the FSSP on human development outcomes that could influence the economic and human capital of girls in the long term. By raising the school enrollment and attendance of girls, gender-targeted CCTs are expected to help close gender gaps in education and, over time, may increase the school attainment, productivity, and employability of girls. Despite these potential dynamic benefits, most programs are evaluated only against the first part of that causal chain, namely their short-term effects on enrollment and attendance. This study seeks to contribute to the understanding of the later steps in the causal chain by addressing the following three questions.

1. ***Average medium-term impacts:*** Does the program improve the educational attainment, early labor market outcomes, and marriage and fertility behavior of beneficiaries? Previous impact evaluation results indicate that girls participating in the program were more likely to be enrolled in middle school (Chaudhury and Parajuli 2008). This study moves one step further along the causal chain to assess whether participant girls are also more likely to progress from grades 6 to 10 and to complete school, and whether they work less while attending school.[19] Finally, considering the social and cultural norms that govern marriage and fertility decisions in Pakistan, this evaluation examines the extent to which increased school participation may lead to delayed marriage and childbirth among beneficiaries.
2. ***Heterogeneity of impacts:*** Do different subgroups of program participants benefit differently? This study tests whether program effects differ for girls with relatively more educated parents, for girls in rural areas (as compared with urban areas) and, for girls in poorer districts. Additionally, this evaluation explores if younger cohorts of girls show higher impacts (possibly due to stronger pulling effects, for example, if more girls become aware of the program or the program improves over time). Knowledge about the distribution of impacts across different participant groups is critical to maximize benefits and efficiency through more effective targeting.

Indirect effects: Does the program have positive or negative indirect effects on the educational outcomes of school-age boys within the eligible households? On the one hand, the transfers and increased school attendance of girls may ease household budget constraints and motivate their parents to keep boys in school. On the other hand, the positive effects of the program on female enrollment may reduce the enrollment and attendance of boys if they have to substitute for girls' time devoted to income-generating activities, household chores, and child caring. This evaluation estimates the net impacts of these opposite effects on school choices, progression, and completion by boys.

Empirical Research Design

Because the FSSP was not randomly allocated, a simple comparison of girls in stipend and nonstipend districts does not accurately identify the causal effect of the program on the outcomes of interest. Due mainly to budgetary constraints, the FSSP was rolled out only in the neediest districts of Punjab (those with literacy rates below 40 percent). In total, 15 out of 34 districts, most of which are located in southern Punjab, were eligible for the program (see appendix A). This targeting may pose a selection problem because stipend districts were chosen based on characteristics (possibly both observed and unobserved) that may be correlated with low educational attainment. Moreover, there are two other issues that could make it difficult to identify the causal effect of the stipend program in the absence of a reliable counterfactual. One is participation bias, namely families that decide to enroll girls in public middle school, and therefore participate in the program, may be different from families with girls of relevant age but not in public school. The other issue is measuring effects on outcomes that are conditioned on the enrollment of girls (school progression and completion). In this case, the samples for analysis may suffer from selection bias if they include lower-ability girls in stipend areas that would not have attended school in the absence of the program.[20]

This evaluation relies on secondary data and quasi-experimental techniques to estimate the impacts of the program. Because the program design and implementation did not plan for an impact evaluation, no specific data were collected for the purpose of evaluation based on a counterfactual framework. This study attempts to overcome some of the issues discussed above with the use of cross-sectional data at the school and household levels, from 2003 through 2009, and quasi-experimental methods. More specifically, the analysis contrasts cohorts of girls in treatment and control districts in difference-in-differences (DD) and regression discontinuity design (RDD) framework.[21] While DD compares the changes in the outcomes of analysis between stipend and nonstipend districts over time, RDD exploits the literacy rate cutoff for program eligibility across districts. The RDD models are also implemented within a DD framework. Therefore, both models compare the trends in stipend and nonstipend districts, which helps address possible selection bias introduced by time-fixed characteristics related to nonrandom program placement (such as access to schools and educational facilities). In order to check the sensitivity of the findings to the issue of participation bias, additional models were estimated with samples that do not condition on enrollment in middle school.[22] The validity of the identifying assumptions of the DD models and the implications on the expected selection on returns to schooling are discussed in chapter 4.

This evaluation constructs treatment and comparison groups of girls at both the school and household levels. At the *school level*, the analysis draws upon administrative data from school censuses to estimate the annual average changes in female enrollment at public schools in stipend districts (treatment schools) for the period between 2004 and 2010, relative to the preprogram level (2003–04 academic year). These changes are then compared with the corresponding annual changes in girls' enrollment for public schools in nonstipend districts (control schools) during the same period. This approach is limited, however, by the fact that schools are the units of analysis, and it is not possible

to tease out other behavioral responses regarding the investments in the human capital of girls.

To overcome this limitation, the *household-level* approach uses household surveys to estimate the disaggregated changes at the individual level. It uses the same identification strategies and assumptions as the school-level analysis. However, the richness of information of the household-level data, as compared with school censuses, allows the analysis to take advantage of the extra variation at the individual level and the better accuracy in measuring girls' exposure to the program. The outcomes of interest at this level include relative changes between baseline (2003) and follow-up (2007–08) in school attainment (middle school completion and high school enrollment and completion), labor market participation, marriage, and childbirth. The comparison is made between cohorts of girls with various levels of exposure to the program in stipend districts (treatment cohorts) and similar cohorts in nonstipend districts (control cohorts).[23] For the outcomes of boys in eligible households, the impacts are estimated by comparing boys in households that have girls exposed to the program in stipend districts with comparable boys in nonstipend districts (in a DD analysis). This difference is then compared with boys in households in stipend districts that do not have girls that are exposed to the program (in a triple difference analysis). Program impact estimates at the school and household levels should be interpreted as impacts on all eligible girls, whether or not they participated in the program (intent-to-treat impacts).[24]

Box 2.1. Quasi-experimental Evaluation Techniques

Impact evaluation is based on the comparison of a particular outcome in situations with an intervention and those without. Since it is not possible empirically to observe the two situations for the same households or individuals simultaneously, a counterfactual similar to the treatment group needs to be constructed. A counterfactual is naturally created when treatment is randomly allocated among eligible beneficiaries because, although the characteristics of individual observations are very different, variances cancel each other out over a large enough sample, and the treatment and control groups are generally similar on average. When randomization is not possible, a quasi-experimental method needs to be used. One method is **difference-in-differences**, which compares the average change in outcomes across a certain period of time for the treatment group (the first "difference") with the average change in the same period of time for the control group. This technique is based on the assumption that the treatment and control groups would have progressed similarly over time without the intervention. Another method is the **regression discontinuity design**, which takes advantage of the cutoff score or selection criteria on which the treatment assignment is based. It assumes that those just above and below the cut-off point have similar characteristics, and a comparison of their outcomes could provide an estimate of the program's local impact around the cutoff.

Source: Authors' compilation.

SCHOOL-LEVEL ANALYSIS

In the school-level analysis, the treatment and comparison cohorts of girls are constructed synthetically based on their enrollment in different grades. The treatment cohorts were enrolled in public middle school in stipend districts, and therefore should have been exposed to the program for at least one year between baseline (2003–04) and subsequent academic years until 2009–10. Based on these criteria, seven distinct cohorts,

A–G, can be identified (table 2.1). The counterfactuals consist of the corresponding cohorts enrolled in public schools in nonstipend districts. For example, cohort C consists of girls in grade 6 in stipend districts in 2004–05 who are expected to progress to grades 7 through 10 in the succeeding years between 2005–06 and 2008–09. If their progress through school is consistent, they should have been exposed to the program for all three years of middle school (grades 6 to 8) between 2004–05 and 2006–07, as well as to the expansion of the stipends to girls in high school (grades 9 and 10), which started in 2006–07. Girls in corresponding cohorts in nonstipend districts should not have been exposed.

Table 2.1. Construction of Cohorts for School-Level Analysis

	Grade 6	*Grade 7*	*Grade 8*	*Grade 9*	*Grade 10*
2003–2004 (baseline)	Cohort B	Cohort A			
2004–2005 (first year of middle school stipend)	Cohort C	Cohort B	Cohort A		
2005–2006	Cohort D	Cohort C	Cohort B	Cohort A	
2006–2007 (first year of high school stipend)	Cohort E	Cohort D	Cohort C	Cohort B	Cohort A
2007–2008	Cohort F	Cohort E	Cohort D	Cohort C	Cohort B
2008–2009	Cohort G	Cohort F	Cohort E	Cohort D	Cohort C
2009–2010		Cohort G	Cohort F	Cohort E	Cohort D

The school-level analysis measures whether the increase in enrollment experienced by a cohort when they first joined the program (the short-term impact) is continued through successive grades, relative to the trend observed for the comparison cohort. A sustained increase in net enrollment could be due to increased transition across grades (the push effect), a reduction in drop-out rates, and/or an influx of new entrants (including both rejoiners and those who switch schools). But, the school-level approach cannot distinguish these channels of impacts because the school censuses do not provide information to measure them separately. However, preprogram and postprogram data from household surveys indicate that the number of rejoiners (new entrants) and grade repeaters at the middle school level is very low.[25] This suggests that net changes in girls' enrollment (if they occur) are largely driven by the combined effects of the program on both transition and dropout rates.

HOUSEHOLD-LEVEL ANALYSIS

The units of observation used for this analysis consist of girls expected to have been exposed to the program for at least one year between 2003–04 and 2007–08. Although the household survey data used for this analysis provide rich information on a wide variety of indicators, they were not collected with the idea of evaluating the impact of the FSSP. The data therefore do not contain explicit information on the treatment status and length of exposure of girls. Therefore, some steps were taken to select the girls who fall within an age range and are likely to have been exposed to the program for at least one year.[26] Appendix C explains more fully the construction of the cohorts for analysis, and addresses the assumptions and concerns underpinning the construction. Girls are included in the treatment cohorts if they live in the stipend districts and in the comparison cohorts if they live in nonstipend districts. In addition, both groups have to

meet one of the following criteria to ensure that the analysis is restricted to the relevant girls, namely those who could enroll in middle school for at least one year during the period of study:

- Girls who were enrolled in grades 7–12 in 2007–08 and/or grades 6–11 in the previous year (2006–07) are assumed to have been exposed for at least one year to the program.
- For the rest of the girls who were enrolled in neither 2007–08 nor 2006–07, only those whose highest grade completed is, at least, grade 6 and, at most, grade 10, and who are likely to have been enrolled in a middle school grade between 2003–04 and 2005–06 (based on the age-grade distribution of the reference cohort at baseline) are assumed to have been exposed for at least one year to the program.

The household data are used to estimate the influence of the program on such medium-term outcomes as educational attainment and labor market, marriage, and fertility decisions. A possible caveat to this analysis is that the girls studied were still young (12 to 19 years old) at the time of the follow-up survey. This means that many had probably not reached their full educational attainment and were not yet married or had children. However, this age range is still relevant for the analysis of program effects on early marriage and childbearing, both of which are prevalent issues for young women in Pakistan.

Data Sources and Samples

This impact evaluation uses various data sources collected at different times in Punjab, representative at different levels, and focusing on different dimensions of the educational process.[27]

SCHOOL-LEVEL ANALYSIS

The main data source used is the Punjab public school census, collected annually in October, from 2003 to 2010, by the provincial education department, including from before and during the program implementation. This survey is part of the institutional strengthening capacity effort within the PESRP and is used for monitoring resource allocation and performance in the sector. It covers all public schools[28] in Punjab and focuses on school information (including village name and date of establishment), infrastructure, teaching staff, enrollment, and some organizational indicators. The enrollment data are reported by gender of the school's students and separated into different grade levels. Schools are also disaggregated by urban and rural areas. The total number of schools is 60,000–62,000, for the period from 2003–04 to 2009–10. The sample used for this analysis is a pool of school-level, cross-sectional data that includes approximately 4,000 middle schools for girls in stipend and nonstipend districts. These are the schools that could be eligible for the program at both baseline and follow-up.[29]

In order to extract contextual information and construct covariates, this analysis uses several household surveys. The household surveys include the Pakistan Social and Living Standard Measurement survey (PSLM), collected in 2004–05; the Punjab Multiple

Cluster Survey (MICS), collected in 2003; and the Pakistan Integrated Household Survey (PIHS), collected in 2001. They use multitopic questionnaires focusing on a comprehensive set of social and economic indicators as well as information on infrastructure supply and quality, thus allowing for construction of some covariates at the district level. These surveys are representative at the district level and disaggregated by urban and rural areas. It is important to note that, although lower-level data provide more precise estimates, the questionnaires are significantly shorter; therefore, some key economic welfare indicators are excluded.

Other sources include educational censuses and administrative data that provide information on the supply of educational services and other programs. This evaluation uses the National Education Census, collected in 2005, on both public and private schools to construct additional covariates. It also relies on the administrative data of other elements of the educational reform implemented in Punjab at the same time as the FSSP, such as the distribution of free textbooks, recruitment of new teachers, construction of school facilities, and improvement of school councils.

HOUSEHOLD-LEVEL ANALYSIS

The main data source for the estimation of impacts at the individual level is two waves of the MICS collected in 2003 and 2007–08. The household survey covers more than 40 social and economic indicators, all repeated in both rounds for comparability. The sampling design follows the sampling frame of the previous population census (1998) and provides estimates for the province, for urban and rural areas, and for each of the districts in Punjab and the canton of Lahore. Sampling for 2003 was stratified by rural, urban, and large city areas. In the follow-up survey in 2007–08 the sample size was vastly enlarged to increase the precision and make it representative at the *tehsil* level (the level below the district). The total number of surveyed households is 30,932 in 2003 and 91,280 in 2007–08. For the empirical models of this paper, the analysis uses pooled cross-sectional data, comprised of girls aged 12–19 selected based on the criteria described above.

The surveys used to construct covariates for the school-level analysis are again used for the household-level analysis. This includes the PSLM, which relies on a similar sampling strategy as the MICS. Most of the classifications and definitions used throughout their documentation resemble those of the population census. Other data come from the PIHS, the National Education Census, administrative data of other programs under the PESRP, and the annual Punjab public school census collected in 2003.

Outcome Variables and Covariates

The medium-term impacts of the program are assessed on the following outcomes:

- *Schooling indicators associated with female enrollment, progression, and completion.* The school-level analysis looks at (1) **enrollment** beyond the initial grades and years by measuring the percentage change in the number of girls enrolled in grades 6 to 10 for each academic year during the period from 2004–05 to 2009–10,

as compared with baseline. The household-level analysis seeks to gauge impacts of the program on the actual accumulation of human capital by measuring changes in: (2) **middle-school completion** — the probability that a girl completes grade 8 provided that it is possible for her to progress to that grade based on her age and grade at baseline; (3) **transition from middle to high school** measured by the enrollments in grade 9 among girls enrolled in grade 8 the year before; and (4) **high-school grade completion** — the probability of completing either grade 9 or 10.

- *Behaviors related to girls' participation in the labor market*. In particular, household data are used to examine the effects of the program on: (1) **labor force participation** — probability of looking for a job and working, including unpaid family work; and (2) **work intensity** — number of days worked per month for girls working at the time of the survey.
- *Decisions of adolescent girls on marriage and childbirth.* These outcomes are also measured at the individual level, and include (1) **marriage decision** — the probability of being married and the age at marriage (for girls reported to be legally married);[30] and (2) **early childbirth,** as proxied by the probability of giving birth and number of births during the three years before the survey.

Given the targeting of the program to low-literacy districts, stipend and nonstipend districts are likely to differ in aspects other than just the treatment, and a baseline comparison confirms this. Because literacy rates are likely correlated with other underlying socioeconomic characteristics, such factors may, in turn, be associated with different education attainment and divergent trajectories in human capital accumulation between households in stipend and nonstipend districts. Indeed, at the community level, households in stipend districts are more likely to be below the poverty line, to be located in a rural village, and to spend a little less on education. The stipend districts also have less access to public transport and to schools of different levels (access is defined as less than a distance of 2 kilometers or travel time of 15 minutes) as well as fewer private middle schools relative to public schools. At the school level, schools in stipend districts tend to have worse facilities, but the differences are very small, except for electricity. However, stipend schools exhibit more favorable student-teacher ratios. Finally, at the household level, households in stipend districts appear to have more dependents and be headed by individuals with approximately 1.6 fewer years of education. They are also less likely to have dwellings connected to gas and electrical services, and to have access to family planning and health centers (appendix E, table E.1).

