Policy Research Working Paper 9094

Effects over the Life of a Program

Evidence from an Education Conditional Cash Transfer Program for Girls

Esha Chhabra
Fatima Najeeb
Dhushyanth Raju

WORLD BANK GROUP
Social Protection and Jobs Global Practice
&
South Asia Region
Office of the Chief Economist
December 2019

Abstract

While most evaluations of education programs in developing countries examine effects one or two years after a program has been introduced, this study does so over an extended duration of a program. Administered in Punjab, Pakistan, the program offers cash benefits to households conditional on girls' regular attendance in secondary grades in government schools. The study evaluates the evolution of the program's effects on girls' secondary school enrollment numbers over roughly a decade of its existence. The program was targeted to districts with low adult literacy rates, a targeting mechanism that provides an observed, numerical program assignment variable and results in a cutoff value. Recent advances in regression discontinuity designs allow the study to appropriately fit key features of the data. The study finds that the program had positive effects on girls' secondary school enrollment numbers throughout the period and that these effects were stable. This pattern is observed despite a loss of more than 60 percent in the real value of the cash benefit over the period. The findings are consistent with potential behavioral explanations, such as the program making girls' education salient to households or catalyzing a shift in social norms around girls' education.

This paper is a product of the Social Protection and Jobs Global Practice and the Office of the Chief Economist, South Asia Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at draju2@worldbank.org.

Effects over the Life of a Program:
Evidence from an Education Conditional Cash Transfer Program for Girls

Esha Chhabra
Fatima Najeeb
Dhushyanth Raju

JEL codes: H52, I25, I28, O15

Keywords: impact evaluation, regression discontinuity design, conditional cash transfer program, education, gender

Esha Chhabra, International Monetary Fund, echhabra@imf.org; Fatima Najeeb, University of Maryland–College Park, fnajeeb@terpmail.umd.edu; Dhushyanth Raju, World Bank, draju2@worldbank.org. We thank Matias Cattaneo and Gonzalo Vazquez-Bare for guidance on the empirical strategy; Narmeen Adeel, Fiaz Ahmed, Freeha Fatima, Maria Qazi, Sohail Raza, and Abdul Samee for assistance on background statistics, data, and information on program design and implementation; and Juan Baron, Maurizio Bussolo, Tazeen Fasih, Koen Martijn Geven, and Hans Timmer for comments.

1. Introduction

How do the effects of a program evolve over many years of its life, over successive cohorts of beneficiaries? We examine this question in relation to an education conditional cash transfer (CCT) program targeted to girls of secondary school age in Punjab province, Pakistan. Specifically, we look at how the effects on girls' secondary school enrollment numbers evolved over a decade of the program's existence, from its launch in 2004 until 2015.

While there is extensive evaluative literature on education CCT programs in developing countries, most studies examine effects on the likelihood of school participation—and, to a lesser extent, the likelihood of class attendance—one or two years after the program has been introduced (Fiszbein et al. 2009; Baird et al. 2014; Garcia and Saavedra 2017). We are unaware of any program evaluations in developing countries, in education or other areas, that examine how effects on immediate outcomes evolve over many years of a program's life.[1]

The question is of interest for two reasons. First, effects on beneficiary cohorts could change over time because of changes in relevant environmental, institutional, or intervention-specific factors. In the CCT program that we examine, a key intervention-specific factor that changed was the real value of the program benefit, which was not indexed to inflation. Indeed, the value of the benefit declined substantially over the study's observation period not only in real terms but relative to average household consumption spending, weakening the program's economic incentive as observed by households. As a result, standard economic theory would predict that program effects should fall over the study's observation period.

[1] This question is distinct from that of the effects of programs on participants over the long run, such as the effects of CCT programs on individuals exposed to the program during childhood on cumulative schooling, employment, and earnings measured in adulthood. Evidence on the latter question is summarized in recent reviews by Bouguen et al. (2019) and Molina Millan et al. (2019).

Second, the effects of a program could change after the first few years of implementation as different agents have time to reoptimize their behavior in response to the program's introduction—in the case of an education program, such agents as households, school staff, and education officials (Glewwe and Muralidaran 2016). Most evaluations that examine effects in the first few years after a program's introduction are unlikely to capture such reoptimizing of behavior, which can dampen or amplify program effects (Das et al. 2013). Our investigation, by looking at the evolution of enrollment effects over the extended life of the program, allows us to converge on the program's "full" effect.[2]

The CCT program that we study, the Punjab Female School Stipend Program, is aimed at stimulating household demand for education where that demand is particularly weak, by targeting cash transfers explicitly on the basis of gender and location and implicitly on the basis of age and household economic status. The program, initiated by the Punjab government in 2004, offers households cash benefits tied to girls' regular attendance in secondary grades in government schools in districts in the province with low adult literacy rates. The school participation rate (the percentage of children attending school) in the province is lower for girls than for boys—and this gender gap is greater in districts with low adult literacy (low literacy is correlated with poor overall socioeconomic status). The gender gap in school participation also widens somewhat when children reach preadolescence. Moreover, poorer households are less likely than richer ones to send their children to school. And when they do send their children to school, they are more likely to send them to government schools, which provide free education, while richer households are more likely to send their children to fee-based private schools (Nguyen and Raju 2015).

[2] Das et al. (2013) and Glewwe and Muralidharan (2016) refer to the "full" effect as the "policy" effect, while they argue that the estimated effect in the early years of a program likely represents the "production function" effect (when other factors have not had the time to adjust).

Assignment of the program is based on a clear rule. Of the 34 districts in Punjab at the time of its launch in 2004, the program was offered to the 15 districts with the lowest adult literacy rates, which results in a cutoff literacy rate of 40 percent. Data from before the program's conception were used for the assignment variable. Thus, there is no risk of manipulation of the assignment variable by program administrators or beneficiaries. And to our knowledge, this variable (or correlated variables) and, importantly, the cutoff have not been used in assigning any other programs.

Exploiting the program assignment process that generates a cutoff, we evaluate program effects on the basis of a sharp regression discontinuity (RD) design. Because the program assignment variable is not continuous but discrete, with 34 mass points (given 34 districts), local randomization-based RD (Cattaneo, Idrobo, and Titiunik, forthcoming a, forthcoming b) is a better fit than the parametric and continuity-based nonparametric methods typically used in RD applications. Local randomization-based RD uses the smallest window that yields two consecutive mass points (two districts), one on either side of the cutoff. We apply this approach to government administrative data from annual censuses of government schools. These data are independent of those used by the government to administer the CCT program. Constructing a school-level panel with the census data, we examine program effects in the RD program district on school-level changes in girls' enrollment in secondary grades across different subperiods from 2003 (before the program) until 2015.

The Punjab Female School Stipend Program has been evaluated in multiple studies (Population Council 2007; Hasan 2010a, 2010b; Chaudhury and Parajuli 2010; Alam, Baez, and Del Carpio 2011). These evaluations have used various sources of data, including annual government school censuses, covering periods extending as far as 2009. The evaluations tend to find significant, positive effects on girls' secondary school enrollment, among other measures. While our research questions mirror those examined in the previous studies, a basic difference is that our study's observation period extends several years longer.

A more important difference is that our RD design improves on the strategies adopted in previous evaluations. Only some of these evaluations exploit the cutoff for RD identification, and those that do have relied on parametric or continuity-based nonparametric estimation methods. These methods are susceptible to bias—because the parametric estimations account for observations far from the cutoff while the nonparametric estimations treat the assignment variable as continuous, though it is not in this case.

Our local randomization-based RD estimations suggest that the program had significant, positive effects on girls' secondary school enrollment throughout the observation period. The average gain in the number of girls per school in the RD program district ranges from 33 to 53, equivalent to an average gain in girls' secondary enrollment at the school level of 25–41 percent. These effects are on the higher end of the range of effects documented in previous evaluations of the program.

Program effects appear to be stable over the study's observation period, though measured with substantial imprecision. The results suggest that effects were robust to a large decline in the real value of the program benefit over the observation period and to the potential re-optimization of behavior by agents in later years of the program. Our calculations of cost-effectiveness, using the estimated program effect over the full observation period (2003–15), indicate that for every $100 spent annually, the program induced the enrollment of 0.2–2.3 beneficiary girls in the RD program district.

