WPS4102

IMPACT EVALUATION SERIES NO. 9

Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Program on Public School Enrollments in Punjab, Pakistan

Nazmul Chaudhury

Dilip Parajuli

The World Bank

Abstract

Instead of mean-tested Conditional Cash Transfer (CCT) programs, some countries have implemented gender-targeted CCTs to explicitly address intra-household disparities in human capital investments. This study focuses on addressing the direct impact of a female school stipend program in Punjab, Pakistan – Did the intervention increase female enrollment in public schools? To address this question, we draw upon data from the provincial school censuses 2003 and 2005. We estimate the *net* growth in female enrollments in grade 6-8 in stipend eligible schools. Impact evaluation analysis, including difference-and-difference (DD), triple differencing (DDD) and regression-discontinuity design (RDD), indicate a modest but statistically significant impact of the intervention. The preferred estimator derived from a combination of DDD and RDD empirical strategies suggests that the average program impact between 2003 and 2005 was an increase of six female students per school in terms of absolute change and an increase of 9 percent in female enrollment in terms of relative change. A triangulation effort is also undertaken using two rounds of a nationally representative household survey before and after the intervention. Even though the surveys are not representative at the sub-provincial level, the results corroborate evidence of the impact using school census data.

World Bank Policy Research Working Paper 4102, December 2006

The Impact Evaluation Series has been established in recognition of the importance of impact evaluation studies for World Bank operations and for development in general. The series serves as a vehicle for the dissemination of findings of those studies. Papers in this series are part of the Bank's Policy Research Working Paper Series. The papers carry the names of the authors and should be cited accordingly. 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.

The authors would like to thank Jishnu Das, Michelle Riboud, Tahseen Sayed and an anonymous referee for their valuable comments and suggestions on the paper.

1. Introduction

Conditional Cash Transfer (CCT) programs targeted to poor households are rapidly becoming a key policy instrument used by developing countries to reduce poverty and increase human capital investments (Janvry and Sadoulet 2006). They have ranged from employing poor adults conditioned upon their labor contributions towards public infrastructure construction (e.g., Maharashtra Employment Guarantee Scheme in India), to giving poor households in-kind food grain distribution conditioned upon sending their children to school (e.g., Food for Education Program in Bangladesh), to giving poor households cash conditioned upon school attendance of their children (e.g., Bolsa Escola in Brazil), to giving poor households cash and nutritional supplements conditioned upon sending their children to school and taking them to health clinics for check-ups (e.g., PROGRESA/Oportunidades in Mexico). Targeting poor households is, however, a non-trivial endeavor in many developing countries which often do not have systematic household/individual level income or consumption data (van de Walle 1998). Proxy means testing can be used as an alternative targeting method but its use so far has been largely limited to middle-income countries because it is administratively challenging (Coady, Grosh, and Hoddinott 2004). Furthermore, in many developing countries households under-invest in female education and health due to complex socio-economic factors, often necessitating gender-targeted interventions (Ezeminari, Chaudhury, and Owens 2000)[1].

Instead of mean-tested CCTs, some countries have implemented gender-targeted CCTs to explicitly address intra-household disparities in human capital investments. One of the most prominent examples of this genre of CCTs is the Female Secondary School Stipend (FSSS) Program in Bangladesh. On one hand it is a typical demand-side CCT which gives money to girls conditioned upon their enrollment in secondary school (among other conditions such as maintaining regular attendance and passing grades; remaining unmarried till 18). On the other hand it goes beyond most CCTs by explicitly factoring in supply-side constraints by inducing the growth of private secondary schools in rural areas by providing government financing conditioned upon the ability of the school to attract female pupils. While it is almost impossible to adequately evaluate the impact of FSSS given the absence of randomization or a properly defined control and treatment group with baseline and follow-up data, (due to rapid scaling-up of the program), the program was associated with a rapid increase in the share of female secondary

[1] Means-tested CCT programs like PROGRESA/Oportunidades pay higher cash transfers for female children, thereby addressing issues of gender disparity within a poverty-targeted framework.

school enrollment. The share of female enrollment went from less than 35% before the program, to more than 50% in less than a decade after the program. The program might actually have been 'too-successful', resulting in a reversal where we now possible have under-investment in male secondary schooling within the household (Khandker and Pitt 2003; Asadullah, Chaudhury, and Dar 2006). The FSSS program in Bangladesh has inspired a similar program in Punjab, Pakistan, with some important differences.

2. Background of the CCT Program and Evaluation Objectives

Unlike Bangladesh where the public education system is highly centralized, Pakistan has essentially devolved responsibility over public education to the provincial level. The focus of our study is Punjab, the largest and wealthiest province in Pakistan, where literacy and enrollment rates are higher and gender differences are lower compared to other provinces. Despite being the leading province in Pakistan in terms of educational outcomes, the provincial government (Government of Punjab, henceforth referred to as GoP) was concerned with a variety of problems plaguing the education sector - insufficient allocation of resources to the education sector, systemic weakness in public sector delivery as a result of over-centralization and inadequate management, and poor performance of the schooling system in terms of quality, access and governance (The World Bank, 2004)[2]. To address these issues, in 2003 GoP initiated the most ambitious reform program ever in Pakistan – the Punjab Education Sector Reforms Program. The Punjab Education Sector Reform Program (PESRP), also financed via World Bank lending, has three pillars: (A) public finance reforms to realign expenditures at the provincial and district level toward education and other pro-poor expenditures; (B) devolution and public sector management reforms; and (C) education sector reforms to improve quality, access, and governance of the education system.

Reforms within the Education Sector

Within the education sector, the reform program has introduced a number of interventions: distribution of free textbooks, upgraded/new school infrastructure, recruitment of regular and contract teachers, increasing entry qualifications of teachers, female stipend program, providing greater authority to school councils, NGO and community management of school councils. Not

[2] Program Document for Education Sector Adjustment Credit for the Government of Punjab Province, January 8, 2004.

all of these reform components have been phased in at the same time; certain components are yet to be fully implemented (e.g., enhanced community involvement in school councils).

Female Secondary School Stipend Component

This component was first implemented in February of 2004. Fifteen of Punjab's 34 districts were selected as program districts on the basis of average literacy rate of population 10 years and older as estimated from Population Census 1998 data. The cut-off literacy rate was 40 percent: 15 selected districts below the cut-off were stipend-eligible and the remaining 19 above the cut-off were not (Table 1 shows district level literacy rates). Randomization of stipend districts was not politically feasible, nor deemed to be equitable. Rather, transparent allocation of this component to the lowest-literacy districts was considered to be the least contentious way to introduce this new stipend program[3].

Under the CCT program the girl receives a stipend conditional on her being enrolled in grade 6-8 in a government girl's school[4] in a target district and conditional on her maintaining average class attendance of at least 80 percent. Eligible student receives Rs.200 per month[5]. For households which were going to send their daughters to middle/secondary school anyway, it is essentially an income transfer. Only for households with out-of-school daughters, is it an incentive payment. Funds are transferred directly to the student (student's household) via postal money order from the Education District Office (EDO), which has a special account at the Provincial Program Monitoring and Implementation Unit (PMIU) of the Education Department. An important difference with the Bangladesh FSSS program is that females are eligible for the stipend *only if they attend public secondary schools*. Indeed the secondary education sector has a dearth of private schools (unlike the primary education sector), however, linking the stipend program to only public schools highlights the initial public focus of the reform agenda. One of the long-run objectives of the stipend program targeted to females is to increase the supply of female teachers and other female service providers (e.g., health care staff), particularly in low literacy and relatively poorer regions of Punjab (most of the stipend districts fall in southern Punjab). Interestingly, Andrabi, Das, and Khwaja (2006), find that female graduates of public high schools

[3] Implementation of the stipend program within the 15 low literacy districts could have been randomly phased in, however, there was no variation in program implementation within the target districts either.
[4] All public secondary schools in Punjab are segregated by gender.
[5] In dollar terms, this is little over $US 3 per month. Household survey data from Punjab province (PIHS 2001-02) estimates that per student out-of-pocket cost of middle school education in a public school averages Rs. 160 per month. Median household consumption expenditure from the same data is Rs. 5,900 per month.

might have been instrumental the in staffing of private primary schools and, thereby contributing to the growth of the private sector.

In the PESRP program document[6], it notes that: a) female middle school enrollment rate increased from 43% (baseline 2003) to 53% in 2005, b) share of female enrollment in government primary and middle school increased from 45% in 2003 to 50% in 2005, and c) female dropout rates between grade 5 and 6 decreased by 25%, and that in middle school decreased by 20%. While these performance indicators point towards an increase in enrollments and decrease in dropouts, none of them specify the intended net impact in a counterfactual framework (in program districts relative to non-program districts, or in girls schools relative to boys schools). Furthermore, given that post and pre reform school census data forms the basis of these figures (rather than household-based data), combined with the fact that school census data do not have individual student identifiers, the dropout rates should be taken with several grains of salt, at best.

Given that PERSP covers the entire province, evaluating the overall impact of the reform on enrollments becomes quite difficult in the absence of a suitable counterfactual. The State Bank of Pakistan (2006)[7], SBP, recently conducted an impact analysis of the Punjab Education Reform program using the entire province as the treated unit and the other three provinces of Pakistan as counterfactuals. The analysis is based on a difference-in-difference estimator of progress in gross and net enrollment indicators reported in the Pakistan Social and Living Standards Measurement survey (2004-05). They conclude that the program contributed significantly to improving gross and net enrollments and that suggest that extension of similar program activities elsewhere in the country would substantially improve the country's education access outcomes. Because SBP analysis attempts to evaluate the overall PERSP program and not any specific intervention such as the stipends program, it is difficult to attribute its usefulness to the impact analysis of girls' stipends program. While their use of relevant data and technique is a noteworthy start, the choice of other provinces as counterfactuals (controls) is problematic because of significant social, political, and institutional differences between Punjab and other provinces. Moreover, their analysis does not include schooling outcome indicators for middle school grades– the relevant indicator for female stipends program. Precise estimates (proper statistical inference such as including standard errors) of these net impacts are missing in their analyses.

