

Evaluating Public Per-Student Subsidies to Low-Cost Private Schools

Regression-Discontinuity Evidence from Pakistan

Felipe Barrera-Osorio

Dhushyanth Raju

The World Bank
South Asia Region
Education Team

&

Human Development Network
Education Team
April 2011



Abstract

This study estimates the causal effects of a public per-student subsidy program targeted at low-cost private schools in Pakistan on student enrollment and schooling inputs. Program entry is ultimately conditional on achieving a minimum stipulated student pass rate (cutoff) in a standardized academic test. This mechanism for treatment assignment allows the application of regression-discontinuity (RD) methods to estimate program impacts at the cutoff. Data on two rounds of entry test takers (phase 3 and phase 4) are used. Modeling the entry

process of phase-4 test takers as a sharp RD design, the authors find evidence of large positive impacts on the number of students, teachers, classrooms, and blackboards. Modeling the entry process of phase-3 test takers as a partially-fuzzy RD design given treatment crossovers, they do not find evidence of significant program impacts on outcomes of interest. The latter finding is likely due to weak identification arising from a small jump in the probability of treatment at the cutoff.

This paper is a product of the Education team, South Asia Region; and the Education Team, Human Development Network. 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://econ.worldbank.org>. The authors may be contacted at fbarrera@worldbank.org and draju2@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. 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.

Evaluating public per-student subsidies to low-cost private schools: Regression-discontinuity evidence from Pakistan

Felipe Barrera-Osorio
Dhushyanth Raju

JEL classification codes: I21; I28; O10

Keywords: education; Pakistan; regression-discontinuity design; subsidies; private schools

Contact information: Barrera-Osorio: World Bank, fbarrera@worldbank.org; Raju: World Bank, draju2@worldbank.org. We thank the staff of the Punjab Education Foundation, in particular Usman Tariq Bhatti, Mian Kashif Ijaz, Shahid Hafiz Kardar, Allah Bakhsh Malik, Haroon Raheem, and Huma Rizvi, for extensive discussions on program design and implementation and assistance with the program administrative data; Nabiha Asghar, Hummayun Javed, Kashif Mahmood, and Sehar Sarah Sikander Shah for assistance with the follow-up data collection; Corinne Saiens for assistance with the National Education Census data; Amit Dar, Michelle Riboud, Sofia Shakil, and the Government of Punjab School Education Department/Project Management and Implementation Unit for their encouragement and support of the research project; and Salman Asim, Elizabeth King, Leigh Linden, and Quynh Nguyen for helpful comments and suggestions. Financial support for the follow-up data collection from the World Bank's Education Program Development Fund and Bank-Netherlands Partnership Program Trust Funds is gratefully acknowledged. The findings, interpretations, and conclusions expressed herein are our own and do not necessarily represent the views of the World Bank, its Board of Directors, or any of its member countries. All remaining errors are our own.

1. Introduction

School participation in Pakistan is generally low in absolute terms, relative to other countries in its region, and relative to developing countries at its level of per capita income. Given the present trend, Pakistan is likely to fall significantly short of the United Nations' Millennium Development Goal of universal primary education by 2015. Thus, measures that produce sharp and sustained gains in participation are viewed as pressing by development policymakers and practitioners in the country.

In this paper, we estimate the causal effects of a recently-instituted public subsidy program to low-cost private schools in the province of Punjab, Pakistan on school enrollment and inputs. Coined the Foundation Assisted Schools (FAS) program and administered by the Punjab Education Foundation (PEF), a semi-autonomous intermediary organization, the rationale behind the program is to leverage, essentially via public financing, the growing low-cost private school system in Punjab to increase equitable access to schooling more efficiently than can be achieved through the province's public school system.

Given the contracting modality, the program falls under the rubric of public-private partnerships (PPPs) in education, a set of interventions which is increasingly perceived in international policy circles as a promising mechanism for attaining key education goals (World Bank 2009). In addition, opportunities for introducing PPP programs of medium to large scale are emerging in several developing countries (e.g., India, Kenya, and Nigeria) as the private education sector matures and becomes an important player in service delivery.

The FAS program was initiated on a pilot basis in November 2005, and rapidly expanded in phases. As of September 2008, the program had completed four phases of expansion and

covered 1,082 low-cost private schools (primary, middle, and secondary levels¹) with roughly 474,000 students across 18 out of 35 districts in Punjab. In the districts with the highest concentration of program schools, we estimate that the program covers approximately one-fifth of all private schools. The program offers a monthly per-student cash subsidy strictly tied to, among other things, free schooling for all enrolled children (thus, pricing program schools in line with public schools which are officially free) and a minimum student pass rate in a specially-designed standardized academic test administered semi-annually by PEF. Program schools have free rein on how to spend the subsidy amount.²

To the best of our knowledge, credible evidence on the impacts of public subsidies to private schools is limited. The study by Kim et al. (1999) is particularly pertinent as they evaluate a public subsidy program in Pakistan which shares some of the design elements of the FAS program. They study the impact of a program that offered a low temporary per-girl student subsidy conditional on free girls' schooling to establish and operate private primary schools in a randomly-selected subset of poor urban neighborhoods lacking public girls' primary schools in the province of Balochistan. Using a difference-in-differences approach, they find that the program substantially increased girls' as well as boys' school participation in treatment neighborhoods, and that these increases were obtained at lower costs than would have been possible through the public school system. For other well-identified evaluation studies of subsidy programs (as well as of other types of PPP interventions), see the recent review by World Bank (2009).

¹ Primary schools are composed of grades 1–5. Middle schools are composed of grades 1–8 or 6–8. Secondary schools are composed of grades 1–10, 6–10, or 9–10.

² In contrast, the standard subsidy program is designed to directly finance educational inputs (Gauri and Vawda 2003).