The stable effects on girls' secondary school enrollment over time despite the large decline in the real value of the program benefit are inconsistent with what standard economic theory would have predicted. The observed pattern does point to potential behavioral explanations, however. For example, as Benhassine et al. (2015) argue in explaining the effects they find in evaluating an unconditional cash transfer program in Morocco, "labeling" a program as support for education can affect perceptions of the importance of schooling. Similarly, the program we evaluate may have signaled the value that the government places on girls' secondary education—making such education salient to households and communities and inducing girls' secondary school enrollment independent of the size of the benefit.

Alternatively, greater exposure of households and communities to girls' secondary education in the early years of the program, induced by the economic incentive of the program benefit, may have generated a durable, positive change in the "taste" for girls' secondary education or helped correct for preexisting biases in beliefs or preferences (Kremer, Rao, and Schilbach 2019). And these adjustments in beliefs and preferences may have catalyzed a positive shift in social norms around girls' secondary education.

2. Related literature

Most evaluations of education CCT programs in developing countries assess program effects on the likelihood of school participation, typically in the first few years after the program has been initiated (Fiszbein et al. 2009; Baird et al. 2014; Garcia and Saavedra 2017). A recent meta-evaluation study of the effects of education CCT programs finds a significant, positive average effect on the likelihood of school participation, at both the primary and secondary levels (Garcia and Saavedra 2017). The study also finds a larger average effect on the likelihood of participation at the secondary level than at the primary level. Effects found by studies included in the meta-evaluation sample vary substantially, however.

Among education CCT programs that, like the one we evaluate, are targeted to girls and condition cash benefits on secondary school participation, almost all those that have been evaluated are in Asia. Here too, evaluations generally find that the programs have significant, positive effects on the likelihood of school participation. For example, the Japan Fund for Poverty Reduction scholarship program in Cambodia, which awarded scholarships to girls from poor households who were in their last grade of primary school, increased the likelihood of school participation by 30 percentage points (Filmer and Schady 2008). Similarly, a CCT program in Turkey increased the likelihood of secondary school participation for girls from very poor households by 11 percentage points (Ahmed et al. 2007). And the Female Secondary School Assistance Project in Bangladesh, targeted to rural girls of secondary school age, increased the likelihood of girls' school participation by 8 percent for each additional year of exposure to the program (Khandker, Pitt, and Fuwa 2003).

Similar evidence is available for Pakistan. For example, Cheema et al. (2016) examine the effects of a CCT program introduced in 2012 for beneficiaries of the Benazir Income Support Program (BISP). The CCT program benefit for households was conditioned on regular school attendance by children of primary school age. Using household survey data collected in 2016, and propensity score matching of BISP households in non-CCT program areas to BISP+CCT program households in CCT program areas, they find that the CCT program increased the likelihood of school participation.

The program that we study also has been the subject of earlier evaluations (Population Council 2007; Chaudhury and Parajuli 2010; Hasan 2010a, 2010b; Alam, Baez, and Del Carpio 2011), including some based on the same data that we use—annual rounds of government school censuses. In contrast to our study, however, these evaluations examine effects on girls' education outcomes only in the early years of the program.

For example, using government school census data and a parametric triple-differences RD approach,[3] Chaudhury and Parajuli (2010) find that the program increased female

[3] The triple-differences approach compares outcomes before and after program onset, between boys' and girls' middle schools (grades 6–8), and between program and nonprogram districts.

enrollment in grades 6–8 in government schools by an average of 6 girls per school (or by 9 percent relative to initial enrollment) over the period 2003–05. Using government school census data and difference-in-differences estimation,[4] Hasan (2010a) finds that the program increased female enrollment in grades 6–8 in government schools by an average of 11 girls per school over the period 2003–08. Hasan also finds that the program increased corresponding male enrollment by an average of 10 boys per school. In another study, Hasan (2010b) uses nonrepresentative panel household sample survey data for three districts (one program district, two nonprogram districts) with adult literacy rates some distance from the cutoff and applies within-family difference-in-differences estimation.[5] In this study, Hasan finds that the program had no effect on the likelihood of government school participation by girls ages 10–14 in 2005 or 2006 but had a significant positive effect of 3 percentage points on this outcome in 2007. Finally, using government school census data from 2003–09 and parametric RD estimation, Alam, Baez, and Del Carpio (2011) find that the program had significant positive effects on girls' enrollment in grades 6–8 from 2005 to 2009. They find average gains in girls' enrollment by grade ranging from 10 to 32 percent.

3. Program context

Pakistan, a lower-middle-income economy with a per capita income in 2018 of $4,928 (in 2011 purchasing power parity U.S. dollars), is the sixth most populous country in the world (World Bank 2019). The country is divided into four large provinces (Balochistan, Khyber Pakhtunkhwa, Punjab, and Sindh) as well as small territories.

Punjab is a major province. It accounted for 53 percent of Pakistan's population of almost 208 million in 2017 (Pakistan Bureau of Statistics 2017a) and for 55 percent of the

[4] The difference-in-differences approach compares outcomes before and after program onset and between program and nonprogram districts.

[5] The difference-in-differences approach compares outcomes before and after program onset and between program and nonprogram districts. The within-family component compares siblings of different genders.

country's annual economic output in 2015/16 (PERI 2017). Sixty-three percent of the province's population lives in rural areas (Pakistan Bureau of Statistics 2017b).

Government schools in the province provide education at five levels: preprimary, primary (grades 1–5), middle (grades 6–8), high (grades 9–10), and intermediate (grades 11–12). In this study, we define *secondary school* as grades 6–10 (middle and high school), and *secondary school age* as ages 11–15. Government schools provide free education.

Punjab tends to perform better than the other provinces in girls' and boys' school participation rates (see figure 1 for net enrollment rates in 2014/15). Nevertheless, within Punjab, there is a large gender gap in school participation rates. In 2014, near the end of our study's observation period, the school participation rate for girls of secondary school age (66 percent) was 7 percentage points lower than that for boys (73 percent) (table 1).

Most students in Punjab attend government schools. In 2014, among students of secondary school age, 65 percent of boys and 67 percent of girls attended government schools, while most of the rest attended private schools.

School participation rates vary substantially across districts within the province. In 2014, for girls of secondary school age, they ranged from a low of 28 percent in Rajanpur to a high of 89 percent in Gujrat (figure 2).

While the levels differ across time, preprogram patterns in school participation are qualitatively similar to the patterns discussed here.

4. Program design

The Punjab government initiated the Female School Stipend Program in 2004 with the aim of increasing the rate of girls' participation in secondary education in parts of the province with the poorest education outcomes.[6] The selection of districts for the program was based on the literacy rate (among those age 10 and above), as measured by the 1998 Pakistan Population and Housing Census. Of the 34 districts that Punjab had at the time, the 15 with literacy rates below 40 percent were selected for the program (table 2). Over time, as the number of districts in Punjab increased from 34 to 36,[7] the number of program districts increased to 16, and the number of nonprogram districts to 20.

Districts selected for the program had poor school participation outcomes (see table 1). Measured in 2003, preprogram school participation rates for children of primary and secondary school ages were lower in program districts than in nonprogram districts, for both boys and girls. Moreover, the gender difference in preprogram school participation rates was greater in program than in nonprogram districts, particularly among children of secondary school age. In program districts, 59 percent of boys of secondary school age were enrolled in school, compared with 41 percent of girls—a difference of 18 percentage points. In nonprogram districts, the corresponding gender difference was only 4 percentage points. By 2014, near the end of our observation period, school participation rates had increased for both boys and girls and in both program and nonprogram districts, but large gender differences in school participation rates remained for children of secondary school age in program districts (figure 3).

When initiated in 2004, the program covered girls enrolled in government schools in grades 6–8. In 2005, it was expanded to also cover girls enrolled in grades 9–10. Under the program, households were offered a cash benefit of 200 Pakistan rupees (PRs) a month ($3.40) for each girl enrolled in grades 6–10 in a government school, conditional on an

[6] The first distribution of cash benefits was initiated in January 2004, based on enrollment and attendance information for the October–December 2003 quarter.
[7] Nankana Sahib was formed from the division of Sheikhupura in 2005, and Chiniot from the division of Jhang in 2009.

attendance rate of at least 80 percent in the preceding academic quarter.[8] At the time of program introduction, the value of the per-girl benefit was 14 percent of average monthly household consumption spending per adult equivalent in Punjab (or 15 percent of the corresponding average in rural Punjab).[9]

Program benefits were transferred to households on a quarterly basis, with benefits for a given quarter delivered in the following quarter. The maximum individual transfer per quarter was PRs 600 (= PRs 200 × 3 months), offered for four quarters in a year.

