DISCUSSION PAPER SERIES

IZA

IZA DP No. 5703

Does Cash for School Influence Young Women's Behavior in the Longer Term? Evidence from Pakistan

Andaleeb Alam
Javier E. Baez
Ximena V. Del Carpio

May 2011

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Does Cash for School Influence Young Women's Behavior in the Longer Term? Evidence from Pakistan

Andaleeb Alam
World Bank

Javier E. Baez
World Bank
and IZA

Ximena V. Del Carpio
World Bank

Discussion Paper No. 5703
May 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 5703
May 2011

ABSTRACT

Does Cash for School Influence Young Women's Behavior in the Longer Term? Evidence from Pakistan*

The Punjab Female School Stipend Program, a female-targeted conditional cash transfer program in Pakistan, was implemented in response to gender gaps in education. An early evaluation of the program shows that the enrollment of eligible girls in middle-school increased in the short term by nearly 9 percentage points. This paper uses regression discontinuity and difference-in-difference analyses to show that five years into the program implementation positive impacts do persist. Beneficiary adolescent girls are more likely to progress through and complete middle school and work less. There is suggestive evidence that participating girls delay their marriage and have fewer births by the time they are 19 years old. Also, girls who are exposed to the program later-on, and eligible for the benefits given in high school, increase their rates of matriculating into and completing high school. The persistence of impacts can potentially translate into gains in future productivity, consumption, inter-generational human capital accumulation and desired fertility. Lastly, there is no evidence that the program has negative spillover effects on educational outcomes of male siblings.

JEL Classification: I25, J13, J21, O15

Keywords: conditional cash transfers, female education, female labor participation, fertility, Pakistan

Corresponding author:

Javier Baez
The World Bank
1818 H Street, NW
MSC 9-006A
Washington, DC 20433
USA
E-mail: jbaez@worldbank.org

* Special thanks go to Mark Sundberg, Emmanuel Skoufias, John Hoddinott, Andrew Warner, Martha Ainsworth and Kathleen Beegle for insightful comments. Tu Chi Nguyen provided superb research assistance. Special recognition goes to operational staff in the World Bank, in particular Dhushyanth Raju, Sofia Shakil and Mariam Qaiser. We also thank various government agencies and research centers in Islamabad and Lahore, Pakistan, especially Muhammad Asif and Sohail Malik (PMIU), Shamim Rafique (Bureau of Statistics), Minhaj ul Haque (Population Council), and Feyza Bhatti (Mahbub Ul Haq Human Development Center). Many thanks go to many others inside and outside the World Bank who provided helpful comments. We acknowledge financial support from the Norwegian Agency for Development Cooperation. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

IZA Discussion Paper No. 5703
May 2011

NON-TECHNICAL 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.

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.

I. Introduction

Many developing and transition countries have implemented some modality of conditional cash transfer (CCT) program to incentivize poor households to make investments in the human capital of their children. These programs are a popular tool for alleviating short-term poverty and reducing the inter-generational transmission of poverty. Though CCTs have some basic elements, they vary in terms of scale, transfer size, conditionalities, eligibility, and many implementation features. In spite of these differences, evidence arising from impact evaluations indicates that most of these programs fulfill their short-term objectives through increases in immediate consumption and use of educational and health services. Indeed, a recent literature review shows that CCTs increased the consumption of participant households and 87 percent raised the school enrollment and attendance of children (World Bank, 2010).

However, little is known about the longer-term impacts of these programs on human capital accumulation and measures of human welfare. Most evaluations to date only assess immediate changes in behaviors related to school attendance, enrollment and visits to health centers. While these investments are critical for the accumulation of human capital of children, such changes say little about whether the accumulation is taking place and whether immediate impacts also influence the future prospects of children in their progress in school, productivity and employability (Filmer and Schady, 2009).[1] Likewise, the evidence is thin regarding indirect effects on outcomes such as marriage and fertility decisions, which have plausible implications for the long term on the welfare of the beneficiaries.

This paper contributes to the literature by focusing on the impacts of a gender targeted CCT program in the longer-term. The program investigated is the *Female School Stipend Program* (FSSP) which was implemented in the province of Punjab, in northeast Pakistan. The evaluation identification strategy relies on a regression discontinuity design. The analysis focuses on examining impacts on school progression, middle and high school completion, early labor market outcomes, marriage and fertility decisions of adolescent girls who have been participating in the program for up to four years. The paper also measures distributional impacts

[1] Evidence from 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.

across various participant groups and explores spillover effects of the program on the enrollment of boys.

Results shows that after four years of implementation, participant girls are three to six percentage points more likely to complete middle school; and younger cohorts of girls who are exposed to the program later-on have a greater likelihood of transitioning to high school (by four to six percentage points) and completing at least one high school grade (by five percentage points). In terms of labor market outcomes, the positive effects of the program on progression and completion seem to be accompanied by a reduction in the labor force participation of beneficiary adolescent girls of four to five percentage points. In addition, the younger cohorts of girls also appear to work at least two to three hour less per day relative to non-participants. Lastly, there is some evidence (marginally significant in statistical terms) that participant girls (between 15 to 19 years of age) tend to delay their marriage by 1.4 years and have 0.3 fewer children.

II. Related literature

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. Evaluations of the program based on a randomized design and multiple rounds of panel data show significant positive impacts on school enrollment and attainment, child labor, immediate consumption, and some health outcomes (Gertler, 2000; Hoddinott et. al., 2000; Behrman et al., 2001; Schultz, 2004). These assessments have been corroborated by subsequent evaluations of similar interventions in Brazil, Honduras, Colombia, Cambodia, Nicaragua, Turkey, Pakistan, Malawi, and El Salvador. For the most part, their 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 et al., 2003; Glewwe and Olinto, 2004; Attanasio and Mesnard, 2005; Attanasio et al., 2005a and 2005b; Filmer and Schady, 2006; Maluccio and Flores, 2005; Ahmed et al., 2007; Chaudhury and Parajuli, 2008; Baird et al., 2009).

In addition to positive effects on school attendance and enrollment, some evaluations also suggest that grade progression improves with CCTs. For instance, de Janvry et al. (2006) shows that the *Bolsa Escola* program in Brazil increases grade advancement by 6 percentage points; an experimental analysis of the Honduras *PRAF II* program finds that the probability of matriculating to subsequent grades increases by two to four percentage points among beneficiary children. Comparable effects have also been documented for CCT programs in Mexico, Nicaragua, and Paraguay (Behrman et al., 2001; Maluccio and Flores, 2005; Veras Soares et al., 2008).

The literature also gives evidence on the indirect effects of conditional transfers on the labor market outcomes and marriage and fertility decisions of adolescents, which are thought 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 treated children not only spend more time in school but also have relatively lower participation in income generating activities and domestic work (Behrman et al., 2005; Maluccio and Flores, 2005; de Janvry et al., 2006; IFS – Econometria - SEI, 2006; Ferreira et al., 2009). 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 et al., 2005; Dubois and Rubio-Codina, 2009). Furthermore, increased time spent in school and improved knowledge could also result in changes in marriage, sexual and fertility decisions. It is shown that CCTs have caused delays in marriage and onset of sexual activity in Malawi and Mexico, reductions in the number of sexual partners in Malawi, and reductions in pregnancy rates in Malawi and Turkey (Baird et al., 2009; Behrman et. al., 2005; Gulemetova-Swan, 2009; Ahmed et. al., 2007).

Evidence regarding the impacts of CCTs designed to explicitly address gender disparities in investments in human capital, such as the program evaluated in this paper, shows that they have succeeded in increasing school enrollment and attendance of girls. At least four interventions that were set up to address these issues achieve positive impacts in the short term. The *Bangladesh Female Secondary Stipend* Program is shown to raise the enrollment of girls in sixth through eighth grade, by eight to twelve percent (Khandker et. al., 2003). Similarly, the *Zomba Cash Transfer* experiment in Malawi more than doubles the re-enrollment of girls between 13 and 22 years old who had dropped out of school before (Baird et. al., 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 most directly related to this analysis, the short-term impact evaluation of the FSSP in Pakistan –the program analyzed in this paper –shows an increase in the enrollment of girls of about nine percentage points between 2003 and 2005 (Chaudhury and Parajuli, 2008). Although the impacts of these studies differ by magnitude, and the empirircal approaches vary, results appear to be consistent across programs and implementation contexts.

Nonetheless, the evidence of the connection between the provision of incentives for school participation and the well-being of adolescent girls in the longer-term remains very thin. Therefore, this paper aims to begin filling this evidence gap by assessing whether girls who benefited from a school stipend program show improved outcomes in the longer-term, or up to four years after their first exposure to the program.

III. Schooling context and project

Schooling in Punjab

The literacy rate in Punjab in 2003 was approximately 54 percent (45 percent in Pakistan as a whole) compared to 56 and 63 percent for South Asia and all low income countries, respectively (Lloyd, 2004).[2] Divided into 36 districts (34 districts at the time the program began), Punjab is home to over 55 percent of the total Pakistani population.

School enrollment in Punjab, particularly in secondary education, is largely constrained by the scarcity of schools. Primary education is supplied by both public and private schools. On the contrary, the secondary level is dominated by the public sector, and not every village has access to middle schools (grades 6 to 8) and high schools (grades 9 and 10).[3] For instance, whereas 84 percent of the households in Punjab report that their children are within walking distance (15 minutes) of a primary school, only 55 percent of the households have access to a middle or high school within the same travel time (Pakistan Social & Living Standards

[2] Estimates for Pakistan and 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.
[3] Around 90 percent of private schools offer only primary classes (Andrabi et al, 2006).

Measurement Survey - PSLM, 2004). Differences in access are even larger in rural areas, making it even far less likely that children in rural villages enroll in schools at these levels (Andrabi et al., 2006; Sathar et al., 2003). In 2003, before the program started, while enrollment in primary school for children aged 6-10 was 58 percent, enrollment at the middle school level for children aged 11-14 was 29 percentage points lower (Multiple Indicator Cluster Survey – MICS, 2003).

Education levels of females are low, both in absolute and compared to 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 non-stipend districts, large differences are observed in stipend districts – up to 8 percentage points (MICS, 2003). Such gaps in enrollment by gender are even wider when the samples are restricted to rural areas of Punjab. In terms of school attainment, numbers for 2001 (before the program started) show that women in Pakistan achieved on average 4.9 and 1.3 years of schooling in urban and rural areas, respectively – about 1.9 and 2.8 fewer years of education than men. These disparities by gender are also evident for Punjab, where the levels of enrollment and educational attainment are systematically lower for women in rural and urban areas (Table 1).

Finally, there are also marked differences in school enrollment and access between districts in Punjab covered and not covered by the GSP program. Enrollment levels in program districts were well below those of non-stipend districts. In 2003, for instance, enrollment in primary school for children aged 6 to 10 was 20 percentage points lower in stipend districts; similarly, only 20 percent of children 11-14 years old were in middle school compared to 34 percent in non-stipend districts (MICS, 2003). As shown in Table 2, access to school was also more difficult in stipend districts before the program began, where higher fractions of households live in villages without a school and, thus, their children had to travel longer distances to go to school (MICS, 2003; PSLM, 2004).

The Female School Stipend Program

The FSSP was designed as part of a broader educational reform of the Provincial Government of Punjab (GoP) to improve the educational attainment among girls and decrease gender inequities, particularly at the middle school level.[4] Prior to the start of the program, as noted before, female enrollment in primary and secondary schools was low, both in absolute terms and relative to that of boys.[5] Additionally, low enrollment for girls is further compounded by low retention and completion rates (Sawada and Lokshin, 2009). A large body of analytical work has identified various community characteristics (access to roads, transport, share of teachers that reside 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 gaps in school entry for girls (Khan, 1993; Sathar and Lloyd, 1994; Lloyd et al. 2005; Das et al. 2006; Lloyd at al. 2007). In response to this, the GoP mobilized resources in 2003 to finance the FSSP seeking to target a clearly disadvantaged group (girls) in districts with lagging literacy rates. None of the ineligible districts were covered by the FSSP but were still eligible for other components of the wider educational reform, such as the distribution of free textbooks, school rehabilitation, and new teachers, among other things.[6]

The main goal of the FSSP was to promote participation in public education for girls in middle school (grades 6 through 8).[7] Program benefits comprise a quarterly stipend of approximately PKR 600 (equivalent to US$10) per female student.[8] Beneficiary girls were targeted based on their district of residence (districts with the lowest literacy rates in the province – below 40 percent), and enrollment in eligible grades (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.