Consistent with the objectives of the FAS program, we ask two questions in this study. (1) What is the causal effect of the program on the number of students in program schools? (2) What is the causal effect of the program on inputs, namely the number of teachers, classrooms, blackboards, and toilets, and on student-teacher and student-classroom ratios in program schools? To answer these questions, using school-level baseline and follow-up data on school characteristics and outcomes obtained from program administrative records and telephone interviews of school administrators, we fit appropriate regression-discontinuity (RD) designs to the treatment assignment mechanism in order to obtain reliable nonexperimental estimates of program impacts.

In the last two entry phases preceding this study, phases 3 and 4, a standardized academic test, called the Short Listing Quality Assurance Test (SLQAT), was administered by PEF as the final step in the program entry screening process. In order for schools to enter the program in these two phases, they had to apply to the program, pass a qualitative physical inspection, and then pass the SLQAT. If the school achieves the stipulated minimum student pass rate (the cutoff) in the SLQAT, then the school becomes eligible for the program, and not otherwise. Furthermore, in practice, virtually all schools that become eligible elect to participate in the program.

At the time of the follow-up data collection in October 2008, as phase 4 was the last entry phase, schools that took the phase-4 SLQAT were either untreated or treated based on their SLQAT pass rate relative to the cutoff (i.e., the probability of treatment jumps from zero to one at the cutoff). This structure allows us to apply a sharp RD design to these data. On the other hand, schools that took the phase-3 SLQAT and failed found they had another opportunity to seek entry when phase 4 was announced—some phase-3 SLQAT “failers” reapplied to phase 4,

releared the physical inspection, retook the SLQAT, and passed it. We find that the probability of treatment for phase-3 SLQAT takers also jumps at the cutoff but by less than one. This structure allows us to apply a special case of a fuzzy RD design which accounts for what can be viewed as equivalent to noncompliance by initially-untreated schools.

Given these designs, the effects of the program are identifiable and estimable. In the case of the sharp RD design, under some mild regularity conditions, the average causal effect of the treatment on the treated at the cutoff is identified. In the case of our fuzzy RD design, under the same regularity conditions, the average causal effect of the treatment on the *untreated* at the cutoff is identified. Under both designs, the respective treatment effects at the cutoff are estimated nonparametrically using local linear regressions.

Our findings on the impacts of the FAS program differ by phase. For phase-3 SLQAT takers, applying a partially-fuzzy RD design, we find no evidence of program impacts at the cutoff on our outcomes of interest. We posit that this is likely due to weak identification arising from a small jump in the probability of treatment at the cutoff. In contrast, for phase-4 SLQAT takers, applying a sharp RD design, we find robust evidence of positive program impacts at the cutoff on the number of students, teachers, classrooms, and blackboards. The program impacts were sizeable both in absolute terms as well as relative to the baseline means for these outcomes for phase-4 SLQAT takers near the cutoff: our most conservative estimates suggest that the program expanded schools by roughly 85 students, and 3–4 teachers, classrooms, and blackboards, which translate into relative increases of 27–47%, depending on the outcome of interest. These impact estimates are particularly impressive when we consider that phase-4 program schools were exposed to only about ten months of treatment before the follow-up data were collected. Our cost-effectiveness analysis suggests that the FAS program is among the

cheapest interventions in developing countries for generating enrollment gains. Notwithstanding, the estimated impact on enrollment needs to be interpreted cautiously as, given our data, we are unable to establish the extent to which the enrollment gains translate into school participation gains given that, among other things, the elimination of school fees in program schools is likely to also attract students from non-program schools in the vicinity of program schools.

The remainder of the paper is organized as follows. Section 2 describes the education context in Pakistan at the time the FAS program was introduced. Section 3 describes the program in detail. Section 4 lays out our identification and estimation strategies. Section 5 describes the data and provides some summary statistics for SLQAT takers. Section 6 presents our impact and cost-effectiveness findings. Section 7 summarizes and interprets our main findings, as well as provides some concluding comments. These comments comprise of a discussion of external validity and two specific potential threats to internal validity: treatment spillovers to non-program low-cost private schools within the sphere of influence of program schools, and behavioral changes in non-program schools in anticipation of future program entry.

2. The education context

The FAS program was conceived and introduced into an education landscape defined by three features. One: equitable access to schooling is an acute, persistent challenge. Two: the public sector, which remains the dominant provider of education, suffers from chronic weaknesses which impair its ability to effectively address this challenge. Three: in the wake of this public sector failure, the private sector has aggressively stepped in in recent years, growing dramatically in size and reach, and now constitutes a major policy opportunity for addressing this challenge.

The education situation of Pakistan is generally poor in absolute terms, relative to other countries in its region, and relative to developing countries at its level of per capita income (see, e.g., United Nations Educational, Scientific and Cultural Organization 2007 for a recent comparative picture). Given the present trend, Pakistan is unlikely to meet the United Nations' Millennium Development Goal of universal primary education by 2015. The education situation of Punjab—Pakistan's most prosperous and populous province and the site for the FAS program—is by and large comparable to the rest of the country. As Panel A in Table 1 shows, estimates using household sample survey data from 2004/05, which is just before the FAS program was initiated, show that the participation rate in formal school (grade 1+) of children ages 6–15 years was 65.7%, with a significantly lower share for children from the poorest (bottom expenditure quintile) households.

The public school system has been hampered in its ability to improve education outcomes due to, in large part, the lack of effective accountability and incentive systems which promote the legitimate and efficient use of allocated resources (Social Policy and Development Center 2003).³ While the public school system has struggled to enroll children and educate them, the private school system has grown dramatically as reflected by the number of institutions and share of enrollment. In addition, responding to the broad demand for greater access and better quality, the system has evolved in character, becoming less elite and more egalitarian, increasingly

³ Recent evidence on relative teacher performance in private and public schools in Punjab suggests this. Andrabi et al. (2008a) find that, while public school teachers are on average significantly better paid and have higher levels of education, training, and experience than their counterparts in private schools, the teacher absenteeism rate is 15% in public schools compared to 8% in private schools. They also find that, whereas private school teacher salaries are increasing in teacher competency (as measured by teacher test scores) and decreasing in teacher absenteeism, public school teacher salaries do not exhibit similar correlation patterns and instead appear to be largely determined by teacher credentials (education, training, and experience). Looking at child education outcomes, Andrabi et al. find that private school teacher salaries are increasing in student test scores whereas public school teacher salaries are not. The responsiveness of private school teacher salaries to these key behaviors and outcomes suggests the influence of market discipline which the public school system is not subject to, at least directly.

reaching low-income and rural households.⁴ Using private school census data from 2000, Andrabi et al. (2008b) find that there was an exponential increase in the number of private schools over the 1990s, with over 50% of existing private schools established in or after 1996. Furthermore, they find that while existing schools set up before 1990 were predominantly in urban areas, the distribution since then has become increasingly rural.