These differences have obvious implications for the strategy to isolate the impact of the program. In order to deal with potential biases which may arise due to these differences, this paper includes econometric specifications that control for other factors at the community, school, and household levels, which could contribute to the same outcomes, as well as age-fixed effects, for all of the empirical models. Control variables that may be affected by the stipend program are added to the models using their mean values at baseline. Furthermore, econometric models in the analysis also add binary indicators to control for the plausible effects of other programs related to improving the education system within the PESRP. The covariates used in the models include:

- *Community-level variables*: (1) **access to schools** (distance to primary, middle, and high schools, and access to public transport),[31] (2) **school supply across education levels** (ratio of primary to middle schools and middle to high schools), (3) **share of the private sector** in total enrollment (private/public enrollment ratio), (4) **dropout rates** at different education levels, (5) **poverty** (poverty headcount ratio), and (6) **share of education in household expenditures**.
- *School-level variables*: (1) **proxy of school quality** (student/teacher ratio),[32] (2) **school facilities** (electricity, drinking water, toilets, and boundary walls).
- *Household-level variables*: (1) **access to basic services** (healthcare, water, electricity and gas), (2) **location** (rural or urban), (3) **household structure** (household size and dependency ratio), (4) **birth order**, and (5) **parental education** (education of household head).
- *Other programs*: (1) distribution of free textbooks, (2) recruitment of contract teachers, (3) construction of school facilities, and (4) strengthening of local school councils.

3. Impacts of the Program

This chapter presents the results from the analyses of the medium-term impacts of the FSSP. The program impacts are estimated based on DD and RDD models implemented with data at the school and household levels. This analysis uses preprogram data (2003–04) and data covering the program implementation period from 2004–05 to 2009–10, for the school-level analysis, and data from baseline (2003) and follow-up (2007–08) surveys for the household-level analysis. This chapter is divided into three sections based on the three evaluative questions. First, the analysis examines if the early effects of the FSSP, in the form of increased female school enrollment, translate into such medium-term impacts as improvements in school completion and labor market, marriage, and fertility decisions.[33] The presentation of the results follows the natural sequence of the causal chain investigated in the paper, namely from the use of school inputs to school progression, then to labor market, marriage, and fertility outcomes. The next section disaggregates the average impacts and presents evidence on the distribution of program impacts across different groups of girls. This evidence will introduce more nuances to the discussion of the impacts of the programs, helping to understand who among the beneficiaries may benefit more.[34] The dimensions investigated include age (young versus old), location (urban versus rural), income (above versus below the median district consumption level), and education of parents. Finally, the analysis examines the spillover effects of the FSSP on school choice, enrollment, progression, and completion by boys who live in households with eligible female siblings.

In general, the evidence shows that after four years of implementation:

- Girls benefiting from the FSSP are more likely to progress through and complete middle school as well as reduce their labor force participation. They also tend to delay marriage and have fewer children by the time they are 19. Farther along the educational path, at the high school level, the FSSP appears to have impacts on enrollment and completion mostly among the younger cohorts, probably because they are exposed to benefits longer, join the program when it is more mature, and are eligible for program benefits at the high school level.
- Many of the positive impacts described above accrue mostly to girls in urban areas and with more educated parents. The poorest beneficiaries also tend to respond more in terms of labor force participation and marriage decisions. Finally, the effects on high school grade completion and marriage decisions increase with the duration of program benefits.
- While there is no indication that the educational outcomes of boys living with eligible girls are affected, they tend to be sent more often to private primary schools, which are growing in number, less expensive, and of higher quality relative to public schools. This may increase gender disparities in learning as girls who benefit from the FSSP start to enroll more in public schools of lower quality.

Average Medium-Term Impacts

IMPACTS ON EDUCATIONAL OUTCOMES

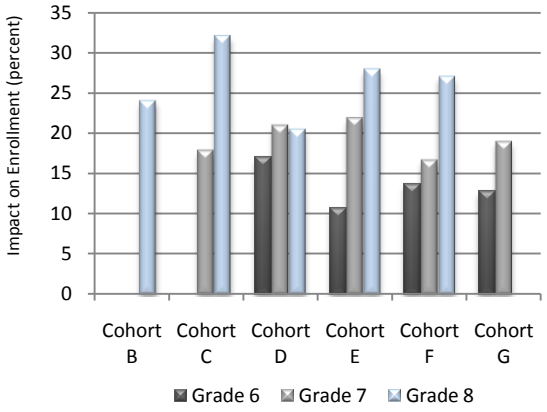
In theory, the effects of a CCT like the FSSP on grade progression, school completion, and educational attainment are ambiguous. The cash transfer has an income effect. This is the only effect relevant for families that would have sent their daughters to middle school, with or without the program. At the same time, for families with daughters out of school, the program's condition on school attendance also induces a substitution effect because it reduces the price of education. Both effects point toward an increase in the demand for education, particularly among credit-constrained households in environments with imperfect capital markets. On the one hand, if this demand is sustained over time, the higher utilization of education services can lead to higher school progression and completion.[35] On the other hand, there might be perverse incentives or constraints that prevent program participants from accumulating more human capital. For example, families may try to keep their daughters in those grades that make them eligible for the cash transfers, in effect preventing them from progressing and graduating. Moreover, assuming no changes in supply, schools in program areas may become overcrowded due to the additional demand, which can hinder progression and completion. The following empirical exercises seek to shed some light on the net effects of these conflicting forces within the context of the FSSP.

Effects on Enrollment

The analysis based on school-level data provides systematic evidence that the program helps girls progress through middle school. The point estimates from DD and RDD models indicate that the program raises enrollment not only in grade 6 but also in grades 7 and 8 for the same cohorts. For instance, cohort C sees higher percentage changes in enrollment in stipend than in nonstipend districts, in both grade 7 in 2005–06 and grade 8 in 2006–07 (the second and third year of exposure to the program). The increases relative to enrollment in stipend versus nonstipend are 18 and 32 percent, respectively. Although the relevant cohorts (cohorts B to G) participate in the program at different times and for different periods, the estimates of the program effects on enrollment, as seen in the matrix, yield consistent results for all of these cohorts and grades over time (positive and significant impacts). The insignificant impacts in 2004–05 are expected because awareness of the program was not widespread during its first year.[36] For subsequent cohorts, program impacts are still large and statistically significant, ranging from 11 to 32 percent (figure 3.1).[37] It seems that the initial push effect in enrollment due to the program at the beginning of middle school persists over the length of exposure (generally three years between 2004–05 and 2009–10).[38] Impacts of comparable magnitude have been found in similar female-targeted CCTs in Bangladesh, Cambodia, and Malawi (Khandker and others 2003; Baird, McIntosh and Ozler2009; Filmer and Schady 2006).

Figure 3.1. Program Impacts on Enrollment through Middle School (percent)

	Grade 6	*Grade 7*	*Grade 8*
2004–05	–0.3	4.3	6.1
	(C)	(B)	(A)
2005–06	10.6**	17.8***	24.0***
	(D)	(C)	(B)
2006–07	16.9***	21.8***	32.0***
	(E)	(D)	(C)
2007–08	13.6***	20.9***	27.9***
	(F)	(E)	(D)
2008–09	12.7***	16.6***	20.5***
	(G)	(F)	(E)
2009–10	15.6***	18.9***	27.0***
	(H)	(G)	(F)



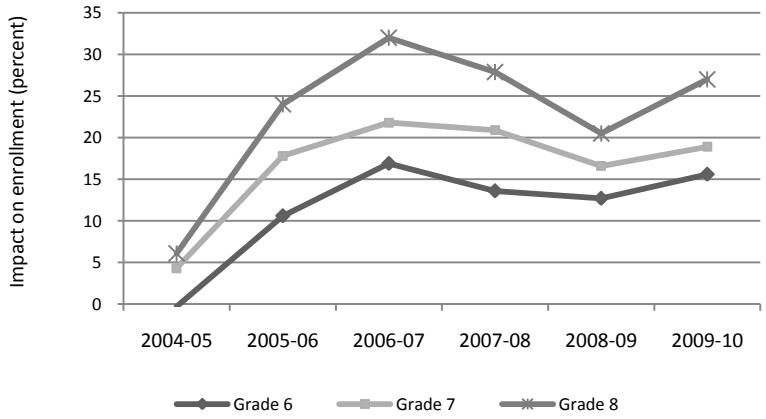
Source: Authors' calculations.
Note: Estimates are from parametric RDD (whole sample). Cohorts are shown in parentheses.
** Significant at the 5 percent level. *** Significant at the 1percent level.

In addition to the consistently positive impacts on female enrollment over time, program impacts appear to increase with school grades. For instance, while enrollment in grade 6 for cohort D in 2005–06 (their first year of participation in the program) increased, on average, by about 11 percentage points in stipend relative to nonstipend districts, in the next two years the same cohort experienced increases of 22 and 28 percentage points in enrollment grades 7 and 8, respectively. Several factors may explain this increasing trend. First, the program might have given incentives to girls who had left the education system to reenter school. However, this does not seem to explain the changes for middle school grades because reentry is not only very low (less than 0.5 percent) but also similar across stipend and nonstipend districts. Second, and more plausible, the program may have differential effects in retaining girls in schools, perhaps more so in the higher grades. Dropout rates in Punjab (as in most developing settings) increase with grades and age. In fact, a simple DD analysis between 2003–04 and 2007–08 using household survey data shows a decrease of 2–3 percentage points in drop-out rates in middle and high school grades in stipend districts, relative to nonstipend districts, while the DD estimates for ineligible grades are similar between stipend and nonstipend districts. Finally, the expansion of the FSSP in 2006–07 to cover grades 9 and 10 may provide additional incentives for girls to finish middle school and transition to high school.

Program impacts on female enrollment are sustained only up to the end of middle school, except when the program is extended to include high school. As noted before, the findings show that the increases in enrollment attributed to the program remain as long as girls are enrolled in middle school. However, such effects are not continued to grades 9 and 10. Point estimates of the effects on enrollment in grade 9 only become positive and statistically significant for some econometric specifications in the academic years 2007–08 and 2009–10, when the expansion of the FSSP to high school was fully implemented. This may suggest that this second phase of the program has positive impacts on enrollment in high school.

The analysis also reveals positive evidence regarding program sustainability. Evaluators and policy makers may also be interested in learning the extent to which the program continues to function beyond its initial implementation cycle, as determined largely by the availability of resources, willingness to participate in the program, and the social and political environment. In general, the findings discussed above suggest that the program indeed remains effective over time in increasing girls' school participation. Figure 3.2 illustrates this point with the increasing impacts on enrollment during the entire period of analysis for grades 6 through 8. Furthermore, the size of the impacts is in line with the estimates of about 9–23 percentage points found in the short-term evaluation conducted in 2006–07 (Chaudhury and Parajuli 2006).

Figure 3.2. Effects of the Program on Enrollment in Grades 6 to 8



Source: Authors' calculations.
Note: Estimates are from parametric RDD (whole sample). All results are significant at least at the 5 percent level, except for 2004–05.

Middle School Completion

In addition to the positive effects of the program on school progression, the evidence suggests that the program also increases the chances that girls will complete middle school. The analysis above shows that girls' enrollment increased over time in stipend districts compared with nonstipend districts. The next question would then be whether girls who benefit from the program are also relatively more likely to complete middle school. Impact estimates based on household data appear to validate the findings obtained with school data. Overall, they suggest that the 12–17 year-old girls who were exposed to the program (for different periods) between 2003–04 and 2007–08 are, on average, around 3.3 percentage points more likely to have completed middle school (grade 8), relative to the control groups. This is equivalent to an increase of 4.5 percent relative to the baseline value of the treatment group. The impacts on middle school completion are larger for girls aged 15–16 (6 percentage points with 5 percent significance) but none for girls older than 16 (table 3.1).[39] Similar impacts have been shown for two CCT programs—*Bolsa Escola* in Brazil and *PRAF II* in Honduras (de Janvry and others 2006; Glewee and Olinto 2004).

Table 3.1. Impact on Middle School Completion

	All (Age 12–19)		*Age 12–17*		*Age 15–16*		*Age 17–18*	
	(DD)	(RDD)	(DD)	(RDD)	(DD)	(RDD)	(DD)	(RDD)
FSSP Impact (percentage points)	1.2	1.21	3.27*	3.30*	5.90**	5.90**	–3.4	–3.4
	(1.3)	(1.3)	(1.7)	(1.8)	(2.8)	(2.8)	(2.7)	(2.7)
Value at baseline (percent)	78.8		74.1		65.3		85.1	

Source: Authors' calculations.
Notes: (1) Standard errors are clustered at the district level. (2) All specifications included household, community and school-level covariates: rural/ urban, birth order, dependency ratio, education of household head, access to drinking water supply, access to school, private/public enrollment ratio, access to public transport, gas connection in the house, mean per capita consumption, ratio of middle to high schools, proportion of schools with boundary wall, electricity and drinking water, and student-teacher ratio. (3) All specifications control for age-fixed effects. (4) See endnote 39 for an explanation of RD estimates. (5) Baseline value is given for stipend districts.
***1% significance level, **5% significance level, *10% significance level. Standard errors are shown in parentheses.

Transition from Middle to High School

There is some evidence that the program helps girls transition from middle to high school, especially those younger than 17. The transition from middle to high school is a critical juncture in schooling attainment in Punjab as well as the rest of Pakistan, particularly in poor rural communities. In fact, rates of transition from middle to high school were low at baseline.[40] This analysis measures the probability that a girl enrolled in grade 8 during the program period progressed to grade 9 or beyond. The average impact for the entire cohort of girls aged 12–19 is not statistically significant. However, for girls aged 15–16, the impacts are positive and significant (5.5 percentage points, relative to a baseline value of 58.5 percent). There are two plausible explanations for this observation. First, it is possible that the younger girls benefit more from the FSSP because they are not only exposed for longer but are also more aware of the program's existence; they joined the program when it was more popular. Second, and perhaps more important, only the younger cohorts have the additional incentives induced by the expansion of the FSSP (2006) to also include conditional transfers to enrollees in high school (table 3.2).

Table 3.2. Middle to High School Transition

	Whole sample (12–19)		*Age 15–16*		*Age 17–18*	
	(DD)	(RDD)	(DD)	(RDD)	(DD)	(RDD)
FSSP Impact (percentage points)	1.16	1.09	5.54**	5.50**	–3.90	–3.96
	(1.8)	(1.8)	(2.6)	(2.6)	(3.2)	(3.2)
Value at baseline (percent)	69.1		58.5		63.2	

Source: Authors' calculations.
Notes: (1) Standard errors are clustered at the district level. (2) All specifications included household, community and school-level covariates: rural/urban, birth order, dependency ratio, education of household head, access to drinking water supply, access to school, private/public enrollment ratio, access to public transport, gas connection in the house, mean per capita consumption, ratio of middle to high schools, proportion of schools with boundary wall, electricity and drinking water, and student-teacher ratio. (3) All specifications control for age-fixed effects. (4) See endnote 39 for an explanation of RDD estimates. (5) Baseline value is given for stipend districts.
***1% significance level, **5% significance level, *10% significance level. Standard errors are shown in parentheses.

High School Grade Completion

Similarly, while the analysis does not provide evidence that the older cohorts of girls exposed to the FSSP are more likely to complete grade 9 or 10, there is significant positive impact for the younger cohorts. For instance, the analysis finds a significant impact for girls aged 15–16; they are 5 percentage points more likely to complete grade 9 than the control group. Again, this is consistent with the reasons discussed above, which might explain why impacts tend to accrue to the younger cohorts of girls. Furthermore, 2007–08, the year of the follow-up survey, may be too early to capture the impacts of the program on high school (grade 10) completion because the majority of girls enrolled in grade 10 at the time of the follow-up survey were aged 15–16. Finally, the stipend provided by the FSSP is the same across middle and high schools and, therefore, does not account for the fact that the opportunity cost of attending school increases with the age of girls (table 3.3).

Table 3.3. High School Grade Completion

	Whole sample (14–19)		*Age 15–16*		*Age 17–18*	
	(DD)	(RDD)	(DD)	(RDD)	(DD)	(RDD)
Grade 9 completion						
FSSP Impact (percentage points)	0.629	0.545	4.95**	4.92**	–3.21	–3.33
	(2.0)	(2.0)	(2.3)	(2.4)	(3.1)	(3.1)
Value at baseline (percent)	54.6		54.5		61.2	
Grade 10 completion						
FSSP Impact (percentage points)	–3.36	–3.36	3.03	2.81	–3.86	–3.72
	(2.2)	(2.2)	(4.9)	(4.9)	(2.6)	(2.6)
Value at baseline (percent)	28.2		54.7		20.2	

Source: Authors' calculations.
Notes: (1) Standard errors are clustered at the district level. (2) All specifications included household, community and school-level covariates: rural/urban, birth order, household size, dependency ratio, education of household head, access to school, private/public enrollment ratio, access to transport, gas connection in the house, mean per capita consumption, ratio of middle to high schools, status of school facilities, student-teacher ratio (3) All specifications control for age-fixed effects. (4) See endnote 39 for an explanation of RD estimates. (5) Baseline value is given for stipend districts.
***1% significance level, **5% significance level, *10% significance level. Standard errors are shown in parentheses.