[6] *Annex B* of Program Document for Education Sector Adjustment Credit for the Government of Punjab Province, January 8, 2004.
[7] State Bank of Pakistan (2006), Special Section: Impact Analysis of Punjab Education Sector Reforms in *First Quarterly Report for FY06, The State of Pakistan's Economy*.

Evaluation and Policy-Relevant Questions

This study represents the first systematic effort to evaluate the impact of the female stipend component of PERSP. There are several policy-relevant research questions regarding this stipend program[8]:

1) What is the impact of the stipend component on female enrollment in public middle/secondary schools?

2) What is the impact on overall female enrollment, public and private schools? (Despite the fact that there are fewer private middle schools in the province, there might have been some switching from private to public school.)

3) What is the impact on enrollments across household welfare quintiles? And was the stipend sufficient to attract a female child at the margin (indifferent between attending and not attending)?

4) Did it have any (negative or positive) externality effect on male schooling (via changes in the explicit/implicit cost of male and female schooling within the household)?

5) What is the impact on quality of learning?

6) Does the stipend program have impacts outside of schooling outcomes such as that on health, in particular, on marriage, fertility and malnutrition?[9]

7) Could the money currently being spent on the stipend component be rather spent on other types of interventions with greater impact?

This study focuses on addressing the first question and most direct impact of the intended intervention – did it increase female enrollment in public schools? To address this question, we draw upon data from the provincial Education Management Information Systems (EMIS) school censuses 2003 and 2005. We estimate the *net* growth in female enrollments in grade 6-8 in stipend eligible schools. Impact evaluation based on counterfactual analysis using methods including difference-and-difference (DD), triple differencing (DDD) and regression-discontinuity design (RDD), indicate modest but statistically significant impact of the intervention. The preferred estimator derived from a combination of DDD and RDD empirical strategies suggests

[8] This is not meant to be an exhaustive set of policy relevant questions.

[9] A recent study from Bangladesh indicates that the female secondary education may have had significant positive effects on fertility choices and their children's malnutrition (A paper presented at a conference on "Towards a Strategy for Achieving the MDG outcomes in Bangladesh" in June 2006 by Qaiser Khan, Syed Rashed Al Zayed, and Yaniv Stopniztky).

that average program impact between 2003 and 2005 was an increase of 6 female students per school in terms of absolute change and an increase of 9 percent female enrollment in terms of relative change. A triangulation effort of our findings using school census data is also corroborated with evidence of the impact of the stipend using two rounds of a nationally representative household survey before and after the intervention.

3. Methodology

Outcome of interest and unit of analysis

Frequently used indicators of access to education such as gross or net enrollment rates at a particular level of schooling (e.g. middle school grade 6-8) or school participation rate for a particular age-group (e.g. 10-14 year olds) by geographic regions (e.g. districts) are best estimated using representative household surveys or population censuses. The last Population Census in Pakistan was conducted in 1998. In terms of large-scale household surveys, the two relevant ones are: Pakistan Integrated Household Survey 2001-2002 (PIHS 2001-02) for the pre-intervention period, and Pakistan Social and Living Standard Measurement survey (PSLM 2004-05) for the post-intervention period. Unfortunately the sample size and statistical representation for sub-provincial level analysis limited in these surveys – hence, we can only draw upon them to supplement our impact analysis.

The primary data for this study available to us come from the PMIU's government school census for 2003 to 2005. These school censuses provide school-level information on enrollments by grade for both pre-program and post-program periods. Since public schools in Pakistan are either boys only or girls only, schools can conveniently form the units of our analysis[10]. This implies that our impact estimate on school enrollments can be measured in terms of average growth in grade 6-8 enrollments in program schools relative to non-program schools. Program effect can be interpreted in the following manner: when eligible schools offered stipends to girls, did these schools record an *extra increase* in enrollments in grade 6-8 relative to non-program schools? We will measure changes in enrollments in terms of both *absolute* and *percent growth*. While absolute growth estimates change in actual size of student enrollment, percent growth would estimate change relative to some initial figure. When there is a wide variation in initial middle

[10] We also estimate program impacts using districts as the unit of observation in one of our sensitivity checks (see section 5).

school enrollment numbers, as is the case for Punjab schools under consideration, percent growth often provides alternative meaningful information. The usual comparison group for girl-schools in program districts would be girl-schools in non-program districts. An alternative counterfactual group would be boy-schools in program-districts since boys do not receive stipends regardless of their school location.

Empirical strategy

In an ideal world of true counterfactual impact analysis, net impact of the stipend program would be the difference in the relevant outcome indicator (e.g. growth in school enrollment) in *presence* of the intervention (treatment) and that in *absence* of the intervention (control). However, one never observes a unit of evaluation (individuals or schools or districts) in both states (presence of and absence of program) at the same time, thus leading to evaluation problem of missing data. A naïve method to address such missing data problem is to focus on changes (or growth) in school enrollments over time *within* stipend districts only. This type of estimator can simply overestimate the impact – enrollments could have increased in those districts even in the absence of the stipend program.

Difference-in-difference approach

A more suitable approach is to measure changes in girl enrollments in both stipend and control districts over time and take the difference in changes between the two groups. This so-called difference-in-difference (DD) (or double difference) estimator assumes enrollment growth in non-program schools as the counterfactual indicator. DD estimator takes care of the time-trend sensitivity present in before-after estimator by sweeping out time-invariant effects on school enrollments and thereby netting out any growth that could have occurred even in the absence of the stipend program (Ravallion, 2005).

There are in fact three sets of potential control groups we can use in our analysis: girl schools in non-program districts (control 1), boy-schools in program districts (control 2) and non-program districts (control 3). We will use all three control groups when using triple-differencing method, DDD). The following matrix in Table A clarifies this point:

Table A:

Comparison groups	Stipend districts	Non-stipend districts
Girl schools	Treated	Control 1
Boy schools	Control 2	Control 3

Outcome indicators	Stipend (s)	Control (c)
Before-program period (b)	Ys,b	Yc,b
After-program period (a)	Ys,a	Yc,a

Net Program impact is then given by $(y_{s,a} - y_{s,b}) - (y_{c,a} - y_{c,b})$ = enrollment growth for stipend (treated) group (*s)* between before year (*b)* and after year (*a)* minus enrollment growth for control (non-stipend) group *(c)*.

For two periods of data, this can be estimated using the convenient regression[11]:

(1) $$\Delta y_i = \beta + \alpha S_i + \varepsilon_i$$

where Δy_i denotes growth in enrollments for school *i*, *S* is the categorical variable that takes a value 1 for stipend-receiving school and 0 otherwise. The coefficient α measures the net impact of the stipend program, while the standard error of the coefficient would provide the statistical significance.

The crucial identification assumption in DD estimator is that outcomes in treated and control groups would have followed a parallel path in the absence of the program. In other words, growth in enrollments would have been the same in both groups has there been no stipend program introduced.

However, this assumption is unlikely to be satisfied, thereby making our simple DD estimator biased. The fact that program was not randomly assigned but targeted to low literacy districts

[11] Alternatively one could use this regression specification $y_{it} = \mu + \gamma S_i + \delta T_t + \alpha S_i * T_t + e_{it}$ where time period dummy is denoted by (T=1 for t=2 and T=0 for t=1). α , now the coefficient on interaction of Time dummy and Stipend dummy, would still be the net impact of the program. Because of panel nature of data and presence of serial correlation, one would need to use cluster (at the individual unit, i.e. school id code in this case) option to get correct standard errors (see Bertrand et al. (2004)).

implies that the targeting criteria itself may influence subsequent growth rates, leading to new form of selection bias (Ravallion, 2005)[12]. Because stipend districts are less literate (targeting basis) and possibly less well-off, their initial conditions may very well result in smaller growth in enrollments compared to non-stipend districts that are relatively better-positioned for higher growths (Jalan and Ravallion (1998) illustrate this point in their impact assessment of poor-area development programs in China). In econometric terminology, equation (1) suffers from an endogeneity problem where stipend status S is correlated with the error term e, making the estimated alpha coefficient inconsistent.

Alternative DD estimator

There are a number of ways to deal with this. One could choose boy schools in stipend-districts as control group for girl schools in stipend-districts. This is still a DD approach, except now that the counterfactual is the growth in enrollment in boy schools (that are not receiving stipends) in stipend districts. This would likely alleviate the problem of initial district welfare and literacy conditions that affect subsequent growth. Having treatment and control groups in the same areas would make them more comparable from this standpoint. However, long standing gender preferences ("discrimination") in the country and subsequent differential improvements in their educational attainments over time makes it very restrictive to assume that enrollment growth of boys and girls would have been similar in the absence of the program. The fact that stipends are targeted to girls implies that their initial (socio-cultural) conditions are not comparable.

Difference-in-difference-in-Difference estimation (DDD)

Because we have data for boys and girls in stipend and non-stipend districts, we can use a method that involves using DD estimate for girls (stipend districts vs non-stipend districts) and DD estimate for boys (stipend districts vs non-stipend districts) and take the difference. In essence, this entails triple differencing (DDD): DD estimate for girls minus DD estimate for boys. Such DDD estimator still allows one to compare girl enrollment growth in stipend relative to non-stipend districts but also provides a cleaner way to "separating out" some of the bias from the differential growth effects that arise due to gaps in initial literacy rates. There are other programs (such as infrastructure improvement) in effect across the entire Punjab province with similar objectives (e.g. to increase overall school participation for both boys and girls) but differential treatment effects of such programs (between low and high literacy districts) will also be

[12] The selection bias would not have been a so much of a problem if it were time-invariant and additive because DD technique sweeps away this bias by taking changes over time between two groups.

accounted for in this triple differencing strategy. DDD is given by $(\Delta y_{FS} - \Delta y_{FC}) - (\Delta y_{MS} - \Delta y_{MC})$ where subscripts FS denotes female school in stipend district, MC male school in non-stipend control districts and so on. $(\Delta y_{FS} - \Delta y_{FC})$ is our usual DD estimator of net growth of female enrollment in stipend vs non-stipend districts, $(\Delta y_{MS} - \Delta y_{MC})$ takes care of differential growth in absence of the program and presence of alternative programs, $(\Delta y_{FS} - \Delta y_{MS})$ sweeps out any common alternative program effect in stipend districts only (as long as these alternative programs are not gender-targeted or have differential effects by gender). This can be estimated by using our convenient regression framework:

2) $$\Delta y_i = \mu + \gamma S_i + \delta F_t + \alpha_{DDD} S_i * F_t + e_{it}$$

Where S is the stipend district (1,0) and F is the female schools (1,0). α_{DDD} , the coefficient on the interaction of Female and Stipend categorical variables, provides us with DDD impact estimate.