In line with this increase in institutions, the share of enrollment in private school has also increased. As Panel B in Table 1 shows, in 2004/05, 15.8% and 18.5% of children ages 6–15 years were enrolled in private schools in all-Pakistan and Punjab, respectively; these shares represent significant increases from 1998/99. The increases were particularly dramatic in rural areas and for households in the poorest expenditure quintile. Similar evidence is provided by Andrabi et al. (2008b) who find that the growth rate in private school enrollment over the 1990s was highest among low-income households nationally, and among middle-income households in rural areas, whereas the growth rate in public school enrollment over the same period was negative in both urban and rural areas and across the household income distribution. They also find that fees in private schools are generally low: median annual fees per student in 2000 in Pakistan were 960 (US\$23.4⁵) and 751 rupees (US\$18.3) in urban and rural areas, respectively, and account for a small percentage of mean annual household expenditure. The corresponding statistics for Punjab were lower, at 828 rupees (US\$20.2) and 600 rupees (US\$14.6).

⁴ It is conceivable that much of the growth and metamorphosis in the private school system is in direct response to the rigidities and shortcomings in the public school system.

⁵ The exchange rate in 2000 was 41 rupees per US dollar.

3. The Foundation Assisted Schools program⁶

Institutional background: The FAS program is administered by PEF, a publicly-funded semi-autonomous statutory organization established in 1991 which serves as the main institutional mechanism for PPPs in education in Punjab. The organization's primary aims are to enable socioeconomically-disadvantaged households to access private education and to raise the quality of education in low-cost private schools. To these ends, it employs an array of instruments such as start-up and operational subsidies, vouchers to households in poor localities, school management and teacher training, and specialist teacher services to promote private education. The FAS program is PEF's largest program. In fiscal year 2007–08, PEF spent 1.1 billion rupees (US\$12.9 million⁷) on FAS program benefits.⁸ This amount accounted for 61% of total expenditures by the organization in that year.

Program coverage and timeline: The FAS program was initiated in November 2005 on a pilot basis in 54 schools in five districts in Punjab. Since then, PEF has expanded the program in phases to cover additional districts as well as more schools within program districts. (See Table 2 for a timeline of the program highlighting the start-dates of each of the phases as well as the dates of other key events.) As of September 2008, the program has completed four entry phases (the pilot phase represents phase 1), and covers 1,082 private primary, middle, and secondary schools in 18 out of the 35 districts in Punjab.⁹ Of these, 945 schools (87%) are located in just seven districts (see Table 3 for a district- and phase-wise disaggregation of the number of program schools as well as Figure 1 for the location of program districts in Punjab). This number

⁶ The information on the program presented in the paper reflects program design and administration until end 2008.

⁷ In this and following sections, the exchange rate used for the conversions is 85 rupees per US dollar (effective March 2011). When the program was introduced in 2005, the exchange rate was roughly 60 rupees per US dollar; the rupee has steadily weakened since then.

⁸ FAS program administrative costs are unavailable as PEF does not disaggregate administrative costs by program. We however know that total administrative costs were less than 1% of total expenditures in fiscal year 2007–08, suggesting that FAS program administrative costs are negligible relative to current FAS program benefit outlays.

⁹ The FAS program also currently covers three higher secondary schools (schools with grades 11–12).

of program schools represents a significant share of all private schools in these seven districts: using the 2005 National Education Census (NEC), a census of schools in Pakistan, we estimate that the program covers 21% of private schools in these districts (see Table 4).¹⁰

Program location: The program was initially designed to be targeted at districts ranked lowest in terms of adult literacy rates based on 2003/04 Multiple Indicators Cluster Survey data (Government of Punjab 2004). However, in phases 1 and 2, this targeting decision was not applied as PEF was limited in its institutional and logistical capacity at that time to effectively administer the program in poorly-ranked districts that had physical environmental challenges and limited transportation and accommodation options (Malik 2007). In contrast, in phases 3 and 4, the targeting decision was effectively applied. Consequently, as Table 3 shows, 51% and 89% of program schools are located in districts ranked among the bottom-quarter and bottom-half in terms of adult literacy rates, respectively.

Program school characteristics: Table 5 presents the distribution of program schools by selected characteristics measured in September 2008. Aggregating the phases together (Column 5), the mean school size is 351 students. The majority of schools are middle level (59%), coeducational (83%), registered with local government authorities (87%), and rural (55%). Disaggregating by phase (Columns 1–4), the distribution of program schools by these characteristics is roughly comparable except for the level of the school: in contrast to the pattern in the aggregate sample, program schools that entered in phases 1 and 2 were mainly secondary schools (65–73%). School size also appears to be monotonically decreasing with phase; the mean size of program schools which entered in phase 1 is 561 students, whereas that of program schools which entered

¹⁰ NEC's coverage of private schools might be incomplete. According to the survey documentation, private schools were identified in the field by interviewers with the assistance of local officials. It is conceivable that private schools that are, for example, unregistered or very small, or operate in obscure locations are more likely to be missed by interviewers. Thus, the estimate of 21% can be considered as an upper-end estimate of FAS program coverage of private schools.

in phase 4 is 242 students. While there may be multiple (competing) explanations for this pattern, length of exposure to the program is among them.