IMPACTS ON LABOR MARKET OUTCOMES

Theoretically, the direction of the effects of the FSSP on working behaviors of the beneficiaries in the medium and long term is unclear. While the program introduces incentives for girls to spend more time in school and work less, if girls are an important source of labor for the household, they may give up leisure time rather than working less. Moreover, girls are expected to obtain better jobs and higher wages as they progress through school and accumulate more education. At the same time, schooling might have diminishing marginal returns and increasing marginal costs. Therefore, girls have more incentives to join the labor market as their ages and years of accumulated

schooling increase. Additionally, because the FSSP seems to facilitate school progression, girls may begin working at earlier ages if they are able to complete their desired schooling sooner. In sum, while it is more possible that the program would decrease working in the short term, the combination of these other factors may have an ambiguous impact on working over the long term. In order to shed light on this question, this section presents results of program impacts on two outcomes related to work for girls aged 12–17: (1) labor force participation and (2) time devoted to working.

Labor Force Participation

The results show that labor force participation falls substantially among adolescent girls exposed to the program. Participation in labor markets (the extensive margin of labor supply) is proxied by the probability of looking for a job and participation in work for pay or unpaid work (for example, unpaid family help and unpaid work outside the home). Results of DD and RDD models indicate that there is a statistically significant reduction in the labor force participation of around 4–5 percentage points among girls in stipend districts, which is largely driven by a reduction in girls' participation in unpaid family work. Point estimates of all empirical models are not sensitive to different specifications, subsamples, or approaches. Furthermore, the decrease in labor force participation is notable. Four years after the program began, the labor force participation rates of girls 12–19 are almost equal between stipend and nonstipend districts, even though they were roughly 4–5 percentage points higher for the former group at baseline. Unfortunately, there are no available data (regarding time use, for example) to further investigate the link between the reduction in labor participation attributed to the program and the extra time spent by girls in school-related activities (table 3.4).

Table 3.4. Labor Force Participation and Work Intensity

	Whole sample (15–19)		*Age 15–16*		*Age 17–18*	
	(DD)	(RDD)	(DD)	(RDD)	(DD)	(RDD)
Labor force participation						
FSSP Impact (percentage points)	–4.94**	–4.90**	–4.03**	–4.01**	–7.86**	–7.62*
	(2.0)	(2.0)	(1.7)	(1.7)	(3.8)	(3.8)
Value at baseline (percent)	10.2		6.8		15.4	
Work intensity						
FSSP Impact (days per month)	–0.611	–0.548	–7.028*	–6.837*	–1.315	–1.260
	(2.037)	(2.077)	(4.104)	(3.922)	(2.612)	(2.697)
Value at baseline (days per month)	25.3		26.3		24.8	

Source: Authors' calculations.
Notes: (1) Standard errors are clustered at the district level. (2) All specifications included household, community and school-level covariates: rural/urban, birth order, household size, dependency ratio, education of household head, access to school, private/public enrollment ratio, access to transport, gas connection in the house, mean per capita consumption, ratio of middle to high schools, status of school facilities, student-teacher ratio. (3) All specifications control for age-fixed effects. (4) See endnote 39 for an explanation of RD estimates. (5) Baseline value is given for stipend districts.
***1% significance level, **5% significance level, *10% significance level. Standard errors are shown in parentheses.

Work Intensity

There is some evidence that the program reduces the number of hours worked for those girls who participate in the labor market. In particular, the analysis examines the effects of the program on the number of days worked per month (the intensive margin of labor supply) among girls who are working for pay at the time of the survey. Samples are restricted not only to girls that work, but also those that are 15–19 years old because no comparable data exist at baseline for younger cohorts. Although the average impact on girls 15–19 is statistically insignificant, the results indicate that cohorts aged 15–16 worked 7 days less per month (around 2.8 hours per day), relative to their counterfactuals in nonstipend districts (significant at the 10 percent level). This is consistent with the hypotheses described earlier about higher impacts for younger cohorts.

IMPACTS ON MARRIAGE AND FERTILITY OUTCOMES

There are multiple channels through which the program could influence marriage and childbirth outcomes. Most of these channels are transmitted through the effects of the program on the decisions about education (and perhaps also work) — as girls have more incentives to stay in school longer, they could delay marriage and sexual activity and alter their decisions about childbirth (such as the timing, birth spacing, and total number of children). Conversely, if the program increases grade progression and helps girls finish school earlier, they might move into marriage and childbearing sooner.[41] Furthermore, social norms and cultural tradition in Pakistan may figure prominently in girls' decisions, making it difficult for girls to resist pressures to marry and have children early. Besides, the local socioeconomic determinants of marriages may mean that the additional schooling brought about by the program could increase the economic status of women and, with it, the chances of getting married. When combined, these different factors could manifest themselves as opposite effects. Therefore, establishing the net effect of the FSSP is, to a large extent, a matter of empirical analysis. To do this, this paper examines the impacts on girls, aged 15 to 19, of two main aspects: (1) the probability of marriage and girl's age at marriage, and (2) whether had previously given birth and the number of births.

Marriage

Although there is no clear impact of the FSSP on marital status, adolescent girls exposed to the program appear to marry at later ages. Estimates from DD and RDD models do not show any impacts of the program on the probability of being legally married. However, when looking at age at marriage, the analysis shows that after the program started, girls in stipend districts were likely to marry about 1.2–1.5 years later. Results are significant in statistical terms and consistent across model specifications and subsamples with different age cohorts.

Fertility

Girls in stipend districts have fewer births at young ages than girls in nonstipend districts. The sample for this part of the analysis consists of girls aged 17 to 19 because these are the ages by which girls are most likely to have been married and possibly

given birth. This group of girls was likely to receive the stipend during the early years of the program (2004–05) but was not likely to be enrolled in middle school at the time of the survey (2007–08). The first outcome indicator measures the probability that girls have given birth. This analysis reveals no differential effects on the probability of girls aged 17–19 having given birth. A second exercise looks at the number of births given for the same sample. The empirical analysis based on this sample indicates that, relative to the control group, girls exposed to the FSSP have, on average, 0.3 fewer children (table 3.5). This result is marginally significant around the 10 percent level.[42] Even though both are partial measures of fertility, given the young ages of the girls, and do not reflect fertility targets, the findings may still signal some changes in fertility that may be sustained over the long term. In the event that the total fertility rates of both treatment and control groups converge over time, the reduction in the number of children by early adulthood still implies that girls in stipend districts tend to delay subsequent births.

Table 3.5. Marriage and Fertility Outcomes

	Main sample (Age 15–19)			*Age 17–19*		
	Value at baseline	(DD)	(RDD)	Value at baseline	(DD)	(RDD)
Married (percentage points)	4.1[T]	0.820	0.825	8.2[T]	0.760	0.709
		(0.8)	(0.8)		(1.6)	(1.6)
Age at marriage (in years)	16.2	1.460**	1.222*	16.2	1.342***	1.262***
		(0.621)	(0.643)		(0.342)	(0.389)
Gave birth (percentage points)				40[T]	–8.08	–6.91
					(17.2)	(17.3)
Number of Children				0.44	–0.329*	–0.349*
					(0.181)	(0.202)

Source: Authors' calculations.
Notes: (1) Standard errors are clustered at the district level. (2) All specifications included household, community and school-level covariates: rural/urban, birth order, household size, dependency ratio, education of household head, access to school, private/public enrollment ratio, access to transport, gas connection in the house, access to Lady Health Worker, access to Family Planning Centre, mean per capita consumption, ratio of middle to high schools, status of school facilities, student-teacher ratio. (3) All specifications control for age-fixed effects. (4) See endnote 39 for an explanation of RDD estimates. (5) Baseline value is given for stipend districts. ***1% significance level, **5% significance level, *10% significance level. Standard errors are shown in parentheses.
[T] baseline value is expressed in terms of percent.

Distribution of Program Impacts

This section presents evidence on the heterogeneity of the FSSP effects across some socioeconomic characteristics of program beneficiaries. So far this paper has provided evidence of positive average intent-to-treat impacts — particularly on measures of partial school attainment and labor outcomes — about four years after the FSSP began. However, knowing who benefits more from the program is as important as knowing its average impacts. From a policy-making perspective, evidence on the distribution of impacts across beneficiaries is critical to increasing the effectiveness and efficiency of the

intervention. This paper disaggregates program impacts by socioeconomic status, parental education, location (rural or urban), age, and length of exposure of the beneficiaries (table E.6 in appendix E). A couple of patterns emerge from this analysis.[43]

Girls who live in rural villages and have parents with no education benefit much less from the program in terms of the probability of completing middle school, but their labor participation also decreases more. The evidence indicates that most educational benefits (as measured by middle school completion) attributed to the program accrue to girls who belong to households in urban centers and have parents with at least primary education. On the other hand, these girls also reduce their participation in the labor market less than girls living in rural areas and with parents that have no education. Unfortunately, with the data available, it is not possible to tease out the underlying reasons that may help reconcile the lower school attainment and lower labor force participation identified for girls in these rural, less-educated households.

Program impacts on labor force participation are larger for the poorest households. Reductions in labor force participation associated with the FSSP, among girls from poorer households (those below the median consumption level of the district), are more than twice those found for the rest of the beneficiaries. There is also some suggestive evidence that girls in the poorest households are more likely to delay marriage.

The effects of the FSSP on educational outcomes vary with the year that the girls join the program and the length of exposure. As noted in the previous section, there is evidence that the younger cohorts of girls who joined the program after two years of implementation are more likely to complete middle and high school. As noted above, the two most plausible reasons of this are an increase in awareness of the program (due to informational campaigns launched in the second year of operation of the program) and the expansion of the FSSP, introduced in 2006–07, to cover grades 9 and 10. Finally, the analysis also shows that girls with more than one year of exposure to the program are more likely than other beneficiaries to complete one grade of high school and less likely to be married at the time of the survey. This provides evidence on the marginal effects of treatment length.

Are There Indirect Effects on the Schooling of Boys?

In theory, educational stipends that are targeted to girls may have indirect effects on households' investments in the education of ineligible siblings. Models of schooling decision show that, in addition to income and substitution effects, school-focused conditional transfers also have a displacement effect. While for eligible children the three effects tend to produce positive impacts on enrollment, for ineligible children the displacement effect often runs in the opposite (negative) direction. For instance, due to the income effect of the transfers, budget-constrained households may be more able to keep all children in school or enroll those out of school. By contrast, the displacement effect may induce parents to reallocate child work (inside and outside the home) away from eligible girls toward boys and other ineligible girls. As a result, the net indirect effects of a program depend on the relative size of these opposing effects.

The existing literature gives mixed evidence on the indirect effects of CCTs on ineligible children. In Colombia, for example, the negative displacement effect was found to offset the positive income effect in the context of a CCT program (Barrera-Osorio and others 2008). While the program increased the enrollment of recipients, ineligible siblings were more likely to drop out of school and enter the labor market. In a contrary example, evidence for a scholarship program in Cambodia that targeted poor children making the transition from primary to lower secondary school shows that the school enrollment of ineligible siblings was not affected by the program (Ferreira and others 2009).[44]

This paper conducts several empirical exercises to assess the spillover effects of the FSSP, particularly on the educational outcomes of ineligible boys. For this purpose, the analysis employs three different identification strategies to examine the spillover effects (table 3.6) on current enrollment, school completion (primary and secondary), and school choice (private or public) for boys of school age (6 to 17 years old). The first strategy implements a DD analysis that compares boys in households with at least one girl exposed to the FSSP in stipend districts, to similar boys in nonstipend districts at baseline and follow-up. The second strategy also looks at relative changes over the medium term but uses boys in households without girls exposed to the FSSP in stipend districts as the comparison group. Finally, the third strategy integrates both of the designs above in a triple difference framework (DDD) to remove other potential sources of bias. The analyses draw on household data from MICS (2003 and 2007–08), and the subsamples used for each of the outcomes are selected based on specific age and grade restrictions.

Most empirical models and research designs do not reveal any secondary effects of the program on the enrollment and school completion of boys. The results indicate that the enrollment and school completion rates of boys with female siblings who have been exposed to the FSSP exhibit statistically similar trends as boys with no eligible siblings in both stipend and nonstipend districts. The only exception is a decrease in enrollment of 3 percentage points in one of the models of the DD analysis that is based on the comparison with boys living with ineligible girls in stipend districts. Yet, this result does not hold when the analysis is extended to DDD models.[45]

Table 3.6. Program Spillover Effects on Boys

	Whole sample (age 6–17)			*Age 6–12*		
	Value at Baseline	**(DDD)**	**(DDD/RDD)**	***Value at Baseline***	**(DDD)**	**(DDD/RDD)**
Primary school completion (percentage points)	91.1[T]	0.898	1.03	94.1[T]	0.00519	0.00519
		(1.8)	(1.8)		(0.048)	(0.048)
Middle school completion (percentage points)	76.3[T]	–1.09	–0.996			
		(3.3)	(3.3)			
Private school enrollment (percentage points)	30.6[T]	3.75**	3.77**	34.0[T]	11.0***	11.0***
		(1.9)	(1.9)		(2.6)	(2.6)
Years of schooling	4.2	–0.00449	0.00182	2.5	–0.0756	–0.0756
		(0.079)	(0.079)		(0.077)	(0.077)

Source: Authors' calculations.
Notes: (1) All specifications included household, school and community covariates: rural/urban, birth order, dependency ratio, education of household head, access to school, private/public enrollment ratio, access to water supply, age of household head, access to public transport, gas connection in the house, mean per capita consumption, initial enrollment rate, ratio of middle to high schools and ratio primary to middle schools, status of school facilities and student-teacher ratio, among others. (2) All specifications control for age-fixed effects. (3) See endnote 39 for an explanation of RDD estimates. (4) Baseline value is given for stipend districts.
***1% significance level, **5% significance level, *10% significance level. Standard errors are shown in parentheses.
[T] baseline value is expressed in percentages.

There is, however, indication of strong indirect effects of the program on school choice among boys. Several econometric models show a statistically significant average increase in enrollment of boys in private schools of nearly 4 percentage points, as compared with a baseline level of 30 percent. Further analysis indicates that this increase in the share of male enrollment in private schools is largely driven by boys 6 to 12 years old who were sent to primary schools. This is consistent with what was observed through qualitative work conducted in the field as part of this impact evaluation. There is a clear pattern that families with daughters eligible for the FSSP tend to respond to the program by enrolling their sons in private schools.[46]

This behavioral response may be explained by reasons related to the supply, cost, and quality of private schools that offer primary-level education. First, there has been a significant increase in the number of private primary schools in rural and urban Punjab in the 2000s. Second, fees in private schools are often low and below the stipend provided by the FSSP (200 rupees per month). In fact, a national census of private schools in 2000 found that the median rural private school charged 60 rupees per month.[47] Third, there are also differences in quality; children who study in private schools score substantially higher in tests in all subjects (Andrabi and others 2006). Therefore, a potential indirect effect of the FSSP may introduce gender disparities in learning as more boys in eligible households are enrolled in better-performing schools. More work, however, is needed to identify this possible indirect effect.

4. Robustness Analysis

Because the research strategy followed in this paper relies on quasi-experimental methods, the empirical analysis was complemented with a number of robustness checks to validate the main results. From an impact evaluation perspective, the ideal analysis would be based on a random experiment so that that the underlying distribution of the relevant observable and unobservable variables would be statistically identical between the treatment and control groups. Unfortunately, the FSSP was not implemented this way, and, therefore, impact evaluations of this program have to rely on quasi-experimental techniques. Whether the estimates of program impacts derived from these methods are unbiased depends, to a large extent, on certain assumptions and issues that could contaminate the design of the evaluation. This chapter shows that the main findings of this evaluation are robust to potential issues, such as (1) differential preprogram trends in the outcomes of interest, (2) measurement error (that is, lack of precision in the measurement of program exposure), (3) endogenous compositional changes in the samples of analysis, (4) the crowding-out effects of the FSSP on female enrollment in private schools, and (5) other issues that may affect the comparability between stipend and nonstipend districts.

Preprogram Trends of Outcomes

The DD analysis is built on the assumption that the outcomes investigated were progressing along similar trends, across stipend and nonstipend districts, before the program began. Therefore, any difference emerging between their trends could be attributed to the program. Program impact estimates based on DD models may be biased if this assumption fails. For example, it could have been the case that the probability of completing middle or high school for girls in stipend districts (who started from a lower base) was converging to the value in control districts even before the program was in place. If this were the case, the impact evaluation would attribute changes in the outcomes to the program, when in fact these changes are explained by differential preprogram trends.