Other Covariates

Some of the bias in the estimated impact can also be alleviated through addition of covariates in the DD or DDD regression specifications of equations (1) and (2) respectively. There are broadly three sets of characteristics that affect growth patterns in public schools.

First is the school attributes themselves. Variables such as location (rural/urban), availability of proper water supply, electricity, toilet facilities, and boundary-walls provide some indication of school infrastructure quality. Student-teacher ratio (total students divided by total teachers at work) can either signal quality or absorption capacity: smaller ratio may mean better learning but also higher intake capacity without putting too much burden on teachers.

Second is the presence of private school (for middle school grades) in the area. Andrabi, Das and Khwaja (2005)[13] report that private schooling (share of private sector in total enrollment) in Pakistan grew significantly in the nineties, especially in Punjab province, and the increase was recorded in rural areas and among middle-class and poor population groups. It is not obvious

[13] Private Schooling: Limits and Possibilities. http://www.economics.pomona.edu/Andrabi/Research/

whether an area with a larger share of private sector in total enrollments (for girls in middle school grades) may have a positive or negative impact on growth in public enrollment when an incentive such as the stipend program comes into effect. But Andrabi et al. observations that private schools are located in larger villages with better infrastructure and in places where there are public schools imply that one can expect some influence of private schools on public schools ability to attract students in areas with the stipend program. However, this should not be a major issue in this context given that private schooling is predominantly in the primary (and tertiary) sector.

Third is the general welfare of the community- households residing in relatively poor communities may respond differently to incentives than those from richer communities when everything else is constant. Controlling for these three types of initial conditions should help us better identify the impact estimates.

Regression discontinuity design

The third way to deal with the selection bias is to use a regression-discontinuity design (RD design) technique that (interestingly) exploits the program design itself. RD design makes use of discontinuities generated by program eligibility criteria such that program assignment is based on a cut-off point of some assignment variable. In the Punjab case, 40 percent literacy is the cut-off point: all districts below this cut-off are eligible for girls stipends while those above are not. This is a case of *sharp RD design*[14], where there is no region of common support in the propensity-score matching language of evaluation literature (Heckman et al. (1997), Rosenbaum and Rubin (1983)). In other words, probability of being exposed to treatment changes from 0 to 1 discontinuously as one crosses the cut-off (see Figure 1). RD design identifies and estimates the program impact in the neighborhood of the cut-off point for selection into the program. Within a small interval of the cut-off, schools on either side of the cut-off have similar values of district literacy rate and can be thought to have similar enrollment growth in absence of stipends or in presence of stipends, and therefore form "very good" comparison groups[15]. Since stipend eligibility is determined such that S=1 for literacy (L)<40% and S=0 L>40%, for some arbitrarily small value h>0, the impact estimate is given by

[14] The other case is called fuzzy RD design where a cut-off of some variable does not fully determine one's treatment exposure.

[15] A more formal treatment of RD designs is found in Hahn et al. (2001).

(3) $$\alpha = E(\Delta y_i \mid L_{40} - h) - E(\Delta y_i \mid L_{40} + h)$$

In the RD literature, there are two broad methods to calculate this net impact: non-parametric and parametric. While non-parametric approach avoids any misspecifications in that of the parametric methods, it effectively uses a smaller sample of observations close to the cut-off point. Parametric specification on the other hand extrapolates from above and below the cut-off point to all observations in our data, and also allows for heterogeneous impact estimates. We will use parametric approach for most of our analysis but we also present results from non-parametric methods in our sensitivity checks.

Parametric approach of RD design

Following Van der Klaauw (2002), our parametric specification is the following:

(4) $$\Delta y_i = \beta + \alpha S_i + \lambda(L_i) + \omega_i$$

Where $\lambda(L)$ is a control function of district literacy rate L. Since L is the only systematic determinant of stipend status (S) in our *sharp RD design*, L will capture any correlation between S and the error term in equation (1), thereby freeing S from any possible endogenous contamination. α, therefore, provides a consistent estimate of the program impact. To allow for heterogeneous impacts across literacy spectrum and non-linear specification of the control function, one can also estimate the following regression specification that includes quadratic[16] of the control function and their interactions with S:

(5) $$\Delta y_i = \beta_0 + \alpha S_i + \beta_1 (L_{40} - L)_i + \beta_2 (L_{40} - L)_i^2 + \lambda_1 (L_{40} - L)_i . S_i + \lambda_2 (L_{40} - L)_i^2 . S_i + \upsilon_i$$

Where L_{40} denotes the cut-off value, 40% literacy rate, for stipend eligibility. In this framework, α still gives us the average impact effect at the cut-off, but now one can also see differential average impacts across various program districts. From a policy point of view, impact effect at the cut-off may be more relevant. If one wants to expand the program to districts above the cut-

[16] We try all specifications in our analysis: linear (only linear terms of main effect and interaction effect), quadratic (linear and quadratic) and cubic (linear, quadratic and cubic) terms of main effects and interaction effects. Use of $(L_{40} - L)$ term in the regression allows us to conveniently read off the coefficient on stipend dummy and interpret it as the impact estimate at the cut-off value.

off, this "marginal" impact estimate will give a picture of what the response may be from those areas.

Our DDD estimator can also be estimated within the RD design by adding a control function of literacy rate in our DDD regression specification:

$$\Delta y_i = \mu + \gamma S_i + \delta F_t + \alpha_{DDD} S_i * F_t + \lambda(L_{40} - L_i) + e_{it} \quad (6)$$

Non-parametric approach of RD design

Non-parametric RD method relaxes those restrictions in parametric approach such as model specification for outcomes and its functional forms. We implement this less restrictive non-parametric approach using local linear regression (LLR)[17] as suggested by Hahn et al. (2001) and Porter (2003) and used in Ludwig and Miller (2005). The LLR can be estimated using the following equation:

$$\Delta y_i = \phi_0 + \alpha S_i + \phi_1 (L_{40} - L)_i + \phi_2 (L_{40} - L)_i . S_i + \nu_i \quad (7)$$

with kernel[18] weights $w_i = K(\frac{L_{40} - L_i}{h})$ for a chosen bandwidth[19] of h. In essence, this is a kernel-weighted local linear regression, where data points close to the cut-off have bigger weights and those further away have smaller weights. α obtained from this non-parametric approach provides us with the average impact effect at the cut-off. Standard errors of this local average treatment effect are calculated using bootstrapping method. We also estimate α_{DDD} using this strategy by including female dummy and interaction term of female and stipend status to equation (7).

[17] Alternatively, one can use standard kernel estimators (such as those from one-sided kernel regressions) but these can suffer from possible boundary problems and bias in impact estimates (Hahn et al 2001).

[18] We use Epanechnikov kernel: K(Z)=(3/4)(1-Zsquared) for |z|<=1, 0 otherwise. This kernel is widely thought to be most optimal for local linear regressions.

[19] Choice of bandwidth is somewhat ad-hoc since there is no widely agreed upon method for optimal choice. We use a range of choices: 4, 8 and 12 for our sensitivity analysis. Larger bandwidth would increase bias while smaller bandwidth increases variance. Choice of bandwidths 4, 8 and 12 give positive kernel weights to approximately 1600, 2600 and 3600 observations out of 5000 sample size of girl middle schools.

4. Data

Having discussed the estimation strategy in the previous section, we briefly describe the data before moving onto the results of our impact analysis. We use school-level census data from Punjab EMIS for 2003-2005. Even though school census data exists for all three years (2003, 2004 and 2005), majority of our analysis is based on data from 2003 and 2005 censuses. The main reason for doing this is the following: despite the fact that stipend program started at the beginning of 2004, the usual lag in both implementation (supply side) and response from the intended recipients (demand side) means that the full meaningful effect could not have taken place until a year later, thereby making 2005 a more suitable post-program data point.

Table 2 presents aggregate enrollments in grades 5 through 9 by year, gender and stipend eligibility of districts. For any given year and grade, girl enrollments in public schools are much lower than boy enrollments in each of stipend and non-stipend districts as a group. Gender bias is more pronounced within low literacy stipend-districts. In terms of trends, growth between 2003 and 2004 is prominent only for grade 6 girl enrollments in stipend districts. For grades 7 and 8, marked changes in girl enrollments are seen only by 2005, suggesting a lag in program impact, justifying our choice of 2005 as the end year of our impact analysis. Enrollment growth among boys are also large in both treatment and control districts, implying there are other forces in effect (not to mention other common reform components) in both areas since stipend are targeted to girls only. Grades 5 and 9 also record notable growth, but more so among girls and in stipend districts suggesting both “pull” and “push” effects of the program. Figures 2 and 3 show levels and growth rates, respectively, for combined enrollments for grades 6-8 by gender and stipend status.

EMIS data also has some useful school-level covariates such as infrastructure variables: rural/urban location, availability of water, electricity, toilet, and boundary walls. We generate yes/no (1,0) categories of each of these variables. We also calculate student-teacher ratio at the school level as the ratio of total enrollments to total number of working teachers. The extent of private schooling is proxied by ratio of private enrollments to public enrollment in a given locality. Data limitation makes us use private enrollment data from 2000 Census of Private Educational Institutions in Pakistan (PEIP 2000) and public enrollment from EMIS 2003. Ratios are calculated at the district level and provide an indicator for initial share of private sector that might affect growth in public enrollment later in time. Economic welfare indicator is measured by

average poverty rate at the district level as estimated from Pakistan Integrated Household Survey (PIHS) 2001-02. Finally, literacy rate - the deterministic variable for program eligibility- comes from the 1998 Population census.