Key steps in program administration began at the government school, where the head teacher was responsible for recording female students' attendance for the quarter in a standard register as well as for determining their eligibility and the amount of their benefit. The teacher also filled in a money order form for each female student beneficiary. The school sent a copy of the completed register and the individual money order forms to the district education department, which transmitted the money orders to the government post service. The post service then delivered cash to beneficiary girls and households, typically at either the girls' school or their home. In addition, district education departments would transmit data and documentation on beneficiary girls and cash delivery to the provincial education department.

A sample-based process evaluation conducted during the first year of the program by a government-contracted consulting firm finds that, while program implementation suffered from delays and was burdensome for different parties (including beneficiaries), all sample teachers and virtually all sample students and parents were aware of the attendance condition

[8] Currency conversion is at the average exchange rate of PRs 58 per U.S. dollar in 2004. The benefit amount was estimated to cover safe travel to a government school with secondary grades, which in rural areas was typically located some distance from the girl's village. While the program benefit is small in absolute terms compared with those in samples of CCT programs in developing countries that have been evaluated (Garcia and Saavedra 2017), the level of the benefit as a percentage of household consumption spending is within the range of benefit sizes in these samples (Baird et al. 2014; Benhassine et al. 2015).

[9] Based on statistics from the 2004/05 Pakistan Household Income and Expenditure Survey. We are unable to estimate the program benefit level as a percentage of average household consumption spending per adult equivalent in program districts because available household sample surveys that gather data on income and consumption spending are not representative at the district level.

for program eligibility, and all sample beneficiary households reported receiving the full benefit amount (Innovative Development Consultants 2005).

In 2017, the Punjab government adjusted the value and delivery of program benefits to households. Renaming the program Khadim-e-Punjab Zewar-e-Taleem, it raised the benefit value, which had remained constant in nominal terms since the program's introduction, to PRs 1,000 a month ($9.50).[10] The new benefit level is equivalent to 18 percent of average household consumption spending per adult equivalent in Punjab (or 21 percent of the corresponding average for rural Punjab).[11] The government also began modifying how program benefits are delivered to households by requiring parents of beneficiary girls to register for an ATM card called the Khidmat Card. This card allows them to withdraw program benefits from branches of the Bank of Punjab. Our study's observation period ends before the introduction of these adjustments to the program.

5. Data and empirical strategy

Data

We use data from the Punjab government's Annual School Census from 2003 to 2015. These data are independent of those used by the government for administering the Punjab Female School Stipend Program. Initiated in 2003, the census is an annual field survey of government schools that gathers basic information on schools and teachers, including student enrollment by grade and gender, with a reference date of October 31.[12] Its main purpose is to monitor the status of government schools.

The provincial education department has primary responsibility for managing all stages of the census and sole responsibility for designing the census questionnaire. At the

[10] At the average exchange rate of PRs 105 per U.S. dollar in 2016.
[11] Based on the 2015/16 Pakistan Household Income and Expenditure Survey.
[12] The government school year runs from April to March.

school level, teachers complete the questionnaires delivered to them. These questionnaires are collected by district government officials and returned to the provincial education department, where the information is digitized.

Using the annual censuses, we construct a school-level panel.[13] On the basis of this panel, we construct the outcome measure—the absolute change in school-level girls' enrollment in secondary grades for six subperiods of different lengths originating from baseline (2003): 2003–05, 2003–07, 2003–09, 2003–11, 2003–13, and 2003–15.

Empirical strategy

We estimate program effects on the basis of a sharp RD design. One of two outcomes is possible for school s, conditional on the district literacy rate X_D. Schools in districts with a literacy rate below 40 percent $(X_D < 40)$ are covered by the program. These schools have a potential girls' secondary school enrollment outcome of $Y_s(1)$. Schools in districts with a literacy rate equal to or above 40 percent $(X_D \geq 40)$ are not covered by the program. These schools have a potential outcome of $Y_S(0)$. In theory, the average observed outcomes, given X_D, are two functions: $\mathrm{E}[Y_S(1) \mid X_D]$ if $X_D < 40$ and $\mathrm{E}[Y_S(0) \mid X_D]$ if $X_D \geq 40$. The sharp RD program effect at the cutoff, τ^{RD}, reflects the difference between $\mathrm{E}[Y_S(1) \mid X_D]$ and $\mathrm{E}[Y_S(0) \mid X_D]$:

(1) $$\tau^{RD} \equiv \mathrm{E}[Y_S(1) - Y_S(0) \mid X_S = 40]$$

[13] While government schools are assigned unique identification (ID) numbers, school ID protocols were adjusted by the provincial and district education departments in certain years and assigned school IDs were not strictly adhered to during data collection, entry, and processing. Consequently, to construct an accurate school-level panel, we worked closely with provincial education department officials with intimate working knowledge of the Annual School Census data and ground-level knowledge of schools.

Parametric methods approximate the global shape of these functions, but observations far from the cutoff can distort the approximation close to the cutoff and misrepresent the RD effect (Gelman and Imbens 2014; Cattaneo, Idrobo, and Titiunik, forthcoming a, forthcoming b).

A common nonparametric solution is a continuity-based RD design, which approximates the conditional expectation functions as polynomials within an optimal bandwidth close to the cutoff. This design is less sensitive to extreme features of the data generation process or outliers far from the cutoff (Cattaneo and Vazquez-Bare 2016; Imbens and Kalyanaraman 2012).

The continuity-based RD design does not fit key defining features of our data, however. The design requires the program assignment variable—the district literacy rate—to be a continuous random variable. Because program assignment is at the district level while our outcome indicator is at the school level, our assignment variable is discrete, as schools within the same district have the same assignment score. With 34 districts at baseline in 2003, the score has that many "mass points."

Given the features of our data, we adopt the local randomization-based RD design as the only practical and valid method for estimation and inference. This design formalizes the idea that treatment is as-if randomly distributed within a small window close to the cutoff, by imposing randomization-type assumptions that are stronger than the assumptions under the continuity-based RD design. In particular, it requires the existence of a window W_0 close to the cutoff where the distribution of the program assignment variable is known and uncorrelated with potential outcomes. Within W_0, potential outcomes do not depend directly on the program assignment variable; they are associated only through program assignment. In other words, the design requires $\mathrm{E}\left[Y_S(1) \mid X_D\right]$ and $\mathrm{E}\left[Y_S(0) \mid X_D\right]$ to be constant functions of X_S within W_0, and the vertical distance between these functions would reflect the local randomization-based RD effect, τ^{LRD} (Cattaneo, Titiunik, and Vazquez-Bare 2017; Cattaneo, Idrobo, and Titiunik, forthcoming a, forthcoming b).

Following Cattaneo, Idrobo, and Titiunik (forthcoming b), because we have only 34 mass points, the window that we select for local randomization is the smallest window, W_{LRD}, with two consecutive mass points, one on either side of the cutoff. The two districts within this window are Khushab (nonprogram) and Khanewal (program) (see table 2).

Within W_{LRD}, we assume complete randomization—or fixed-margins randomization—under which the numbers of program and nonprogram schools are fixed. Using this mechanism, we apply two inference methods. The first is Fisherian inference, which tests the sharp null hypothesis of no program effect for any given school and leads to correct inference even when the number of observations within W_{LRD} is very small. Fisher's *p*-value is calculated by using randomization inference through the permutation of the program assignment status of observations under the sharp null hypothesis of no program effect.

The second is Neyman inference, which is based on the large-sample approximate behavior of the statistic and tests the null hypothesis of no program effect. While the Fisherian framework focuses on testing the sharp null hypothesis, one of the main purposes of the Neyman framework is to estimate the sample average program effect. The difference in average outcomes between program and nonprogram observations is an unbiased estimator of this parameter, assuming complete randomization within W_{LRD}. Because our sample of schools within W_{LRD} is sufficiently large, we use the Neyman method to obtain unbiased estimates of the sample average program effect.

Manipulation of program assignment status can undermine the identification of program effects. In our case, program assignment is based on district literacy rates obtained from 1998 data, while the program was initiated in 2004. Thus, we can consider districts to be plausibly exogenously sorted around the cutoff of 40 percent.