There is no evidence that the medium-term impacts of the program identified in this paper are driven by differences in preprogram trends. Two main checks were conducted to investigate the existence of differential trends between stipend and nonstipend districts. The first check uses prebaseline and baseline data from PIHS (2001) and MICS (2003) to run DD models on schooling, labor market, marriage, and fertility outcomes for the same cohorts of girls.[48] In principle, one should not expect to see impacts of the FSSP since the program began after the period covered by this "placebo test." In fact, results of this analysis show that all outcomes examined in the paper were progressing in similar trends for stipend and nonstipend districts before the program. The second check consists of DD models using data from MICS (2007–08) to examine the trends of school attainment for two slightly older cohorts of girls (ages 20–22 and 23–25) that, by definition, finished school before the program started and, therefore, did not benefit directly from it. Again, results of these exercises do not reveal any differential

preprogram trends that may have contaminated the main findings of the paper. The assumption for the DD analysis, therefore, seems to hold (table E.7 in appendix E).

Measurement Error

The empirical analysis of this paper is based on a proxy measure of the level of exposure of girls to the program. Consequently, program duration for the synthetic cohorts used in the analysis is likely measured with some error. This measurement error in the indicator of treatment could produce imprecise or biased estimates of program impacts.[49] If this is the case, the parameters that measure the effects of the FSSP on the medium-term outcomes would be biased toward zero (also known as "attenuation bias"). Although there is no sensible way to correct for this type of bias in the context of this evaluation, the results of the analysis are still very informative because they can be interpreted as lower bounds of the true impacts.

Endogenous Compositional Changes

An endogenous compositional change may potentially arise if the FSSP induces the self-selection of lower-ability female students into middle schools. The marginal girl who is most likely to be brought into school by the stipend program may be, on average, relatively poorer, less motivated, and spend less time on schoolwork than the girls who are already enrolled in school.[50] Thus, there is a possibility that at least some of these new entrants are drawn from the left-hand side of the ability distribution. As a result, there may be a selection bias on the expected returns because the new entrants are expected to benefit less from schooling, as compared with the girls who were already attending school regularly before the program started. This type of selection bias may be important for evaluations like this one, that investigate the impacts on educational indicators such as progression and completion, which condition on school enrollment (itself an outcome of the program). Although the findings of this paper are not adjusted for the self-selection of lower-ability students, the sign of the bias suggests that they still provide lower bounds of the actual impacts.[51]

There is no sign the program causes changes to the age-grade distribution in stipend districts and causes compositional issues in this aspect. In addition to bringing new entrants with relatively lower expected returns into school, the program could have motivated overage girls to go back to school. This may have added extra "noise" to the construction of the synthetic cohorts and the measures of treatment status. However, comparisons of the age-grade distributions at baseline and follow-up, within stipend districts as well as between stipend and nonstipend districts, do not indicate that overage girls were more likely to enroll in schools in stipend districts after the FSSP began. Furthermore, various model specifications allow for cohorts that are wide enough to capture overage girls.

There is no evidence that migration could affect the composition of the samples of analysis. A recurrent concern for evaluations of this type is that the program could either attract migrants from control regions to stipend regions or stimulate permanent migration among program participants. The decision to migrate permanently may be

influenced by factors that also determine participation in the program and the outcomes of interest (that is, endogenous migration). Both issues may alter the composition of the samples in a way that compromises the internal validity of the analysis. Yet, there are several reasons why this may be less of an issue in this study. First, the modest size of the transfer (nearly $2.5 per month) is not expected to encourage households to move from richer nonstipend districts to stipend districts. Second, out-of-province family migration in both stipend and nonstipend districts is rather uncommon, particularly in rural areas. Instead, most migration is in the form of temporary out-migration of one member of the household — often a working man. It is very unlikely that this migration was affected by the stipend program. Finally, some simple econometric models give no evidence that the FSSP affect either type of migration after controlling for most of the covariates used in the analysis.

Finally, the findings do not appear to be driven by crossover effects between contiguous stipend and nonstipend districts. The conditional stipend could motivate households located in nonstipend districts adjacent to stipend districts to send their girls there for school. This would increase the demand for education in stipend districts but reduce it in nonstipend districts, leading to no changes in net enrollment across the province. In order to check this, econometric models of educational outcomes were estimated on a subsample of schools and households located in treatment and comparison districts that do not have common borders.[52] The direction, magnitude, and significance of the findings using these samples are, for the most part, consistent with the results obtained with the full sample, with the exception of some models that employ school-level data for which the findings are qualitatively similar but not statistically significant.

Does the FSSP Crowd Out Private Enrollment?

A natural concern is that the stipend program led to a reduction in female enrollment in private schools. There is, however, no evidence of this crowding out effect. A possible response of families to the FSSP is to move their daughters who were previously enrolled in private schools to public schools, in order to become eligible for the stipend. As a result, the observed increase in female enrollment in the public system may be offset (partially or completely) by a fall in enrollment in private schools, hence reducing the overall effectiveness of the program. Nevertheless, the main set of results of this paper uses household-level data and defines progression and completion irrespective of the type of school in which the girl is enrolled. Therefore, the DD and RDD models using these data provide estimates of the net impacts of the FSSP that already account for the possible switch from private to public schools induced by the program.[53]

Other Issues That Can Affect the Comparability of Districts

In order to check for potential issues of selection in participation, empirical models for all the outcomes were estimated on the sample of girls eligible for the program, irrespective of their participation. The household-level analysis was also run with an

alternative cohort, which included girls aged 12–19 that were exposed to the program, as well as girls ages 12–19 that were eligible for the program but never enrolled in grades 6–8 for the duration of the program. For instance, a girl who completed grade 5 but never enrolled in grade 6 is included under this alternative definition. Although impacts estimates based on this larger cohort are closer to the intent-to-treat parameter, the results of this analysis are likely to get diluted because the units of observations include a significant number of girls that did not participate; these girls left school before the program started and never returned.[54] Yet, some of the results from the base models remain. For middle school/grade 9/grade 10 completion, and for middle to high school transition outcomes, average impacts on cohort age 12–17 are not statistically significant but have a positive sign across alternative definitions. However, for the cohort age 15–16, schooling outcomes are positive and statistically significant for some definitions. Impacts on labor outcomes are robust to the inclusion of eligible but not girls exposed to the FSSP, but with lower significance levels (5–10 percent). Results for marriage and fertility outcomes (with the exception of age at marriage) have a negative sign across alternative definitions, but are not statistically significant.

The evidence available does not suggest that changes on the supply side were systemically different between stipend and nonstipend districts over the time period covered by this analysis. One concern for internal validity is that certain impacts on educational outcomes could be mistakenly attributed solely to the program when they are actually the combined effects of the stipend (a demand-side intervention) and changes on the supply side. This may be the case if more schools — particularly public schools — were built, at the time of the program, in stipend than in nonstipend districts.[55] There is, however, no evidence of such systematic changes in school supply. In fact, data from the annual Punjab school census (Education Management Information System) for the period 2003–10 shows that the number of public schools was more or less unchanged. For instance, during 2004–07, the number of schools in stipend and nonstipend districts increased by only 0.34 and 0.15 percent, respectively, as compared with the number of schools at baseline.[56] Furthermore, the trends in the school supply of stipend and nonstipend districts are very similar in the period before and during the program, 2000–07 (figure E.4 in appendix E).

Similarly, the findings do not appear to be driven by incidences of aggregate and idiosyncratic income shocks (positive or negative). Although there is no available quantitative information to investigate this possibility, there are several reasons why it seems implausible. First, there is no evidence of aggregated income shocks that systematically benefited or harmed stipend districts, as a whole, more than nonstipend districts. Second, yield losses or gains due to crop shocks, if they happened, are supposed to be local and seasonal and should be controlled in the statistical analysis by the district-year fixed effects. Furthermore, program impacts are found for various age cohorts that benefited from the program at different times, which rules out the potential confounding effect of transitory shocks. Finally, the incidence of idiosyncratic shocks (such as health and employment) that are not partially determined by household characteristics already included in the regression analysis can hardly be argued to be systematically correlated with treatment status.

5. Conclusion

This report aims to fill the evidence gap on the medium-term impacts of CCTs by evaluating a female-targeted CCT program in Pakistan, the Female School Stipend Program. The FSSP was part of the Punjab Educational Sector Reform and was implemented together with other measures to enhance the provision and quality of educational services. As a CCT targeting girls, the program aim to address the disparities in educational outcomes between boys and girls, and across districts within the region. Early impact evaluation results show that the FSSP increased school enrollment of girls in stipend districts by up to 23 percentage points. However, the literature remains unclear about whether these immediate improvements are sustained over time and translated into better human capital, productivity, and employability. This study attempts to answer some of these questions by investigating: (1) the program impacts on educational attainment, early labor market outcomes, and choices about marriage and childbirth by the beneficiaries; (2) the heterogeneity of program impacts across various participant groups; and (3) any spillover effects on the educational outcomes of boys in participating households.

The evidence is encouraging, indicating that impacts are sustained beyond the short term and could help reduce gender gaps in schooling. Indeed, after four years of implementation, the FSSP is found to help girls in stipend districts progress through and complete middle school. The latter impact, however, is distributed more favorably toward those who live in urban areas and have more educated parents. Furthermore, the younger cohorts of girls in stipend districts, in particular, are also more likely to matriculate into and complete at least one high school grade. This matches with the expectation that the younger cohorts join the program when it is more mature and stay in the program longer, indicating some marginal benefits of the length of exposure. Another hypothesis is that these girls also benefit from the expansion of the program (started in 2006) to include high school enrollees. In terms of labor force participation and marriage and childbirth decisions, the program encourages girls — particularly those who live in poorer households — to work less, delay marriage, and have fewer children.

The FSSP may indirectly increase the gender gap in learning, at least at the primary level. Boys residing in the same household as girls in participating districts do not seem to be affected in terms of school enrollment and completion. However, they are more likely to be enrolled in private primary schools. These schools are growing in number, are more affordable (school fees are lower than the stipend), and perform better in terms of students' learning, which means that boys are getting a better education than their FSSP-eligible siblings. However, further research is needed to investigate whether learning disparities may arise as a result of this change in the school choice of boys. If this is the case, the evidence should perhaps motivate actions to either address the discrepancy in quality between private and public schools or expand the FSSP benefits to include private schools so that girls have similar school choices as boys.

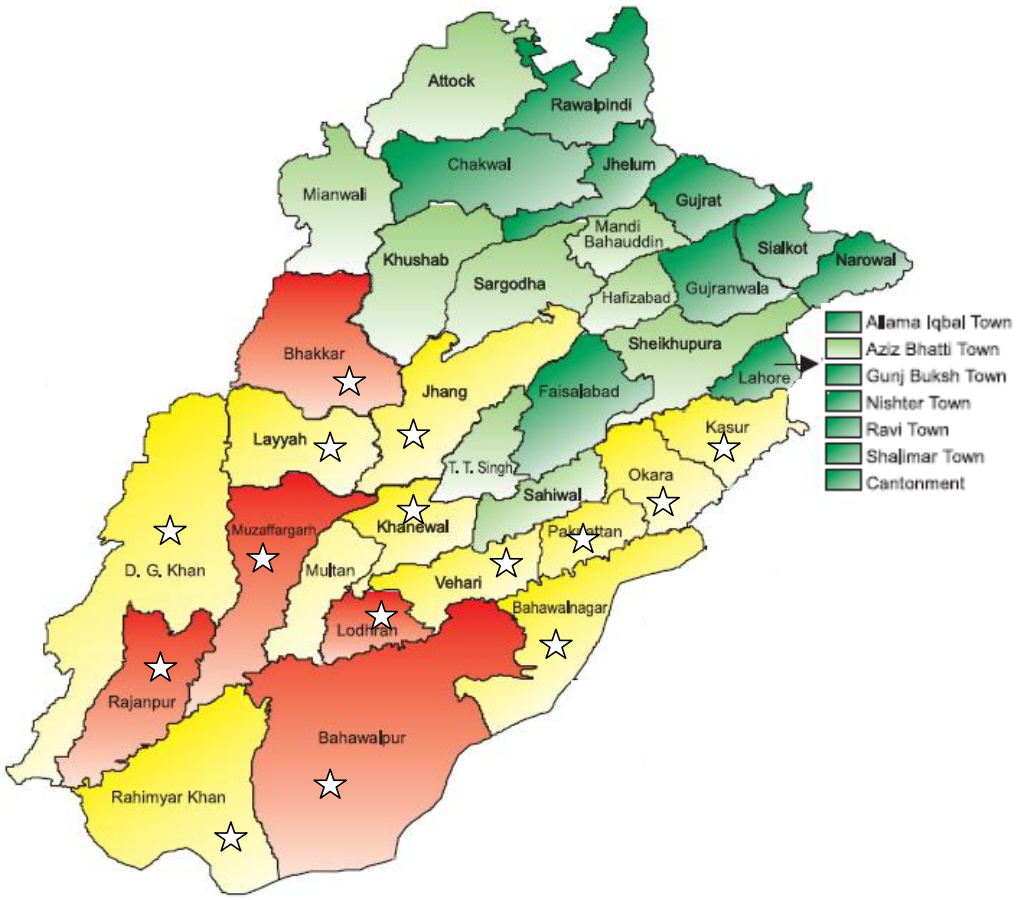
The impacts on educational attainment may have important implications for future productivity and the welfare of beneficiaries. As in many other developing countries,

the gradient between education and poverty is evident in Pakistan. Women who complete middle and high school, for instance, live in households that enjoy nearly 14 and 29 percent (equivalent to $5.3 and $10.8) higher monthly consumption per capita, respectively, as compared with women with a less than middle school education. Much of the reason is that education provides women with more marriage opportunities. Women with a middle school education or higher marry men who, on average, have 4–5 more years of education than that of the husbands of women with less than a middle school education, nearly double the years. Another reason is that higher schooling also enables women to attain higher earnings. Simple "back-of-the-envelope calculations" estimate that the wage premiums for females with a middle or high school education (relative to lower than a middle school education) are 24 and 157 percent, respectively. These improvements represent an annual increase in earnings of $16 and $100, respectively, which may more than compensate for the annual stipend cost of the program ($34).

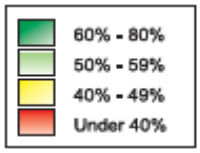
In addition, the impacts of the FSSP may have further dynamic effects on other dimensions. First, women who have higher education and/or marry at a later age tend to have fewer children. In fact, women in Punjab with a middle or high school education have around 1.8 fewer children than those with a lower than middle school education by the end of their reproductive life. Simple extrapolations also indicate that the 1.4-year delay in marriage by beneficiaries associated with the program could lead to 0.4 fewer births by the end of their childbearing years. Furthermore, more educated women are expected to not only have fewer children, but also to invest more in their children's human capital, which may lead to positive intergenerational effects. For example, children aged 6–16 of mothers with a middle and high school education in Punjab are 34 percentage points more likely to enroll in school, relative to those of less-educated mothers.

These results could provide critical information for policy makers in assessing the effectiveness and efficiency of programs like the FSSP. However, more needs to be done in future research to explore the sustainability of program impacts, the effects on learning, and the role of program design and implementation in explaining the impacts. First of all, many of the beneficiaries are still young (12–19 years old), so they might not have reached their full potential in terms of educational attainment. Moreover, little is known about their future decisions on marriage and childbirth. Therefore, the impacts can only be considered as early outcomes. Second, it would be useful to understand whether the impacts on schooling are translated into improvements in cognitive development (for example, looking at test scores). Third, because the transfer amount is the same for all beneficiaries across grades and income levels, there is no available information to investigate the marginal benefits of the transfer size. This knowledge would contribute to a better understanding of the adequacy of program benefits. Fourth, while this analysis accounts for simultaneous educational interventions on the supply side that may affect the outcomes, it is unclear how these instruments may or may not complement the impacts of the FSSP. Finally, more information on the monitoring and evaluation framework would be helpful for further research, to identify whether factors in the implementation process and context of the program may interfere with its effects.