We construct two main outcome variables: 1) absolute change in enrollment in grade 6-8 between 2003 and 2005 and 2) percent change in enrollment in grade 6-8 between the two periods. For secondary program effects, we also analyze enrollments in grade 5 and in grade 9 separately to gauze any possible *pull* and *push* effects. As mentioned earlier, schools are our primary units of analysis. For a balanced comparison, we restrict our impact analysis to those schools that have non-zero enrollments in each of grade 6, 7 and 8 for both pre-program and post-program periods.

Descriptive statistics of outcome variables and other covariates are presented in Table 3. Not only are there fewer girl schools for middle schooling level in stipend districts relative to both boys and relative to non-stipend districts, their average enrollments are lower than all other groups in both 2003 and 2005. But female middle schools in stipend districts (relative to three other groups) reported the largest increase in enrollments both in absolute level and percent terms. It should be noted that while the majority of schools saw a positive growth in enrollments, there were a sizable number of schools that reported decrease in enrollments in grades 6, 7 and 8. For instance, about 20% female middle schools and 25% of male middle schools in non-program districts recorded negative changes. On the other hand, only 7% of girl schools and 13% of boy schools in stipend districts recorded such negative growth.

In terms of other covariates, average literacy rates, not surprisingly, are lower and consumption-based poverty rates are higher in stipend-districts. While the ratio of private schooling to public schooling is higher among non-stipend districts relative to stipend districts, it is higher for girls than for boys in both control and treatment districts. School infrastructure variables, in general, are more favorable in non-program districts. As one would expect, boundary walls and toilet facilities are available in higher proportion of girl schools. In terms of changes from 2003 to 2005, on average additional 8% of girl stipend schools built toilets; the comparable figures were 9% for female non-stipend schools, 19% for both types of male schools. And finally, student-teacher ratios are lower in stipend program districts, suggesting there is some space in their absorption capacity.

5. Impact Estimates

We present the results of impact estimates obtained from different empirical strategies in tables 4-11. Table 4 shows impact estimates of absolute growth in enrollments in middle school grades 6-8 combined using basic double differencing (DD), DD combined with regression discontinuity design (RDD) and inclusion of other covariates. Table 5 does the same using triple differencing (DDD) and its combination with RDD and covariates. Tables 6 and 7 report impact estimates of percent growth using the specifications used in tables 4 and 5 respectively. Sensitivity analysis is presented in tables 8-10 and district level estimates in tables 11.

Absolute change in enrollments

Average treatment effect of the stipend program is estimated to be an increase of approximately 12 students per school between 2003 and 2005 when using DD approach. Estimates change very little when we use regression-discontinuity design and add other covariates that are deemed to affect our estimates (table 4). Our preferred[20] specification is column (4) in the table: DD combined with RDD and other covariates produces very similar estimate. The last column (column 5) in the table controls for the initial size of enrollment in a school.

On the other hand, impact estimates from triple differencing (DDD) strategy are roughly half of those obtained from DD method. In this approach, the average program impact is an increase of 6 students per school (column 4 of table 5, the preferred model specification) and statistically significant at the 99% confidence level. Within DDD approach, results do not vary that much across different specifications.

If DDD strategy is taken to be more suitable than DD strategy in separating out the net impact of the stipend program for girls (as we argued in our methodology section), program impacts are overestimated using DD method even when we combine it with RDD approach and add other control variables.

While our main objective is to report impact estimates from the stipend program, results from our specification with other covariates (see columns 4 in both tables 4 and 5) provide useful information on whether the initial level of covariates under consideration have significant

[20] Preferred because we believe it produces the best set of casual impact estimates after sufficiently accounting for selection bias and omitted variable bias.

association with change in enrollments. Lower the literacy rate higher is the change in enrollment (significant in DDD approach). Poverty rate, for instance, has a positive correlation. Share of private schooling too has a positive association. Infrastructure variables generally have positive correlations, with strongest effects for electricity and toilet facilities. Rural location seems to have a negative association.

Percent change in enrollments

As reported in tables 6 and 7, size of impact estimates for percent change in enrollments vary across specifications within and between DD and DDD approaches. Basic DD estimate for percent change is over 23 percent (column 2, table 6), but it goes down to 10 percent when using RDD and covariates (column 4, table 6). Basic DDD estimate is 11 percent and that in our preferred specification (column 4, table 7) is 9 percent. Interestingly, the impact estimate obtained from DDD and covariates is statistically insignificant when initial level of enrollment is accounted for (column 5, table 7).

Illiteracy rate, poverty rate and share of private schooling all have positive associations with percent change in enrollments, while school infrastructure variables now have negative correlations. Interaction effects suggest that impact estimates on percent change in enrollments are stronger for rural areas but school-level infrastructure variables do not indicate differential effects.

Sensitivity analysis

We present three sets of sensitivity analyses: i) applying our preferred specifications (RDD and covariates) to both DD and DDD for each of grades 6, 7 and 8 separately; ii) adding possibilities of heterogeneity impacts and non-linear functional form to our preferred RDD approach, and iii) using non-parametric local linear regression in RDD approach. Sensitivity of impact estimates are reported in tables 8-10.

While DD estimates for absolute increase in enrollments are significant for all grades, magnitude is the largest for grade 6 not surprisingly (table 8). But our DDD estimate for the same reports no impact for grade 6 but positive impact for grades 7 and 8. For percent change outcome indicator, impacts are seen only for grade 8 using DD and only for grades 7 and 8 using DDD.

As seen from table 9, impact estimates for absolute change do not vary significantly when using interaction of program status with literacy to account for heterogeneity and adding quadratic and cubic terms to account for non-linearity (column 2 and 3). Those for percent change, however, are sensitive to alternative specifications. For instance, none of the estimates using DDD are significant. Failure of some of these estimates to hold up with possible non-linearity and heterogeneity need not be concerning, it simply implies we do not need to introduce extra terms in our specifications.

Finally, use of non-parametric approach in RD design strategy produces interesting results. Close to the program eligibility cut-off point (e.g. using smaller bandwidth), impact estimates in terms of absolute changes are larger in magnitude (table 10). For instance, within 4 percentage points of the cut-off literacy rate (between 36 and 44), average impact estimate is an increase of 20 students for DD and 10 students for DDD estimators. Increasing the bandwidth to 12 percentage points on each side, estimates are reduced by half for each estimator. Estimates for percent change vary less across both DD and DDD and across choice of bandwidth, ranging from 16 to 21 for DD and 14 to 17 percent for DDD.

District level estimates

Table 11 reports program impacts in terms of percent change in enrollments at the district level. Results, in general, support those obtained from our school level analysis for grades 6-8 individually and combined. While district-level analyses suffer from problems associated with grouped data, they provide an alternative check to our overall results. Moreover, the results for grades 5 and 9 indicate that DDD estimates are statistically insignificant implying absence of "pull" and "push" effects at early stages of the stipend program.

6. Supplementary Evidence from Repeated Cross-sectional Household Surveys

We use child-level data from PIHS 2001-02 and PSLM 2004-05 household surveys to provide an additional set of estimates for the impact of girls-stipend program on school participation rate for 10-14 year old children. Because of limited sample size, the estimates of gross enrollment and net enrollment rates are not representative at the district-level (the surveys were meant to be representative up to the province level only). This does not allow us to do an in-depth impact analysis at the district-level. However, the aggregate impact estimates of the stipend program on school attendance that we provide in this section are statistically reliable assuming the sample

statistics are representative at least two groups of districts in Punjab: 15 stipend districts together as one treatment group and 19 non-stipend districts as one control group. Repeated cross-sectional surveys, namely PIHS 2001-02 and PSLM 2004-05, are the only other independent sources of data available to us to compare our results from the provincial school census data.

The official age-group for middle school grades in Pakistan is 10-12 years. Age-grade profile from the survey data indicates that there are a large number of 13 and 14 year olds attending middle school grades because of either delayed entry to schools or higher repetition rates. For this reason, we focus on children of ages 10 to 14 years inclusive. Note that individual data from household surveys cover both public and private schools and the estimates on school attendance we present are for all types of schools. Descriptive statistics on the proportion of school attendance[21] for 10-14 year olds are shown in table 12. Levels and trends in these proportions follow similar patterns to those from the school census data. Girls in stipend districts have the lowest attendance rate in both periods, but they record the highest increase from 29 percent in 2002 to 48 percent in 2005. There is an increase of 7 percentage points for girls in non-stipend districts. Comparable figures for boys in stipend and boys in non-stipend are increases of 6 and 7 percentage points respectively.

We employ the same empirical techniques to estimate the impact of stipend program. Rather than schools as units of analysis, now the units of observations are individual children aged 10-14 years old. Because the surveys are repeated cross-sections and not pure panel at individual level, our regression specification includes interaction terms of stipend dummy and post-program time dummy in case of difference-in-difference (girls only) and a triple interaction term of stipend, post-program period and female dummy in case of triple differencing (girls *net* of boys).

DD specification with RDD and covariates takes the following form:

8) $$A_{it} = \mu + \gamma S_i + \delta T_t + \alpha_{DD} S_i * T_t + \lambda(L_{40} - L_i) + \beta X_{it} + e_{it}$$

[21] One can also present results on an alternative indicator: proportion of out-of-school children. But as long as we measure progress in terms of percentage points change (and not percent change), results remain the same. For instance, increase in school participation rate from 29% in 2002 to 48% in 2005 implies proportion of out-of-school children decreased from 71% in 2002 to 52% in 2005. The change is 19 percentage points in both cases. In terms of percent change, one would be an increase of 66% while the other would be a decrease of 27%.

Where for each child i and time-period t, A is the probability of child's student attendance (1=yes, 0=no), S is a stipend dummy (1=yes, 0=no), T is the post-program time dummy (1=yes, 0=no), $(L_{40} - L)$ is the stipend-eligibility cut-off minus the actual district literacy rate to reflect regression-discontinuity design, X is the set of child/household specific control variables, and α_{DD} the impact estimate.