[4] The Punjab Education Sector Reform (PERSP) seeks 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.

[5] The gender disparity in Pakistan is well established in the literature; for example, authors such as Khan (1993), Behrman and Schneider (1993), Alderman et al. (1995), Alderman et al. (2001), Holmes (2003), Lloyd et al. (2005), and a gender assessment by the World Bank (2005) make this point clearly.

[6] Parts of these activities were either phased in throughout the province over the years based on need or were implemented only in places where demand was most salient. In this respect, the evaluation accounts for three components that could potentially affect the outcome of the *GSP*: school improvements[6], the recruitment and allocation of new (probably better qualified) teachers, and the phased distribution of free textbooks to all primary school children.

[7] Since 2006 a public subsidy to low cost private schools was put in place; this program known as stipend to Foundation Assisted Schools covers girls in approximately 50 percent of the GSP program. Additionally, an education voucher scheme was also implemented throughout Punjab, but targeting and modality are orthogonal to the program.

[8] It is estimated that the costs of schooling per quarter are approximately $9, leaving some funds left over for the family to use for other needs

The FSSP started in late 2003 and covered 15 eligible districts. Although planning for the program began in early 2003, stipends were not distributed until the first quarter of 2004, with compensation for attendance in the fourth quarter of 2003, as reported and accounted by the monitoring unit. During the first quarter of the program, about 156,000 girls received the stipend. In 2006, the FSSP was extended to include high school (grades 9 and 10); however, the impact of this extension is not the main focus of this paper. The FSSP is still active in Punjab, and as of 2007, it covered 245,000 beneficiaries in middle school, which correspond to a take-up rate of 51 percent.[9] According to actual quarterly stipend data, the program spent around US$7.3 million in 2007 on stipends for grades 6 through 8.

A previous impact evaluation of the FSSP investigated the short-term impacts of the program on female enrollment. The evaluation uses difference-in-differences (DD) and Regression Discontinuity Design (RDD) models on provincial school censuses to estimate net growth in female enrollments in grades 6-8 (middle school) from 2003 (before the program started) to 2005 (two years into the program) and in stipend public schools relative to non-stipend public schools. Other models estimated compare the differences in female growth enrollment relative to male growth enrollment in a triple-difference framework. Even though the magnitude of program impact estimates varies across different models, this evaluation found that the FSSP increased the number of females enrolled in grades 6-8, on average, by 9-23 percentage points (Chaudhury and Parajuli, 2008).

3. Data

This paper estimates the impact of the program with information at the school and household levels. The main data source for the first approach is the Punjab public school census collected annually in October from 2003 to 2009 by the provincial education department, including data from before and during the program implementation. It covers all public schools in Punjab and focuses on school information (including village name and date of establishment), infrastructure,

[9] The take-up rate is difficult to calculate because eligibility to the program is determined by the grade, and not the age, of girls in stipend districts. For simplicity, this analysis estimates the take-up 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 grade 6-8 in private schools and those whose highest grade is 5-7 but are not enrolled at the time of the survey).

teaching staff, enrollment, and some organizational indicators.[10] The enrollment data are reported by gender of the school 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 from 2003-04 to 2009-10, however the sub-sample for this analysis is a pool of cross-sectional data that includes approximately 4,000 middle schools for girls in stipend and non-stipend districts. These are the schools that could be eligible for the program at both baseline and follow-up.[11]

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 multi-topic 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.[12]

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 distribution of free textbooks, recruitment of new teachers, construction of school facilities, and improvement of school councils.

[10] This survey is part of the institutional strengthening capacity effort within the PERSP and is used for monitoring resource allocation and performance in the sector. The 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.

[11] The sample of eligible schools 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, as these would not have been eligible for the stipend at both baseline and follow-up. Of the remaining schools, only those schools that had non-zero enrollment in grades 6-8 at both baseline and follow-up were retained. Finally, although in theory all public schools in Punjab are single-gender, the data shows that a small number of mixed schools exist. These schools were also excluded from the final sample.

[12] It is important to note that, although lower-level data provide more precise estimates, the questionnaires are significantly shorter; thus, some key economic welfare indicators are excluded.

The main data source for the estimation of impacts at the individual level is two waves of the MICS collected in 2003 (pre-program) and 2007-08 (post-program). This household survey covers a wide range of 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 Cantonment 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 (a 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. 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.

Given the targeting of the program to low-literacy districts, stipend and non-stipend districts are likely to differ in aspects other than the treatment and a baseline comparison confirms this. Since literacy rates are likely correlated with other underlying socio-economic characteristics, such factors may, in turn, be associated with different education attainment and divergent trajectories in human capital accumulation between households in stipend and non-stipend districts. Indeed, as shown in Table 3, 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 (defined as a distance of two 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 school facilities, but the differences are very small, except for electricity. On the other hand, 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.

In order to deal with potential biases that may rise due to these differences, we run econometric specifications which control for factors at the community, school, and household levels that could contribute to the outcomes, as well as district and cohort-fixed effects. Control variables that may be endogenous to 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 PERSP.[13]

4. Empirical Strategy

As the FSSP was not randomly allocated, a simple comparison of girls in stipend and non-stipend districts is unlikely to 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 (Annex A). This targeting may pose an obvious selection problem since stipend districts were chosen based on characteristics (possibly both observed and unobserved) that may be correlated with low educational attainment. Moreover, there are two further issues that could make it difficult to identify the causal effect of the program in the absence of a reliable counterfactual. One is participation bias, namely families which decide to enroll girls in public middle school and therefore participate in the program may be different from other families with girls in relevant ages but not in public school. The other issue is the focus on outcomes that are conditioned on the enrollment of girls. In this case, the samples for analysis may suffer from negative selection

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

bias if they include lower-ability girls in stipend areas that would not have attended school in the absence of the program.[14]

Since 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. We attempt 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 double-difference and RDD frameworks.[15] While double-differencing compares the changes in the outcomes of analysis between stipend and non-stipend districts overtime, RDD exploits the literacy rate cut-off for program eligibility across districts. The RDD models are also implemented within a DD framework as done by Chaudhury and Parajuli (2008). Therefore, both models compare the trends in stipend and non-stipend districts, which helps address possible selection bias introduced by time-fixed characteristics related to nonrandom program placement. 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.[16] The validity of the identifying assumptions of the DD models and the implications on the expected selection on returns to schooling are discussed in the robustness section.

We construct 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 total female enrollment at public schools in stipend districts (treatment schools) for the period between 2004 and 2010 relative to the pre-program level (2003-04 academic year). These changes are then compared with the corresponding annual changes in female enrollment observed for public schools in non-stipend districts (control

[14] 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 non-stipend districts (through DD method), and the exogenous variation in program eligibility around the literacy cut off (through RDD).

[15] 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 —that 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.

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

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 that, we also use household surveys to estimate changes at the individual level. The research design for this approach follows for the most part the same identification strategies and assumptions as the school-level analysis. However, the richness of information of the household-level data, as compared to 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 partial measures of school attainment (middle school completion, and high school enrollment and completion), labor market participation, marriage, and fertility. 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 non-stipend districts (control cohorts). 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).[17] In what follows, both approaches are explained in more detail.

School-level analysis

The treatment and comparison cohorts of girls for this analysis are constructed synthetically based on their enrollment at different grades. The treatment cohorts were enrolled in public middle school in stipend districts; thus, they should have been exposed to the program for at least one year between baseline (2003-04) and subsequent academic years up to 2009-10. Based on these criteria, seven distinct cohorts can be identified (Table 4). The counterfactuals consist of the corresponding cohorts enrolled in public schools in non-stipend districts. For example, a cohort that consists of girls in grade 6 in stipend districts in 2004-05 is expected to progress to grade 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, and to the expansion of the stipends

[17] 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 non-stipend districts even though not all girls in the stipend districts participated.

to girls in high school (grades 9 and 10) which started in 2006-07. Girls in the corresponding cohort but in non-stipend districts should not have been exposed.

The outcome of interest for this analysis is the percentage change in grade level enrollment (grades 6-10 each), which is estimated using the following specification:

$$\Delta Y_i = \alpha + \beta_1 * X_i + \beta_2 * FSSP_i + \varepsilon_i \qquad (1)$$

Where i indexes schools, X includes baseline socioeconomic characteristics, *FSSP* is a binary variable that takes the value 1 for a stipend-eligible school and 0 otherwise. β_2 measures the net impact of the stipend program.

Therefore, 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), reduction in drop-out rates, and/or influx of new entrants (including both rejoiners and those who switch schools). Yet, the school-level approach cannot distinguish these channels of impacts since the school censuses do not provide information to measure them separately. However, preprogram and post-program data from household surveys indicate that female rejoining (new entrants) and grade repetition at the middle school level is very low.[18] This suggests that net changes in female enrollment (if they happen) 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 be 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 do not contain explicit information on the treatment status and length of exposure of girls to the FSSP.

[18] 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 non-stipend 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 re-entry, repetition and school switching accounted for some 5 percent of total enrollment in 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 non-stipend districts between 2003-04 and 2007-08.

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. 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[19] Annex B 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 non-stipend 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: (a) 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; (b) 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.

We use the household data to estimate the influence of the program on such longer-term outcomes as educational attainment and labor market, marriage, and fertility decisions. The general specification of the base regression equations for the DD analysis is as follows:

$$Y_{it} = \alpha + \beta_1 * FSSP_i + \beta_2 * T_t + \beta_3 * FSSP_i * T_t + \beta_4 * X_{it=0} + \beta_5 * P_{it} + \varepsilon_{it} \qquad (2)$$

where *i* indexes individuals, *t* is the time subscript, *T* is a dummy variable taking a value of 1 for post-program year and a value of 0 for the pre-program year, *FSSP* identifies the location of the girl between stipend and non-stipend districts, and X is a vector of baseline household demographics, community-level characteristics, and measures of education market place and school quality. *P* is a vector of dummy variables to capture other programs that were introduced

[19] The previous impact evaluation of the FSSP (Chaudhury and Parajuli 2008) 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.

by the GoP along with the FSSP and sought to build school facilities, provide free textbooks, recruit contract teachers, and revitalize school councils.[20] The outcomes of interest (Y) measure educational outcomes (middle school completion, middle to high school transition, high school completion), labor force participation (extensive and intensive margins) marriage and fertility.[21] The coefficient β_3, the interaction of time and treatment status, gives the average net impact of the program. Additional specifications of this model also include district and cohort fixed effects. Because observations within a district could be serially correlated, standard errors are clustered at the district level.

We also estimate program impacts in a RDD framework, exploiting the fact that assignment to the stipend district (beneficiaries) is determined by a threshold base on the average literacy rate of the district (Chaudhury and Parajuli, 2008). Thus, girls in stipend districts with literacy rates of 40 percent or less are eligible for FSSP, with all girls in districts above the 40 percent literacy threshold comprising the control group. The RDD specification includes a control function L for distance of district literacy rate (subscript D) from the cutoff. It is assumed that the control function is linear, making this a parametric specification. The identifying assumption here is that making treatment status conditional on literacy would imply treatment to be independent of the error term given the literacy rate. The general specification of the RDD is as follows:

$$Y_{it} = \alpha + \beta_1 * FSSP_i + \beta_2 * T_t + \beta_3 * FSSP_i * T_t + \lambda * (L_D) + \varepsilon_{it} \qquad (3)$$

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 equation (3). To avoid misspecification issues arising from the functional form, we also estimated non-parametric RDD models with different bandwidths and weights on either side of the discontinuity.

[20] These programs were *per se* 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.

[21] 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 have 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.

5. Empirical Results

Educational Outcomes

Theory provides ambiguous results on the effects of this sort of program on grade progression, school completion, and educational attainment. 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 since 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. 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 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. We seek to shed some light on the net effects of these conflicting forces in the context of the FSSP.