Identification of program effects can also be undermined if the program assignment variable and cutoff were also used for other programs. To our knowledge, no other government program, in education or in any other sector, has used district literacy rates and this program's cutoff for assignment.

Finally, while the enrollment data in government school censuses may not be fully accurate,[14] we have no reason to suspect that the quality of data varies systematically between program and nonprogram observations within W_{LRD}.

6. Results

Validity of the local randomization-based RD window

The validity of local randomization-based RD estimation requires that there be at least one covariate whose preprogram values are correlated with the program assignment variable (the district literacy rate) everywhere except within our selected window W_{LRD} around the cutoff. This test is comparable to the baseline balance test in randomized controlled trials.

We select three important school-level covariates whose preprogram values should be associated with the district literacy rate: the number of teachers, the number of classrooms, and the condition of the school building. We test the null hypothesis of no effect on the program assignment variable for each covariate. The window would be invalid if the null hypothesis is rejected for at least one covariate within it.

[14] Reported enrollment data may not be accurate if school enrollment records are not up to date as of the reference date for the government school census or if reported enrollment data become high stakes (and thus subject to manipulation) because they are used by government education authorities for decisions affecting the interests of stakeholders in the public education sector.

We fail to reject the null hypothesis within the smallest window W_{LRD} of 0.6 percentage points (as well as within the next smallest window of 2.6 percentage points), which contains our RD program and nonprogram districts, Khanewal and Khushab (table 3). The results indicate that the preprogram averages for the covariates are balanced within the selected small window around the cutoff but not in larger windows.

Program effects on girls' secondary school enrollment

The local randomization-based RD results are presented in table 4. We find that the program produced significant gains in girls' secondary school enrollment across all subperiods between 2003 and 2015, whether we apply Fisherian or Neyman inference. Program effects over the different subperiods range from an average gain of 33 to 53 girls per school. Given an average preprogram secondary school enrollment of 130 girls per school in the RD program district, the effects are equivalent to an average gain of 25–41 percent in girls' secondary school enrollment. These effects are on the higher end of the range of effects previously documented for the program we examine, which may reflect, among other things, differences across the studies in the data, sample, and estimation strategy.

We also aimed to examine differences in program effects between rural and urban areas.[15] While the sample of schools in rural areas within W_{LRD} is large enough to apply the Neyman approach to estimate the program effect, the corresponding sample of schools in urban areas is too small.[16] The Fisherian approach could provide valid inference for such a small sample, but not the program effect. Therefore, we do not examine program effects in

[15] On the one hand, program effects may be larger in rural communities than in urban areas because the economic incentive (in terms of benefit size relative to average household consumption spending) is marginally stronger in rural communities. On the other hand, program effects may be smaller in rural communities because government secondary schools that girls can attend are fewer and on average farther away, or because social stigma may be associated with travel by socially disadvantaged girls to schools outside their own rural community (Jacoby and Mansuri 2015).

[16] There were 29 schools in urban areas within the selected RD window.

urban areas. In rural areas, results show that the average gain in girls' secondary school enrollment is significant across all subperiods and ranges from 10 to 31 girls per school (see column 2 of table 4).

While overall program effects differ somewhat in size, they are statistically similar across all subperiods, on the basis of an examination of Fisherian confidence intervals (figure 4). The size and statistical significance of the program effects are sustained over the study's observation period despite a decline in the real value of the annual program benefit of more than 60 percent over the period, from PRs 2,400 in 2004 to PRs 860 in 2015 (figure 5). As a result of this decline in the real value of the benefit, along with an increase in average real household income over the period, the relative benefit level decreased from an amount equivalent to 14 percent of average household consumption spending per adult equivalent in Punjab to 4 percent.[17]

As a falsification test, we estimate local randomization-based RD effects near two arbitrarily selected false cutoffs of the program assignment variable, at 30 percent and 50 percent. The results are presented in table 5. While the estimated effects (overall and rural only) are sometimes large, they are almost always statistically insignificant. The only exception is the estimated effect for the subperiod 2003–11 measured near the false cutoff of 30 percent. Significant at the 5 percent level, the effect indicates a decline in girls' secondary school enrollment in the false RD program district.

Note that, because of data limitations, we are not able to evaluate the extent to which the estimated program effects on girls' government school secondary enrollment numbers reflect gains in the likelihood of girls' secondary school participation. If the program mainly induced girls who would have joined secondary schools without the program to choose government schools over private schools, the gains in girls' secondary school participation

[17] Based on statistics from the 2004/05 and 2015/16 Pakistan Household Income and Expenditure Surveys.

would be less than commensurate. While multiple rounds of district-representative household sample surveys and private school censuses are available over the study observation period, we are unable to construct panels at the household or school levels using these data to examine this question.[18] Estimates of local randomization-based RD program effects using single cross-sections would be biased.[19] We would also not be able to apply local randomization-based RD if these data are used to construct a panel at the district level—with 34 districts—but must rely on parametric RD estimation which potentially introduces bias as discussed in section 5.

Notwithstanding, we consider the extent to which the estimated program effects are driven by displacement or diversion of students from private schools to government schools to be contained by the somewhat low share of private secondary school enrollment in the RD districts. Household sample survey data for 2014 show that only one of every five female secondary students from poorer backgrounds attended private schools in the RD districts.[20]

Program cost-effectiveness

We assess the program's cost-effectiveness in relation to girls' secondary school enrollment in the RD program district, following Dhaliwal et al. (2013) and costing guidelines from the Abdul Latif Jameel Poverty Action Lab (J-PAL n.d.). Per beneficiary, apart from the cost of the benefit transfer (PRs 2,400 a year), the government paid a fee to the government post service of PRs 50 per money order (which amounts to PRs 200 a year given quarterly benefit transfers) and what the government refers to as an "incidental expense" of 1 percent of the

[18] The various rounds of the private school censuses lack a common school ID system.
[19] This also applies to the examination of school attainment (years of schooling) using data from the household sample surveys.
[20] Based on data from the 2014 Punjab Multiple Indictor Cluster Survey.

benefit amount (PRs 24 a year).[21] These values remained the same over the study's observation period. Thus, the total annual per-beneficiary cost was PRs 2,624. This per-beneficiary cost almost perfectly matches the value from official government accounting records of actual annual outlays and beneficiaries for the program.

Note that the PRs 2,624 value reflects recurrent costs and does not capture fixed costs or costs incurred in designing and establishing the program. Moreover, for recurrent costs, the value does not capture the cost of time and other resources spent by provincial and district education department officials or schoolteachers in carrying out their program responsibilities. Nor does it capture any costs borne by households in ensuring that they were present when and where needed to receive the benefit transfer or in pursuing any grievances with the authorities relating to problems with receiving the benefit. While such omissions from cost calculations are not unusual in impact evaluation studies for developing countries (Dhaliwal et al. 2013), the PRs 2,624 value should be considered an underestimate of the actual annual per-beneficiary cost of the program.

In 2015, there were 28,559 girls in secondary grades in government schools in the RD program district. Comparing enrollment numbers from the school census rounds and program beneficiary numbers from government administrative data, we find that about 90 percent of enrolled girls receive program benefits.[22] Using these numbers, we estimate that the government spent PRs 67,496,341 on the program in the RD program district in 2015.

For the calculation of cost-effectiveness, we arbitrarily choose the program effect for the subperiod 2003–15 (that is, the full observation period), an average gain of 41 girls per school in the RD program district. This effect, which lies roughly in the middle of the range

[21] This maximum administrative cost of PRs 224 per beneficiary per year is about 9 percent of the maximum benefit transfer per beneficiary per year. This percentage is on par with the modal ratio of per-beneficiary annual administrative cost to benefit transfer in a sample of CCT programs in developing countries that have been evaluated (Garcia and Saavedra 2017).

[22] That the number of beneficiary girls is lower than girls' enrollment is presumably due to the application of the regular attendance condition for program benefit eligibility.

of program effects across subperiods, is equivalent to 32 percent of average preprogram girls' enrollment in secondary grades in government schools in the RD program district. Given a total preprogram enrollment of 25,676 girls in secondary grades in the RD program district, this effect translates into 8,109 girls being induced by the program to enroll in school.