Similarly, DDD specification with RDD and covariates takes the following form:

$$9)\ A_{it} = \mu + \gamma S_i + \delta T_t + \psi F_i + \omega F_i * T_t + \rho F_i * S_i + \kappa S_i * T_t + \alpha_{DDD} S_i * F_t * T_t + \lambda(L_{40} - L_i) + \beta X_{it} + e_{it}$$

In this framework, α_{DDD}, the coefficient on triple interaction of female dummy F, Stipend S and post-program period T, provides the impact estimate.

Impact estimates from basic DD and DDD specifications and those from DD and DDD combined with RDD and child-specific covariates are presented in table 13. These results use linear probability models. We also use probit models[22] and the results are similar to those from linear regression models.

Average treatment effect of the stipend program on the treated is estimated to be 12 percentage points increase in school attendance for 10-14 year old girls using the basic DD approach. When DD is combined with regression-discontinuity design and controls for individual characteristics, this impact estimate is 9.6 percentage points. In the case of DDD approach, average treatment effect is an increase of 13 percentage points in girls' school participation rate *net* of boys' school participation rate. When RDD and covariates are added in this triple-differencing approach, the treatment effect remains high at 12 percentage points. All of these impact estimates are statistically significant at the 95% confidence level.

[22] Compared to linear probability models, probit models are generally used for (0,1) categorical dependent variables such as the student attendance for three reasons: i) they ensure that predicted probabilities are between 0 and 1; ii) they allow and correct for heteroskedasticity present in (0,1) dependent variable; iii) they allow for heterogeneity in marginal effects of explanatory variable of interest. It should be noted that interaction effects and their standard errors in probit-type models from standard commands in statistical packages such as STATA (e.g., dprobit, mfx) are not correct. Results in our models are corrected using appropriate formulae for nonlinear combination of estimators (Wooldridge 2002).

Because stipends are given to all girls attending public middle school grades (despite the fact the target districts in the stipend program have the lowest literacy rates), it may be of general interest to look at equity implication of these conditional cash transfers. As seen in table 14, the school participation rates for these 10-14 year old girls vary significantly across per capita household consumption quintiles. More importantly, there was a large improvement in this indicator in stipend districts between 2002 and 2005, particularly for poorer quintiles. For instance, participation rate among 10-14 year old girls in the poorest quintile rose impressively from a 10% to 37% in that period. Regarding impact heterogeneity across welfare quintiles, the estimates indicate that the net program effects were stronger for the bottom two (poorest 40%) and the highest (richest 20%) quintiles, suggesting some kind of a U-shaped effect where stipends may have had a real effect on the poorer quintiles and the impact on the richest quintile is just picking up its disproportionate share and growth in private school attendance.

In terms of benefit incidence, we also see a noteworthy change from pre-stipend period to post-stipend period. Aggregating all public middle school grade girl enrollments in the stipend program districts, PIHS 2001/02 and PSLM 2004/05 indicate that the current stipend transfers are neither pro-poor nor pro-rich. 50% of all stipends (and all other public resources in those public schools) go to girls from the poorest 50% of the population and the rest to the richest 50% (see table 15)[23]. Prior to the stipend program, the poorest 50% of girls were receiving only 38% of the government subsidy.

7. Conclusions

This study evaluates the impact of a stipend program on female enrollments in government middle schools in the Punjab province of Pakistan. It makes use of a panel dataset of schools constructed from the provincial school census for 2003 (before the program) and 2005 (after the program). Because stipends for enrollees in grades 6-8 are targeted to girls in the 15 lowest literacy districts (out of 34) in the province, any simple naïve impact estimator (such as female enrollment grew by some y percent) is likely to suffer from a substantial program placement bias. The study considers estimates obtained using various empirical approaches including double-differencing (DD) and triple-differencing (DDD) in combination with regression-discontinuity design (RDD) and controlling for covariates. The results, not surprisingly, vary with

[23] Sample size does not allow dis-aggregation by quintiles.

methodological choice. The preferred estimator derived from a combination of DDD and RDD empirical strategies suggests that the average program impact between 2003 and 2005 was an increase of 6 girl students per school in terms of absolute change and an increase of 9 percent in terms of relative change. These are modest but statistically significant program effects. We also make use of data from repeated cross-section household surveys (usually thought to be more independent and objective relative to administrative data such as the school census). Given that these household surveys are statistically representative only to the provincial level, any district-level analysis should be tempered with statistical caution. However, for what it is worth, we find an average treatment effect on proportion of school attendance for 10-14 year old girls ranging from 10 to 13 percentage points. Taken together, these results, strongly suggest that the stipend is attaining its objective of increasing female enrollment in public schools in Punjab. The evidence from the household survey data also points that the stipend program is helping children from poorer households to attend schools. Incidence of stipend subsidy also shows more equal distribution across rich-poor population groups.

Additional sources of data will help to test the robustness of our results (which essentially hinge on the veracity of the school census data), and more importantly, help address other pertinent issues besides enrollment increases. For example, Andrabi, Das, and Khwaja (2006) point out that the majority of Punjabi villages do not have middle/high schools. Sathar, Lloyd, Mete, and Haque (2000) have shown that access to school is one of the key determinants of enrollment (and other subsequent non-schooling outcomes such as contraceptive use) in rural Punjab. An immediate policy-relevant question then is what happens in villages in stipend districts which do not have a public middle/high school? To address this and other relevant questions we mentioned earlier, we are in the process of collecting relevant household and schooling data via primary surveys (instead of relying upon existing administrative data and limited household data).

References

Andrabi, Tahir, Jishnu Das and Asim Ijaz Khwaja (2005), "Private Schooling: Limits and Possibilities", processed, Department of Economics, Pomona College, USA.

Andrabi, Tahir, Jishnu Das and Asim Ijaz Khwaja (2006), "Students Today, Teachers Tomorrow? The Rise of Affordable Private Schools", processed, Department of Economics, Pomona College, USA.

Andrabi, Tahir, Jishnu Das, Asim Ijaz Khwaja, and Tristan Zajonc (2006), "Report on Learning and Educational Attainments in Pakistani Schools", processed, Department of Economics, Pomona College, USA.

Asadullah, Mohammad N,. Nazmul Chaudhury, and Amit Dar (2006), "Conditional Cash Transfer and Educational Gender Gaps: Insights from Bangladeshi households", processed, Oxford University and World Bank.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan (2004), "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119(1):249-275

Coady, David, Margaret Grosh, and John Hoddinott (2004), *Targeting of transfers in developing countries: review of lessons and experience*, The World Bank.

Ezemenari, Kene, Nazmul Chaudhury, and Janet Owens (2002), "*Gender and Social Risk Management"*, Social Protection/Human Development Network, World Bank.

Federal Bureau of Statistics, Government of Pakistan (2005), *Pakistan Social and Living Standards Measurement Survey 2004-05*.

Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw (2001), "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica* 69(1): 201-209

Heckman, James, Hidehiko Ichimura and Petra Todd (1997), "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies* 64(4): 605-654.

Jalan, Jyotsna and Martin Ravallion (1998), "Are there Dynamic Gains from a Poor-Area Development Program?" *Journal of Public Economics*, 67(1): 65-86.

de Janvry, Alain, and Elisabeth Sadoulet (2006), "Making Conditional Cash Transfer Programs More Efficient: Designing for Maximum Effect of the Conditionality", World Bank Economic Review, Vol. 20, No.1.

Khandker, Shahidur and Mark M. Pitt (2003) "Subsidy to Promote Girls' Secondary Education: The Female Stipend Program in Bangladesh", Annual Meeting of the Population Association of America.

Ludwig, Jens and Douglas L. Miller (2005), "Does Head Start Improve Children's Life Chances: Evidence from a Regression-Discontinuity Design," Working Paper 11702, National Bureau of Economics Research (NBER).

Porter, Jack (2003), "Estimation in the Regression-Discontinuity Model", Working Paper, Harvard University Economics Department.

Ravallion, Martin (2005), Evaluating Anti-Poverty Programs. *Handbook of Agricultural Economics* vol. 4. Robert E. Evenson and T. Paul Schultz. North-Holland.

Rosenbaum, Paul and Donald Rubin (1983), "The Central Role of the Propensity Score in Observational Studies of Casual Effects," *Biometrika*, 70: 41-55.

Sathar, Zeba, Cynthia B. Lloyd, Cem Mete, and Minhaj ul Haque (2000). "Schooling opportunities for girls as a stimulus for fertility changes in rural Pakistan," *Policy Research Division Working Paper* no. 143. New York: Population Council.

State Bank of Pakistan (2006), Special Section: Impact Analysis of Punjab Education Sector Reforms in *First Quarterly Report for FY06, The State of Pakistan's Economy.*

Van der Klaauw, Wilbert (2002), "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach," *International Economic Review*, 43(4): 1249-1287.

van de Walle, Dominique, 1998 " Targeting Revisited", World Bank Research Observer, Vol.13, No.2.