Program benefits: The main program benefit is an enrollment-related subsidy: the school receives a monthly per-student cash subsidy amount of 300 rupees (US\$3.5) up to a maximum of 750 students (i.e., the total amount is capped at 225,000 rupees or US\$2,647).¹¹ Given a mean school size of 215 students in phase-3 and phase-4 program schools at the time of application to the program, the mean monthly subsidy payment is roughly 64,500 rupees (US\$759) at program entry. Enrollment information for determining the subsidy amounts is submitted by program schools to PEF on a monthly basis using standardized reporting forms; if enrollment has increased by 50 students or more over one month, PEF visits the school to verify the information before raising the subsidy amount. PEF indicates that, when schools report large changes in enrollment, these reports tend to be at the start of the academic year in April.

PEF reports that the subsidy level was set low for two reasons. First, it confines the attractiveness of the program to low-cost private schools. Second, it raises the political palatability of the program as the per-student subsidy amount is less than half of the estimated per-student expenditure in the public school system at the time the program was introduced.¹² The subsidy benefit is paid for all twelve months of the year. To facilitate timely and regular payments, starting in August 2007, the benefit amounts have been transferred electronically to the bank accounts of program schools.

¹¹ The program also offers two cash bonus benefits. The first is a teacher bonus for a high level of school test performance: once every academic year, a maximum of five teachers in each program school where at least 90% of students in tested classes obtain a score of 40% or higher in the QAT receive an award of 10,000 rupees (US\$118) each. The second is a competitive school bonus for top school test performance: once every academic year, the program school in each of the seven main program districts which has the highest share of students with a score of 40% or higher in the QAT is awarded 50,000 rupees (US\$588).

¹² The exact subsidy amount was heavily guided by a survey conducted by PEF in 2005 in selected districts which showed that the vast majority of private schools that operate in rural areas and disadvantaged urban neighborhoods charge between 50–400 rupees per month (US\$0.6–4.7) in fees. Based on this information, the subsidy amount was set at the upper-segment of this price range (Malik 2007).

Initial benefit eligibility rules: School entry into the program follows a three-step process. In step 1, schools apply to the program when PEF issues a call for applications in newspapers. Application eligibility is restricted to existing private primary, middle, and secondary schools with a minimum enrollment of 100 students from the districts listed in the call. Except in a few cases, only schools that submit properly filled-in applications by the announced deadline are considered for step 2.

In step 2, PEF inspection teams visit schools unannounced to verify the data provided in the applications as well as assess the local reputation of the school and the quality of the physical infrastructure and schooling environment. PEF provides a points and weighting scheme to its inspection teams—this scheme primarily helps ensure that the same school attributes are considered across schools and by different inspection teams. However, the screening exercise is largely subjective: whether a school qualifies for step 3 depends principally on qualitative impressions gathered by the inspection team on a given set of indicators.

In step 3, which was introduced starting with phase 3, all students present in selected grades in the school on the day of the physical inspection screening are offered a 55-minute, written, curriculum-based test called the Short Listing Quality Assurance Test (SLQAT).¹³ The SLQAT is a pared-down version of the Quality Assurance Test (QAT) which is used to determine continuing benefit eligibility once schools enter the program. The SLQAT tests knowledge and comprehension in three subjects: English, Urdu, and mathematics. It is designed by the testing department of PEF, the Academic Development Unit (ADU). Each school is tested on its own premises. Depending on the level of the school, two to three grades among grades 3, 4, 6, 7, and 9 are tested. Which grades are tested in a given school are not disclosed in advance and are randomly selected by ADU. Furthermore, multiple test papers are prepared for each

¹³ For schools that entered the program in phases 1 and 2, step 2 was the final entry step.

grade; ADU randomly selects which specific test papers are offered in a given school. The SLQAT is transported in sealed envelopes and opened only in the presence of the school administrator and teachers. At that time, both the PEF inspection team and the school learn which grades are selected for the SLQAT. Test invigilation is coordinated by the ADU member assigned to the PEF inspection team. The completed tests are transported back to PEF headquarters and graded by ADU. The school passes the SLQAT if at least 67% of tested students score 33% or higher.

Figure 2 presents the numbers of schools which passed each step of the program entry process. Out of the 1,070 and 1,430 schools that had submitted properly filled-in applications and were inspected in phases 3 and 4, respectively, 799 (75%) and 872 (61%) schools were offered the SLQAT. Of the schools that were offered the SLQAT, 514 (64%) and 431 (49%) schools achieved the minimum pass rate and became eligible for the program.¹⁴ In terms of program take-up, PEF reports that 482 (94%) and 425 (98%) of the schools that passed the SLQAT in phases 3 and 4 respectively signed the formal program participation agreement.

Continuing benefit eligibility rules: The program participation agreement stipulates several conditions for maintaining benefit eligibility. The conditions that are stringently applied by PEF are (1) schooling is offered to students without charging them any fees (and displaying the free-schooling status prominently on a PEF-issued signboard outside the school gate) and (2) participation of the program school in the QAT and that at least 67% of the tested students score 40% or higher on the QAT. A one-time violation of these conditions typically results in a warning and the capping of enrollment figures for the subsidy payment until the next QAT round. A second violation results in the permanent disqualification of the school with immediate

¹⁴ Our own tallies based on the SLQAT data deviate slightly from the above numbers: we find that 796 and 856 schools took the SLQAT and, out of these, 511 and 432 schools achieved the minimum pass rate in phases 3 and 4, respectively.

effect. With the QAT-related condition, the school is permanently disqualified if it fails to achieve the minimum pass rate in the QAT in two consecutive attempts.

The QAT is a 65-minute, written, curriculum-based test designed by ADU, and administered twice a year in October–November (in the first term) and February–March (in the second term). It tests knowledge, comprehension, application, analysis, and synthesis. The subject areas covered in the tests are English, Urdu, mathematics, and science (general science in grades 1–8 and biology, chemistry, and physics separately in grades 9–10). The general test administration procedures previously delineated with respect to the SLQAT are also followed with the QAT.¹⁵ By the time of the follow-up data collection, PEF had conducted five QATs (see Table 2 for the dates). Phase-3 program schools had been subject to two QATs and phase-4 program schools one.