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 (Table 5).[22] For instance, cohort C sees higher percentage changes in enrollment in stipend than in non-stipend 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

[22] The school-level approach measures changes in absolute enrollment between stipend and non-stipend districts. These changes, however, may be driven by a plausible mechanical relationship between 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 pre-program period and check if changes in enrollment in stipend district occurred even before the program began. Yet, the empirical models include a number of socio-economic covariates at their baseline values which are also likely to explain differences in fertility between stipend and non-stipend 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.

versus non-stipend 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, estimates of program effects on enrollment yield consistent results for all of these cohorts and grades over time (positive and significant impacts). The insignificant impacts in 2004-05 are expected since awareness of the program was not widespread during its first year.[23] For subsequent cohorts, program impacts are still large and statistically significant, ranging from 11 to 32 percent.[24] Impacts of comparable magnitude have been found in similar female-targeted CCTs in Bangladesh, Cambodia, and Malawi (Khandker and others 2003; Baird and others 2009; Filmer and Schady 2006).

An interesting result is that program impacts on female enrollment increase with school grades for the same cohort. Several factors may be behind this. First, the program might have incentivized girls who had left the education system to re-enter school. However, this does not seem to explain the changes for middle school grades since re-entry is not only very low (less than 0.5 percent) but also similar across stipend and non-stipend districts. Second, and more plausible, the program may have differential effects in retaining girls in schools, perhaps more in 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 rate in middle and high school grades in stipend districts, relative to non-stipend districts, while there are not changes in dropout rates for noneligible grades. 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.

The findings also 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 year 2007-08 and

---

[23] 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 at the second year of the program.

[24] The parameter of interest for the models discussed here is changes in the number of girls relative to 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.

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 evidence on school progression based on household-level data seems to validate the findings obtained with school data as it also suggests that the program increases the chances that girls complete middle school. Overall, results summarized in Table 6 indicate that 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 cohorts. This is equivalent to an increase of 4.5 percent relative to the baseline value of the treatment group.[25] 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.

We also find some evidence that the program helps girls transition from middle to high school and complete grade 9, especially those younger than 17.[26] The first outcome analyzed is the probability that a girl enrolled in grade 8 during the program period progressed to grade 9 or beyond. [27] The average impact for the entire cohort of 12-19 years old girls is not statistically significant (Table 6). 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). This cohort is also 5 percentage points more likely to complete grade 9 than the control group (Table 7). One reason for this is that these younger girls benefit more from the FSSP as they are not only exposed longer but they joined the program when there was more awareness of its existence. Second, and perhaps more important, only the younger cohorts have the additional incentives induced by the expansion of the FSSP, which began in 2006, to also provide conditional transfers to enrollees in high school.

[25] Evidence of similar impacts has been also found for Bolsa Escola and PRAF II, two CCT programs in Brazil and Honduras (de Janvry and others 2006; Glewee and Olinto 2004).

[26] 2007-08, the year of the follow-up survey, may be too early to capture the impacts on this younger cohort (that is, age 15-16) of the program on high school (that is, grade 10) completion, as the majority of girls enrolled in grade 10 at the time of the follow-up survey are aged 15-16

[27] 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. 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 non-stipend districts.

Labor Market Outcomes

Theoretically, the direction of the effects of the FSSP on working behaviors of the beneficiaries in the longer-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. Thus, girls have more incentives to join the labor market as their ages and years of accumulated schooling increase. Additionally, since 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, this section presents results of program impacts on two outcomes related to work for girls aged 12-17: labor force participation and time devoted to working.

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 (unpaid family helper 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, sub-samples, or approaches. Furthermore, the decrease in labor force participation is notable. Four years after the program started, the labor force participation rates of girls 12-19 are almost equal between stipend and non-stipend 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 whether this reduction in labor participation is fully or partly explained by the extra time spent by beneficiaries in school-related activities (Table 8).

We also examine the effects of the program on the intensive margin of the labor supply for those girls who are working for pay at the time of the survey. Samples are restricted to not only girls that work, but also those that are 15-19 years old since 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 non-stipend districts (significant at the 10 percent level). This is also consistent with the hypotheses described earlier about higher impacts for the younger cohorts (Table 8).

Marriage and Fertility Outcomes

There are multiple channels through which the program could influence marriage and fertility 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 their marriages and sexual activity and alter their decisions about fertility (such as timing of childbearing, birth spacing, and total fertility rates). On the other hand, if the program increases grade progression and help girls finish school earlier, they might move into marriage and childbearing sooner.[28]

Furthermore, social norms and cultural tradition in Pakistan may figure prominently in marriage and fertility decisions, such as making it difficult for girls to resist the pressure of marriage and child bearing. 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.

We first investigate the possibility that the program changes the probability of marriage and age at marriage for girls aged 15 to 19. Overall, 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 identifies that after the program started girls in stipend

[28] Women in Pakistan typically move into childbearing shortly after marriage.

districts are 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.

The analysis then turns to the impacts of the program on fertility outcomes. The sample for this part of the analysis consists of girls aged 17 to 19, since these are the ages when 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 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 17-19 having given birth. A second exercise looks at the number of births given for the sub-sample of girls who became mothers during the period under study. 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. However, this result is marginally significant around the 10 percent level.[29] Even though both are partial measures of fertility given the young age 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 their subsequent births (Table 9).

Heterogeneity of Program Impacts

This section presents evidence on the heterogeneity of the FSSP effects across some socioeconomic characteristics of program beneficiaries such as socioeconomic status, parental education, location (rural or urban), age, and length of exposure of the beneficiaries.[30] We first find that 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

[29] 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 non-stipend districts. Moreover, the problem of missing values does not appear to be systematically correlated with some determinants of fertility.

[30] 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.

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 (Table 10).

Program impacts on labor force participation also appear to be 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 participant girls in the poorest households are more likely to delay marriage.

There is evidence in the data that 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, the results show 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 the 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 may provide suggestive evidence of positive marginal effects of treatment length.

Impact on Male Siblings

In theory, educational stipends that are targeted to girls could 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 ones the displacement effect often runs in the opposite (negative) direction. For instance, due to the income effect of the transfers, budgetconstrained households may be more able to keep all children in school or enroll those out of school. In contrast, the displacement effect may incentivize 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 forces.

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 et al. 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 et al. 2009).[31]

We exploit three different identification strategies to examine the spillover effects of the FSSP effects on current enrollment, school completion (primary and secondary), and school choice (private or public) for boys in school age (6 to 17 years old). The first strategy implements a DD analysis that compares boys in households with at least one exposed girl in stipend districts to similar boys in non-stipend districts at baseline and follow-up. The second strategy also looks at relative changes over the longer term but uses boys in households without exposed girls in stipend districts as the comparison group. Finally, the third strategy integrates both of the designs above in a triple difference framework (DDD). 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 gender-specific age and grade restrictions.

None of the empirical models estimated for each of the research designs reveal any spillover effects of the program on the enrollment and school completion of boys. The results, summarized in Table 11, 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

[31] The CESSP Scholarship Program in Cambodia 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).

with no eligible siblings in both stipend and non-stipend 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.[32]

Consistent with what was observed through qualitative work conducted in the field as part of this impact evaluation, there is, however, indication of strong indirect effects of the program on school choice among boys. We find that families with daughters eligible to the FSSP tend to respond to the program by enrolling their sons in private schools. In fact, several econometric models show a statistically significant average increase in enrollment of boys in private schools of nearly 4 percentage points compared to 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.[33]

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 Rs. per month). In fact, a national census of private schools in 2000 found that the median rural private school charged 60 Rs. per month.[34] 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 better understand this possible indirect effect.

6. Robustness Checks

As eligibility to the FSSP was not assigned randomly, the analysis conducted in this paper has to rely on quasi-experimental techniques and their corresponding identifying assumptions. To assess the internal validity of this evaluation, we discuss the robustness of the results to a few

[32] 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 difference in of age-grade distributions.
[33] Aslam (2009) also finds that boys are more likely than girls to enroll in private schools in Punjab.
[34] School fees in the private sector may increase by more than three-fold at the secondary level.

issues that could contaminate the design of the evaluation. We start with the underlying assumption of the double-difference analysis, namely that the outcomes investigated were progressing along similar trends across stipend and non-stipend districts before the program began. Program impact estimates based on double-difference models may be biased if either of the outcomes 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.

Yet, there is no evidence that longer-term impacts of the program identified in this paper are driven by differences in pre-program trends. Two main checks were conducted to directly investigate this. The first uses pre-baseline and baseline data from PIHS (2001) and MICS (2003) to run double-difference models on schooling, labor market, marriage, and fertility outcomes for the same cohorts of girls.[35] 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 non-stipend districts at the preprogram time. The second check consists of double-difference 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, thus, did not benefit directly from it. Again, results of these exercises do not reveal any differential pre-program trends that may contaminate the main findings of the paper (Table 12).

Another important concern stems from the fact that the empirical analysis 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. If this is the case, the parameters that measure the effects of the FSSP on the longer-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

[35] 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 non-stipend 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 non-stipend districts).

results of the analysis are still very informative as they can be interpreted as lower bounds of the true impacts.

There are various problematic compositional changes in the sample that could take place. An example is that the FSSP likely 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 school work than the girls who are already enrolled in school.[36] 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 since the new entrants are expected to benefit less from schooling compared to 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 a 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.[37]

Similarly, there is no sign the program causes changes to the age-grade distribution in stipend districts. 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 add 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 non-stipend districts, do not indicate that over-age girls are 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 over-age girls. We do not find evidence that migration could affect the composition of the samples of analysis.

An additional 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

[36] 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.
[37] 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 Lee (2002) and Angrist and others (2004).

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 paper. First, the modest size of the transfer (nearly $2.5 per month) is not expected to encourage households to move from richer non-stipend districts to stipend districts. Second, out-of-province family migration in both stipend and non-stipend districts is rather uncommon, particularly in rural areas. Instead, most migration is in the form of temporary out-migration of a member of the household — often of working men. 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.

A final compositional issue could arise if the conditional stipend motivates households located in non-stipend districts but adjacent to stipend districts to send their girls to schools there. This would increase the demand for education in stipend districts but reduce it in non-stipend districts, leading to no changes in net enrollment across the province. In order to check the possibility of crossover effects between contiguous stipend and non-stipend districts, 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.[38] 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.

A natural issue for a program like the FSSP focused exclusively to public schools is families could have moved 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.

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

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 crowding out effect induced by the program.[39]

In order to check for potential issues of selection in participation, we also run empirical models for all the outcomes on the sample of girls eligible to the program irrespective of their likely participation in the program The sample for this includes girls ages 12-19 that were eligible for the program even though they were never enrolled in grades 6-8 during the duration of the program. However, although impacts estimates based on this larger cohort are closer to the intent-to-treat parameter, a problem of this approach is that the results of this analysis are likely to get diluted as the units of observations may include a significant number of girls that did not participate; they could have left school before the program started and never rejoined.[40] Yet, some of the results from the base models remain.[41]

A final 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 same time as the program in stipend than non-stipend districts. There is, however, no evidence of such systematic change in school supply. In fact, data from the annual Punjab school census (EMIS) 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 non-stipend districts increased by only 0.34 and 0.15 percent, respectively, compared to the number of schools at baseline. Furthermore, the trends in the school supply of stipend and non-stipend districts are very similar in the period before and during the program, 2000-07 (Figure 1).

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

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

[41] For middle school/grade 9/grade 10 completion, and 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 exposed girls, but with lower significance levels (5%-10%). 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.

6. Concluding Remarks

This paper aims to fill the evidence gap on longer-term impacts of CCTs by evaluating a female-targeted CCT program in Pakistan, the Female School Stipend Program. The evidence indicates 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. 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. In terms of labor force participation and marriage and fertility decisions, there is some evidence (although in some cases marginally significant in statistical sense) that the program encourages girls to work less, delay their marriage, and have fewer children.

We also find that, as a result of the stipend, boys residing in the same household as girls in participating districts are more likely to be enrolled in private primary schools. These schools are growing in number, more affordable — school fees are lower than the stipend — and perform better in terms of students' learning, which means the FSSP may indirectly increase the gender gap in learning, at least at the primary level. 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 choice as boys.