Figure 1: Probability of Stipend eligibility among girl schools

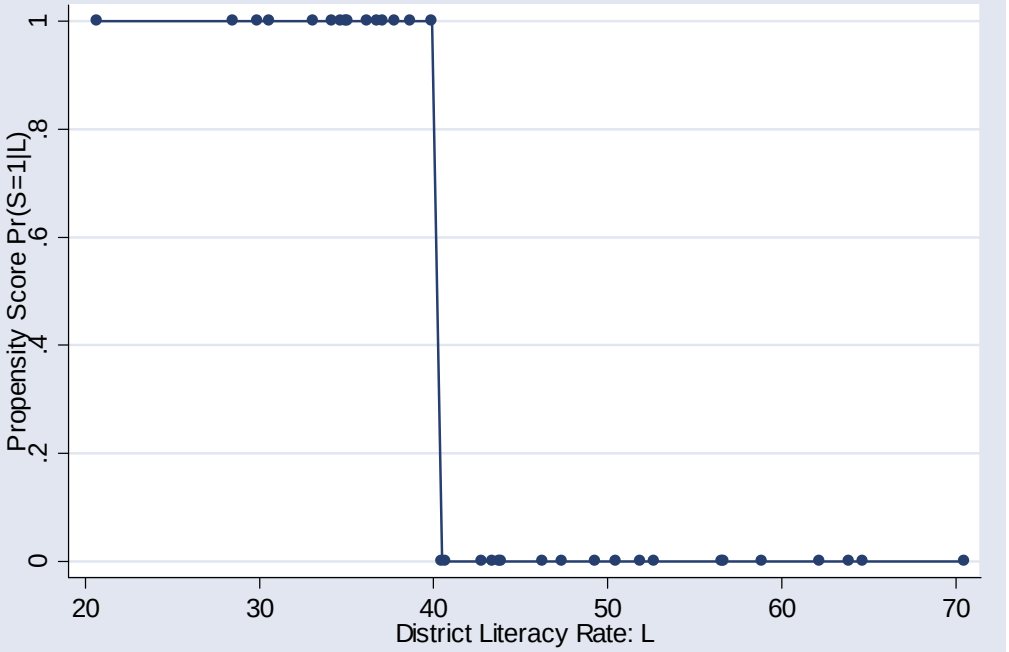


Figure 2. Total enrollment in grades 6-8 combined

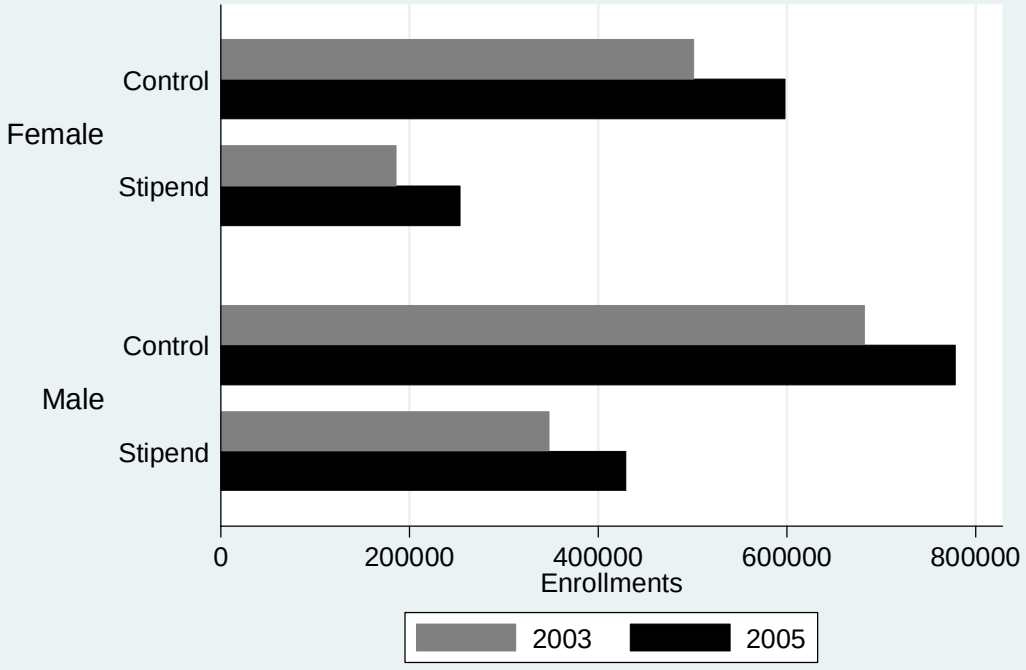


Figure 3. Growth in enrolment in grades 6-8 combined

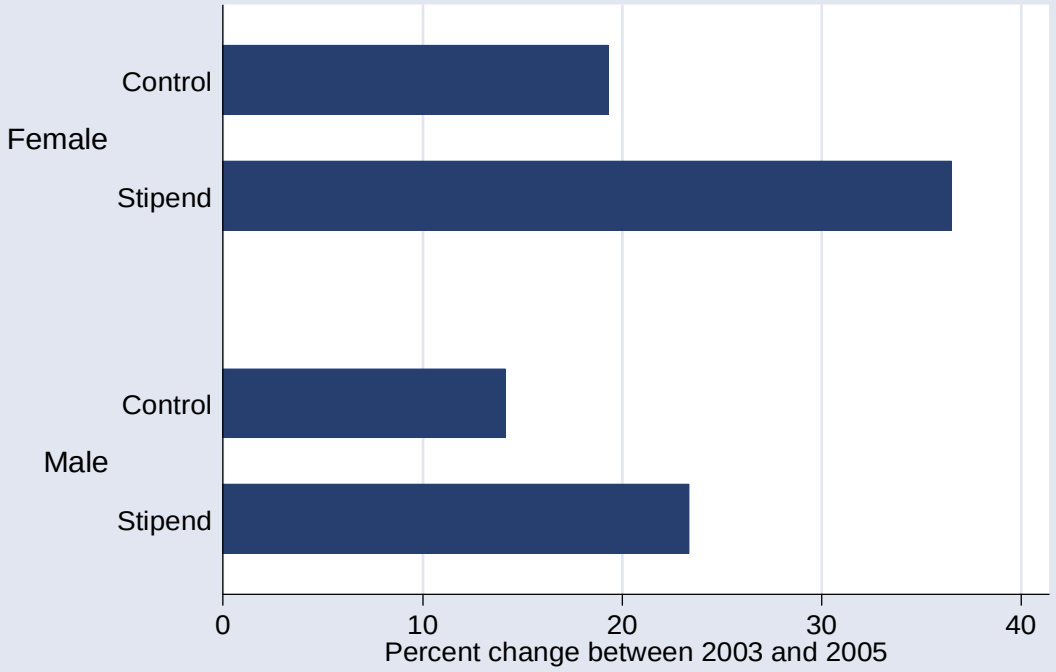


Figure 4: Average enrollment in grades 6-8 combined

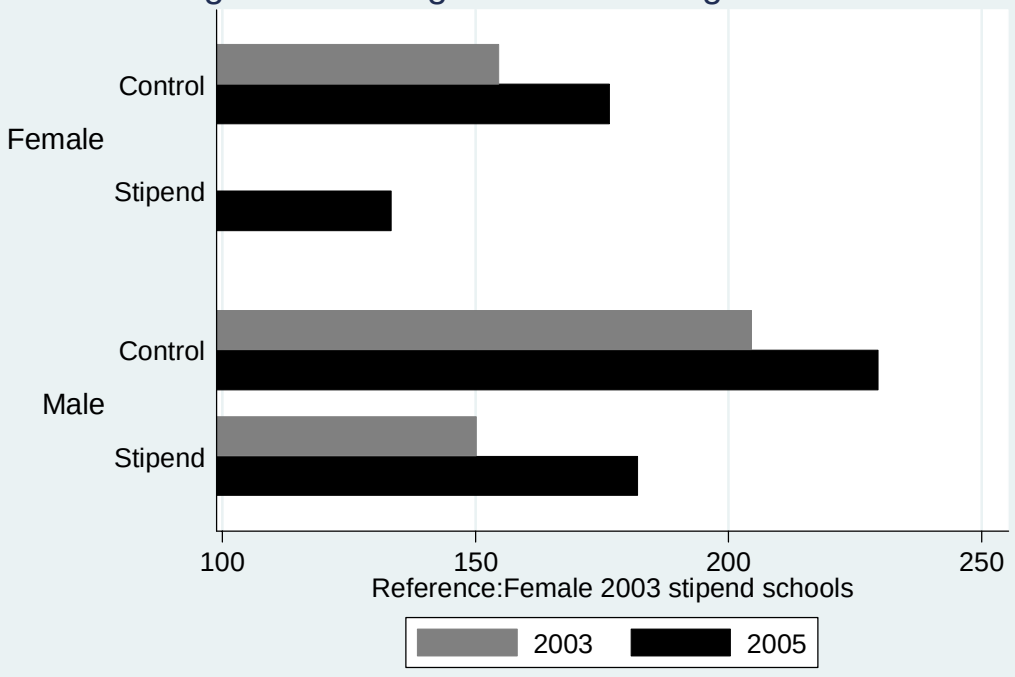


Figure 4b: Average enrollment in grade 6

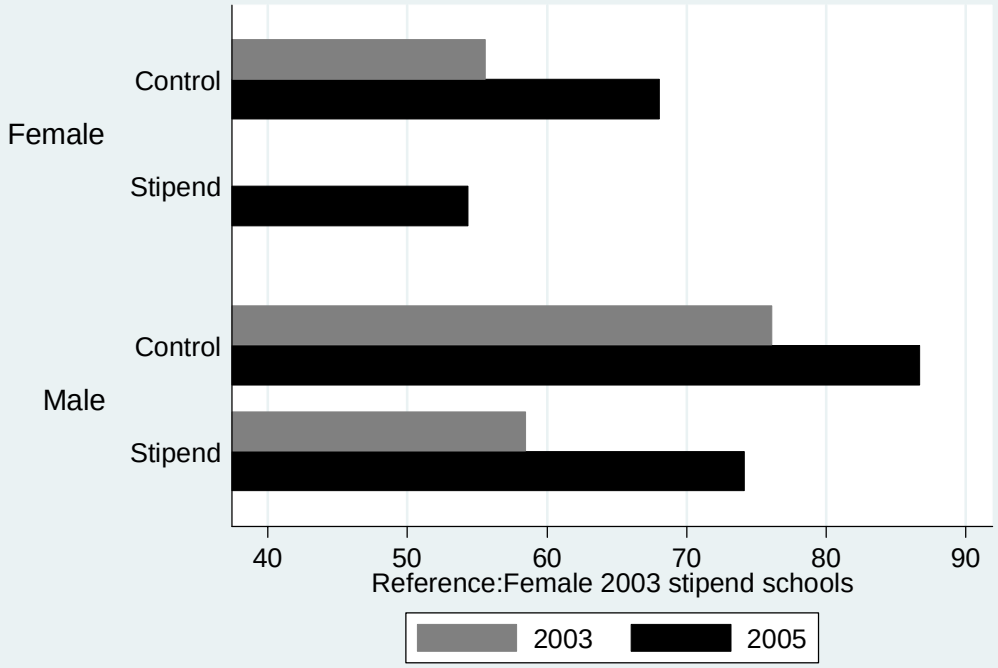
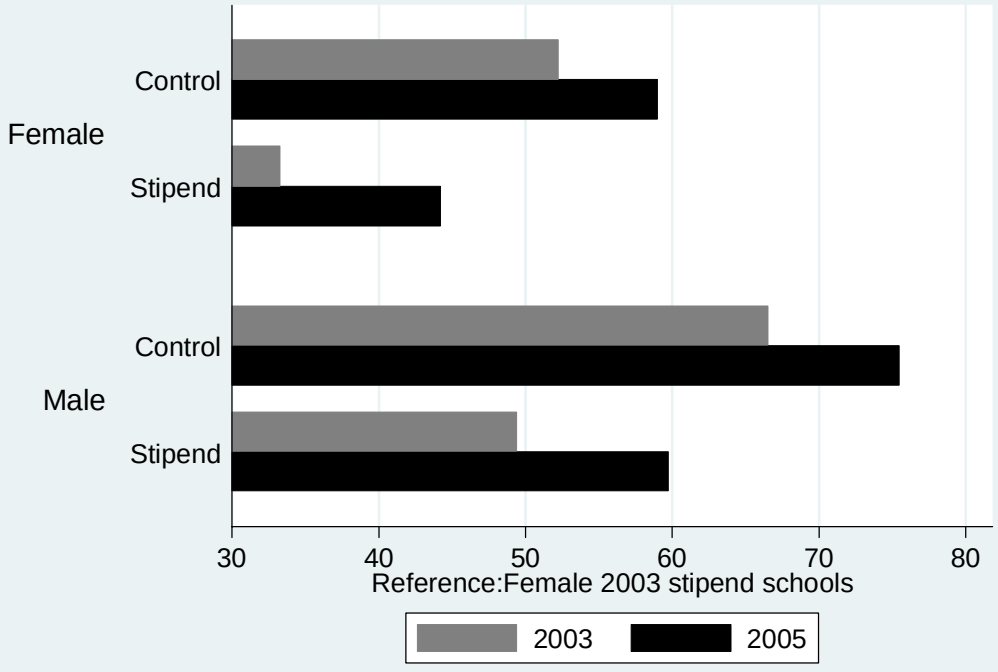


Figure 4c: Average enrollment in grade 7



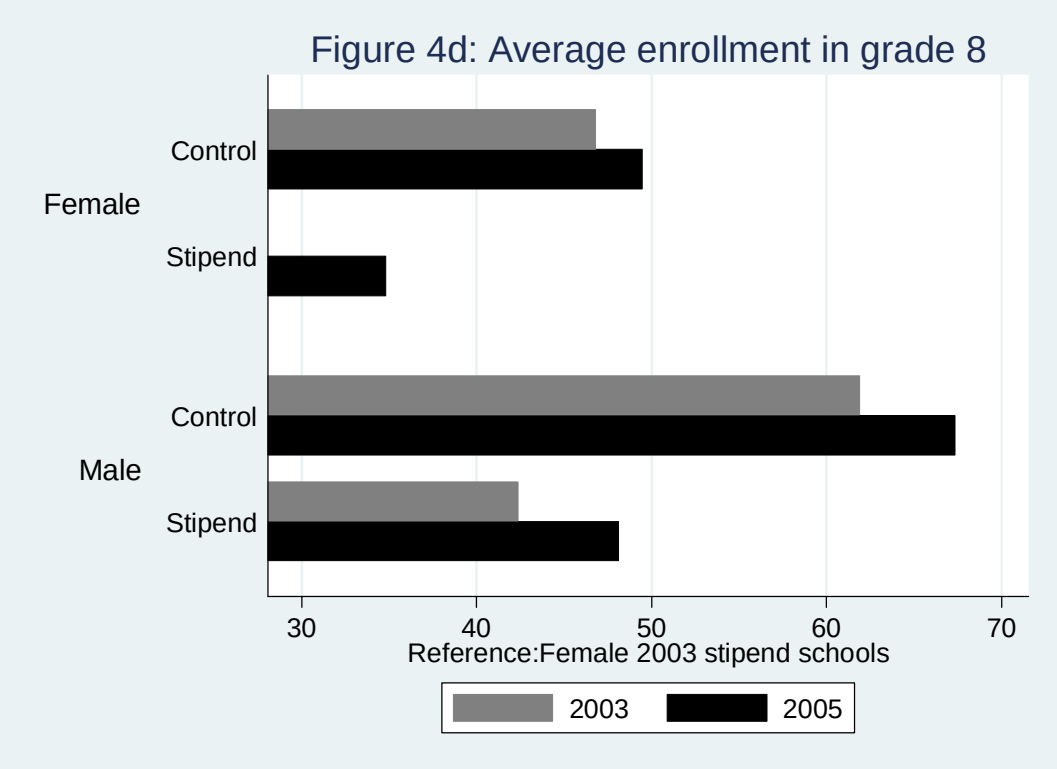

Figure 4d: Average enrollment in grade 8
Control
Female
Stipend
Control
Male
Stipend
30
40
50
60
70
Reference:Female 2003 stipend schools
2003
2005

Table 1. District Literacy rates (%) in Punjab Province, Pakistan.

District	Literacy Rate (Popul. 10 years and older)
RAJANPUR	20.7
MUZAFFARGARH	28.5
LODHRAN	29.9
D.G. KHAN	30.6
RAHIMYAR KHAN	33.1
BHAKKAR	34.2
PAKPATTAN	34.7
BAHAWALPUR	35.0
BAHAWALNAGAR	35.1
KASUR	36.2
VEHARI	36.8
JHANG	37.1
OKARA	37.8
LAYYAH	38.7
KHANEWAL	39.9
KHUSHAB	40.5
HAFIZABAD	40.7
MIANWALI	42.8
MULTAN	43.4
SHIEKHUPURA	43.8
SAHIWAL	43.9
SARGODHA	46.3
MANDI BAHA UD DIN	47.4
ATTOCK	49.3
T.T.SINGH	50.5
FAISALABAD	51.9
NAROWAL	52.7
GUJRANWALA	56.6
CHAKWAL	56.7
SIALKOT	58.9
GUJRAT	62.2
JHELUM	63.9
LAHORE	64.7
RAWALAPINDI	70.5

Source: Population Census, 1998.

Table 2: Aggregate Enrollment Trends by gender and stipend eligibility of districts

		Female			Male		
Year		2003	2004	2005	2003	2004	2005
Grade 5	Control	220385	232671	253268	285914	296639	307613
	Stipend	102016	115156	125047	195121	207418	215066
Grade 6	Control	180722	213063	233493	254399	261749	295834
	Stipend	71150	89311	105341	135908	148499	176956
Grade 7	Control	169012	169687	199895	221838	230326	256445
	Stipend	62690	68253	84299	114623	120267	140781
Grade 8	Control	151812	153820	164953	205935	208175	226401
	Stipend	52775	57591	65121	98323	101653	112528
Grade 9	Control	122966	126982	139534	183349	158895	188380
	Stipend	38300	47013	51809	82165	86048	89980

Source: EMIS 2003, 2004, 2005