There are also other conditions for maintaining benefit eligibility. These include (1) registering the school with the District Registration Authority within one year of joining the program; (2) conducting only one class in a classroom in any period; (3) maintaining or upgrading the quality of the school’s physical infrastructure (e.g., adequate classroom space, properly-constructed rooms and buildings, sufficient ventilation, and sufficient artificial and natural light); (4) acquiring and maintaining adequate furniture and teaching tools (e.g., benches, desks, and blackboards); (5) providing monthly reports to PEF on enrollment counts; (6) keeping student-teacher and student-classroom ratios below 35:1; (7) keeping enrollment above 100 students; and (8) not holding after-hours classes or tutoring services at the school. These additional conditions are applied more leniently; typically, when PEF detects a violation among this subset of conditions, schools are provided with a warning and a grace period within which to

¹⁵ One important difference is that, unlike with the SLQAT, the school receives formal advance notices of the date of the QAT, and at least 80% of its students are expected to be in school on the day of the test.

comply. As of the follow-up data collection, no program schools have been disqualified for repeated violations of these conditions.

Table 6 presents the number of program schools that were disqualified, disaggregated by phase. The data are up until the time of the follow-up data collection. The number of schools that have been disqualified for any reason is negligible: only 28 out of the 1,111 program entrants (2.5%) have exited the program over time. In addition, in most cases, the reason for disqualification is two consecutive failures to achieve the minimum pass rate on the QAT. Importantly, for our analysis, only three schools that entered the program in phases 3 and 4 have exited the program for any reason.¹⁶ Thus, program dropout has not been an issue with phase-3 and phase-4 program schools.

Direct impact channels: The structure of the program can be expected to have a positive effect on enrollment via several direct channels. For example, setting the monthly subsidy as a linear function of the number of children enrolled incentivizes program schools to draw in additional students. Tying the receipt of program benefits to the elimination of school fees—which puts the program school’s price at an advantage over other local non-program competitor private schools and at par with public schools—is likely to raise the attractiveness of program schools, particularly among households for which school fees serves as a major constraint to sending their children to private school.¹⁷

The structure of the program can also be expected to directly affect investments in the quantity and quality of school inputs and resources. For example, increases in enrollment

¹⁶ PEF reports that the three schools were problematical cases from the outset; the schools were ejected from the program for general noncompliance and nonperformance.

¹⁷ An increase in enrollment in program schools may not translate into a commensurate increase in participation among children in areas where program schools operate. The lower price of schooling in program schools is likely to induce both displacement and diversion effects: some share of the new enrollment in the program school will likely come from students already enrolled in other schools or children that where initially *considering* enrolling elsewhere.

induced by the per-student subsidy have to be met by increases in the number of classrooms and teachers if the stipulated maximum student-teacher and student-classroom ratio conditions are to be complied with. The maximum ratios encourage program entrants with preexisting ratios in excess of the maximums to invest in additional classrooms and teachers. The physical infrastructure and learning environment quality conditions encourage program schools to ensure the proper design and construction of infrastructural expansions, and to invest in teaching tools (e.g., blackboards) and basic facilities (e.g., toilets), in step with enrollment growth. These input-related conditions encourage schools to schedule investments in school inputs and resources to either lead or accompany enrollment increases. However, given that these conditions are not stringently applied, it suggests that PEF might tolerate investments in school inputs and resources that lag enrollment increases, although reportedly not by too long.

4. Empirical strategy

Identification

Following the exposition in van der Klaauw (2008) and Todd (2007), let y_i denote the outcome of interest (e.g., enrollment) in school i , and let the indicator variable $d_i \in \{0,1\}$ denote treatment assignment, where one denotes that the school is covered by the FAS program (treated), and zero if not (untreated). In addition, let y_{0i} and y_{1i} denote the potential outcomes of school i in the untreated and treated states, respectively. The actual outcome observed for school i is given by

$$y_i = d_i y_{1i} + (1 - d_i) y_{0i} = y_{0i} + [y_{1i} - y_{0i}] d_i = y_{0i} + \alpha_i d_i, \quad (1)$$

where α_i denotes the treatment effect for school i .

Unless treatment is randomly assigned, simply comparing mean outcomes of treated schools and untreated schools, $E[y_{1i} | d_i = 1] - E[y_{0i} | d_i = 0]$, would not yield an unbiased estimate of the average causal effect of the treatment: $E[\alpha_i]$. In general,

$$E[y_{1i} | d_i = 1] - E[y_{0i} | d_i = 0] = E[\alpha_i | d_i = 1] - \{E[y_{0i} | d_i = 1] - E[y_{0i} | d_i = 0]\}. \quad (2)$$

The first term on the right-hand side of (2) is the average treatment effect on the treated (ATT) and the second term (in braces) is the selection bias arising from differences in potential average untreated outcomes between treated and untreated schools. Selection is clearly plausible in our case. For example, if, ex ante, program schools have higher levels and better quality inputs and resources relative to non-program schools, then the selection bias term will be positive and generate an upward bias in our estimate of the ATT.

The potential selection bias problem can however be overcome by using the institutional feature that program eligibility is ultimately determined by the student pass rate obtained by the school in the SLQAT relative to the known pass rate cutoff of 67%. Given that virtually all schools that become eligible to participate in the program also choose to participate in the program, in practice, the cutoff determines actual program participation. Thus, program participation status is assigned based on the decision rule

$$d_i(z_i) = 1\{z_i \geq c\}, \quad (3)$$

where z_i denotes school i 's pass rate which is perfectly observed (z will be more generally referred to as the assignment variable), c the known distinct cutoff pass rate, and $1\{\cdot\}$ an indicator function. This design for treatment assignment based on a completely deterministic function is referred to as a sharp regression-discontinuity (RD) design.