Appendix A
Geographical Coverage of the Program

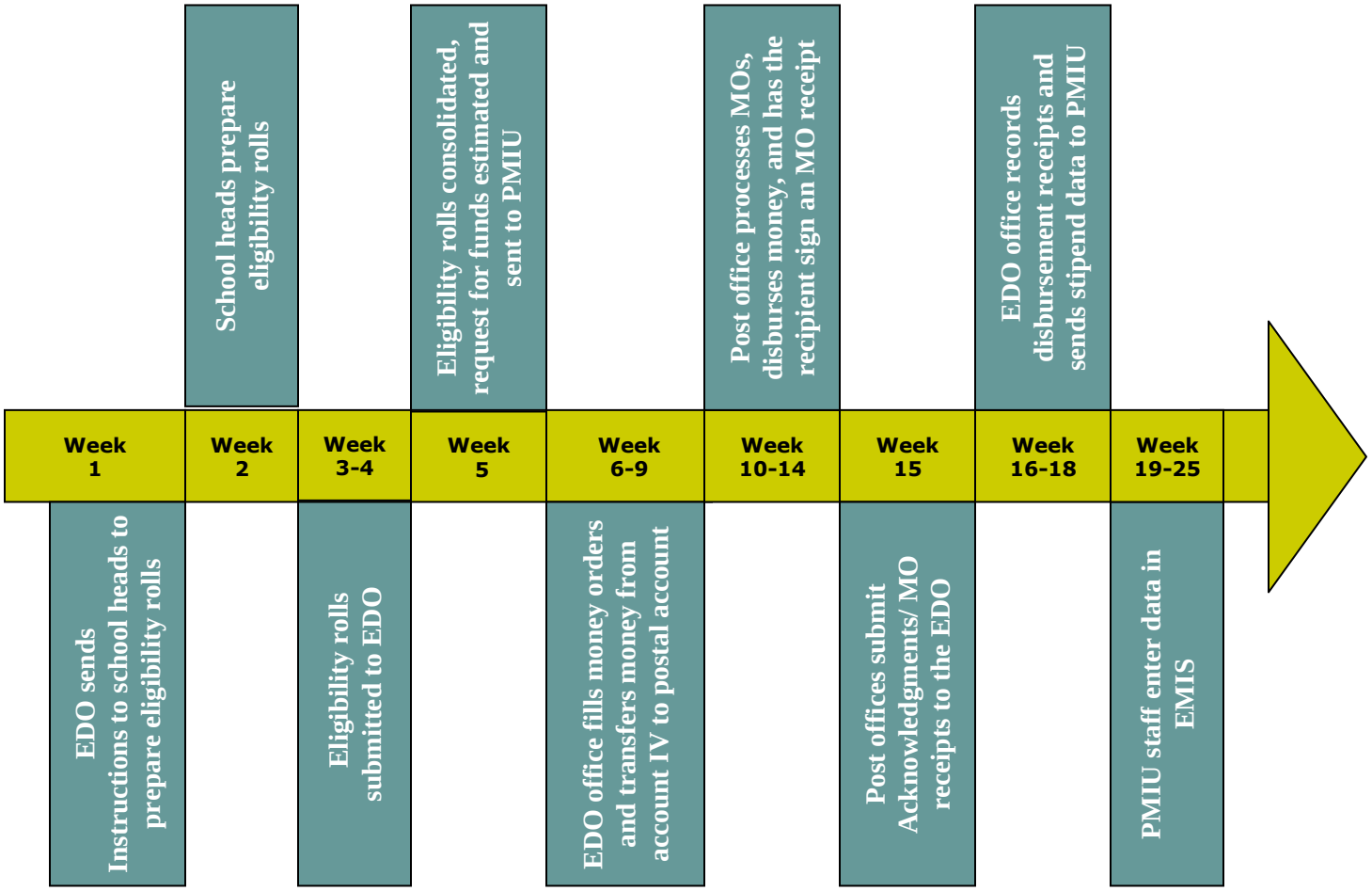


Notes: (1) The districts with stars are covered by the stipend program. (2) The shaded areas represent district-level literacy rate amongst population 10 years and older (see legend below)



Source: MICS 2003 report

Appendix B
Implementation Process of the Program



Source: Compiled by authors from PESRP Web site.
Note: MO= Money Order; EMIS=Education Management Information System; EDO=Executive District Officer; Account IV is maintained by the provincial government for making inter-government transfers (from province to district).

Appendix C
Construction of the Samples for Analysis

The household-level analysis mostly draws on two rounds of the Multiple Indicators Cluster Survey (MICS). The first round was undertaken in 2003 and the second in 2007–08; both are cross-sectional. Neither of these rounds was intended to collect data on the program, nor did they ask any questions regarding participation in the FSSP. However, both rounds of the MICS had questions on: (1) age; (2) district of residence; (3) gender; (4) school enrollment in the current year; (5) school enrollment in the previous year; (6) grade level in the current year; (7) grade level in the previous year; and (8) highest grade completed.

Girls eligible to receive the stipend must be enrolled in grades 6–8, in a public school, in any of the 15 low-literacy districts. In the absence of any identifying information in the MICS or any other household survey on program participation, the analysis relies on the questions on age and enrollment in the MICS to construct synthetic cohorts of the relevant group—girls that were likely to have been exposed to the program for at least one year since 2003–04. The motivation for constructing synthetic cohorts based on age AND grade criteria, as opposed to only an age-based criterion, was to arrive at a cohort that is more reflective of actual participants.[57] For instance, at baseline, 49 percent of girls aged 12–19 in stipend districts had never attended school, as opposed to 21 percent in nonstipend districts. Under an exclusively age-based criterion, these girls would have been misidentified as having been exposed to FSSP, although it was impossible for these girls to participate in the program since its implementation.[58]

This appendix discusses the three criteria used to construct the cohorts of possible participant girls. In addition, the appendix presents detailed examples of different combinations of age and schooling that also need to be addressed to ensure that the educational histories of girls are attributed correctly. The criteria are the following:

1. The **first** criterion applies to girls who were enrolled in either 2007–08 and/or 2006–07.

   a. If a girl was enrolled in 2006–07 in grades 6–8, she is considered as having been exposed to the FSSP. Enrollment and grade level in 2006–07 allow us to determine exposure for two possible enrollment scenarios: (i) the girl was enrolled in 2006–07 as well as in 2007–08, and (ii) the girl was enrolled in 2006–07 but not enrolled in 2007–08.

   b. However, there is a third possible enrollment scenario, in which a girl was not enrolled in 2006–07 but was enrolled in 2007–08. In this scenario, a girl who was enrolled in 2007–08 in grades 7–8 is considered as having been exposed to the FSSP.[59]

However, girls exposed to the program in its early years will no longer be in grades 6–8 in either 2006–07 or 2007–08. For example, a girl who was in grade 7 in the academic year 2003–04 (and exposed to the program) would be in grade 11 in 2007–08. Therefore, if a girl was enrolled in grades 9–11 in 2006–07, she is considered exposed to the

program. As before, this addresses the two enrollment scenarios: (i) the girl was enrolled in both 2006–07 and 2007–08, and (ii) the girl was enrolled in 2006–07 but not in 2007–08.

The underlying assumption here is that these girls progressed through their school years without any interruption. For instance, a girl enrolled in grade 11 in 2006–07 is assumed to have been exposed to the program, the assumption being that she was enrolled in grade 8 in 2003–04, and then moved on to grade 9 in 2004–05, grade 10 in 2005–06, and finally grade 11 in 2006–07. The two exceptions to this assumption are: (1) grade repetition, and (2) reentry. To illustrate each of these exceptions, two examples are presented:

a. If a girl was enrolled in grade 11 in 2006–07, she is considered exposed to the program because it is assumed that she was in grade 8 in 2003–04 (year 1 of the program). However, this girl may have repeated grade 9, in which case she would have been in grade 9 in 2003–04 and not in grade 8, and in actuality was not exposed to the program.

b. If a girl was enrolled in grade 11 in 2006–07, she is considered exposed to the program because it is assumed that she was in grade 8 in 2003–04 (year 1 of the program). However, this girl may have finished grade 8 in 2002–03, and rejoined a year later in 2004–05 in grade 9, and so was not exposed to the program.

In both of the above examples, the girls are misidentified as having been exposed to the program when they had not been in grades 6–8 since the start of the program. However, baseline data suggest that both repetition and reentry at the secondary school level are very low: around 0.5 percent of the girls enrolled in secondary grades are repeaters or rejoiners. Not only is the proportion of such girls very low, there are no statistically significant differences in their proportions between stipend and nonstipend districts at baseline or over time.

2. The **second** criterion is based on the age-grade distribution and applies to girls who were not enrolled in either 2006–07 or 2007–08, and only their highest grade completed is observed, but NOT the year of completion. Before using the age-grade distribution to determine program exposure, two restrictions were imposed to define the sample of girls for whom program exposure is a possibility:

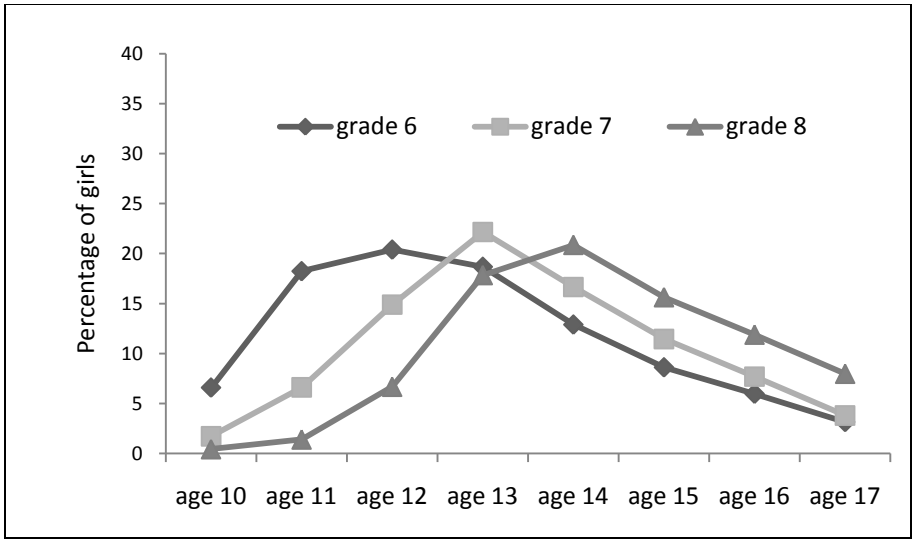
   a. Girls whose highest grade completed is grade 5 or below were excluded. This is because for a girl to be exposed to the program, she must have at least completed grade 6.

      - One may argue that it is possible that a girl was enrolled in grade 6 but did not pass, in which case her highest completed grade would be reported as grade 5. However, baseline data indicate that only 0.7 percent of girls enrolled in grade 6 did not pass, so the exclusion error in the case of this restriction is very small.

   b. Girls whose highest grade completed is grade 11 or above were excluded. For instance, the latest year that a girl exposed to the program, but not enrolled in 2006–07 or 2007–08, could have finished grade 11 was in 2005–06, in which

case she would have been in grade 9 in 2003–04 (year 1 of the program), and therefore not exposed to the program.

FigureC.1. Age-Grade Distribution at Baseline



The age-grade distribution at baseline is then used to determine if any of these girls who were not in school in either 2006–07 or 2007–08, and whose highest grade completed was between 6 and 10, were exposed to the program for at least one year between 2003–04 and 2005–06. The graph above shows the distribution of ages among grades 6–8 at baseline.[60] To illustrate the construction process, an example is presented:

a. A girl aged 16 in 2007–08 and whose highest grade completed is grade 7 would have been exposed to the program only if she was in grades 6 or 7 in at least 2003–04. The only way she could not have been exposed to the program was if she completed grade 7 before the program began, say in 2002–03. In 2002–03, this girl was 11 years old, and from the age-grade distribution it is evident that an 11-year old is not very likely to be in grade 7, but is more likely to be in a lower grade. Hence, this girl is assumed to have completed grade 7 after the program began and is therefore considered exposed to the program. Similarly, a girl who is 19 years old and whose highest grade completed is grade 6 is not considered exposed to the program because she would have been 15 years old when the program began, and the likelihood of her being in grade 6 in 2003–04, at age 15, is very low; she is assumed to have finished grade 6 prior to the program.

3. The **third** criterion was to further restrict the sample cohorts to ages 12–19. After the first two criteria were implemented, girls between ages 8–24 were identified to be in the treatment group. The sample was then further restricted to girls 12–19, who constituted approximately 95 percent of all the girls identified as exposed to the program after imposing the first two criteria. This is akin to the

treatment of outliers, and a safeguard against misreporting. For instance, a girl who is 8 years old is highly unlikely to have completed grade 6.

Appendix D
Model Specifications

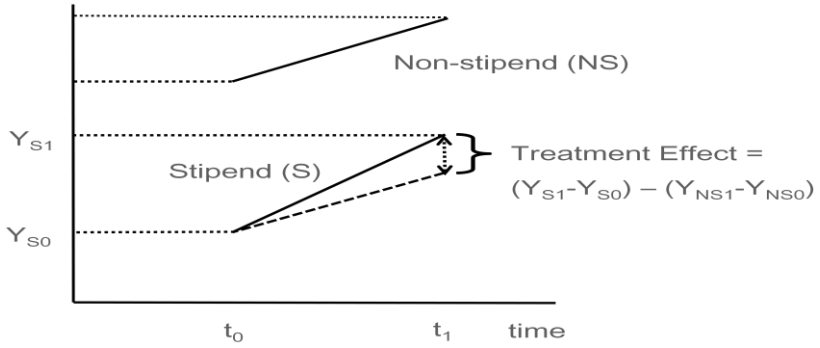
The main goal of this study is to estimate the impact of the FSSP on schooling, early labor market participation, and fertility outcomes in the medium term. Given that one cannot observe the effects of the treatment and nontreatment on the same group at the same time, a comparable (control) group must be identified to serve as the counterfactual group for the analysis. Because all girls eligible to be enrolled in grades 6–8 were potentially eligible for the program, there is no similar cohort WITHIN stipend districts that could serve as a counterfactual. Therefore, girls with the same age and grade enrollment profiles in nonstipend districts were identified as a control group.

Tests described in appendix E show that there are significant differences between stipend and nonstipend districts, in terms of baseline household demographics, access to facilities and community infrastructure, and quality of schools, among others. However, preprogram trends between stipend and nonstipend were shown to be statistically insignificant (chapter 4), which allows the use of difference-in-difference (DD) estimation. The DD setup is fairly simple: outcomes are observed for two groups at two periods. One of the groups (stipend) is exposed to the treatment in the second period but not in the first period, while the second group (nonstipend) is not exposed to the program in either period. The average impact of the program is measured as the average gain in the first (stipend) group, net of the average gain in the second group (nonstipend). This removes biases in second period comparisons between the stipend and nonstipend groups that could be due to permanent differences between the two groups, as well as biases from comparisons over time in the stipend group that could be due to factors other than the stipend. Mathematically,

$$\Delta_i = (Y_{Si,1} \; Y_{Si,0}) \; (Y_{NSi,1} \; Y_{NSi,0}) ,$$

where i= unit of analysis (individual or school), S=stipend group, NS=nonstipend group, 0=preprogram year, 1=postprogram year, and Y=outcome

Graphically,



In the case of *school-level analysis,* the treatment group is girls' schools offering middle-level grades in the stipend districts, while the control group is girls' schools offering middle-level grades in the nonstipend district. The outcome of interest is the percentage

change in grade-level enrollment (grades 6–10 each), which is estimated using the regression:

$$\Delta 0_i = ß + \alpha * FSSP_i + \varepsilon_i \qquad (1),$$

where FSSP is a binary variable that takes the value 1 for a stipend-eligible school and 0 otherwise. α measures the net impact of the stipend program.

In the case of *household-level analysis*, the treatment group is defined as girls exposed to the program since its inception until 2007–08 in stipend districts (see appendix C for synthetic cohort construction), while the control group consists of girls with the same age and grade-enrollment profile as the ones in the treatment group, but who are situated in nonstipend districts. For this analysis, the study uses two cross-sectional Multiple Indicators Cluster Surveys (MICS), conducted before the program (2003) and then again in 2007–08. The simple DD regression specification used here is:

$$Y_{it} = \mu + \gamma * FSSP_i + \delta * T_t + \alpha * T_t * FSSP_i + \varepsilon_{it} \qquad (2),$$

where i is the individual, t is the time subscript, T is a dummy variable taking a value of 1 for the postprogram year and a value of 0 for the preprogram year and FSSP identifies household in stipend villages. The outcomes of interest (Y) in this case are (i) probability of completing middle school; (ii) probability of middle to high school transition; (iii) probability of completing high school; (iv) labor force participation; (v) number of days worked per month; (vi) probability of being married; (vii) age at marriage; (viii) probability of having given birth; and (ix) number of children. The coefficient α, the interaction of time and treatment status, gives the average net impact of the program. Because observations within a district could be serially correlated, standard errors are clustered at the district level.