Table 3. School Level descriptive statistics of selected variables.

	Female		Male	
	Control	Stipend	Control	Stipend
Number of Districts	19	15	19	15
Number of Schools (N)	3156	1779	3265	2247
Number of Schools with negative enrollment growth in grade 6-8, 2003-2005	622	129	798	292
Enrollment grade 6-8, 2003	154	99	205	150
Enrollment grade 6-8, 2005	177	133	230	182
Absolute change, 2003-2005	22.3	34.7	24.4	31.6
Percent change, 2003-2005 (%)	31.0	53.5	20.1	31.6
District Literacy rate, 1998 (%)	53.2	35.2	53.0	34.7
District Poverty Rate, 2002 (%)	26.8	33.8	27.2	35.1
District Private Enrollment ratio, 2000	0.36	0.30	0.27	0.20
Rural Location (0,1)	0.78	0.85	0.80	0.87
Water availability (0,1)	0.95	0.94	0.96	0.97
Electricity (0,1)	0.82	0.75	0.85	0.80
Toilet facility (0,1)	0.78	0.78	0.50	0.53
Net increase in toilet facilities, 2003-2005	0.09	0.08	0.19	0.19
Boundary wall (0,1)	0.84	0.85	0.58	0.43
Net increase in boundary walls, 2003-2005	0.05	0.04	0.07	0.10
Student to teacher ratio	28.7	21.6	28.3	24.4

Note: 1. All Infrastructure variables are for 2003.
2. Private schooling ratio is the ratio of grade 6-8 enrollments in private schools to public schools.

Table 4: Impact estimates of absolute changes in grades 6-8 enrollments using DD strategy and RDD approach

	M1	M2	M3	M4
Stipend	**12.45*** **	**10.87*** **	**11.66*** **	**11.88*** **
(t-statistic)	**(7.71)**	**(4.19)**	**(4.44)**	**(4.61)**
(40-Literacy rate)		0.09	0.20	0.37*
Poverty Rate			0.38***	0.33***
Private/Public ratio			14.97***	15.65***
Rural Location (0,1)			-9.88***	-1.55
Student to teacher ratio			0.49***	0.20*
Water availability (0,1)			3.62	1.21
Electricity (0,1)			5.99***	2.13
Toilet facility (0,1)			3.89**	3.10*
Boundary wall (0,1)			2.84	1.19
Initial level of enrollment (log)				10.86***
Constant	22.25***	23.41***	-10.80*	-47.47***

*** *denotes statistically significant at 99.9% confidence level, ** at 99%, * at 95%*
Note: Column 2 is basic DD, column 3 combines basic DD with RDD, column 4 adds covariates to column 3, column 5 controls adds initial enrollment size. Standard errors corrected for clustering.