The impacts on educational attainment may have important implications for future productivity and 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, relative to women with less than middle school education. A big part of the reason is that education provides women with more marriage opportunities. Another reason is that higher schooling also enables women to attain higher earnings. Simple back-of-the-envelope calculations estimate that the female wage premiums for middle and high school education (relative to lower than middle school education) are 24 and 157 percent, respectively. These improvements represent an annual increase in earnings of $16 and $100, respectively, which might more than compensate 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 later age tend to have fewer children. In fact, women in Punjab with middle and high school education have around 1.8 fewer children than those with lower than middle school education by the end of their reproductive life. Simple extrapolations also indicate that the 1.4 year delay in marriage of 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 invest more in their human capital, which may lead to positive intergenerational effects.

The results of this paper 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 more fully explore the dynamics of program impacts and the effects on learning. First of all, many of the beneficiaries observed in this paper are still young (12-19 years old), so they might not have reached their full potential in terms of human capital. In addition to continue investigating the impacts on final school attainment and fertility, more needs to be done in subsequent works to identify the effects of the program on future productivity and wages. Second, and related to this, it would be useful to understand whether the impacts on schooling are also translated into improvements in learning. Knowledge on these areas is critical to establish whether CCTs and comparable programs also contribute to poverty reduction in the long-term.

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." IFPRI: Washington, DC, in collaboration with the AGRIN Co. Ltd.: Ankara, Turkey.

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 Papers 1970, Iowa State University, Department of Economics.

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 D. van de Walle and K. Nead (eds.) Public Spending and the Poor: Theory and Evidence. Baltimore, Maryland: Johns Hopkins University Press.

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

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

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.

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. Washington, DC: World Bank.

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

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

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

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

Behrman, Jere 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.

Behrman, J. R. and J. Hoddinott. 2001. “An Evaluation of the Impact of PROGRESA on Pre-school Child Height.” FCND Discussion Paper No. 104. Washington, DC International Food Policy Research Institute.

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.

Chaudhury, N. and D. Parajuli. 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*.

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.” World Bank Policy Research Working Paper No. 4102. Washington, DC: World Bank.

Das, J., P. Pandey, and T. Zajonc. 2006. “Learning Levels and Gaps in Pakistan” World Bank Policy research Working Paper 4067. Washington, DC: World Bank.

Dubois, P. and M. Rubio-Codina. 2009. “Child Care Provision: Semiparametric Evidence from a Randomized Experiment in Mexico.” CEPR Discussion Paper No. DP7203.

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 Series 5001. Washington, DC: World Bank.

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

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

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

Glewwe, P. and P. Olinto. 2004. “Evaluating the Impact of Conditional Cash Transfers on Schooling: An Experimental Analysis of Honduras.”Mimeo, University of Minnesota: Minneapolis, Minnesota.

Gulemetova-Swan, M. 2009. "Evaluating the Impact of Conditional Cash Transfer Programs on Adolescent Decisions about Marriage and Fertility: The Case of Oportunidades" Dissertations available from http://repository.upenn.edu/dissertations/AAI3363363

Hoddinott, J., E. Skoufias, and R. Washburn. 2000. “The Impact of PROGRESA on Consumption.” Washington, DC: International Food Policy Research Institute.

Holmes, J. 2003. "Measuring the Determinants of School Completion in Pakistan: Analysis of Censoring and Selection Bias." *Economics of Education Review* 22: 249-64.

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

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

Khandker, Sh., M. M. Pitt, and N. Fuwa. 2003. "Subsidy to Promote Girls' Secondary Education: The Female Stipend Program in Bangladesh." Mimeo. Washington, DC: World Bank.

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: IFPRI.

Khan, S. R.1993. "South Asia." Chapter 6 in King and Hill (eds.) Women's Education in Developing Countries: Barriers, Benefits and Policy. Baltimore, Maryland: John Hopkins University Press.

Lloyd, C.B., C. Mete, and M. Grant. 2007. "Rural Girls in Pakistan: Constraints of Policy and Culture." In Maureen Lewis and Marlaine Lockheed (eds.) Exclusion, Gender and Education: Case Studies from the Developing World (pp. 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*, vol. 2.

Lloyd, C. 2004. "The Changing Transitions to Adulthood in a Comparative Perspective: The Case of Pakistan". *Pakistan Development Review* 43 (4):441-67.

Olinto, P., S.S. Morris, R. Flores, and A. Veiga. 2003. "The Impact of the Bolsa Alimentacao Program on Food Consumption." Paper presented at Annual Meetings of the International Association of Agricultural Economists: Durban, South Africa.

Sathar, Z., C. Lloyd, C. Mete, and Minhaj ul Haque. 2003. "Schooling Opportunities for Girls as a Stimulus for Fertility Change in Rural Pakistan." *Economic Development and Cultural Change*, 51 (3): 677-98.

Sathar, Z. and C. Lloyd. 1994. "Who Gets Primary Schooling in Pakistan: Inequalities Among and Within Families", *Pakistan Development Review*, 33(2): 103-34.

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." Publications 3, International Policy Centre for Inclusive Growth.

World Bank. 2010. “Evidence and Lessons Learned from Impact Evaluations in Social Safety Nets.” Manuscript. Washington, DC: The World Bank.

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

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

Table 1. Educational Outcomes at Baseline

	School enrollment (percent)						Years of schooling [c]					
	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
Non-stipend	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[a]	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

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

Table 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%
Non-stipend	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.

Table 3. Summary Statistics of Covariates at Baseline

COVARIATES	Non-stipend	Stipend	Difference	SE
<u>Community-level Characteristics</u>				
Rural	0.643	0.818	0.175***	(0.005)
Poverty Headcount	0.436	0.672	0.236***	(0.008)
Access to Primary School^	0.886	0.752	-0.134***	(0.013)
Access to Middle School^	0.606	0.383	-0.223***	(0.019)
Access to High School^	0.479	0.232	-0.247***	(0.018)
Access to Public Transport^	0.746	0.59	-0.156***	(0.018)
Middle School Private/Public Enrollment Ratio	0.345	0.254	-0.0908***	(0.022)
Share of Education in Total Expenditure (%)	7.762	6.692	-1.070***	(0.126)
<u>School-level Characteristics</u>				
Boundary wall - girls middle school^^	0.841	0.848	-0.0067	(0.010)
Drinking water - girls middle school^^	0.926	0.95	-0.0245***	(0.007)
Electricity - girls middle school^^	0.744	0.852	-0.108***	(0.012)
Toilet - girls middle school^^	0.780	0.785	-0.005	(0.011)
Student Teacher Ratio - girls middle public schools^^	32.443	26.108	-6.335***	(0.423)
<u>Household-level Characteristics</u>				
Household Size	6.569	6.663	0.0935**	(0.044)
Dependency Ratio	0.454	0.495	0.0407***	(0.003)
Education of HH Head~	4.774	3.133	-1.641***	(0.080)
Gas Connection	0.293	0.047	-0.246***	(0.008)
Electricity Connection	0.926	0.677	-0.249***	(0.011)
Visit from Health Worker	0.353	0.338	-0.0150	(0.013)
Access to Family Planning Center^	0.341	0.209	-0.132***	(0.017)
Access to Health Center/Hospital^	0.466	0.292	-0.174***	(0.019)
Access to Drinking Water Supply^	0.957	0.9753	0.0183***	(0.006)

Source: MICS 2003; ^Source: PSLM 2004; ^^ Source: Annual Public School Census 2003
~Those who have never attended school are taken to have completed zero years of schooling; SE=standard error

Table 4. 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 H	Cohort G	Cohort F	Cohort E	Cohort D

Table 5. Program Impacts on Grade Level Enrollment (percentage)

	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10
2004-05	-0.28	4.33	6.12		
	(C)	(B)	(A)		
2005-06	10.56**	17.83***	23.98***	-10.51	
	(D)	(C)	(B)	(A)	
2006-07	16.96***	21.81***	31.99***	-0.42	-28.20
	(E)	(D)	(C)	(B)	(A)
2007-08	13.56***	20.92***	27.93***	20.48*	-5.08
	(F)	(E)	(D)	(C)	(B)
2008-09	12.69***	16.57***	20.54***	12.76	5.43
	(G)	(F)	(E)	(D)	(C)
2009-10	15.59***	18.96***	27.00***	30.07**	8.80
	(H)	(G)	(F)	(E)	(D)

Source: Authors' calculation
Note: Estimates from parametric RDD (whole sample). ** Significant at the 5 percent level. *** Significant at the 1% level. Cohort names are in parentheses

Table 6. Average Program Impact on Educational Outcomes - I

	Whole Sample (12-19)					*Cohort 15-16*					*Whole Sample (12-19)*					*Cohort 15-16*				
	DD	DD	RDD	RDD	Non-Parametric RDD	DD	DD	RDD	RDD	Non-Parametric RDD	DD	DD	RDD	RDD	Non-Parametric RDD	DD	DD	RDD	RDD	Non-Parametric RDD
	(1)	(2)	(3)	(4)	(1)	(1)	(2)	(3)	(4)	(1)	(1)	(2)	(3)	(4)	(1)	(1)	(2)	(3)	(4)	(1)
	Outcome A. Middle Scool Completion										*Outcome B. Middle to High School Transition*									
FSSP*Time	0.00564	0.0120	0.00509	0.0121	0.0416	0.0503*	0.0590**	0.0500*	0.0590**	0.1040*	-0.00713	0.0108	-0.00819	0.0102	0.0556	0.0350	0.0555**	0.0346	0.0554**	0.1272**
	(0.015)	(0.013)	(0.015)	(0.013)	(0.0302)	(0.026)	(0.028)	(0.026)	(0.028)	(0.0547)	(0.020)	(0.018)	(0.019)	(0.018)	(0.0353)	(0.023)	(0.026)	(0.023)	(0.026)	(0.0566)
FSSP	-0.0481***	0.00149	-0.0157	0.000589	-0.0640*	-0.0914***	0.0122	-0.0420	0.0155	-0.1467**	-0.0633***	-0.0187	-0.00373	-0.00703	-0.1078***	-0.0940***	-0.00350	-0.0317	0.0291	-0.1654***
	(0.015)	(0.014)	(0.020)	(0.019)	(0.0338)	(0.024)	(0.023)	(0.029)	(0.029)	(0.0596)	(0.023)	(0.023)	(0.030)	(0.026)	(0.0396)	(0.022)	(0.032)	(0.027)	(0.033)	(0.0616)
Time	-0.00974	-0.0179***	-0.00943	-0.0179***	-0.0282	-0.0424***	-0.0605***	-0.0423***	-0.0604***	-0.0772**	0.0144	0.000991	0.0150	0.00126	-0.0430*	-0.0391***	-0.0583***	-0.0390***	-0.0576***	-0.1138***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.0207)	(0.010)	(0.012)	(0.011)	(0.012)	(0.0380)	(0.009)	(0.009)	(0.009)	(0.009)	(0.0246)	(0.012)	(0.012)	(0.012)	(0.012)	(0.0397)
L40-L	No	No	-0.00156**	0.0000874	-0.0002	No	No	-0.00240***	-0.000330	-0.0008	No	No	-0.00286**	-0.00143	-0.0037	No	No	-0.00303**	-0.00418*	-0.0054
			(0.001)	(0.001)	(0.0067)			(0.001)	(0.002)	(0.0125)			(0.001)	(0.001)	(0.0079)			(0.001)	(0.002)	(0.0130)
Covariates	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No
Age Fixed Effects	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No
No. of obs	22289	20826	22289	20826	6076	9396	8962	9396	8962	2573	22237	20773	22237	20773	6062	9391	8956	9391	8956	2573

***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 water, 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) Non-parametric RD estimates are presented for bandwidth size=4 unless otherwise stated. Bandwidth 4 was the most frequently recurring optimal bandwidth (calculated using Imbens and Kalyanamaran's method (2009)) followed by 9 in our school level data, and so results have been presented for bandwidth 4 unless otherwise specified. Results for other bandwidth choices are available upon request (5) Non-parametric RD results are presented for models that use kernel weights, not sampling weights. Results for models using both kernel and sampling weights are available upon request.