Covariates

There are significant differences (such as poverty, access to school, and school quality) between stipend and nonstipend groups at baseline. Also, program allocation was not randomized but targeted to districts with low literacy. This implies that both targeting criteria and differences in baseline conditions may influence subsequent growth rates, leading to selection bias (Ravaillon 2005). To alleviate this endogeneity problem, covariates on household demographics (education of household head, household size, dependency ratio, birth order); community-level characteristics (poverty, access to water, access to gas, access to public transport, access to middle school, access to high school, school supply), education marketplace (public-private enrollment ratio), and school quality (status of school facilities, including boundary walls, toilet, electricity, drinking water, and the student-teacher ratio) were added to base specifications. For fertility outcomes, access to a family planning center and female health workers were included in addition to the other covariates. The revised regression specifications with covariates for school and household-level analysis are:

$$\Delta Y_i = ß + \alpha * FSSP_i + \mu * COV_i + \varepsilon_i \quad \text{(3a – School Level)},$$

$$Y_{it} = ß + \gamma * FSSP_i + \delta * T_t + \alpha * T_t * FSSP_i + \mu * COV_{it} + \varepsilon_{it} \quad \text{(3b – Household Level)}.$$

New Programs

Another source of bias is other programs since the introduction of FSSP that were implemented differentially in girls schools/targeted at girls in stipend districts, and that were geared toward improving outcomes of interest. While FSSP was in place, the government of Punjab, under the rubric of the Education Sector Reform Program, introduced programs for building missing school facilities, free textbooks, contract teacher recruitment, and revitalization of school councils. These programs, in and of themselves, were not targeted at stipend districts but could have been implemented differentially. For instance, girls' schools in stipend districts were worse off and it is possible that any program on rehabilitation of school infrastructure would focus on schools with many missing facilities, many of which happen to be in stipend districts.

$$\Delta Y_i = ß + \alpha * FSSP_i + \mu * COV_i + \gamma * New_i + \varepsilon_i \quad \text{(4a – School Level)},$$

$$Y_{it} = ß + \gamma * FSSP_i + \delta * T_t + \alpha * T_t * FSSP_i + \mu * COV_{it} + \lambda * New_{it} + \varepsilon_{it} \quad \text{(4b – Household Level)}.$$

Regression Discontinuity Design

Another way to address the endogeneity issue is to use a regression discontinuity design (RDD), which exploits the program design itself (Chaudhury and Parajuli 2006). The basic idea behind the RDD approach is that assignment to the stipend district (beneficiaries) is determined by a threshold (otherwise called the cutoff point), established due to administrative resource constraints, which is set before the program begins (Imbens and Lemieux 2008). The designated cutoff point is the predictor of whether a group is on one side of the discontinuity or the other; assignment to the benefited (or control) group is, therefore, a deterministic function of the literacy ratio (*L*);

$$L=1[X \leq 40],$$

where 1[.] is the indicator function (equal to one in a sharp discontinuity design) if a district has a literacy rate of 40 percent or less, and zero otherwise (Imbens and Wooldridge 2009). In other words, girls in stipend districts with literacy rates of 40 percent or less are eligible for FSSP; all girls in districts above the 40 percent literacy threshold are in the nonstipend or control group. About 15 districts were assigned to the treatment group and 19 districts were assigned to the control group.

The following specification is a case of RDD wherein program impact is estimated in the neighborhood of the cutoff point for selection into the program. The RDD specification includes a control function for distance of district literacy rate (subscript D) from the cutoff (L). It is assumed that the control function is linear, making this a parametric specification. Here, L is included in the specification because literacy rate is the only

determinant of treatment status. That is, making treatment status conditional on literacy would imply treatment to be independent of the error term given the literacy rate.

$$\Delta Y_i = \beta + \alpha FSSP_i + \lambda(L_D) + \epsilon_i \qquad \text{(5a – School Level)},$$

$$Y_{it} = ß + \gamma * FSSP_i + \delta * T_t + \alpha * T_t * FSSP_i + \lambda(L_D) + \varepsilon_{it} \qquad \text{(5b – Household Level)}.$$

The parametric RDD specification extrapolates from above and below the cut-off point to all observations in the data, and allows for use of covariates. A batch of RDD specifications is run adding covariates and new programs to (5a) and (5b).

A disadvantage of the parametric RDD specification is that there is no justification for the assumption that the model is linear. Some of the RDD literature finds that the consequences of using the incorrect functional form is a concern because running a linear regression could minimize specifications globally while the specification errors at different points, including the cutoff, could be big (Lee and Lemieux 2009). If this is so, there may be a bias in the RDD estimates in the paper. Because misspecification of the functional form is a serious issue in RDD, its estimation has generally been viewed as a nonparametric estimation problem. This involves running a linear regression on either side of the discontinuity within some specified bandwidth h[61] on either side of the cutoff point. Moreover, observations on either side of the cutoff and within bandwidth h are kernel-weighted based on distance from the cutoff; in this way, observations further from the cutoff get weighted less than observations closest to the cutoff. The following specifications represent the nonparametric estimate of the program impact, for school and household-level analysis, for bandwidths h,

$$\Delta Y_i = \Phi_0 + \alpha * FSSP_i + \Phi_1(L_D) + \nu_i \qquad \text{(6a – School Level)},$$

$$Y_{it} = \Phi_0 + \alpha * FSSP_i + \Phi_1(L_D) + \delta * T_t + \alpha * T_t * FSSP_i + \nu_{it} \qquad \text{(6b – Household Level)}.$$

Boys

The estimation strategy for boys uses a difference-in-difference-in-difference (DDD) framework. The treatment group is defined as boys aged 6–17 living in households where at least one girl has been exposed to the program (FSSP). There are two control groups: (1) boys aged 6–17 in nonstipend districts and (2) boys aged 6–17 in stipend districts, in households where no girl has been exposed to the program. The outcomes of interest are: (i) probability of primary school completion, (ii) probability of middle school completion, (iii) school type, a dummy variable with value 1 for private school attendance, and 0 otherwise. The DDD specification is as follows:

$$Y_{it} = ß + \gamma * FSSP_i + \delta * T_t + \alpha * T_t * FSSP_i * B_i T_t + \mu * B_i T_t + \lambda * B_i T_t * T_t + \sigma * B_i T_t * S_i + \varepsilon_{it} \qquad (7),$$

where BT has a value of 1 if the boy is in a household in a stipend district, and 0 if the boy is in a household in a nonstipend district. The coefficient α gives the average net impact of the program on boys in stipend districts. More specifications are run adding household, community infrastructure, and school-level covariates to the base specification above.

Appendix E
Additional Data

Table E.1. Summary Statistics of Covariates at Baseline

COVARIATES	*Nonstipend*	*Stipend*	*Difference*	*Standard error*	*number of observations*
Community-level Characteristics					
Rural	0.643	0.818	0.175***	(0.005)	202,947
Poverty Headcount	0.436	0.672	0.236***	(0.008)	202,694
Access to Primary School^	0.886	0.752	–0.134***	(0.013)	31,621
Access to Middle School^	0.606	0.383	–0.223***	(0.019)	31,621
Access to High School^	0.479	0.232	–0.247***	(0.018)	31,621
Access to Public Transport^	0.746	0.59	–0.156***	(0.018)	31,621
Middle School Private/Public Enrollment Ratio	0.345	0.254	–0.0908***	(0.022)	3,888
Share of Education in Total Expenditure (%)	7.762	6.692	–1.070***	(0.126)	18,848
School-level Characteristics					
Boundary wall - girls middle school^^	0.841	0.848	–0.0067	(0.010)	5,648
Drinking water - girls middle school^^	0.926	0.95	–0.0245***	(0.007)	5,654
Electricity - girls middle school^^	0.744	0.852	–0.108***	(0.012)	5,362
Toilet - girls middle school^^	0.780	0.785	–0.005	(0.011)	5,624
Student Teacher Ratio - girls middle public schools^^	32.443	26.108	–6.335***	(0.423)	5,654
Household-level Characteristics					
Household Size	6.569	6.663	0.0935**	(0.044)	30,758
Dependency Ratio	0.454	0.495	0.0407***	(0.003)	202,908
Education of Household Head[a]	4.774	3.133	–1.641***	(0.080)	30,383
Gas Connection	0.293	0.047	–0.246***	(0.008)	30,758
Electricity Connection	0.926	0.677	–0.249***	(0.011)	30,758
Visit from Health Worker	0.353	0.338	–0.0150	(0.013)	24,655
Access to Family Planning Center^	0.341	0.209	–0.132***	(0.017)	31,621
Access to Health Center/Hospital^	0.466	0.292	–0.174***	(0.019)	31,621
Access to Drinking Water Supply^	0.957	0.9753	0.0183***	(0.006)	31,621

Sources: MICS 2003; ^ = PSLM 2004; ^^ = Annual Public School Census 2003.
a. Those who have never attended school are taken to have completed zero years of schooling. ***1% significance level, **5% significance level, *10% significance level.

Table E.2. Average Program Impact on Educational Outcomes – I

	Whole Sample (12-19)				*Cohort 15-16*				*Whole Sample (12-19)*				*Cohort 15-16*			
	DD (1)	**DD (2)**	**RDD (3)**	**RDD (4)**	**DD (1)**	**DD (2)**	**RDD (3)**	**RDD (4)**	**DD (1)**	**DD (2)**	**RDD (3)**	**RDD (4)**	**DD (1)**	**DD (2)**	**RDD (3)**	**RDD (4)**
	Outcome A. Middle School Completion								***Outcome B. Middle to High School Transition***							
FSSP*Time	0.00564	0.0120	0.00509	0.0121	0.0503*	0.0590**	0.0500*	0.0590**	-0.00713	0.0108	-0.00819	0.0102	0.0350	0.0555**	0.0346	0.0554**
	(0.015)	(0.013)	(0.015)	(0.013)	(0.026)	(0.028)	(0.026)	(0.028)	(0.020)	(0.018)	(0.019)	(0.018)	(0.023)	(0.026)	(0.023)	(0.026)
FSSP	-0.0481***	0.00149	-0.0157	0.000589	-0.0914***	0.0122	-0.0420	0.0155	-0.0633***	-0.0187	-0.00373	-0.00703	-0.0940***	-0.00350	-0.0317	0.0291
	(0.015)	(0.014)	(0.020)	(0.019)	(0.024)	(0.023)	(0.029)	(0.029)	(0.023)	(0.023)	(0.030)	(0.026)	(0.022)	(0.032)	(0.027)	(0.033)
Time	-0.00974	-0.0179***	-0.00943	-0.0179***	-0.0424***	-0.0605***	-0.0423***	-0.0604***	0.0144	0.000991	0.0150	0.00126	-0.0391***	-0.0583***	-0.0390***	-0.0576***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.010)	(0.012)	(0.011)	(0.012)	(0.009)	(0.009)	(0.009)	(0.009)	(0.012)	(0.012)	(0.012)	(0.012)
L40-L	No	No	-0.00156**	0.0000874	No	No	-0.00240***	-0.000330	No	No	-0.00286**	-0.00143	No	No	-0.00303**	-0.00418*
			(0.001)	(0.001)			(0.001)	(0.002)			(0.001)	(0.001)			(0.001)	(0.002)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Age Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
No. of obs	22289	20826	22289	20826	9396	8962	9396	8962	22237	20773	22237	20773	9391	8956	9391	8956

***1% significance level **5% significance level *10% significance level. Standard errors are in parentheses.
Source: Authors' calculation
Notes: (1) Standard errors are clustered at the district level. (2) Covariates include household, community and school-level characteristics: rural/urban, birth order, dependency ratio, education of household head, access to drinking water supply, access to school, private/public enrollment ratio, access to public transport, gas connection in the house, mean per capita consumption, ratio of middle to high schools, proportion of schools with boundary wall, electricity and drinking water, and student-teacher ratio. (3) See endnote 39 for an explanation of RD estimates.

Table E.3. Average Program Impact on Educational Outcomes – II

	Whole Sample (12-19)				*Cohort 15-16*				*Whole Sample (12-19)*				*Cohort 15-16*			
	DD (1)	**DD (2)**	**RDD (3)**	**RDD (4)**	**DD (1)**	**DD (2)**	**RDD (3)**	**RDD (4)**	**DD (1)**	**DD (2)**	**RDD (3)**	**RDD (4)**	**DD (1)**	**DD (2)**	**RDD (3)**	**RDD (4)**
	Outcome C. Grade 9 Completion								***Outcome D. Grade 10 Completion***							
FSSP*Time	-0.0146	0.00563	-0.0156	0.00495	0.0264	0.0494**	0.0263	0.0495**	-0.0545**	-0.0323	-0.0560**	-0.0339	0.0301	0.0285	0.0264	0.0249
	(0.022)	(0.020)	(0.022)	(0.020)	(0.022)	(0.024)	(0.022)	(0.024)	(0.025)	(0.022)	(0.025)	(0.022)	(0.051)	(0.048)	(0.051)	(0.049)
FSSP	-0.0444*	-0.00577	0.0148	0.00656	-0.0628**	-0.0216	-0.0372	-0.00142	0.00276	0.0722***	0.0641*	0.0928***	-0.0640	0.0652	0.0134	0.144**
	(0.026)	(0.026)	(0.037)	(0.027)	(0.027)	(0.036)	(0.045)	(0.036)	(0.031)	(0.023)	(0.034)	(0.030)	(0.050)	(0.069)	(0.073)	(0.067)
Time	-0.0215	-0.0360**	-0.0210	-0.0357**	-0.0766***	-0.0955***	-0.0766***	-0.0951***	0.0669***	0.0609***	0.0677***	0.0615***	-0.0779**	-0.0707**	-0.0756**	-0.0677*
	(0.013)	(0.013)	(0.013)	(0.014)	(0.015)	(0.017)	(0.015)	(0.017)	(0.010)	(0.011)	(0.010)	(0.011)	(0.036)	(0.034)	(0.036)	(0.033)
L40-L	No	No	-0.00285**	-0.00151	No	No	-0.00124	-0.00258	No	No	-0.00296*	-0.00241	No	No	-0.00374	-0.00941**
			(0.001)	(0.002)			(0.002)	(0.003)			(0.002)	(0.002)			(0.003)	(0.004)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Age Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
No. of obs	19915	18482	19915	18482	8988	8570	8988	8570	12831	11623	12831	11623	2387	2202	2387	2202

***1% significance level **5% significance level *10% significance level. Standard errors are in parentheses.
Source: Authors' calculation
Notes: (1) Standard errors are clustered at the district level. (2) Covariates include household, community and school-level characteristics: rural/urban, birth order, dependency ratio, education of household head, access to drinking water supply, access to school, private/public enrollment ratio, access to public transport, gas connection in the house, mean per capita consumption, ratio of middle to high schools, proportion of schools with boundary wall, electricity and drinking water, and student-teacher ratio. (3) See endnote 39 for an explanation of RD estimates.

Table E.4. Average Program Impact on Labor Outcomes

	Whole Sample (12-19)				*Cohort 15-16*				*Whole Sample (15-19)*				*Cohort 15-17*			
	DD (1)	**DD (2)**	**RDD (3)**	**RDD (4)**	**DD (1)**	**DD (2)**	**RDD (3)**	**RDD (4)**	**DD (1)**	**DD (2)**	**RDD (3)**	**RDD (4)**	**DD (1)**	**DD (2)**	**RDD (3)**	**RDD (4)**
	Outcome E. Labor Force Participation								***Outcome F. Work Intensity (days per month)***							
FSSP*Time	-0.0467**	-0.0494**	-0.0465**	-0.0490**	-0.0452***	-0.0403**	-0.0452***	-0.0401**	0.897	-0.611	0.417	-0.548	0.0744	-6.088*	-0.392	-6.137*
	(0.020)	(0.020)	(0.020)	(0.020)	(0.016)	(0.017)	(0.016)	(0.017)	(1.671)	(2.037)	(1.727)	(2.077)	(2.494)	(3.107)	(2.873)	(3.139)
FSSP	0.0471***	0.0427*	0.0391*	0.0309	0.0493***	0.0528**	0.0549***	0.0431*	-0.632	-3.405**	-3.465**	-4.046**	-0.309	3.631	-2.296	3.342
	(0.016)	(0.022)	(0.023)	(0.023)	(0.014)	(0.024)	(0.019)	(0.025)	(1.192)	(1.292)	(1.481)	(1.644)	(1.530)	(4.731)	(2.176)	(5.234)
Time	0.0193	0.0188	0.0192	0.0186	0.0222*	0.0226*	0.0222*	0.0224*	-0.711	1.393	-0.351	1.294	-1.699	3.380	-1.351	3.361
	(0.013)	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(1.281)	(1.828)	(1.306)	(1.841)	(2.253)	(2.514)	(2.566)	(2.533)
L40-L	No	No	0.000387	0.00202**	No	No	-0.000270	0.00180**	No	No	0.178**	0.0711	No	No	0.145	0.0390
			(0.001)	(0.001)			(0.001)	(0.001)			(0.072)	(0.163)			(0.160)	(0.305)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Age Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
No. of obs	27448	26037	27448	26037	9090	8672	9090	8672	292	273	292	273	125	120	125	120

***1% significance level **5% significance level *10% significance level. Standard errors are in parentheses.
Source: Authors' calculation
Notes: (1) Standard errors are clustered at the district level. (2) Covariates include household, community and school-level characteristics: rural/urban, birth order, dependency ratio, education of household head, access to drinking water supply, access to school, private/public enrollment ratio, access to public transport, gas connection in the house, mean per capita consumption, ratio of middle to high schools, proportion of schools with boundary wall, electricity and drinking water, and student-teacher ratio. (3) See endnote 39 for an explanation of RD estimates.