The SLQAT pass rate z might be correlated with our outcomes of interest. Thus, the treatment assignment mechanism is clearly not random and comparing schools that receive the treatment ($d_i(z_i) = 1$) to schools that do not ($d_i(z_i) = 0$) will yield a biased estimate of the ATT. If, however, we consider schools with pass rates near the cutoff to be comparable, then the treatment assignment mechanism in the neighborhood of the cutoff can be viewed as if it was “almost random”.

More formally, let $e > 0$ denote an arbitrarily small number. Comparing the outcomes of schools with pass rates just below the cutoff (marginal failers) with the outcomes of schools with pass rates at or just above the cutoff (marginal passers), yields

$$E(y_{1i} | z_i = c + e) - E(y_{0i} | z_i = c - e) = E[\alpha_i | z_i = c + e] - \{E[y_{0i} | z_i = c + e] - E[y_{0i} | z_i = c - e]\}. \quad (4)$$

Under the assumptions that (1) the limit $\lim_{e \downarrow 0} E[y_{0i} | z_i = c + e]$ is well defined; (2) $E[y_{0i} | z_i = c]$ is continuous in the assignment variable z at the cutoff (i.e., the conditional expectations of the outcome variable exhibits local smoothness at the cutoff in the absence of the treatment); and (3) the density of the assignment variable z is positive in the neighborhood of the cutoff, the difference in the mean outcomes between marginal passers and marginal failers identifies

$$E[\alpha_i | z_i = c] = \lim_{e \downarrow 0} E(y_{1i} | z_i = c + e) - \lim_{e \downarrow 0} E(y_{0i} | z_i = c - e). \quad (5)$$

This object is interpreted as the ATT at the cutoff (Hahn et al. 2001; Todd 2007).

A sharp RD design neatly fits the phase-4 SLQAT taker data, as the school’s pass (eligibility) versus fail (ineligibility) status remains fixed given that phase 4 was the last entry phase before the follow-up data collection. A variant of the above framework will however be required for phase-3 SLQAT takers given that some phase-3 failers reapplied to the program in phase 4 and passed the SLQAT at that time. As a result, our set of phase-3 SLQAT takers can be

divided into three distinct subsets: (1) phase-3 failers who did not enter the program any time later; (2) phase-3 failers who entered the program later in phase 4 (crossover schools); and (3) phase-3 passers who entered the program at that time (as noted before, the treatment take-up rate by passers was virtually 100%, and there have been no dropouts to date). While the pass rate cutoff rule strictly determines program eligibility status for phase-3 SLQAT takers, the processes and factors behind the phase-4 application decision of phase-3 failers are unobservable to us and likely to be nonrandom. Given this, the eventual participation status of this group of schools is also considered to be driven partly by selection on unobservables.

In our data, we find that the probability of program participation as a function of the pass rate in the phase-3 SLQAT, $E(d_i | z_i) = \Pr[d_i = 1 | z_i]$, consists of a positive but smaller jump than one at the cutoff. The probability of program participation will still be discontinuous in the assignment variable z at the cutoff. This feature implies that the phase-3 SLQAT taker data are more appropriately fitted with a fuzzy RD design. Under this design, apart from the identifying assumptions for the sharp design, if we assume that treatment d_i is nondecreasing in the assignment variable z at the cutoff (local monotonicity) and that potential outcomes are independent of treatment conditional on the assignment variable z at the cutoff (local conditional independence), then the average causal effect of the treatment is identified by

$$E[\alpha_i | z_i = c] = \frac{\lim_{e \downarrow 0} [y_{1i} | z_i = c + e] - \lim_{e \downarrow 0} [y_{0i} | z_i = c - e]}{\lim_{e \downarrow 0} [d_{1i} | z_i = c + e] - \lim_{e \downarrow 0} [d_{0i} | z_i = c - e]}, \quad (6)$$

where the denominator of (6) is nonzero given the discontinuity in the probability of treatment at the cutoff (Hahn et al. 2001; Todd 2007).

This treatment parameter represents the RD analog to the local average treatment effect (LATE) (Imbens and Angrist 1994; Angrist et al. 1996). It can be interpreted as the average

treatment effect for schools in the neighborhood of the cutoff who are induced to take up treatment when their pass rates are at or above the cutoff ($d_i(z_i)=1$) but not so when they are below ($d_i(z_i)=0$).

The general fuzzy design setup above allows for partial (or imperfect) compliance for both treatment-eligible and treatment-ineligible groups; that is, it allows for both crossovers (treatment-ineligibles taking up treatment) and no-shows (treatment-eligibles not taking up treatment) (Bloom 1984). In our case, the problem is essentially one of one-way partial compliance: we have the equivalent of crossovers over time from treatment-ineligible to treatment-eligible status, with all treatment-eligible schools choosing to take up treatment. Hence, the probability of program participation for phase-3 SLQAT passers is one, while the probability for failers lies between zero and one.

Battistin and Rettore (2008) describe the other one-way partial compliance case, where treatment-ineligible agents are not exposed to treatment and treatment-eligible agents self-select into treatment. They formally show that the LATE at the cutoff can be identified simply under the conditions required for identifying the ATT at the cutoff under the sharp RD design, thus labeling their framework as a “partially” fuzzy RD design. Their framework straightforwardly applies to our case of one-way partial compliance as well. However, whereas in their specific setup, they identify the ATT at the cutoff ($\lim_{e \downarrow 0} E[\alpha_i | z_i = c + e]$), by applying the smoothness condition to obtain the counterfactual average outcome for the untreated ($\lim_{e \downarrow 0} E[y_{i1} | z_i = c - e]$), we instead identify the average treatment effect on the *untreated* ($\lim_{e \downarrow 0} E[\alpha_i | z_i = c - e]$), by applying the smoothness condition to obtain the counterfactual average outcome for the treated ($\lim_{e \downarrow 0} E[y_{i0} | z_i = c + e]$) (Duflo et al. 2007).