Table 7. Average Program Impact on Educational Outcomes - II

	Whole Sample (12-19)					*Cohort 15-16*					*Whole Sample (12-19)*					*Cohort 15-16*				
	DD	DD	RDD	RDD	Non-Parametric RDD	DD	DD	RDD	RDD	Non-Parametric RDD	DD	DD	RDD	RDD	Non-Parametric RDD	DD	DD	RDD	RDD	Non-Parametric RDD
	(1)	(2)	(3)	(4)	(1)	(1)	(2)	(3)	(4)	(1)	(1)	(2)	(3)	(4)	(1)	(1)	(2)	(3)	(4)	(1)
	Outcome C. Grade 9 Completion										*Outcome D. Grade 10 Completion*									
FSSP*Time	-0.0146	0.00563	-0.0156	0.00495	0.0229	0.0264	0.0494**	0.0263	0.0495**	0.1206**	-0.0545**	-0.0323	-0.0560**	-0.0339	0.0044	0.0301	0.0285	0.0264	0.0249	0.2538**
	(0.022)	(0.020)	(0.022)	(0.020)	(0.0407)	(0.022)	(0.024)	(0.022)	(0.024)	(0.0614)	(0.025)	(0.022)	(0.025)	(0.022)	(0.0450)	(0.051)	(0.048)	(0.051)	(0.049)	(0.1123)
FSSP	-0.0444*	-0.00577	0.0148	0.00656	-0.0307	-0.0628**	-0.0216	-0.0372	-0.00142	-0.0655	0.00276	0.0722***	0.0641*	0.0928***	-0.0455	-0.0640	0.0652	0.0134	0.144**	-0.1663
	(0.026)	(0.026)	(0.037)	(0.027)	(0.0459)	(0.027)	(0.036)	(0.045)	(0.036)	(0.0685)	(0.031)	(0.023)	(0.034)	(0.030)	(0.0508)	(0.050)	(0.069)	(0.073)	(0.067)	(0.1286)
Time	-0.0215	-0.0360**	-0.0210	-0.0357**	-0.0573**	-0.0766***	-0.0955***	-0.0766***	-0.0951***	-0.1491***	0.0669***	0.0609***	0.0677***	0.0615***	0.0420	-0.0779**	-0.0707**	-0.0756**	-0.0677*	-0.2412***
	(0.013)	(0.013)	(0.013)	(0.014)	(0.0291)	(0.015)	(0.017)	(0.015)	(0.017)	(0.0447)	(0.010)	(0.011)	(0.010)	(0.011)	(0.0327)	(0.036)	(0.034)	(0.036)	(0.033)	(0.0816)
L40-L	No	No	-0.00285**	-0.00151	0.0092	No	No	-0.00124	-0.00258	0.0168	No	No	-0.00296*	-0.00241	-0.0013	No	No	-0.00374	-0.00941**	0.0047
			(0.001)	(0.002)	(0.0091)			(0.002)	(0.003)	(0.0136)			(0.002)	(0.002)	(0.0104)			(0.003)	(0.004)	(0.0274)
Covariates	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No
Age Fixed Effects	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No
No. of obs	19915	18482	19915	18482	5425	8988	8570	8988	8570	2444	12831	11623	12831	11623	3519	2387	2202	2387	2202	666

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

[a] Results are positive and statistically significant using wider bandwidths and therefore larger sample sizes

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 water , 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) Non-parametric RD estimates are presented for bandwidth size=4 unless otherwise stated. Bandwidth 4 was the most frequently recurring optimal bandwidth (calculated using Imbens and Kalyanamaran's method (2009)) followed by 9 in our school level data, and so results have been presented for bandwidth 4 unless otherwise specified. Results for other bandwidth choices are available upon request (5) Non-parametric RD results are presented for models that use kernel weights, not sampling weights. Results for models using both kernel and sampling weights are available upon request.

Table 8. Average Program Impact on Labor Outcomes

	Whole Sample (12-19)					Cohort 15-16					Whole Sample (15-19)					Cohort 15-17				
	DD	DD	RDD	RDD	Non-Parametric RDD	DD	DD	RDD	RDD	Non-Parametric RDD	DD	DD	RDD	RDD	Non-Parametric RDD[b]	DD	DD	RDD	RDD	Non-Parametric RDD[b]
	(1)	(2)	(3)	(4)	(1)	(1)	(2)	(3)	(4)	(1)	(1)	(2)	(3)	(4)	(1)	(1)	(2)	(3)	(4)	(1)
	Outcome E. Labor Force Participation										*Outcome F. Work Intensity (days per month)*									
FSSP*Time	-0.0467**	-0.0494**	-0.0465**	-0.0490**	-0.0154	-0.0452***	-0.0403**	-0.0452***	-0.0401**	-0.0403*	0.897	-0.611	0.417	-0.548	-11.5698***	0.0744	-6.088*	-0.392	-6.137*	-17.0783***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.0143)	(0.016)	(0.017)	(0.016)	(0.017)	(0.0228)	(1.671)	(2.037)	(1.727)	(2.077)	(2.1609)	(2.494)	(3.107)	(2.873)	(3.139)	(3.0777)
FSSP	0.0471***	0.0427*	0.0391*	0.0309	0.0157	0.0493***	0.0528**	0.0549***	0.0431*	0.0230	-0.632	-3.405**	-3.465**	-4.046**	26.9624***	-0.309	3.631	-2.296	3.342	26.8686***
	(0.016)	(0.022)	(0.023)	(0.023)	(0.0167)	(0.014)	(0.024)	(0.019)	(0.025)	(0.0261)	(1.192)	(1.292)	(1.481)	(1.644)	(1.6559)	(1.530)	(4.731)	(2.176)	(5.234)	(2.1383)
Time	0.0193	0.0188	0.0192	0.0186	0.0151	0.0222*	0.0226*	0.0222*	0.0224*	0.0304*	-0.711	1.393	-0.351	1.294	10.6635***	-1.699	3.380	-1.351	3.361	14.6414***
	(0.013)	(0.012)	(0.013)	(0.012)	(0.0102)	(0.012)	(0.012)	(0.012)	(0.012)	(0.0158)	(1.281)	(1.828)	(1.306)	(1.841)	(1.6270)	(2.253)	(2.514)	(2.566)	(2.533)	(2.4932)
L40-L	No	No	0.000387	0.00202**	-0.0037	No	No	-0.000270	0.00180**	-0.0006	No	No	0.178**	0.0711	-4.2563***	No	No	0.145	0.0390	-2.9914***
			(0.001)	(0.001)	(0.0035)			(0.001)	(0.001)	(0.0058)			(0.072)	(0.163)	(0.3983)			(0.160)	(0.305)	(0.6369)
Covariates	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No
Age Fixed Effects	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No
No. of obs	27448	26037	27448	26037	7464	9090	8672	9090	8672	2479	292	273	292	273	189	125	120	125	120	92

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

[b]Results for days worked are presented for bandwidth 9. There were 54 observations only for bandwidth 4. Note that results are consistently negative and statistically significant for bandwidths lower and higher than 9

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 school,water supply, 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) Non-parametric RD estimates are presented for bandwidth size=4 unless otherwise stated. Bandwidth 4 was the most frequently recurring optimal bandwidth (calculated using Imbens and Kalyanamaran's method (2009)) followed by 9 in our school level data, and so results have been presented for bandwidth 4 unless otherwise specified. Results for other bandwidth choices are available upon request (5) Non-parametric RD results are presented for models that use kernel weights, not sampling weights. Results for models using both kernel and sampling weights are available upon request.

(

Table 9. Average Program Impact on Marriage & Fertility Outcomes

	Whole Sample (15-19)				
	DD	DD	RDD	RDD	Non-Parametric RDD
	(1)	(2)	(3)	(4)	(1)
	Outcome G. Probability of Marriage				
FSSP*Time	0.00953	0.00820	0.00940	0.00814	-0.0146
	(0.009)	(0.008)	(0.009)	(0.008)	(0.0141)
FSSP	0.00962	0.0132	0.0185	0.0143*	0.0125
	(0.008)	(0.008)	(0.011)	(0.008)	(0.0143)
Time	-0.00695*	-0.00506	-0.00689*	-0.00503	0.0043
	(0.003)	(0.004)	(0.004)	(0.004)	(0.0088)
L40-L			-0.000430	-0.000164	0.0007
			(0.000)	(0.000)	(0.0032)
Covariates	No	Yes	No	Yes	No
Age Fixed Effects	No	Yes	No	Yes	No
No. of obs	19177	17761	19177	17761	5251

	Whole Sample (15-19)				
	DD	DD	RDD	RDD	Non-Parametric RDD c
	(1)	(2)	(3)	(4)	(1)
	Outcome H: Age at Marriage				
FSSP*Time	-0.151	1.460**	-0.0767	1.222*	-0.1044
	(0.388)	(0.621)	(0.458)	(0.643)	(0.6967)
FSSP	-0.202	-0.0353	-0.588	0.230	0.5494
	(0.391)	(0.439)	(0.865)	(0.460)	(0.8259)
Time	1.689***	0.803	1.610***	0.697	1.2851**
	(0.285)	(0.478)	(0.359)	(0.488)	(0.5388)
L40-L			0.0173	-0.0538*	0.1891**
			(0.028)	(0.031)	(0.0909)
Covariates	No	Yes	No	Yes	No
Age Fixed Effects	No	Yes	No	Yes	No
No. of obs	339	296	339	296	171

	Whole Sample (17-19)				
	DD	DD	RDD	RDD	Non-Parametric RDD c
	(1)	(2)	(3)	(4)	(1)
	Outcome I. Probability of Giving Birth				
FSSP*Time	-0.0114	-0.0808	-0.0257	-0.0691	-0.0387
	(0.121)	(0.172)	(0.122)	(0.173)	(0.1353)
FSSP	0.0232	0.0825	0.129	0.0396	0.4586***
	(0.099)	(0.181)	(0.154)	(0.197)	(0.0949)
Time	-0.0188	-0.152	0.00204	-0.157	0.0620
	(0.072)	(0.130)	(0.074)	(0.131)	(0.0908)
L40-L	No	No	-0.00511	0.00789	0.0390**
			(0.006)	(0.012)	(0.0157)
Covariates	No	Yes	No	Yes	No
Age Fixed Effects	No	Yes	No	Yes	No
No. of obs	392	342	392	342	187

	Whole Sample (17-19)				
	DD	DD	RDD	RDD	Non-Parametric RDD c
	(1)	(2)	(3)	(4)	(1)
	Outcome J. Number of Children				
FSSP*Time	0.0962	-0.329*	0.0663	-0.349*	-0.0386
	(0.150)	(0.181)	(0.149)	(0.209)	(0.1780)
FSSP	-0.0387	0.305*	0.182	0.358*	0.4532***
	(0.122)	(0.158)	(0.182)	(0.209)	(0.1247)
Time	-0.0613	0.0967	-0.0180	0.112	0.1333
	(0.091)	(0.088)	(0.088)	(0.130)	(0.1186)
L40-L	No	No	-0.0106	-0.00840	0.0329
			(0.008)	(0.016)	(0.0205)
Covariates	No	Yes	No	Yes	No
Age Fixed Effects	No	Yes	No	Yes	No
No. of obs	392	342	392	342	186

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

c Results for days worked are presented for bandwidth 9. There were around 100 observations for bandwidth 4.

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) Non-parametric RD estimates are presented for bandwidth size=4 unless otherwise stated. Bandwidth 4 was the most frequently recurring optimal bandwidth (calculated using Imbens and Kalyanamaran's method (2009)) followed by 9 in our school level data, and so results have been presented for bandwidth 4 unless otherwise specified. Results for other bandwidth choices are available upon request (6) Non-parametric RD results are presented for models that use kernel weights, not sampling weights. Results for models using both kernel and sampling weights are available upon request.