Table E.5. Average Program Impact on Marriage and Fertility Outcomes

	Whole Sample (15-19)				*Whole Sample (15-19)*				*Whole Sample (17-19)*				*Whole Sample (17-19)*			
	DD (1)	DD (2)	RDD (3)	RDD (4)	DD (1)	DD (2)	RDD (3)	RDD (4)	DD (1)	DD (2)	RDD (3)	RDD (4)	DD (1)	DD (2)	RDD (3)	RDD (4)
	Outcome G. Probability of Marriage				*Outcome H: Age at Marriage*				*Outcome I. Probability of Giving Birth*				*Outcome J. Number of Children*			
FSSP*Time	0.00953	0.00820	0.00940	0.00814	-0.151	1.460**	-0.0767	1.222*	-0.0114	-0.0808	-0.0257	-0.0691	0.0962	-0.329*	0.0663	-0.349*
	(0.009)	(0.008)	(0.009)	(0.008)	(0.388)	(0.621)	(0.458)	(0.643)	(0.121)	(0.172)	(0.122)	(0.173)	(0.150)	(0.181)	(0.149)	(0.209)
FSSP	0.00962	0.0132	0.0185	0.0143*	-0.202	-0.0353	-0.588	0.230	0.0232	0.0825	0.129	0.0396	-0.0387	0.305*	0.182	0.358*
	(0.008)	(0.008)	(0.011)	(0.008)	(0.391)	(0.439)	(0.865)	(0.460)	(0.099)	(0.181)	(0.154)	(0.197)	(0.122)	(0.158)	(0.182)	(0.209)
Time	-0.00695*	-0.00506	-0.00689*	-0.00503	1.689***	0.803	1.610***	0.697	-0.0188	-0.152	0.00204	-0.157	-0.0613	0.0967	-0.0180	0.112
	(0.003)	(0.004)	(0.004)	(0.004)	(0.285)	(0.478)	(0.359)	(0.488)	(0.072)	(0.130)	(0.074)	(0.131)	(0.091)	(0.088)	(0.088)	(0.130)
L40-L			-0.000430	-0.000164			0.0173	-0.0538*	No	No	-0.00511	0.00789	No	No	-0.0106	-0.00840
			(0.000)	(0.000)			(0.028)	(0.031)			(0.006)	(0.012)			(0.008)	(0.016)
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Age Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
No. of obs	19177	17761	19177	17761	339	296	339	296	392	342	392	342	392	342	392	342

***1% significance level **5% significance level *10% significance level. Standard errors are in parentheses.

Source: Authors' calculation

Notes: (1) Standard errors are clustered at the district level. (2) Covariates include household, community and school-level characteristics: rural/urban, birth order, household size, dependency ratio, education of household head, access to school, private/public enrollment ratio, access to transport, gas connection in the house, mean per capita consumption, ratio of middle to high schools, status of school facilities, student-teacher ratio. (3) Additional covariates for fertility outcomes include access to Lady Health Worker, and access to Family Planning Centre (4) See endnote 39 for an explanation of RDD estimates.

Table E.6. Heterogeneity of Impacts

	Rural	*Poverty*	*Parental Education*		*Exposure*	*Age*	
			None	*Primary*	*One*	*12–14*	*15–16*
Middle-School Completion	–0.0558*	–0.0594**	–0.0813**	0.0161	–0.0144	0.0434	0.0910**
	(0.030)	(0.027)	(0.039)	(0.051)	(0.049)	(0.036)	(0.037)
Middle to High Transition	–0.0707*	–0.0462	–0.0508	0.000510	–0.0318	0.0507	0.0834**
	(0.039)	(0.037)	(0.049)	(0.046)	(0.059)	(0.045)	(0.034)
Grade-9 Completion	–0.0944*	–0.0419	–0.0417	0.0403	–0.0727	0.0490	0.0796**
	(0.048)	(0.046)	(0.069)	(0.059)	(0.070)	(0.071)	(0.031)
Grade-10 Completion	–0.0820**	–0.0577	0.00530	–0.0983	–0.0847	0.0309	0.0849
	(0.033)	(0.038)	(0.052)	(0.091)	(0.060)	(0.173)	(0.054)
Labor Force Participation	–0.0605**	–0.0390	–0.0168	0.0118	–0.0361*	0.0658	0.0401
	(0.027)	(0.038)	(0.024)	(0.023)	(0.019)	(0.042)	(0.035)
Work Intensity	–3.807	4.403	–0.0298	0.00912	–4.956		–5.776*
	(4.660)	(3.494)	(0.032)	(0.037)	(6.414)		(3.316)
Married	–0.00610	0.0188	–0.0410**	–0.000156	0.0550***		0.000657
	(0.015)	(0.015)	(0.016)	(0.023)	(0.017)		(0.015)
Age at Marriage	1.441	–2.822*	0.110	1.355	0.204		
	(1.410)	(1.402)	(0.738)	(0.810)	(0.984)		
Gave birth	–0.0471**	–0.0291	0.00114	–0.0336	0.0660***		
	(0.022)	(0.019)	(0.018)	(0.030)	(0.021)		
Number of Children	–0.531	–0.378	0.330	–0.406	0.239		
	(0.451)	(0.377)	(0.257)	(0.506)	(0.369)		

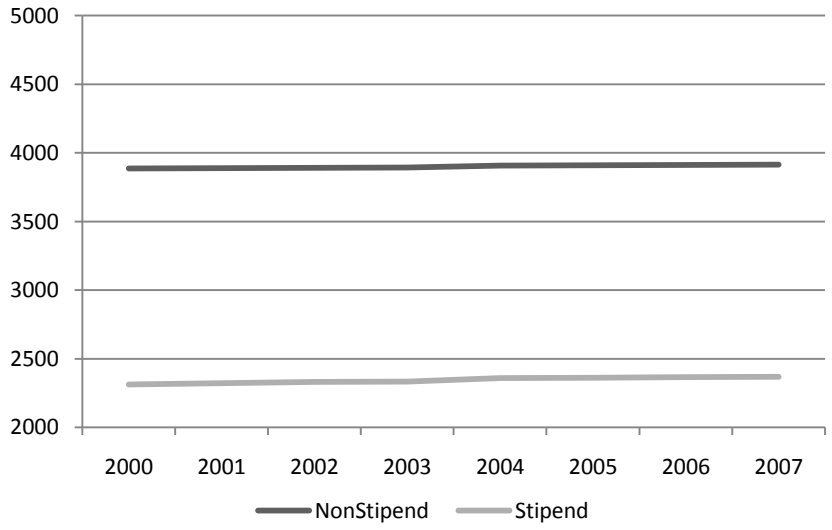
Source: Authors' calculations
Note: (1) Standard errors for urban, birth order, household size, are clustered at the district level. (2) All specifications included household, community, and school-level covariates: rural/dependency ratio, education of household head, access to drinking water supply, access to school, private/public enrollment ratio, access to public transport, gas connection in the house, mean per capita consumption, ratio of middle to high schools, proportion of schools with boundary wall, electricity and drinking water, and student-teacher ratio, among others.
***1% significance level, **5% significance level, *10% significance level. Standard errors are shown in parentheses.

Table E.7. Robustness Check: Preprogram Trends

	Prebaseline (PIHS 2001) vs.bBaseline (MICS 2003)				
	Middle-school completion	Grade-9 completion	High-school completion	Labor force participation	Married
Interaction	0.0228	–0.103	–0.0271	–0.0593	–0.0127
	(0.066)	(0.063)	(0.060)	(0.041)	(0.018)
	Older cohort comparison				
	Middle-school completion	Grade-9 completion	High-school completion		
Interaction	–0.00638	–0.0441	–0.0418		
	(0.026)	(0.044)	(0.045)		

Source: Authors' calculations
Note: Coefficients shown correspond to the interaction between the dummy variable for stipend districts and the dummy t for time (t=1 for observations at baseline, t=0 at prebaseline)
***1% significance level, **5% significance level, *10% significance level. Standard errors are shown in parentheses.

Figure E.1. Total Number of Public Schools in Stipend and Nonstipend Districts, 2000–07



Source: Punjab Public School Census (2003–07).

Endnotes

1. Evidence from two CCTs in Mexico and Cambodia, for instance, suggests that the programs increased grade progression but did not have impacts on scores in mathematics and language tests (Behrman and others 2005b; Filmer and Schady 2009).

2. Information on learning outcomes (test scores) is available and will be assessed in future research to examine the impacts of the Girls' Stipend Program (GSP) on the academic achievement of girls in the program.

3. Estimates for Pakistan and the region were obtained from published documents "Pakistan at a Glance" 2004 available online; Punjab estimates were derived from the MICS 2003–04 Report, table titled "Major Summary Results."

4. Around 90 percent of private schools offer only primary classes (Andrabi and others 2006).

5. Main sources for this section are: Pakistan Education Sector Reform, Action Plan 2000–2003 [[; World Bank 2007.

6. The gender disparity in Pakistan is well established in the literature; for example, authors such as Khan (1993), Behrman and Schneider (1993), Alderman, Behrman, and others (1995), Alderman, Orazem, and Paterno (2001), Holmes (2003), Lloyd, Mete, and Sathar (2005), and a gender assessment by the World Bank (2005) make this point clearly.

7. Only some of these determinants are relevant for boys schooling participation.

8. Some examples of specific activities planned and implemented within each pillar of the PESRP include the revitalization of school councils, merit-based new teacher recruitment and professional development, distribution of free text books to all public school students, provision of support for low-cost private education, and monetary incentives in the form of school stipends to girls living in low-literacy districts. There are other activities related to improving schooling infrastructure in order to link the supply of schools to the existing student demand; for example, classrooms and toilets, often missing from schools, are among the facilities addressed under the reform. Another critical activity undertaken was to strengthen data collection for monitoring progress and evaluating ongoing results. This activity is critical for measuring impacts throughout the implementation of the reform and after its completion.

9. The government's implementation completion report (World Bank 2007) reported that at least 32,000 schools were rehabilitated in some way during the early part of the reform period. The schools were selected based on need.

10. Since 2006 a public subsidy to low-cost private schools has been in place; this program is known as Stipend to Foundation Assisted Schools and covers girls in approximately 50 percent of the GSP program. Additionally, an education voucher scheme was implemented throughout Punjab, but targeting and modality are independent of the program.

11. It is estimated that the costs of schooling per quarter are approximately $9, leaving some funds for the family to use for other needs (Chaudhury and Parajuli 2006). Moreover, after the program's start, free textbooks have been provided to students in middle school and higher.

12. In theory, all public schools in Punjab are single gender; however, the data show some mixed schools exist, an issue that is accounted for in the estimation approach.

13. The up-take rate is difficult to calculate because program eligibility is determined by the grade, and not the age, of girls in stipend districts. For simplicity, this analysis estimates the up-take rate as the ratio of participants (245,000 girls covered by the program in 2007) relative to the sum of the

actual and potential participants (those enrolled in grades 6–8 in private schools and those whose highest grade is 5-7 but were not enrolled at the time of the survey).

14. This is calculated based on the number of beneficiaries and the average quarterly stipend. It does not include administrative or transaction costs, which are difficult to disaggregate from the general costs of the wider educational reform.

15. This method is referred to as difference-in-difference estimation (DD).

16. This method compares the difference-in-difference for female enrollments with the difference-in-difference for male enrollments, resulting in a triple difference estimator (DDD). While the DD estimator has the advantage of eliminating the influence of time-invariant characteristics in both the stipend and nonstipend districts that otherwise might affect enrollments, it does not eliminate the influence of time-varying factors. The DDD method "provides a cleaner way to 'separate out' some of the bias from the differential growth effects that arise due to gaps in initial literacy rates" and also eliminates the influence of the differential effects of other program interventions (Chaudhury and Parajuli 2006).

17. In addition to DD and DDD, the authors used a regression discontinuity design (RDD) model that exploits the literacy criterion cut-off of 40 percent for distinguishing between stipend and nonstipend districts to estimate the impact of the female secondary school stipends on enrollment.

18. Based on the simple model specifications.

19. Including domestic work, work in family business, and work outside of the home.

20. Another potential concern for identification strategy is that school enrollment/attendance and labor force participation are joint decisions. However, this is problematic for analyses that focus on the relationship between the use of school inputs (enrollment and attendance) and labor market participation. In contrast, this paper exploits the extra incentive provided by the stipend and the conditionality to send girls to school, which this analysis argues to be exogenous after controlling for some observables at their baseline values, time-invariant unobserved factors between stipend and nonstipend districts (through DD method), and the exogenous variation in program eligibility around the literacy cutoff (through RDD).

21. It was not possible to also use matching techniques with the pool of cross-sectional data available. The only common level of aggregation between baseline data—which would be used to predict program participation and construct the propensity score for each household—and follow-up data was the district. This implies that matching is only possible at the district level, and the province has only 34 districts.

22. These samples are restricted to girls who finished primary school (having at least five years of schooling). Some econometric exercises show that the GSP program did not have indirect effects on either enrollment in grades 4 and 5 or completion of primary school.

23. The school-level approach measures changes in absolute enrollment between stipend and nonstipend districts. These changes, however, may be driven by a potential relationship among higher fertility, population growth, and the number of girls enrolled, given that stipend districts are poorer. Unfortunately, there are no school-level data available to do the estimation for a preprogram period and check if changes in enrollment in stipend district occurred even before the program began. Yet, the empirical models include a number of socioeconomic covariates at their baseline values, which are also likely to explain differences in fertility between stipend and nonstipend districts. Furthermore, the analysis at the household level looks at changes in progression and completion rates, which, by definition, are not affected by the size of the cohorts.

24. In general, an intent-to-treat parameter measures the impact of a program comparing the relative changes of an outcome between people eligible for the treatment (regardless of their up-take, subsequent withdrawal, or deviation from the treatment protocol) and the comparison group. In the context of the GSP, the intent-to-treat analysis measures changes in female enrollment of public schools in stipend districts relative to public schools in nonstipend districts, even though not all girls in the stipend districts participated.

25. Both rejoiners and repeaters constituted around 0.5 percent of total enrollment in middle school at baseline. Not only were these rates similar between stipend and nonstipend districts in 2003–04, there has been no differential growth in these rates between 2003–04 and 2007–08. School switching is another source of compositional change within the cohort. At baseline, around 2 percent of all children enrolled in middle schools switched from private to public schools, while a similar proportion switched from public to private (the difference is statistically insignificant). So, at baseline, compositional shifts stemming from reentry, repetition, and school switching accounted for some 5 percent of total enrollment at the middle school level. There are no comparable data on school switching in the 2007–08 survey, but the change in repetition and rejoining rate at the middle school level remained comparable between stipend and nonstipend districts, between 2003–04 and 2007–08.

26. The previous impact evaluation (Chaudhury and Parajuli 2006) used cohorts of children 10–14 years old from cross-sectional household surveys to triangulate the findings derived from school data. However, an examination of the age-grade distribution of girls in middle school in Punjab indicates that this cohort has errors of inclusion and exclusion. More specifically, the youngest girls of this cohort (10–11 years old) have a high probability to remain in primary school, whereas a substantial fraction of girls 15–17 who attend school are enrolled in middle school.

27. These data are also collected in other provinces in Pakistan; however, each province has a distinct source of funding and organizational arrangement, depending on its characteristics. In Punjab, UNESCO (United Nations Educational, Scientific, and Cultural Organization) and the UNDP (United Nations Development Programme) financially support the survey.

28. This survey does not cover private and religious schools in Punjab; a private school census was undertaken in 1999–2000, on the basis of which it was estimated that over one-quarter of all students were enrolled in private education. A National Education Census, also conducted in 2006, enumerated all public and private schools.

29. The sample of eligible schools used in the analysis includes only girls' schools that offered grades 6–8 at both baseline and follow-up (less than 1 percent of middle schools are mixed). Schools that were nonfunctional, upgraded to middle level or above, or downgraded to primary or lower level during the program implementation period were excluded, because these would not have been eligible for the stipend at both baseline and follow-up. Of the remaining schools, only those schools that had nonzero enrollment in grades 6–8 at both baseline and follow-up were retained.

30. Cohabitation is very uncommon and usually disapproved of by parents of the couple.

31. Safety concerns are an important driver of program participation. Indeed, parents are more willing to enroll adolescent girls in school if it is relatively safe for the girls to travel to school. Safety in the context of adolescent girls' participation in the program can be determined by multiple factors, such as access to school, access to public transport, distribution and quality of road network, frequency of public transport service, and the safety of the road itself. The empirical models control for access (distance) to school and public transport, but there is no information available on the other three factors.