Estimation

Choice of estimator: Given that we are interested in estimating the treatment effects at a single point using observations in its neighborhood, one suitable approach is local smoothing using nonparametric regression. Following Hahn et al. (2001), we opt for local linear regression (a local polynomial of order one). In the case of the sharp RD design, which applies to our phase-4 SLQAT takers, the application of local linear estimation entails individually estimating the conditional expectations of the outcome y at the cutoff from below (denoted by $\hat{\alpha}_y^-$) and above the cutoff (denoted by $\hat{\alpha}_y^+$) and then subtracting the two estimates. Under the sharp RD design, a consistent estimate of the ATT at the cutoff is given by

$$\hat{\alpha}_y = \hat{\alpha}_y^+ - \hat{\alpha}_y^- . \quad (7)$$

In the case of the fuzzy RD design, the application of local linear regression entails estimating the ratio of two differences. Given our particular problem related to phase-3 SLQAT takers, we individually estimate the conditional expectations of the outcome y at the cutoff c from above and below ($\hat{\alpha}_y^+$ and $\hat{\alpha}_y^-$, respectively), and the conditional expectation of treatment d from below only (denoted by $\hat{\alpha}_d^-$). Note that the conditional expectation of treatment d from above is one by definition. Under our particular partially-fuzzy RD design, a consistent estimate of the LATE at the cutoff is given by

$$\frac{\hat{\alpha}_y}{\hat{\alpha}_d} = \frac{\hat{\alpha}_y^+ - \hat{\alpha}_y^-}{1 - \hat{\alpha}_d^-} . \quad (8)$$

To perform statistical inference, empirical standard errors for (7) and (8) are obtained using the standard i.i.d. nonparametric bootstrap with paired sampling (y_i, z_i) , resampling from the original data 500 times.

Choice of kernel and bandwidth: Implementation of local linear estimation requires the specification of the kernel k , the weighting function, and bandwidth $h > 0$, the window width in which the kernel function is applied. We opt for the triangular kernel given that it is boundary optimal and thus well suited to RD problems (Cheng et al. 1997).¹⁸

The choice of bandwidth is a relatively more important decision given the trade-off between estimation bias and variance. Consensus in the literature on valid and robust methods for bandwidth selection for RD problems is lacking.¹⁹ Given this, we discount the specific method used to select the optimal bandwidth and instead emphasize checking the robustness of our findings to reasonable deviations in bandwidth from the optimal bandwidth choice.

To select the optimal bandwidth, we rely on the default bandwidth choice in Nichols' (2007) `rd` program for `Stata` which assigns positive weight to at least 30 observations on each side of the cutoff, and applies the same bandwidth for the local linear estimators of the conditional mean outcomes (and probabilities of treatment) above and below the cutoff (i.e., $h^+ = h^- = h$). In our case, the default choice translates to a bandwidth of three or four percentage points for our RD impact estimations. We believe that this is a conservative choice, and

¹⁸ As Imbens and Lemieux (2008) note, while more sophisticated kernels are available, they do not provide any significant gain in asymptotic bias reduction. Furthermore, they note that, in general, parameter estimates appear to be robust to the choice of kernel.

¹⁹ This is, in large part, due to two reasons. First, standard automated procedures for bandwidth selection for nonparametric regression have been developed with the estimation of functions at interior points in mind; the nature of the trade-off between bias and variance at interior points may differ from the nature of the trade-off at boundary points (Ludwig and Miller 2005). Second, the literature on the development of bandwidth selectors customized to the RD context is an emergent one (e.g., McCrary and Royer 2006; Imbens and Lemieux 2008; Imbens and Kalyanaraman 2009). In addition, the relative finite sample performances of the proposed bandwidth selection methods for RD designs have yet to be rigorously assessed.

acceptably balances restricting estimation to a local neighborhood around the cutoff against having sufficient statistical power to yield informative estimates of treatment effects. Notwithstanding, we examine whether our RD impact estimates and statistical inference results are sensitive to increasing the optimal bandwidth by 50% and 100%, potentially trading off increased bias for reduced variance.

5. Data and sample

Data

Baseline data: The baseline data come from the program application records and SLQAT test records collected and maintained by PEF. In phase 3, properly filled-in application forms received by PEF before the announced deadline of April 2007, as well as forms filled out by schools solicited by PEF inspection teams on their way to inspect applicant schools were considered. In phase 4, only properly filled-in forms received by PEF before the announced deadline of July 2007 were considered. PEF collected 1,070 and 1,430 properly filled-in application forms in phases 3 and 4, respectively.²⁰

PEF constructed a school-level application electronic database for each phase which contains information on school characteristics (location, gender type, level, physical infrastructure, and registration status), total school enrollment separately by gender, total number of teachers and administrative staff separately by gender, and the minimum and maximum monthly teacher salaries in the school.²¹ These databases serve as the source of baseline data on the following outcomes measured at the school-level: number of students, teachers, classrooms,

²⁰ The total number of unique applications received by PEF is unknown as all rejected applications were discarded.

²¹ As part of this inspection, the data provided by the school in the application form are verified. These inspection data would have been useful for checking the accuracy of the application data; they were however collected in paper form and not entered into an electronic database. Consequently, these data are unavailable for the purposes of this study.

blackboards, and toilets. The school-level outcomes of student-teacher and student-classroom ratios are constructed by us using the data on total number of students, teachers, and classrooms. These databases also serve as the source of baseline data on school-level covariates, namely location, gender type, level, and registration status.