Table 10. 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)		

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

Notes: (1) Standard errors 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

Table 11. Program Spillover Effects on Boys

	Whole sample (age 6-17)				*Age 6-12*			
	Value at Baseline	(DDD)	(RD)	(Non-Parametric RD)	*Value at Baseline*	(DDD)	(RD)	(Non-Parametric RD)
Primary school completion (percentage points)	91.1 T	0.898	1.03	-2.38	94.1 T	0.00519	0.00519	-3.29
		(1.8)	(1.8)	(3..2)		(0.048)	(0.048)	(6.13)
Middle school completion (percentage points)	76.3 T	-1.09	-0.996	-8.3				
		(3.3)	(3.3)	(6.5)				
Private school enrollment (percentage points)	30.6 T	3.75**	3.77**	-10.6**	34.0 T	11.0***	11.0***	-9.3*
		(1.9)	(1.9)	(3.8)		(2.6)	(2.6)	(4.9)
Years of schooling	4.2	-0.00449	0.00182	0.2299	2.5	-0.0756	-0.0756	0.0811
		(0.079)	(0.079)	(0.231)		(0.077)	(0.077)	(0.196)

***1% significance level **5% significance level *10% significance level. Standard errors are in parentheses.
T baseline value is expressed in percentages
Source: Authors' calculation
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 (5) Non-parametric RD estimates are presented for bandwidth size=4. Results for other bandwidth choices are available upon request (5) Non-parametric RD results are presented for models that use kernel weights, not sampling weights. Results for models using both kernel and sampling weights are also available upon request

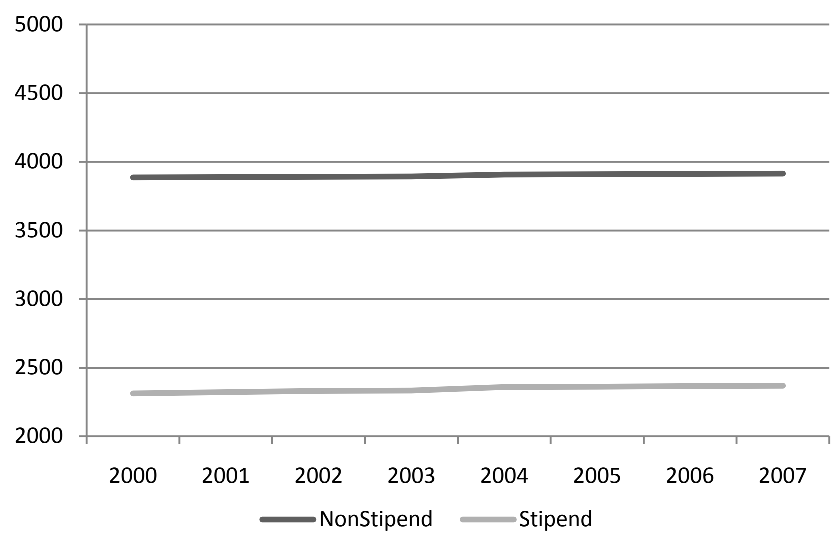
Table 12. Robustness Check - Pre-Program Trends

	Pre-baseline (PIHS 2001) vs. Baseline (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)

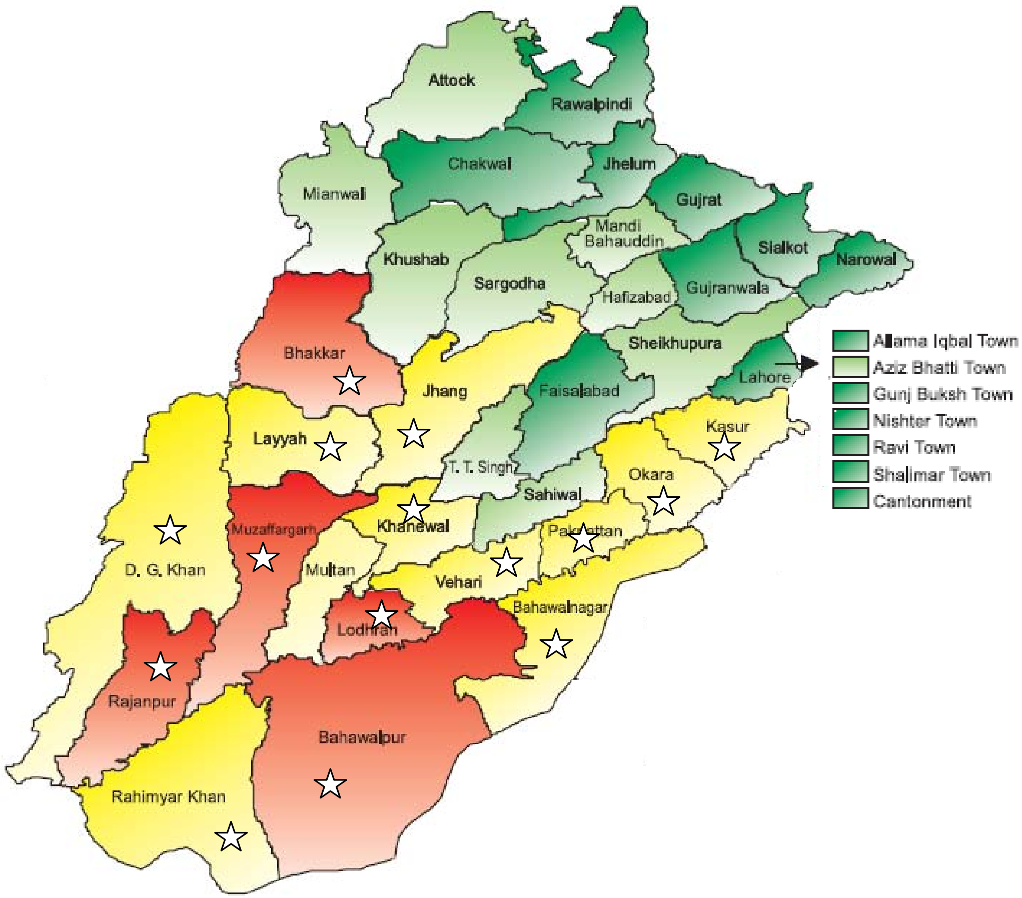
Coefficients shows correspond to the interaction between the dummy for stipend districts and the dummy t for time (t=1 for observations at baseline, t=0 at pre-baseline) ***1% significance level **5% significance level *10% significance level. Standard errors are in parentheses.

Figure 1. Total Number of Public Schools in Stipend and Nonstipend Districts during 2000-07

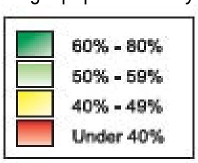


Source: Punjab Public School Census (2003-07)

Annex 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

Annex B. Construction of the Samples for Analysis

The household-level analysis draws mostly 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 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 (see above) in the MICS to construct synthetic cohorts of the relevant group - girls that were likely to be 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.[42] For instance, at baseline, 49% of girls age 12-19 in stipend districts had never ever attended school as opposed to 21% in non stipend districts. Under an exclusively age-based criterion, these girls would have been misidentified as exposed to FSSP, although it was impossible for these girls to participate in the program since its implementation.[43]

This annex discusses the three criteria used to construct the cohorts of possibly participant girls. In addition to this, the annex 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 imputed correctly. The criteria are the following:

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

   a. If a girl was enrolled in 2006-07 in grades 6-8, she is considered 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 exposed.[44]

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 (so exposed to the program) would be in grade 11 in 2007-08. Therefore, if a girl was enrolled in

[42]. 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.
[43]. We also do not find evidence of increase in grade 4-5 enrollment (enrollment in these grade levels more likely to be affected by the program) or primary school completion in stipend districts relative to non-stipend districts over time for girls age 12-17
[44]. Since the 2007-08 survey is conducted two months after the beginning of the academic year 2007-08, the girls enrolled in grade 6 are yet to receive a stipend and are not considered exposed.

grades 9-11 in 2006-07, she is considered exposed. 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 break. For instance, a girl enrolled in grade 11 in 2006-07 is assumed to have been exposed, the assumption being that she was enrolled in grade 8 in 2003-04, and then moved onto 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 as 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 grade 8, so in actuality was not exposed.
b. If a girl was enrolled in grade 11 in 2006-07, she is considered exposed as 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, so was not exposed to the program.

In both the above examples, the girls are misidentified as exposed when they have 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 non-stipend districts at baseline or over time.

2. The **second** criterion is based on 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 for who program exposure is a possibility:

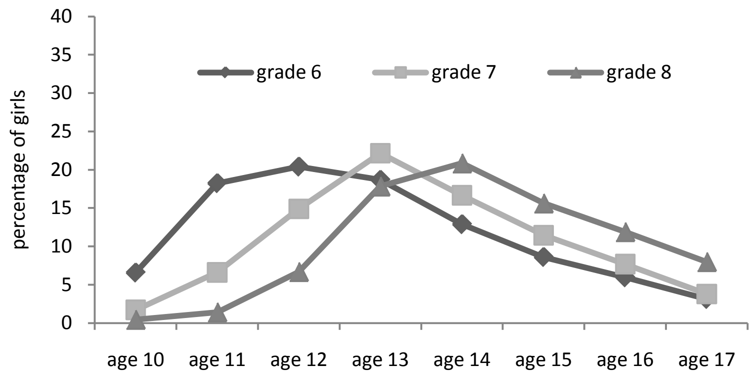
   a. Girls whose highest grade completed is grade 5 or below were excluded. This is because for a girl to be exposed to the program, she must have at least completed grade 6.

      - 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 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), hence not exposed.

Figure 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.[45] To illustrate the construction process, an example is presented:

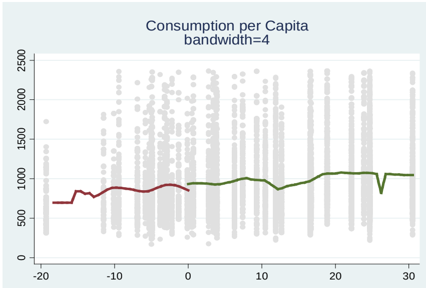
a. A girl whose age was 16 in 2007-08 and whose highest grade completed is grade 7 would have been exposed to the program only if she was in grades 6 or 7 in at least 2003-04. The only way she could not have been exposed to the program was if she completed grade 7 before the program began, say in 2002-03. 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. Similarly, a girl who is 19 years old and her highest grade completed is grade 6 is not considered exposed as 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, and 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 the 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 after imposing the first two criteria. This is akin to treatment of outliers, and a safeguard against misreporting. For instance, a girl who is 8 years of age is highly unlikely to have completed 6.

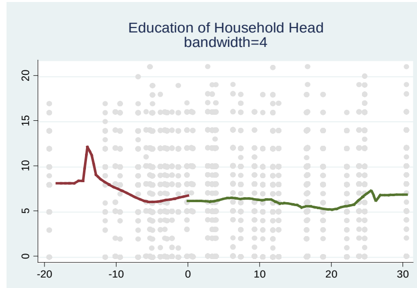
[45]. The age-grade distribution at baseline is used for two reasons: (1) The program impacts in the first few years were low as there was very little awareness of the program, the program had not been scaled up, and implementation maturity was yet to be reached so the age-grade distribution during 2004-05 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 has been no major shifts in age-grade distribution during the time