Table 5: Impact estimates of absolute changes in grades 6-8 enrollments using DDD strategy and RDD approach

	M5	M6	M7	M8
Female*Stipend	**5.25***	**5.36***	**6.19****	**7.40*** **
(t-statistic)	**(2.34)**	**(2.39)**	**(2.74)**	**(3.29)**
Female	-2.18	-2.13	-4.95**	0.00
Stipend	7.21***	1.80	0.70	0.49
(40-Literacy rate)		0.30***	0.47***	0.58***
Poverty Rate			0.17**	0.15*
Private/Public ratio			13.58***	15.66***
Rural Location (0,1)			-9.28***	-2.60
Student to teacher ratio			0.32***	0.05
Water availability (0,1)			4.38*	2.12
Electricity (0,1)			5.48***	2.11
Toilet facility (0,1)			3.33**	1.88
Boundary wall (0,1)			2.46*	1.73
Initial level of enrollment(log)				10.08***
Constant	24.44***	28.27***	8.41*	-31.64***

*** *denotes statistically significant at 99.9% confidence level, ** at 99%, * at 95%*
Note: Column 2 is basic DDD, column 3 combines basic DDD with RDD, column 4 adds covariates to column 3, column 5 controls adds initial enrollment size. Standard errors corrected for clustering.

Table 6: Impact estimates of percent changes in grades 6-8 enrollments using DD strategy and RDD approach

	M9	M10	M11	M12
Stipend	**22.48*** **	**23.89*** **	**10.03***	**9.38***
(t-statistic)	**(6.47)**	**(4.28)**	**(2.56)**	**(2.42)**
(40-Literacy rate)		-0.08	0.42*	-0.09
Poverty Rate			0.35*	0.50**
Private/Public ratio			49.70*	47.64*
Rural Location (0,1)			4.82	-20.12
Student to teacher ratio			-0.57	0.28*
Water availability (0,1)			-3.49	3.72
Electricity (0,1)			-8.27*	3.28
Toilet facility (0,1)			-1.50	0.86
Boundary wall (0,1)			-0.16	4.76
Initial level of enrollment(log)				-32.50***
Constant	31.04***	30.00***	33.69***	143.42***

*** *denotes statistically significant at 99.9% confidence level, ** at 99%, * at 95%*
Note: Column 2 is basic DD, column 3 combines basic DD with RDD, column 4 adds covariates to column 3, column 5 controls adds initial enrollment size. Standard errors corrected for clustering.

Table 7: Impact estimates of percent changes in grades 6-8 enrollments using DDD strategy and RDD approach

	M13	M14	M15	M16
Female*Stipend	**11.03****	**11.11****	**8.66***	**5.66**
(t-statistic)	**(2.97)**	**(2.96)**	**(1.96)**	**(1.19)**
Female	10.90***	10.94***	8.16***	-4.11**
Stipend	11.45***	7.28	-2.22	-1.69
(40-Literacy rate)		0.23	0.72***	0.44***
Poverty Rate			0.16*	0.21**
Private/Public ratio			42.80**	37.65*
Rural Location (0,1)			3.89	-12.67*
Student to teacher ratio			-0.50**	0.18*
Water availability (0,1)			-3.00	2.60
Electricity (0,1)			-6.30**	2.06
Toilet facility (0,1)			-1.59	1.99
Boundary wall (0,1)			-0.27	1.56
Initial level of enrollment(log)				-25.01***
Constant	20.15***	23.10***	33.58***	132.92***

*** *denotes statistically significant at 99.9% confidence level, ** at 99%, * at 95%*
Note: Column 2 is basic DDD, column 3 combines basic DDD with RDD, column 4 adds covariates to column 3, column 5 controls adds initial enrollment size. Standard errors corrected for clustering.

Table 8: Sensitivity analysis: Application of preferred model for each middle school grade separately

Impact estimate and t-statistics in parenthesis

	M3 DD absolute change	M7 DDD absolute change	M11 DD percent change	M15 DDD percent change
Grade 6	5.25***	0.04	6.91	0.04
	(4.42)	(0.03)	(1.20)	(0.01)
Grade 7	2.97*	2.69*	10.54	20.32***
	(2.11)	(2.48)	(1.66)	(3.60)
Grade 8	3.44**	3.46***	20.90*	29.50**
	(2.71)	(3.50)	(2.14)	(3.06)

**** denotes statistically significant at 99.9% confidence level, ** at 99%, * at 95%*

Table 9: Sensitivity analysis: Adding interaction and non-linear terms in parametric RDD approach to preferred models

Impact estimate and t-statistics in parenthesis

	M3 DD absolute change	M7 DDD absolute change	M11 DD percent change	M15 DDD percent change
Adding interaction of (40-Literacy) and program status				
	9.36**	6.34**	12.72***	8.11
	(3.27)	(2.80)	(3.42)	(1.77)
Adding squared terms and its interaction				
	12.49**	6.50**	18.88**	8.33
	(3.08)	(2.87)	(3.23)	(1.84)
Adding cubic terms and its interaction				
	17.91**	6.73**	11.59	8.41
	(3.26)	(2.98)	(1.60)	(1.83)

**** denotes statistically significant at 99.9% confidence level, ** at 99%, * at 95%*

Table 10: Sensitivity analysis: Using non-parametric local linear regressions in RDD approach

Impact estimate and bootstrapped standard errors in italics

	DD absolute change	DDD absolute change	DD percent change	DDD percent change
Kernel Bandwidth: 4%				
	19.96	10.22	16.96	16.75
	7.44	*4.66*	*5.97*	*4.10*
Kernel Bandwidth: 8%				
	16.77	6.66	20.77	14.97
	5.30	*3.50*	*4.81*	*3.38*
Kernel Bandwidth: 12%				
	10.16	6.54	15.88	14.14
	4.27	*2.92*	*4.48*	*3.06*
Limited sample schools**				
	8.21	4.92	17.41	13.57
	2.71	*3.70*	*3.59*	*4.10*

**Our basic difference-in-difference estimated only in districts within 35-45 literacy rates range.

Table 11: District Level impact estimates of percent change in public enrollments
(associated t-statistics in parentheses)

	Grade 6-8	Grade 6	Grade 7	Grade 8	Grade 5	Grade 9
Basic DD	19.28***	19.28***	20.66***	17.14***	9.15*	21.31*
	(5.80)	(5.80)	(4.10)	(4.17)	(2.56)	(2.63)
DD with RDD and covariates	16.57**	16.57**	13.77	20.49**	2.06	31.53*
	(3.22)	(3.22)	(2.03)	(3.01)	(0.29)	(2.59)
Basic DDD	9.83**	9.83**	12.56**	12.92**	5.71	10.17
	(2.91)	(2.91)	(2.90)	(3.46)	(1.76)	(1.05)
DDD with RDD and covariates	9.35*	9.35*	11.85*	12.74**	4.76	10.30
	(2.59)	(2.59)	(2.57)	(3.31)	(1.44)	(1.08)

Note: 1.DD denotes growth in girl enrollments in stipend vs non-stipend
2.DDD denotes DD girls minus DD boys
**** denotes statistically significant at 99.9% confidence level, ** at 99%, * at 95%*

Table 12: Percentage of 10-14 year old children attending (any) school in Punjab

	Control districts		Stipend districts		Both	
	2002	**2005**	**2002**	**2005**	**2002**	**2005**
Male	72%	79%	60%	66%	68%	74%
Female	66%	73%	29%	48%	53%	63%
Total	69%	76%	45%	57%	61%	68%

Source: PIHS (2001-02), PSLM(2004-05)

Table13: Impact estimates on school attendance of 10-14 year old children (all schools)

Linear Probability Models (t-stats in parentheses)	Basic DD	DD, RDD with covariates	Basic DDD	DDD, RDD with covariates
	(girls only)		(girls *net* of boys)	
Female*Stipend*Post-program			**0.131**	**0.119**
			(2.60)**	**(2.50)***
Stipend*Post-program	**0.122**	**0.096**	-0.008	-0.01
	(2.70)**	**(2.34)***	(-0.21)	(-0.28)
Female*Stipend			-0.252	-0.238
			(-6.94)***	(-7.09)***
Female*Post-program			-0.001	-0.002
			(-0.03)	(-0.08)
Post-program (2005)	0.071	0.092	0.072	0.089
	(2.58)**	(3.72)***	(3.41)***	(4.32)***
Stipend	-0.375	-0.143	-0.123	0.05
	(-11.84)***	(-3.66)***	(-4.21)***	-1.53
Female			-0.061	-0.057
			(-2.99)**	(-2.99)**
(40- District Literacy rate)		-0.007		-0.006
		(-5.44)***		(-6.11)***
Child age		-0.061		-0.058
		(-13.15)***		(-17.56)***
Household size		0.009		0.007
		(3.49)***		(3.98)***
Deciles of pc household consumption		0.04		0.04
		(13.22)***		(18.03)***
Rural Location (0,1)		-0.145		-0.051
		(-7.08)***		(-3.33)***
Constant	0.661	1.058	0.722	1.054
	(33.29)***	(14.99)***	(45.39)***	(20.68)***
N	5164	5164	10560	10560

Notes: 1. * p<0.05, ** p<0.01, *** p<0.001; 2. Standard errors corrected for clustering.
Source: Estimates using PIHS (2001-02), PSLM(2004-05)

Table 14: Percentage of 10-14 year old girls attending school in Punjab by quintile and stipend status

	Control districts		Stipend districts		Both	
Per Capita Consumption Quintile	**2002**	**2005**	**2002**	**2005**	**2002**	**2005**
Poorest	44%	59%	10%	37%	29%	47%
Second	58%	64%	28%	43%	44%	55%
Third	62%	75%	29%	51%	50%	65%
Fourth	72%	78%	41%	53%	63%	67%
Richest	85%	88%	52%	79%	79%	86%
Total	**66%**	**73%**	**29%**	**48%**	**53%**	**63%**

Source: PIHS (2001-02), PSLM(2004-05)

Table 15. Benefit Incidence of Girls Stipend in grades 6-8 in public schools

	2002	**2005**
Consumption Groups	(Pre-program)	(Post-Program)
Poorest 50%	37.5	50.1
Richest 50%	62.5	49.9
Total	100	100

Source: PIHS (2001-02), PSLM(2004-05)