From the total estimated outlay and the total number of girls induced to enroll in 2015, we estimate an annual cost per beneficiary girl induced to enroll of PRs 8,324 (or \$81).[23] Accounting for imprecision in the estimated program effect, we find that this annual cost ranges from PRs 57,106 (\$554) at the lower bound of the 90 percent confidence interval for the estimate of the program effect to PRs 4,450 (\$43) at the upper bound for the same interval.[24] As an alternative measure of cost-effectiveness, we estimate that for every \$100 spent annually, the program induced the enrollment of 1.2 beneficiary girls (or, after accounting for the imprecision in the estimated program effect, 0.2–2.3 beneficiary girls).[25]

7. Conclusion

Most rigorous evaluations of development interventions, including of CCT programs, examine effects in the first few years of a program's implementation. Yet a program's effects on later cohorts could differ from those on early cohorts—because agents have had the time to re-optimize their behavior in response to the program or because there may have been changes in key features of the program's design or implementation.

This study contributes to the evaluative literature for developing countries by examining the evolution of effects of a public education CCT program targeted to girls of

[23] At the average exchange rate of PRs 103 per U.S. dollar in 2015.

[24] The lower- and upper-bound estimates of program effects are 6 and 77 girls per school.

[25] Although we use the standardized approach proposed by Dhaliwal et al. (2013), we are unable to compare the cost-effectiveness values for this program with those for other CCT programs in developing countries because of differences in outcome indicators. The available cost-effectiveness values are based on program effects on school participation rates, attendance rates, dropout rates, and education attainment (see, for example, Garcia and Saavedra 2017; and J-PAL 2019a and 2019b), while our outcome measure is school enrollment numbers.

secondary school age over the first decade of the program's life, from its inception in 2004 until 2015. This examination is made possible by a sharp RD design based on an observed, numerical program assignment variable (the district adult literacy rate) and government administrative data from annual censuses of government schools that capture school-level data on girls' enrollment in secondary grades. The local randomization-based RD approach that we apply to estimate program effects improves on the parametric and nonparametric empirical approaches used in previous evaluations of the program by better fitting the structure of the program assignment variable.

We find that the program had significant, large positive effects on girls' enrollment in secondary grades in government schools across cohorts throughout the period of observation—and that the effects were of similar size. Moreover, these program effects were observed even though the program's economic incentive effect on households weakened substantially over the observation period, as a result of a marked decline in the real value of the program benefit (which was not indexed to inflation) accompanied by a rise in average real household income. The effects were also observed despite the potential re-optimization of behavior by various agents in later years of the program. (Note that the documented RD effects, which relate to the RD program district of Khanewal, may not apply to program districts more generally.)

The finding of sustained program effects on girls' secondary school enrollment over the study's observation period despite a loss in economic incentive suggests potential behavioral explanations. The program may have generated durable changes in household and community behavior relating to girls' secondary education. For example, independent of the value of the benefit, the program may have made girls' secondary education salient to households, leading to a change in perceptions by households and communities about the acceptability and value of such education. More generally, greater exposure to girls' secondary education induced by the program may have helped to correct for preexisting biases in beliefs and preferences of households and communities—or in their decision-making processes—in relation to girls' secondary education. Relatedly, consistent with a multiple-equilibria characterization of the setting, program benefits may have functioned as an initial catalytic agent, triggering communities to ratchet up from a low girls' secondary education equilibrium to a higher one (a shift in social norms). Unfortunately, we do not have appropriate data on parental beliefs regarding girls' education to empirically test our conjectures.

We see two relatively straightforward directions for further evaluative research on the program. The first is to examine the effects on girls' enrollment of the higher benefit amount (along with the new method for delivering benefits) introduced in 2017. Relative to average household consumption spending per adult equivalent in the province, the benefit amount increased more than fourfold, from 4 to 18 percent. Does the restored—and enhanced—economic incentive induce greater gains in girls' secondary school enrollment? In due course, enough rounds of government Annual School Censuses should be available to examine the early effects of the program modification on girls' secondary school enrollment.

The second direction is to examine whether the program improves girls' academic achievement as measured by test scores, evidence that has been lacking for this program.[26] Under the management of the Punjab Examination Commission, the provincial and district education departments have been administering standardized exams to students in grades 5 and 8 annually. Validity and reliability appear to be a concern for the existing rounds of exams. But with enough improvement in the design and administration of the exams, future

[26] While scarce, the available rigorous international evidence generally suggests that the effects of CCT programs on student academic achievement are small at best (Baird et al. 2014).

rounds may serve as a credible source of data for analyzing the program's effects on girls' academic achievement.[27]

[27] A caveat: Greater secondary school enrollment induced by the program may change the average composition of students, if, for example, the program attracts girls from lower down in the ability distribution to enroll in school. As a result, using exam score data only on students—instead of on children irrespective of their school enrollment status, such as through testing at home—can produce a downward bias in the potential program effect on girls' academic achievement.

References

Ahmed, Akhter U., Michelle Adato, Ayse Kudat, Daniel Gilligan, Terence Roopnaraine, and Rafik Colasan. 2007. "Impact Evaluation of the Conditional Cash Transfer Program in Turkey: Final Report." International Food Policy Research Institute (IFPRI), Washington, DC; and AGRIN, Ankara, Turkey.

Alam, Andaleeb, Javier E. Baez, and Ximena V. Del Carpio. 2011. "Does Cash for School Influence Young Women's Behavior in the Longer Term? Evidence from Pakistan." Policy Research Working Paper 5669, World Bank, Washington, DC.

Baird, Sarah, Francisco H. G. Ferreira, Berk Özler, and Michael Woolcock. 2014. "Conditional, Unconditional and Everything in Between: A Systematic Review of the Effects of Cash Transfer Programmes on Schooling Outcomes." *Journal of Development Effectiveness* 6 (1): 1–43.

Benhassine, Najy, Florencia Devoto, Esther Duflo, Pascaline Dupas, and Victor Pouliquen. 2015. "Turning a Shove into a Nudge? A 'Labeled Cash Transfer' for Education." *American Economic Journal: Economic Policy* 7: 86–125.

Bouguen, Adrien, Yue Huang, Michael Kremer, and Edward Miguel. 2019. "Using RCTs to Estimate Long-Run Impacts in Development Economics." *Annual Reviews* 11: 523–61.

Cattaneo, Matias D., Nicolas Idrobo, and Rocio Titiunik. Forthcoming a. *A Practical Introduction to Regression Discontinuity Designs.* Vol. 1. Elements in Quantitative and Computational Methods for Social Science. Cambridge, UK: Cambridge University Press.

———. Forthcoming b. *A Practical Introduction to Regression Discontinuity Designs.* Vol. 2. Elements in Quantitative and Computational Methods for Social Science. Cambridge, UK: Cambridge University Press.

Cattaneo, Matias D., Rocio Titiunik, and Gonzalo Vazquez-Bare. 2016. "Inference in Regression Discontinuity Designs under Local Randomization." *Stata Journal* 16 (2): 331–67.

———. 2017. "Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality." *Journal of Policy Analysis and Management* 36 (3): 643–81.

Cattaneo, Matias D., and Gonzalo Vazquez-Bare. 2016. "The Choice of Neighborhood in Regression Discontinuity Designs." *Observational Studies* 2: 134–46.

Chaudhury, Nazmul, and Dilip Parajuli. 2010. "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Program on Public School Enrollments in Punjab, Pakistan." *Applied Economics* 42: 3565–83.

Cheema, Iftikhar, Martina G. Asia, Simon Hunt, Sarah Javeed, Tanya Lone, and Sean O'Leary. 2016. "Benazir Income Support Program: Evaluation of Waseela-e-Taleem Conditional Cash Transfer." Oxford Policy Management, London.

Dhaliwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch. 2013. "Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications to Education." In *Education Policy in Developing Countries*, edited by Paul Glewwe. Chicago: University of Chicago Press.

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2013. "School Inputs, Household Substitution, and Test Scores." *American Economic Journal: Applied Economics* 5: 29–57.

Filmer, Deon, and Norbert Schady. 2008. "Getting Girls into School: Evidence from a Scholarship Program in Cambodia." *Economic Development and Cultural Change* 56 (3): 581–617.

Fiszbein, Ariel, Norbert Schady, Francisco H. G. Ferreira, Margaret Grosh, Niall Keleher, Pedro Olinto, and Emmanuel Skoufias. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty*. World Bank Policy Research Report. Washington, DC: World Bank.

Garcia, Sandra, and Juan E. Saavedra. 2017. “Educational Impacts and Cost-Effectiveness of Conditional Cash Transfer Programs in Developing Countries: A Meta-Analysis.” *Review of Educational Research* 87 (5): 921–65.