Continuity Checks for Individual and School Level Variables

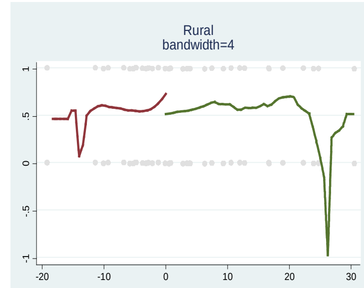
Consumption per Capita



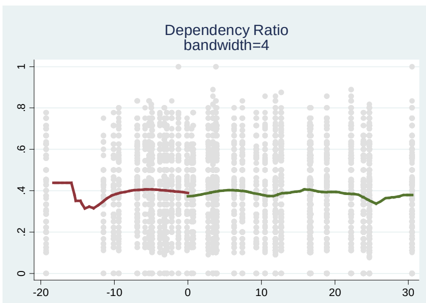
Education of Household Head



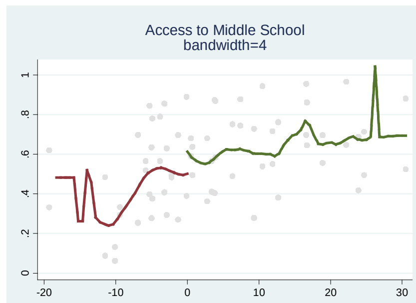
Rural



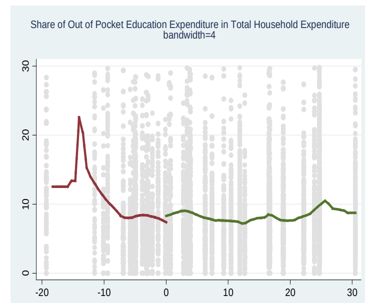
Dependency Ratio



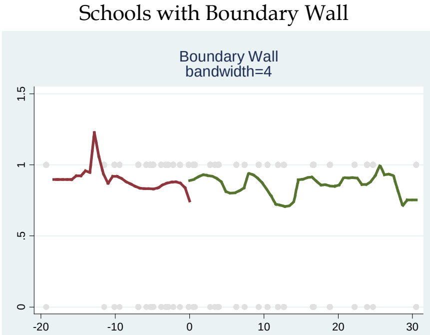
Access to Middle School



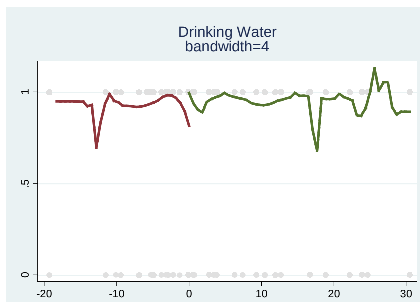
Share of Education in HH Expenditure



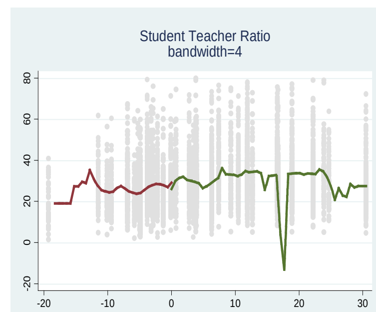
Schools with Boundary Wall



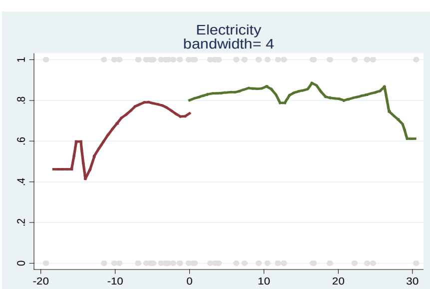
Schools with Drinking Water



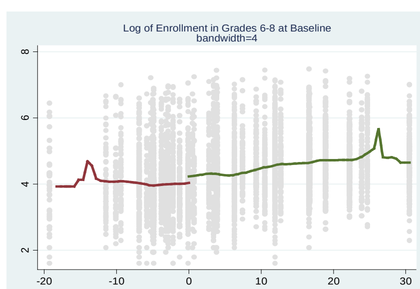
Student Teacher Ratio



Schools with Electricity

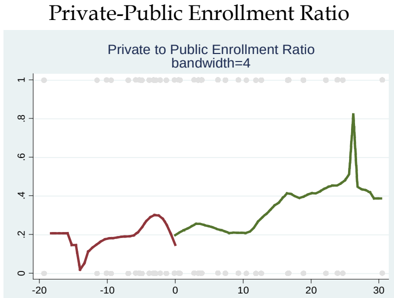


Log of Enrollment in Grades 6-8

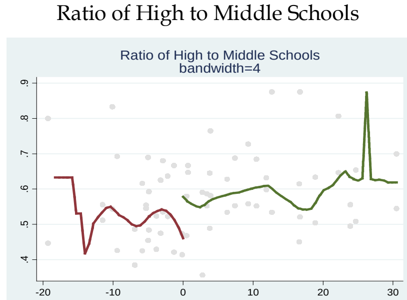


Schools with Toilets

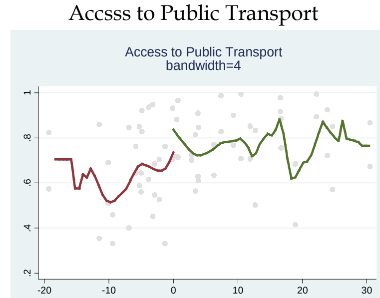
Private-Public Enrollment Ratio



Ratio of High to Middle Schools

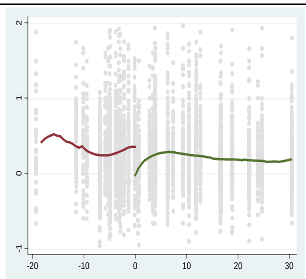
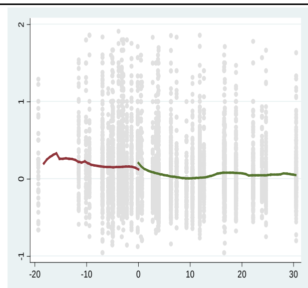
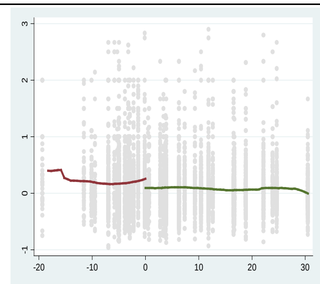
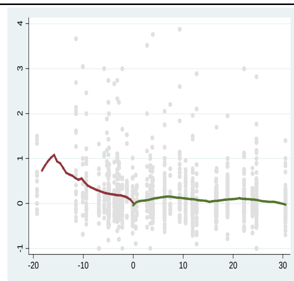
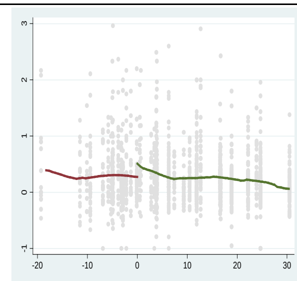
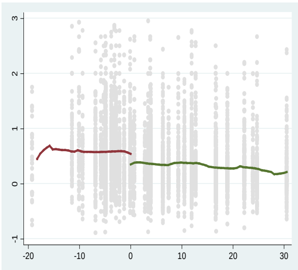
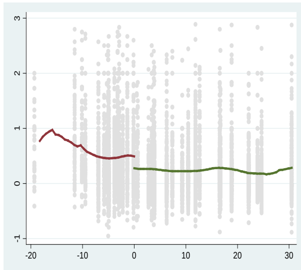
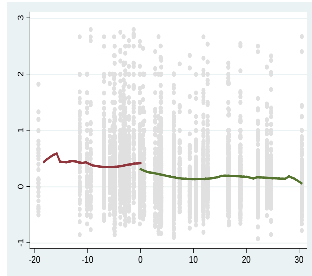
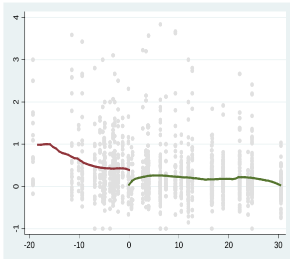
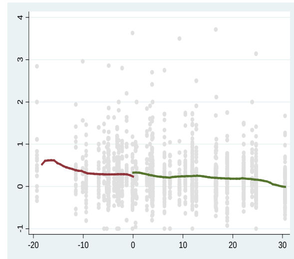
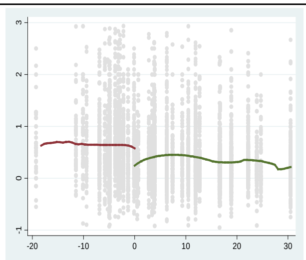
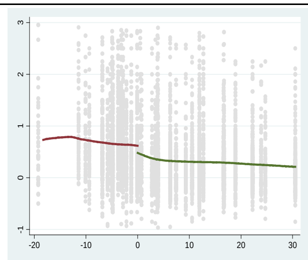
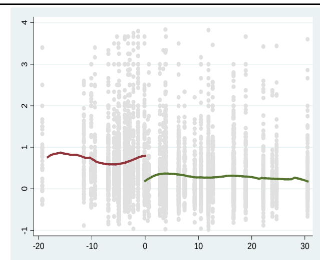
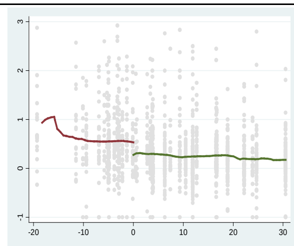
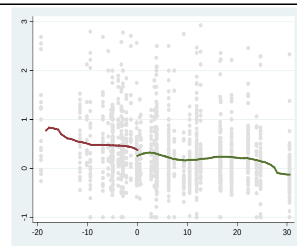


Accsss to Public Transport



School Level RD Analysis of % Changes in Enrollment between 2004 and 2009 relative to 2003

Bandwidth=4

	% Change in Grade 6 Enrollment	% Change in Grade 7 Enrollment	% Change in Grade 8 Enrollment	% Change in Grade 9 Enrollment	% Change in Grade 10 Enrollment
2004					
2005					
2006					

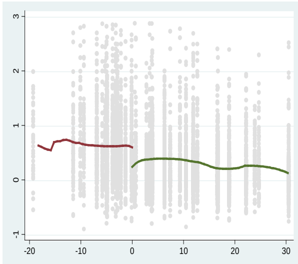
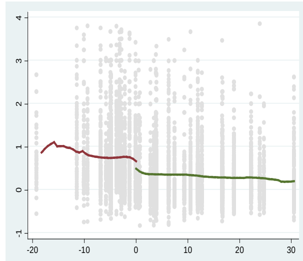
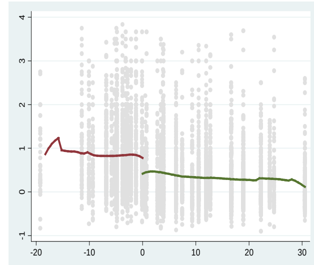
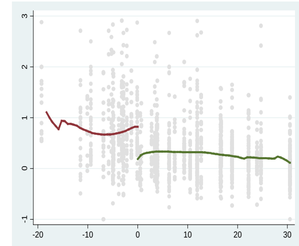
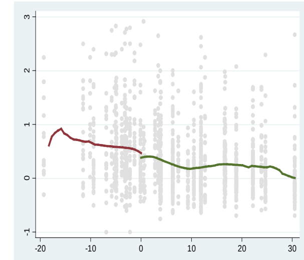
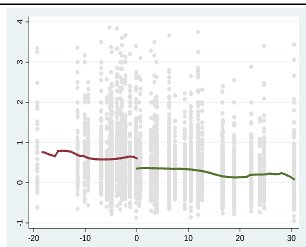
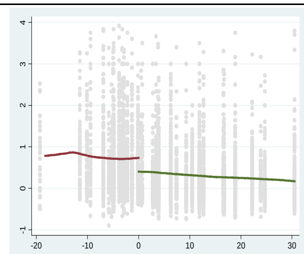
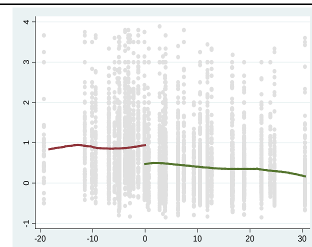
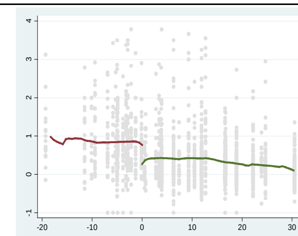
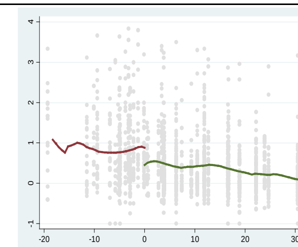
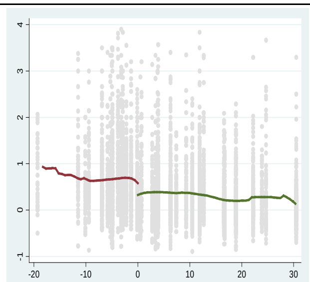
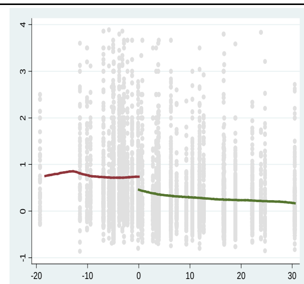
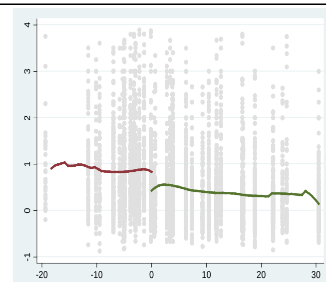
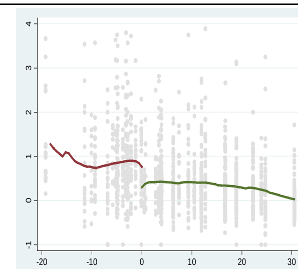
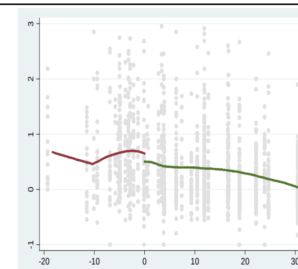
2007					
2008					
2009					