32. This is an imperfect measure of school quality but no better data were available. Regression models also include specifications without this covariate.

33. Indicators that measure the probability of completing middle and high schools are only partial measures of school attainment.

34. For instance, impacts are likely to vary depending on the grade levels of the girls when the program started.

35. A recent synthesis of the impact evaluation literature on safety nets shows that 80–88 percent of the CCT programs rigorously evaluated have positive impacts on grade progression, repetition, and dropout rates (World Bank 2010).

36. Qualitative evidence based on interviews with officials in charge of implementing the program indicates that campaigns to raise the awareness of the existence of the program began only in the second year of the program.

37. The parameter of interest for the models discussed here is changes in the number of girls relative to the number of girls enrolled at baseline. For instance, if the numbers of girls enrolled increased by five (from 16 to 21), this corresponds to a 32 percent change in enrollment. This, however, should not be confused with a change in enrollment rate in percentage points. Suppose that the universe of girls that should be enrolled in middle school is 100, the increase in enrollment rate in percentage points for the above scenario would be around five, from 16 percent to 21 percent.

38. Baseline data shows that repetition and rejoining rates are low for grades 7 and 8.

39. The estimates from the RDD models shown in tables 4 to 8 were conducted in a DD framework for the whole sample, with a control function of district literacy rates that determine program eligibility. Results based on pure RDD parametric and nonparametric models and outside of a DD framework, using optimal bandwidth to select the subsample of analysis, are available from the authors upon request.

40. The distribution of school attainment (measured as the number of years of schooling) for adolescents and adults exhibits a significant drop between grades 8 and 9 in both stipend and nonstipend districts.

41. Women in Pakistan typically begin having children shortly after marriage.

42. A caveat of this analysis is that the follow-up survey has a significant number of missing values for the number of births. Further calculations show that this issue is equally prevalent in both stipend and nonstipend districts. Moreover, the problem of missing values does not appear to be systematically correlated with some determinants of fertility.

43. Estimates from the DD and RDD models shown in tables 4 to 8 were obtained from a DD framework for the whole sample, with a control function of district literacy rates that determine program eligibility. The fact that the results are very similar may suggest that there is low impact heterogeneity across districts. Yet, this does not imply that there is not impact heterogeneity for different types of subgroups, defined based on some socioeconomic characteristics.

44. The scholarship program in the Cambodia Education Sector Support Project provides modest transfers to eligible girls that are equivalent to approximately 2–3 percent of the total expenditures of the average recipient households. The relative size of the GSP in Pakistan is similar. The monthly stipend represents 3.4 percent of median household expenditures of the recipient households in 2004 (Ferreira and others 2009; Chaudhury and Parajuli 2006).

45. Results that look at the average years of schooling completed do not reveal any effects either. Comparisons for school-age boys at baseline and follow-up indicate that this result is not driven by systematic differences in age-grade distributions.

46. Aslam (2009) also found that boys are more likely than girls to enroll in private schools in Punjab.

47. School fees in the private sector may increase more than three-fold at the secondary level.

48. A limitation of this analysis is that PIHS 2001, the household survey used for the prebaseline period, is representative only at the province level. The survey was structured around three main domains: rural, major cities, and other urban areas. The first two domains (rural and major cities) were stratified by district and, thus, can be used to construct the treatment status of girls. The "other urban" domain was stratified by division, a higher administrative level that may include both stipend and nonstipend districts. Therefore, the DD analysis is restricted to the "rural" and "major cities" domains and assumes that the sample statistics based on PIHS 2001 are representative for the two groups of analysis as a whole (that is, stipend and nonstipend districts).

49. This happens in the case of classic measurement error where the difference between the actual value of the explanatory variable and the observed is pure random noise.

50. Filmer and Schady (2009) find evidence of this type of selection to explain why a scholarship program targeted to poor students in Cambodia does not appear to have positive effects on academic tests.

51. An alternative to correct for this type of selection is to construct either parametric or nonparametric bounds on quantile-specific program impacts, based on symmetric truncation of the distribution of the treatment and control groups. See Angrist and others (2004).

52. The sample size of these models is reduced in some cases by more than 50 percent.

53. At baseline, there were no statistically significant differences between stipend and nonstipend districts in the proportion of girls switching from public to private vs. private to public at the middle school level

54. This bigger cohort could be problematic because it may include girls who completed grades 5, 6, and 7 but stopped going to school well before the program started.

55. Girls are eligible for the program only if they enroll and attend a public school.

56. Only 14 new girls' middle/high/higher secondary schools were built during 2004–07, of which 8 were located in stipend districts.

57. As mentioned before, participation in the program is based on enrollment in grades 6–8; moreover, there is no restriction on the girls being a certain age for them to enroll in grades 6–8.

58. We also do not find evidence of increase in grades 4–5 enrollment (enrollment in these grade levels are more likely to be affected by the program) or primary school completion in stipend districts, relative to nonstipend districts over time for girls age 12–17

59. Because the 2007–08 survey was conducted two months after the beginning of the academic year 2007–08, the girls enrolled in grade 6 have yet to receive a stipend and are not considered exposed.

60. The age-grade distribution at baseline is used for two reasons: (1) The program impacts in the first few years were low because there was very little awareness of the program, the program had not been scaled up, and implementation maturity had yet to be reached so the age-grade distribution during 2004–05 and 2005–06 is better proxied by age distribution at baseline than at follow-up. (2) The age-grade distribution at follow-up is fairly similar at baseline, meaning that there have been no major shifts in age-grade distribution during the time period.

61. Optimal bandwidth was determined using a jackknife, cross-validation procedure.

References

Ahmed, A.U., M. Adato, A. Kudat, D. Gilligan, T. Roopnaraine, and R. Colasan. 2007. "Impact Evaluation of the Conditional Cash Transfer Program in Turkey: Final Report," International Food Policy Research Institute, Washington, DC, in collaboration with the AGRIN Co., Ankara, Turkey.

Alderman, H., J. Behrman, S. Khan, D. Ross, and R. Sabot. 1995. "Public Schooling Expenditures in Rural Pakistan: Efficiently Targeting Girls in a Lagging Region." In*Public Spending and the Poor: Theory and Evidence,* eds. D. van de Walle and K. Nead. Baltimore, Maryland: Johns Hopkins University Press.

Alderman, H., P. Orazem, and E. M. Paterno. 2001. "School Quality, School Cost, and the Public/Private School Choices of Low-Income Households in Pakistan." Staff General Research Paper 1970, Department of Economics, Iowa State University, Ames, Iowa.

Andrabi, T., J. Das, and A. I. Khwaja. 2006. "Students Today, Teachers Tomorrow?: The Rise of Affordable Private Schools." Department of Economics, Ponoma College, Claremont, California.

Angrist, J., E. Bettinger, and M. Kremer. 2004. "Long-Term Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia." NBER Working Paper 10713, National Bureau of Economic Research, Cambridge, Mass.

Aslam, M. 2009. "The Relative Effectiveness of Government and Private schools in Pakistan: Are Girls Worse Off?" *Education Economics*, 17 (3): 329–354.

Attanasio, O., E. Fitzsimmons, and A. Gómez. 2005. "The Impact of a Conditional Education Subsidy on School Enrollment in Colombia." Report Summary: Familias 01. London: Institute for Fiscal Studies.

Attanasio, O., L. C. Gómez, P. Heredia, and M. Vera-Hernández. 2005. "The Short-Term Impact of a Conditional Cash Subsidy on Child Health and Nutrition in Colombia." Report Summary: Familias 03. London: Institute for Fiscal Studies.

Attanasio, O., and A. Mesnard. 2005. "The Impact of a Conditional Cash Transfer Programme on Consumption in Colombia." Report Summary: Familias 02. London: Institute for Fiscal Studies.

Baird, S., E. Chirwa, C. McIntosh, and B. Ozler. 2009. "The Short-Term Impacts of a Schooling Conditional Cash Transfer Program on the Sexual Behavior of Young Women," Policy Research Working Paper 5089, World Bank, Washington, DC.

Baird, S., C. McIntosh, and B. Ozler. 2009. "Designing Cost-Effective Cash Transfer Programs to Boost Schooling among Young Women in Sub-Saharan Africa," Policy Research Working Paper 5090, World Bank, Washington, DC.

Barrera-Osorio, F., M. L. Bertrand, L. Linden, and F. Perez-Calle. 2008. "Conditional Cash Transfers in Education: Design Features, Peer and Sibling Effects Evidence from a

Randomized Experiment in Colombia," Policy Research Working Paper 4580, World Bank, Washington, DC.

Behrman, J. R., J. Gallardo-García, S. W. Parker, P. E. Todd, and V. Velez-Grajales. 2005a. "How Conditional Cash Transfers Impact Schooling and Working Behaviors of Children and Youth in Urban Mexico."Unpublished.

Behrman, J. R., and J. Hoddinott. 2001. "An Evaluation of the Impact of PROGRESA on Pre-School Child Height." Food Consumption and Nutrition Division Discussion Paper No. 104, International Food Policy Research Institute, Washington, DC.

Behrman, J. R., S. W. Parker, and P. E. Todd. 2005b. "Long-Term Impacts of the *Oportunidades* Conditional Cash Transfer Program on Rural Youth in Mexico." Discussion Paper 122, Ibero-America Institute for Economic Research, Göttingen, Germany.

Behrman, J. R., and T. Schneider. 1993. "An International Perspective on Pakistani Human Capital Investments in the Last Quarter Century." *The Pakistan Development Review* 32 (1): 1–68.

Behrman, J. R., P. Sengupta, and P. E. Todd. 2005. "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico." *Economic Development and Cultural Change* 54 (1): 237–75.

Chaudhury, N. and D. Parajuli. 2006. "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Program on Public School Enrollments in Punjab, Pakistan." Policy Research Working Paper 4102, World Bank, Washington, DC.

———. 2008. "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Program on Public School Enrollments in Punjab, Pakistan." Journal of Applied Economics 42 (28): 3565–3583.

Das, J., P. Pandey, and T. Zajonc. 2006. "Learning Levels and Gaps in Pakistan," World Bank Policy Research Working Paper 4067, World Bank, Washington, DC.

de Janvry, A., F. Finan, and E. Sadoulet. 2006. "Evaluating Brazil's *Bolsa Escola* Program: Impact on Schooling and Municipal Roles," Department of Agricultural and Resource Economics, University of California, Berkeley.

Dubois, P., and M. Rubio-Codina. 2009. "Child Care Provision: Semiparametric Evidence from a Randomized Experiment in Mexico." CEPR Discussion Paper No. DP7203, Centre for Economic Policy Research, London. Available at SSRN: http://ssrn.com/abstract=1356432.

Ferreira, F. H. G., D. Filmer, and S. Norbert. 2009. "Own and Sibling Effects of Conditional Cash Transfer Programs: Theory and Evidence from Cambodia," Policy Research Working Paper 5001, World Bank, Washington, DC.

Filmer, D. and N. Schady. 2006. "Getting Girls into School: Evidence from a Scholarship Program in Cambodia." Policy Research Working Paper 3910, World Bank. Washington, DC.

———. 2009. "School Enrollment, Selection and Test Scores." Policy Research Working Paper Series 4998. Washington, DC: World Bank.

Gertler, P. 2000. "Final Report: The Impact of PROGESA on Health," International Food Policy Research Institute, Washington, DC.

Glewwe, P. and P. Olinto. 2004. "Evaluating the Impact of Conditional Cash Transfers on Schooling: An Experimental Analysis of Honduras," Final report for US Agency for International Development. http://pdf.usaid.gov/pdf_docs/PNADT588.pdf.

Government of Pakistan. 2001. "Pakistan Education Sector Reform, Action Plan 2000–2003", http://siteresources.worldbank.org/PAKISTANEXTN/Resources/Pakistan-Development-Forum/EduSectorReforms.pdf

Government of Punjab. 2004. "District Based Multiple Indicators Cluster Survey 2003-04 Punjab", survey report prepared by Government of Punjab in collaboration with the Federal Board of Statistics and United Nations Children's Fund, pp. 7

Gulemetova-Swan, M. 2009. "Evaluating the Impact of Conditional Cash Transfer Programs on Adolescent Decisions about Marriage and Fertility: The Case of *Oportunidades*." Dissertation available from ProQuest, Paper AAI3363363. http://repository.upenn.edu/dissertations/AAI3363363. Hoddinott, J., E. Skoufias, and R. Washburn. 2000. "The Impact of PROGRESA on Consumption: A Final Report," International Food Policy Research Institute, Washington, DC.

Holmes, J. 2003. "Measuring the Determinants of School Completion in Pakistan: Analysis of Censoring and Selection Bias." *Economics of Education Review* 22 (3): 249–264.

Imbens, G,. T. Lemieux. 2008. "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, vol. 142(2), pages 615-635, February

Imbens, G., J. M. Wooldridge. 2003. ""Recent Developments in the Econometrics of Program Evaluation," NBER Working Paper 14251

Institute for Fiscal Studies (IFS), Econometría, and Sistemas Especializados de Información (SEI). 2006. "Evaluación del Impacto del Programa Familias en Acción–Subsidios Condicionados de la Red de Apoyo Social." Bogotá, Colombia: National Planning Department.

Khan, S. R.1993. "South Asia." In *Women's Education in Developing Countries: Barriers, Benefits and Policy,* eds. E. M. King and M. A. Hill. Published for the World Bank. Baltimore, Maryland: John Hopkins University Press.

Khandker, S., M. M. Pitt, and N. Fuwa. 2003. "Subsidy to Promote Girls' Secondary Education: The Female Stipend Program in Bangladesh." Unpublished, University Library of Munich, Germany. MPRA Paper 23688.

Lee, D., T. Lemieux. 2009. "Regression Discontinuity Designs in Economics," NBER Working Paper

Lloyd, C. 2004. "The Changing Transitions to Adulthood in a Comparative Perspective: The Case of Pakistan." *Pakistan Development Review* 43 (4): 441–467.

Lloyd, C.B., C. Mete, and M. Grant. 2007. "Rural Girls in Pakistan: Constraints of Policy and Culture." In *Exclusion, Gender and Education: Case Studies from the Developing World*, eds. Maureen Lewis and Marlaine Lockheed, 99–118. Washington DC: Center for Global Development.

Lloyd, C., C. Mete, and Z. Sathar. 2005. "The Effect of Gender Differences in Primary School Access, Type, and Quality on the Decision to Enroll in Rural Pakistan." *Economic Development and Cultural Change* 53 (3): 685–710.

Maluccio, J. A., and R. Flores. 2005. "Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social." Research Report 141. Washington, DC: International Food Policy Research Institute.

Olinto, P., S.S. Morris, R. Flores, and A. Veiga. 2003. "The Impact of the Bolsa Alimentacao Program on Food Consumption." Inter-American Development Bank Discussion Paper S-236, IADB, Washington, DC.

Ravallion, M. 2005, "Evaluating Anti-Poverty Programs." Handbook of Agricultural

Economics vol. 4. Robert E. Evenson and T. Paul Schultz

Sathar, Z. and C. Lloyd. 1994. "Who Gets Primary Schooling in Pakistan: Inequalities Among and Within Families." *Pakistan Development Review* 33 (2): 103–134.

Sathar, Z., C. Lloyd, C. Mete, and M. ul Haque. 2003. "Schooling Opportunities for Girls as a Stimulus for Fertility Change in Rural Pakistan." *Economic Development and Cultural Change* 51 (3): 677–698.

Sawada, Y. and M. Lokshin. 2009. "Obstacles to School Progression in Rural Pakistan: An Analysis of Gender and Sibling Rivalry Using Field Survey Data." *Journal of Development Economics* 88 (2): 335–347.

Schultz, T. P. 2004. "School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program." *Journal of Development Economics* 74 (1): 199–250.

Veras-Soares, F., R. Perez-Ribas, and G. I. Hirata, 2008. "Achievements and Shortfalls of Conditional Cash Transfers: Impact Evaluation of Paraguay's Tekoporã Programme." IPC Evaluation Note 3, International Policy Centre for Inclusive Growth, United Nations Development Programme. http://www.ipc-undp.org/pub/IPCEvaluationNote3.pdf.

World Bank. 2005. "Improving Gender Outcomes: The Promise for Pakistan." Washington, DC: The World Bank.

———. 2005. "Pakistan At a Glance", http://siteresources.worldbank.org/SOUTHASIAEXT/Resources/223546-1189611264671/pak_aag.pdf

2007. "Implementation Completion and Results Report for the Education Sector Development Policy Credit." Washington, DC: World Bank.

———. 2010. “Evidence and Lessons Learned from Impact Evaluations in Social Safety Nets.” Presentation, World Bank, Washington, DC. http://go.worldbank.org/D3O7LZS1L0.