Gelman, Andrew, and Guido Imbens. 2014. “Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs.” NBER Working Paper 20405, National Bureau of Economic Research, Cambridge, MA.

Glewwe, Paul, and Karthik Muralidharan. 2016. “Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications.” In *Handbook of the Economics of Education*, vol. 5, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, 653–743. Amsterdam: Elsevier.

Hasan, Amer. 2010a. “Gender-Targeted Conditional Cash Transfers: Enrollment, Spillover Effects and Instructional Quality.” Policy Research Working Paper 5257, World Bank, Washington, DC.

———. 2010b. “Time Allocation in Rural Households: The Indirect Effects of Conditional Cash Transfer Programs.” Policy Research Working Paper 5256, World Bank, Washington, DC.

Imbens, Guido, and Karthik Kalyanaraman. 2012. “Optimal Bandwidth Choice for the Regression Discontinuity Estimator.” *Review of Economic Studies* 79 (3): 933–59.

Innovative Development Consultants. 2005. “Third Party Validation for Stipend to Girl Students: Final Report.” Innovative Development Consultants, Lahore.

Jacoby, Hanan G., and Ghazala Mansuri. 2015. “Crossing Boundaries: How Social Hierarchy Impedes Economic Mobility.” *Journal of Economic Behavior & Organization* 117: 135–54.

J-PAL (Abdul Latif Jameel Poverty Action Lab). 2019a. “Increasing Student Enrollment and Attendance: Impacts by Gender.” J-PAL Policy Insights. Last modified February 2019. https://doi.org/10.31485/pi.2262.2018.

———. 2019b. “Reducing Costs to Increase School Participation.” J-PAL Policy Insights. Last modified February 2019. https://doi.org/10.31485/pi.2264.2018.

———. n.d. “J-PAL Costing Guidelines.” J-PAL, Cambridge, MA. https://www.povertyactionlab.org/sites/default/files/resources/J-PAL%20Costing%20Guidelines%202016.05.pdf. Accessed September 2019.

Khandker, Shahidur, Mark Pitt, and Nobuhiko Fuwa. 2003. “Subsidy to Promote Girls’ Education: The Female Stipend Program in Bangladesh.” Paper 23688, Munich Personal RePEc Archive.

Kremer, Michael, Gautam Rao, and Frank Schilbach. 2019. “Behavioral Development Economics.” In *Handbook of Behavioral Economics*, edited by Douglas Bernheim, Stefano DellaVigna, and David Laibson. Amsterdam: Elsevier.

Molina Millan, Teresa, Tania Barham, Karen Macours, John A. Maluccio, and Marco Stampini. 2019. “Long-Term Impacts of Conditional Cash Transfers: Review of the Evidence.” *World Bank Research Observer* 34(1): 119–59.

Nguyen, Quynh T., and Dhushyanth Raju. 2015. “Private School Participation in Pakistan.” *Lahore Journal of Economics* 20: 1–46.

Pakistan Bureau of Statistics. 2016. *PSLM 2014–15: Pakistan Social and Living Standards Measurement Survey (2014–15): National/Provincial/District*. Islamabad: Pakistan Bureau of Statistics.

———. 2017a. “Tehsil Wise Provisional Results of Census—2017.” Pakistan Bureau of Statistics, Islamabad. http://www.pbs.gov.pk/sites/default/files/PAKISTAN%20TEHSIL%20WISE%20FOR%20WEB%20CENSUS_2017.pdf.

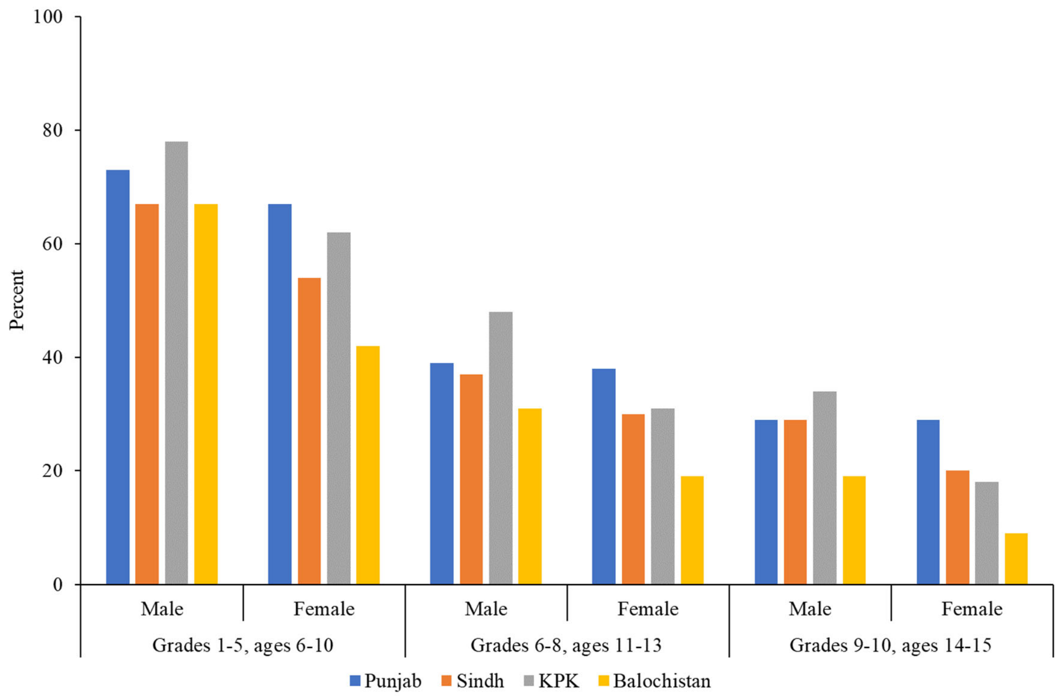
———. 2017b. “District Wise Provisional Results of Census—2017.” Pakistan Bureau of Statistics, Islamabad. http://www.pbs.gov.pk/sites/default/files/DISTRICT_WISE_CENSUS_RESULTS_CENSUS_2017.pdf.

PERI (Punjab Economic Research Institute). 2017. *Punjab Economic Report 2017*. Lahore: Punjab Planning and Development Department.

Population Council. 2007. “Evaluation of the Punjab Female School Stipend Program.” Population Council, Islamabad.

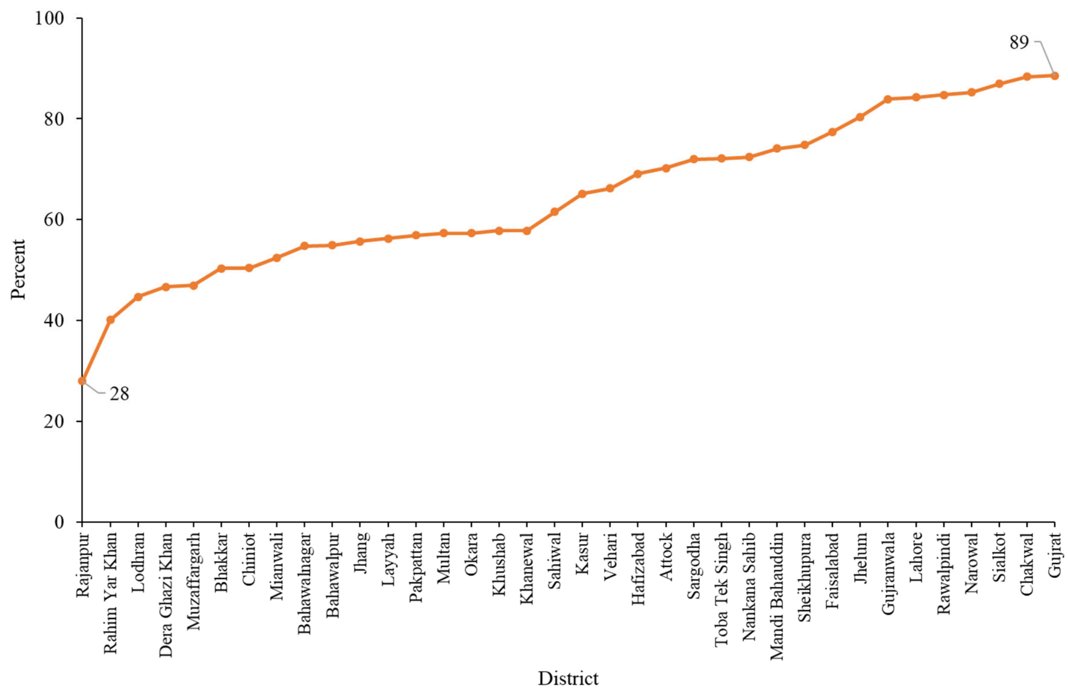
World Bank. 2019. *World Development Indicators Databank*. Washington, DC: World Bank. https://databank.worldbank.org/source/world-development-indicators. Accessed August 2019.