Table 1: Reduced Form Estimates for Analysis of % Change in Enrollment at the School Level

	2004			2004			2004			2004			2004		
	% Change in Grade 6 Enrollment			% Change in Grade 7 Enrollment			% Change in Grade 8 Enrollment			% Change in Grade 9 Enrollment			% Change in Grade 10 Enrollment		
	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD
FSSP	0.0077	0.0090	0.3606***	0.0717**	0.0618*	-0.1625	0.0277	0.0704	0.1876**	-0.0764	-0.0048	0.7291	-0.0867	-0.1007	-0.1303
	(0.0319)	(0.0343)	(0.0722)	(0.0315)	(0.0331)	(0.1209)	(0.0450)	(0.0467)	(0.0904)	(0.0792)	(0.1317)	(0.5911)	(0.0586)	(0.0678)	(0.1575)
Li-L40	No	0.0003	0.0799***	No	-0.0021	-0.0538	No	0.0090***	0.0164	No	0.0090	0.0883***	No	-0.0018	-0.0592
		(0.0019)	(0.0213)		(0.0025)	(0.0367)		(0.0031)	(0.0254)		(0.0104)	(0.0300)		(0.0051)	(0.0386)
Covariates	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Additional Programs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Number of Observations	4960	4960	1589	4960	4960	1589	4960	4960	1589	1534	1534	410	1506	1506	409

	2005			2005			2005			2005			2005		
	% Change in Grade 6 Enrollment			% Change in Grade 7 Enrollment			% Change in Grade 8 Enrollment			% Change in Grade 9 Enrollment			% Change in Grade 10 Enrollment		
	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD
FSSP	0.1457***	0.1350***	0.2187**	0.1681***	0.1844***	0.2782***	0.1083**	0.1222**	0.1335	-0.1893*	-0.1291	0.7467**	-0.1719**	-0.2203**	-0.0942
	(0.0370)	(0.0393)	(0.0868)	(0.0386)	(0.0405)	(0.1019)	(0.0544)	(0.0540)	(0.1305)	(0.1041)	(0.0996)	(0.3304)	(0.0867)	(0.1022)	(0.2731)
Li-L40	No	-0.0023	0.0060	No	0.0035	0.0123	No	0.0030	-0.0208	No	0.0076	0.1236***	No	-0.0061	-0.0517
		(0.0022)	(0.0233)		(0.0029)	(0.0282)		(0.0037)	(0.0376)		(0.0075)	(0.0412)		(0.0064)	(0.0729)
Covariates	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Additional Programs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Number of observations	4944	4944	1587	4944	4944	1587	4944	4944	1587	1516	1516	406	1489	1489	405

***1% significance level **5% significance level *10% significance level

Source: Authors' calculations

Notes: (1) Non-parametric RD estimates are done for bandwidth four. Optimal bandwidth was calculated using the method suggested by Imbens and Kalyanamaran (2009). The most frequently recurring optimal bandwidth for outcome variables across all years was in the neighborhood of 4, so non-parametric RD estimates are presented for this bandwidth. Results unqiue to each outcome's optimal bandwidth as well as for ad hoc bandwidth choices (2 to 12) are available upon request. (2) In the case of grade 8 enrollment in 2006, grade 10 enrollment in 2007, and grade 10 enrollment in 2010, adding a quadratic term (for Li-L40) gives statistically significant estimates (3) Difference and Difference and Parametric RD estimates without any controls, as well as estimates with baseline covariates but no additional program controls are available upon request (4) Covariates include status of school facilities (boundary wall, electricity, toilet, drinking water), student teacher ratio, locaton of school (rural), log of initial enrollment, private to public enrollment ratio. and other community characteristics (e.g. poverty, access) (5) Additional programs include contract teacher recruitment, program for re-building missing facilities and upgrading school infrastructure, and new school construction or upgradation.

	2006			2006			2006			2006			2006		
	% Change in Grade 6 Enrollment			% Change in Grade 7 Enrollment			% Change in Grade 8 Enrollment			% Change in Grade 9 Enrollment			% Change in Grade 10 Enrollment		
	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD
	(iii)	(iii)	(i)	(iii)	(iii)	(i)	(iii)	(iii)	(i)	(iii)	(iii)	(i)	(iii)	(iii)	(i)
FSSP	0.2126***	0.1924***	0.4846***	0.2626***	0.2347***	-0.0824	0.0948	0.1160*	0.5094***	-0.1396	-0.1053	0.3502**	-0.3711*	-0.3683	0.2250
	(0.0418)	(0.0460)	(0.0965)	(0.0444)	(0.0452)	(0.1304)	(0.0673)	(0.0657)	(0.1461)	(0.1055)	(0.1022)	(0.1436)	(0.2032)	(0.2257)	(0.2546)
Li-L40	No	-0.0043	0.0679***	No	-0.0060*	-0.0867**	No	0.0045	0.0375	No	0.0044	0.1224***	No	0.0004	-0.0105
		(0.0030)	(0.0228)		(0.0032)	(0.0371)		(0.0044)	(0.0414)		(0.0065)	(0.0468)		(0.0091)	(0.0617)
Covariates	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Additional Programs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Number of observations	4894	4894	1593	4894	4894	1593	4894	4894	1593	1458	1458	408	1432	1432	407

	2007			2007			2007			2007			2007		
	% Change in Grade 6 Enrollment			% Change in Grade 7 Enrollment			% Change in Grade 8 Enrollment			% Change in Grade 9 Enrollment			% Change in Grade 10 Enrollment		
	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD	Difference in Difference	Parametric RD	Non Parametric RD
FSSP	0.2058***	0.1551***	0.4479***	0.3823***	0.3613***	0.1838	0.2708***	0.2636***	0.3292**	-0.0342	0.0054	1.0996***	-0.2282	-0.1654	0.4164
	(0.0397)	(0.0430)	(0.1001)	(0.0486)	(0.0502)	(0.1216)	(0.0619)	(0.0609)	(0.1365)	(0.1057)	(0.1025)	(0.3547)	(0.1675)	(0.1929)	(0.3112)
Li-L40	No	-0.0108***	0.0515**	No	-0.0045	-0.0422	No	-0.0015	0.0004	No	0.0050	0.1189***	No	0.0079	0.0339
		(0.0026)	(0.0258)		(0.0033)	(0.0340)		(0.0042)	(0.0387)		(0.0083)	(0.0405)		(0.0096)	(0.0656)
Covariates	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Additional Programs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Number of observations	4903	4903	1597	4903	4903	1597	4903	4903	1597	1458	1458	407	1433	1433	407

***1% significance level **5% significance level *10% significance level

Source: Authors' calculations

Notes: (1) Non-parametric RD estimates are done for bandwidth four. Optimal bandwidth was calculated using the method suggested by Imbens and Kalyanamaran (2009). The most frequently recurring optimal bandwidth for outcome variables across all years was in the neighborhood of 4, so non-parametric RD estimates are presented for this bandwidth. Results unqiue to each outcome's optimal bandwidth as well as for ad hoc bandwidth choices (2 to 12) are available upon request. (2) In the case of grade 8 enrollment in 2006, grade 10 enrollment in 2007, and grade 10 enrollment in 2010, adding a quadratic term (for Li-L40) gives statistically significant estimates (3) Difference and Difference and Parametric RD estimates without any controls, as well as estimates with baseline covariates but no additional program controls are available upon request (4) Covariates include status of school facilities (boundary wall, electricity, toilet, drinking water), student teacher ratio, locaton of school (rural), log of initial enrollment, private to public enrollment ratio. and other community characteristics (e.g. poverty, access) (5) Additional programs include contract teacher recruitment, program for re-building missing facilities and upgrading school infrastructure, and new school construction or upgradation.

	2008			2008			2008			2008			2008		
	% Change in Grade 6 Enrollment			% Change in Grade 7 Enrollment			% Change in Grade 8 Enrollment			% Change in Grade 9 Enrollment			% Change in Grade 10 Enrollment		
	Difference in Difference	**Parametric RD**	**Non Parametric RD**	**Difference in Difference**	**Parametric RD**	**Non Parametric RD**	**Difference in Difference**	**Parametric RD**	**Non Parametric RD**	**Difference in Difference**	**Parametric RD**	**Non Parametric RD**	**Difference in Difference**	**Parametric RD**	**Non Parametric RD**
FSSP	0.1703***	0.1243***	0.3080***	0.2644***	0.2160***	0.1390	0.2931***	0.2731***	0.4069***	-0.0495	-0.0391	0.9130**	-0.2292	-0.2132	0.7551**
	(0.0409)	(0.0436)	(0.1028)	(0.0510)	(0.0518)	(0.1399)	(0.0648)	(0.0670)	(0.1549)	(0.1124)	(0.1164)	(0.4608)	(0.3001)	(0.3545)	(0.2934)
Li-L40	No	-0.0096***	0.0211	No	-0.0101***	-0.0102	No	-0.0042	0.0128	No	0.0013	0.1128**	No	0.0019	0.0200
		(0.0028)	(0.0282)		(0.0035)	(0.0431)		(0.0042)	(0.0453)		(0.0100)	(0.0447)		(0.0151)	(0.0546)
Covariates	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Additional Programs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Number of observations	4915	4915	1601	4915	4915	1601	4915	4915	1601	1462	1462	408	1436	1436	407

	2009			2009			2009			2009			2009		
	% Change in Grade 6 Enrollment			% Change in Grade 7 Enrollment			% Change in Grade 8 Enrollment			% Change in Grade 9 Enrollment			% Change in Grade 10 Enrollment		
	Difference in Difference	**Parametric RD**	**Non Parametric RD**	**Difference in Difference**	**Parametric RD**	**Non Parametric RD**	**Difference in Difference**	**Parametric RD**	**Non Parametric RD**	**Difference in Difference**	**Parametric RD**	**Non Parametric RD**	**Difference in Difference**	**Parametric RD**	**Non Parametric RD**
FSSP	0.1583***	0.1397***	0.3444***	0.2838***	0.2481***	0.0533	0.2481***	0.2339***	0.4395**	-0.0427	0.1175	1.0675***	-0.3233	-0.1720	0.6104
	(0.0433)	(0.0460)	(0.1181)	(0.0480)	(0.0511)	(0.1285)	(0.0707)	(0.0692)	(0.1976)	(0.1294)	(0.1360)	(0.3800)	(0.3537)	(0.4065)	(0.4059)
Li-L40	No	-0.0039	0.0175	No	-0.0075**	-0.0582	No	-0.0030	0.0134	No	0.0198*	0.1196**	No	0.0186	0.0159
		(0.0034)	(0.0292)		(0.0034)	(0.0366)		(0.0047)	(0.0556)		(0.0103)	(0.0472)		(0.0164)	(0.0781)
Covariates	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Additional Programs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Number of observations	4874	4874	1585	4874	4874	1585	4874	4874	1585	1446	1446	402	1420	1420	401

***1% significance level **5% significance level *10% significance level

Source: Authors' calculations

Notes: (1) Non-parametric RD estimates are done for bandwidth four. Optimal bandwidth was calculated using the method suggested by Imbens and Kalyanamaran (2009). The most frequently recurring optimal bandwidth for outcome variables across all years was in the neighborhood of 4, so non-parametric RD estimates are presented for this bandwidth. Results unqiue to each outcome's optimal bandwidth as well as for ad hoc bandwidth choices (2 to 12) are available upon request. (2) In the case of grade 8 enrollment in 2006, grade 10 enrollment in 2007, and grade 10 enrollment in 2010, adding a quadratic term (for Li-L40) gives statistically significant estimates (3) Difference and Difference and Parametric RD estimates without any controls, as well as estimates with baseline covariates but no additional program controls are available upon request (4) Covariates include status of school facilities (boundary wall, electricity, toilet, drinking water), student teacher ratio, locaton of school (rural), log of initial enrollment, private to public enrollment ratio. and other community characteristics (e.g. poverty, access) (5) Additional programs include contract teacher recruitment, program for re-building missing facilities and upgrading school infrastructure, and new school construction or upgradation.