Figure 1. Net Enrollment Rate by Province, Pakistan, 2014/15



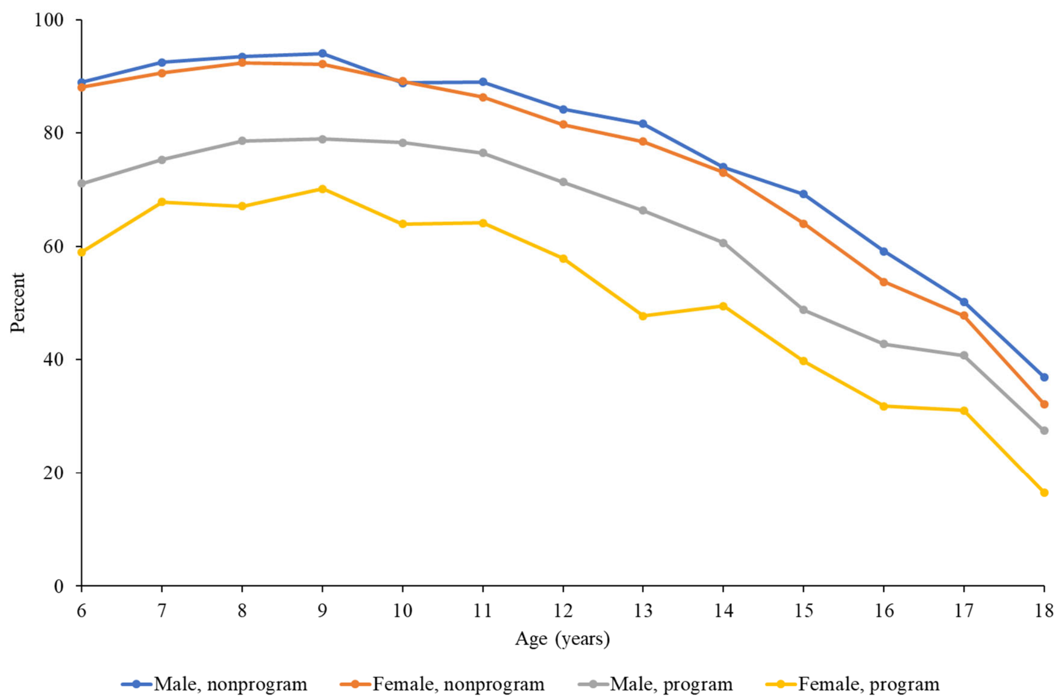
Source: Data from Pakistan Bureau of Statistics (2016).
Note: KPK = Khyber Pakhtunkhwa. The net enrollment rate for a given set of grades is the percentage of children in the corresponding age group who attend those grades.

Figure 2. School Participation Rate for Girls' Ages 11–15 by District, Punjab, 2014



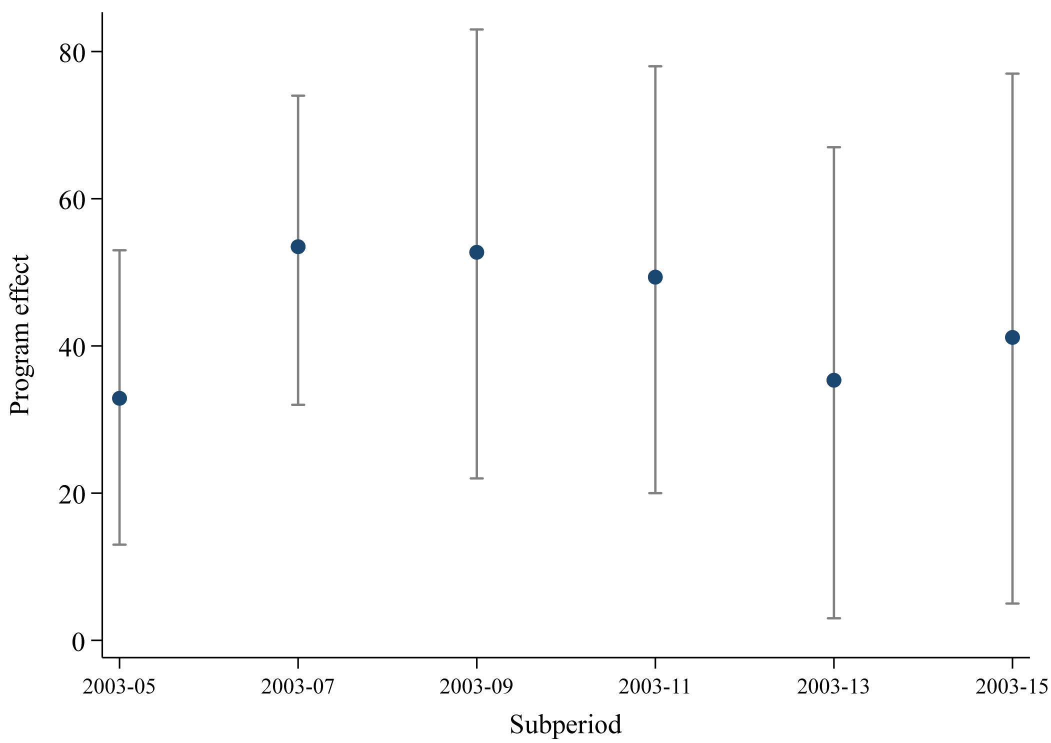
Source: Own estimates based on data from the 2014 Punjab Multiple Indicator Cluster Survey.
Note: The school participation rate for a given age group is the percentage of children in that age group who attend school. Estimates are adjusted for survey sampling weights.

Figure 3. School Participation Rate by Age, Gender, and District Program Status, Punjab, 2014



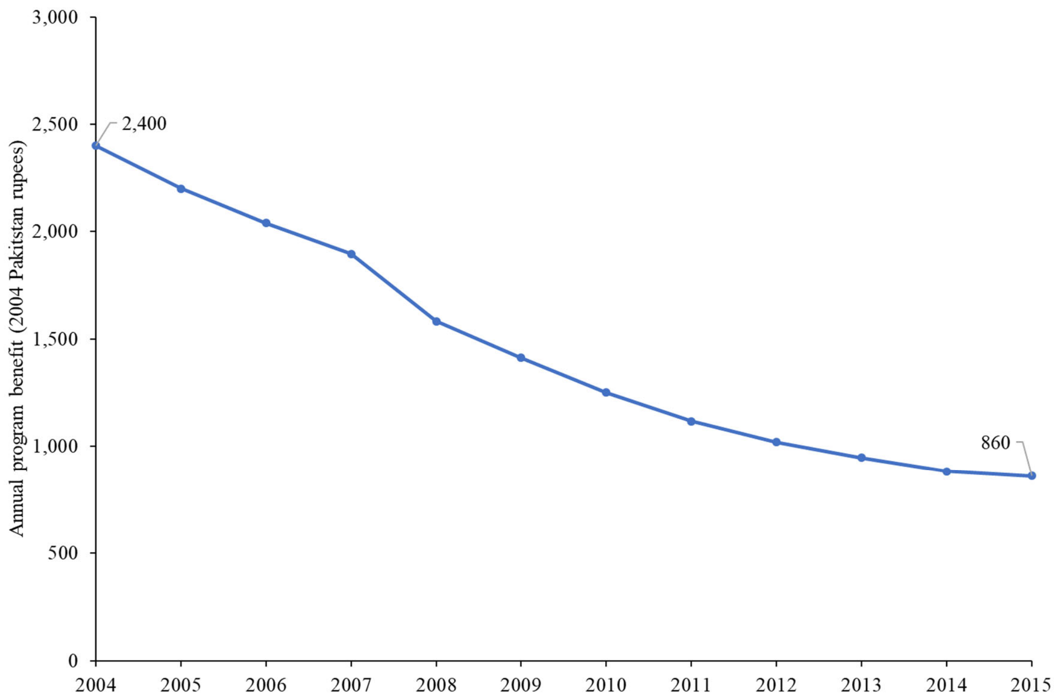
Source: Own estimates based on data from the 2014 Punjab Multiple Indicator Cluster Survey.
Note: The school participation rate for a given age is the percentage of children of that age who attend school. Estimates are adjusted for survey sampling weights.

Figure 4. Program Effects on Girls' Secondary School Enrollment, Alternative Subperiods



Source: Own estimates based on data from the Punjab government's Annual School Census for various years.
Note: Fisherian 95 percent confidence intervals are obtained using the Stata software package `rdrandinf` developed by Cattaneo, Titiunik, and Vazquez-Bare (2016).

Figure 5. Real Value of Annual Program Cash Benefit, 2004–15



Source: Yearly consumer price index data from Global Economic Monitor, World Bank, https://datacatalog.worldbank.org/dataset/global-economic-monitor. Accessed: May 15, 2018.
Note: The base year for deflation is 2004, the year the program was launched. The program's cash benefit remained fixed in nominal terms from 2004 to 2015.

Table 1. School Participation Rate by Gender, Age Group, and District Program Status, Punjab, 2003 and 2014

Districts	Gender	School participation rate (percent) 2003 (1)	2014 (2)
Secondary school age (ages 11–15)			
All	Male	67	73
	Female	57	66
Program	Male	59	65
	Female	41	52
Nonprogram	Male	72	80
	Female	68	77
Primary school age (ages 6–10)			
All	Male	72	85
	Female	66	79
Program	Male	63	76
	Female	51	65
Nonprogram	Male	81	92
	Female	76	90

Source: Own estimates based on data from the 2003 and 2014 Punjab Multiple Indictor Cluster Surveys.
Note: The school participation rate for a given age group is the percentage of children in that age group who attend school. Estimates are adjusted for survey sampling weights.

Table 2. District Program Assignment Status

District	Adult literacy rate in 1998 (percent) (1)	Recentered literacy rate (percentage points) (2)	Covered by program? (3)
Rajanpur	20.7	19.3	Yes
Muzaffargarh	28.5	11.5	Yes
Lodhran	29.9	10.1	Yes
Dera Ghazi Khan	30.6	9.4	Yes
Rahim Yar Khan	33.1	6.9	Yes
Bhakkar	34.2	5.8	Yes
Pakpattan	34.7	5.3	Yes
Bahawalpur	35.0	5.0	Yes
Bahawalnagar	35.1	4.9	Yes
Kasur	36.2	3.8	Yes
Vehari	36.8	3.2	Yes
Jhang	37.1	2.9	Yes
Okara	37.8	2.2	Yes
Layyah	38.7	1.3	Yes
Khanewal	39.9	0.1	Yes
Khushab	40.5	−0.5	No
Hafizabad	40.7	−0.7	No
Mianwali	42.8	−2.8	No
Multan	43.4	−3.4	No
Sahiwal	43.9	−3.9	No
Sargodha	46.3	−6.3	No
Mandi Bahauddin	47.4	−7.4	No
Sheikhupura	47.8	−7.8	No
Attock	49.3	−9.3	No
Toba Tek Singh	50.5	−10.5	No
Faisalabad	51.9	−11.9	No
Narowal	52.6	−12.6	No
Gujranwala	56.5	−16.5	No
Chakwal	56.7	−16.7	No
Sialkot	58.9	−18.9	No
Gujrat	62.2	−22.2	No
Jhelum	63.9	−23.9	No
Lahore	64.7	−24.7	No
Rawalpindi	70.4	−30.4	No

Source: District adult literacy rates from the 1998 Pakistan Population and Housing Census.
Note: The recentered literacy rate is computed by subtracting the district's adult literacy rate from 40 percent, the value we set as the program assignment cutoff.

Table 3. Validity Check of Window Length

Window length (1)	Minimum balance test p-value (2)	Nonprogram schools n (3)	Program schools n (4)
0.6	0.602	74	197
2.6	0.325	123	492
4.6	0	550	992
6.6	0	850	1,554
8.6	0	1,162	1,756
10.6	0	1,490	1,909
12.6	0	2,025	2,016
14.6	0	2,025	2,016
16.6	0	2,242	2,016
18.6	0	2,398	2,016

Source: Own estimates based on data from the Punjab government's Annual School Census for 2003.

Note: The table presents balance test results using preprogram values for three covariates: number of teachers, number of classrooms, and the condition of the school building. Column 1 presents the window length in which the balance test is performed. A balance test is performed using each of the three covariates within each window to test their association with the program assignment variable. Column 2 presents the minimum *p*-value across the three covariates. Column 3 presents the number of schools that are in nonprogram districts, and column 4 the number in program districts, for each window length. Balance tests were conducted using the Stata software package `rdrandinf` developed by Cattaneo, Titiunik, and Vazquez-Bare (2016).

Table 4. Program Effects on Girls' Secondary School Enrollment, Alternative Subperiods

	Overall (1)	Rural (2)
2003–05		
Program effect	32.9	10.3
Fisher's *p*-value	0.000	0.007
Large-sample *p*-value	0.009	0.017
2003–07		
Program effect	53.5	26.3
Fisher's *p*-value	0.000	0.000
Large-sample *p*-value	0.000	0.000
2003–09		
Program effect	52.7	30.9
Fisher's *p*-value	0.000	0.000
Large-sample *p*-value	0.001	0.000
2003–11		
Program effect	49.3	24.0
Fisher's *p*-value	0.000	0.002
Large-sample *p*-value	0.007	0.000
2003–13		
Program effect	35.3	12.9
Fisher's *p*-value	0.031	0.065
Large-sample *p*-value	0.108	0.066
2003–15		
Program effect	41.2	18.4
Fisher's *p*-value	0.020	0.019
Large-sample *p*-value	0.072	0.019
Nonprogram schools *n*	74	64
Program schools *n*	197	178

Source: Own estimates based on data from the Punjab government's Annual School Census for various years.

Note: Program effects are obtained using local randomization methods with two mass points, assuming a complete randomization mechanism. Randomization *p*-values are obtained using 1,000 permutations. Fisherian *p*-values correspond to a test of the sharp null hypothesis of no program effect. Large-sample *p*-values are based on Neyman inference methods and correspond to a test of the null hypothesis of no program effect. Local randomization results are generated using the Stata software package `rdlocrand` developed by Cattaneo, Titiunik, and Vazquez-Bare (2016).

Table 5. Effects on Girls' Secondary School Enrollment at False Cutoffs, Alternative Subperiods

	Cutoff	
	30 percent (1)	50 percent (2)
2003–05		
Program effect	1.3	10.5
Fisher's *p*-value	0.905	0.260
Large-sample *p*-value	0.893	0.319
2003–07		
Program effect	−5.6	8.5
Fisher's *p*-value	0.651	0.462
Large-sample *p*-value	0.636	0.521
2003–09		
Program effect	−16.5	6.2
Fisher's *p*-value	0.266	0.668
Large-sample *p*-value	0.229	0.703
2003–11		
Program effect	−36.3	26.2
Fisher's *p*-value	0.043	0.241
Large-sample *p*-value	0.042	0.219
2003–13		
Program effect	−23.0	22.5
Fisher's *p*-value	0.290	0.232
Large-sample *p*-value	0.278	0.259
2003–15		
Program effect	−14.1	18.9
Fisher's *p*-value	0.586	0.327
Large-sample *p*-value	0.539	0.348
Nonprogram schools *n*	74	210
Program schools *n*	197	199

Source: Own estimates based on data from the Punjab government's Annual School Census for various years.

Note: For the estimation results reported in column (1), the RD program district is Lodhran (adult literacy rate in 1998 = 29.9 percent) and the RD nonprogram district is Dera Ghazi Khan (30.6 percent), while for the estimation results reported in column (2), the RD program district is Attock (49.3 percent) and the RD nonprogram district is Toba Tek Singh (50.5 percent). Effects are obtained using local randomization methods with two mass points, assuming a complete randomization mechanism. Randomization *p*-values are obtained using 1,000 permutations. Fisherian *p*-values correspond to a test of the sharp null hypothesis of no effect. Large-sample *p*-values are based on Neyman inference methods and correspond to a test of the null hypothesis of no effect. Local randomization results are generated using the Stata software package `rdlocrand` developed by Cattaneo, Titiunik, and Vazquez-Bare (2016).