The SLQAT serves as the source of data on the school's student pass rate, our treatment assignment variable z . Separate electronic databases have been constructed by PEF for phase-3 and phase-4 SLQAT takers. These databases contain the total score for each student that took the test. In the databases, student test scores are organized by school and, within school, by grade. The school identification information provided in the databases comprise of the school's name (at times, with some address elements) and location (tehsil²² and district). Although PEF constructed the SLQAT pass rate for each school, we used the student test score data in order to construct our own measure of school SLQAT pass rates. The match rate between their calculation of school SLQAT pass rates and ours' in both phases is nearly perfect at 99.5%. While it makes virtually no difference, we use our measure of school SLQAT pass rates as the treatment assignment variable

Constructing a single electronic database for our analysis required that the application data are linked with the SLQAT data at the school level. Unfortunately, the same unique school identification variable was not used across databases. Consequently, we linked the application databases to the SLQAT databases using an iterative visual-matching process. First, schools were matched across databases using the district name and school name variables.²³ Exact matching failed in a number of cases as PEF did not maintain consistency in the spelling, word ordering,

²² Tehsil is the spatial unit of government administration one tier below district. There are 127 tehsils in Punjab.

²³ Although information on the school's tehsil was also available in both the application and test databases, this information was error-ridden. Consequently, we did not use this information in the cross-database matching exercise.

and completeness of the school's name across the two databases; hence, matching on school name frequently required matching on key words and word patterns. In cases where we suspected that the combination of district name and school name (even with keyword and word pattern matching) did not yield a unique school record in a database, where possible, the set of matching variables was extended to include school address. This extension helped resolved a number of cases where we were tentative. On the basis of this exercise, 94% and 97% of school records in the SLQAT databases were linked with school records in the application databases for phases 3 and 4, respectively. Only the linked school records are used in our analysis, which yield sample sizes of 747 and 830 schools in phases 3 and 4, respectively. We refer to these samples as the SLQAT samples.

Follow-up telephone interview data: We collected follow-up outcome data via telephone in October 2008 from phase-3 and phase-4 schools with SLQAT pass rates between ± 15 percentage points of the cutoff. This information was gathered by a small team of independent interviewers. We were able to collect the data using this mode as schools were required to provide the telephone number of the school owner or administrator when they applied to the program, and the information was included in the school application electronic databases for both phases.

We obtained data on the following outcomes: the number of students, teachers, classrooms, blackboards, and toilets (the data on the last two outcomes were collected only for the phase-4 schools).²⁴ We also obtained information on the program participation status of the schools, and crosschecked this information against PEF's records of program schools as of September 2008. The school sample sizes within the selected pass rate range from the cutoff are

²⁴ We are unable to speculate on the accuracy of these data vis-à-vis the same data collected through, for example, face-to-face interviews in the field, as we are unclear what context factors might lead to differences in accuracy across survey modes.

268 and 319 schools in phases 3 and 4, respectively. These numbers constitute 36% and 38% of the respective SLQAT samples. We refer to these samples as the cutoff neighborhood samples.

The period of the follow-up data collection was roughly 14 and 10 months after the receipt of the first subsidy payment by phase-3 and phase-4 program schools, respectively. The treatment exposure period under investigation roughly covers the second half of the 2007/08 academic year and first half of the 2008/09 academic year.

Unit nonresponse analysis: Despite repeated attempts by the interviewers, not all schools in the cutoff neighborhood samples were contactable over the telephone. Unit nonresponses were mainly due to wrong numbers or unattended telephone calls; when reached, refusals were rare. There were no cases of item nonresponse. Roughly 28% and 22% of schools have missing data due to unit nonresponse in the phase-3 and phase-4 cutoff neighborhood samples, respectively. These large percentages are expected to reduce the study's statistical power (which does pose some concern given that our study is low-powered to begin with) but do not necessarily introduce sample selection bias.

To investigate the presence of systematic bias, we estimate simple bivariate correlations between unit nonresponse and selected covariates, separately by phase. The covariates that we examine comprise of indicator variables for (initial) treatment assignment (i.e., whether the school's SLQAT pass rate is above the cutoff) as well as baseline measures of the school's registration status, level, gender type, and location. Table 7 presents nonresponse counts and rates (Panel A) and the estimated correlations (Panel B). It appears that the probability of nonresponse is not significantly associated with treatment assignment: the estimated correlation coefficients are .03 and .07 and the associated standard errors are .06 and .05 for the phase-3 and phase-4 cutoff neighborhood samples, respectively. In addition, as Columns 3 and 6 in the table

show, barring a few exceptions—which could arise simply due to random chance—the probability of nonresponse does not appear to be significantly associated with the selected school covariates. Thus, the evidence suggests that we cannot reject the null hypothesis that nonresponse rates are generally comparable for marginal failers and marginal passers across the selected dimensions measured at baseline.

Even if nonresponse rates are similar between marginal failers and passers, it is still possible that the selection process for nonresponding marginal failers differs from the selection process for nonresponding marginal passers (Duflo et al. 2007). Table 7 also reports bivariate correlations between unit nonresponse and selected covariates measured at baseline separately for marginal failers and marginal passers (Columns 1 and 2 for phase 3 and Columns 4 and 5 for phase 4). The evidence generally suggests that we cannot reject the null hypothesis that the underlying selection processes for nonresponse for both marginal failers and marginal passers are random. Based on these collective findings, we assume that the nonresponse rate is locally smooth in the assignment variable at the cutoff.

Sample

Table 8 presents statistics on the distribution of schools by selected characteristics measured at baseline, separately by phase and for the SLQAT and cutoff neighborhood samples. We highlight two patterns. First, across both phases and samples, the distributional patterns across characteristics are broadly comparable. For example, the majority of schools are (1) middle schools (63–72%), (2) coeducational (82–87%), (3) officially-registered (81–88%) and (4) rural (53–59%). Second, comparing the phase-3 SLQAT sample to the phase-4 SLQAT sample (Columns 1–3), pairwise *t*-tests of the equality of proportions show significant differences in the

distribution of schools across several characteristics. It appears that, among other things, the shares of secondary schools, single-sex schools, urban schools, and registered schools are larger in the phase-3 SLQAT sample than in the phase-4 SLQAT sample. These distributional differences across phases lose their statistical significance when we restrict our attention to the cutoff neighborhood samples (Columns 4–6). Some part of the explanation for the loss in significance is likely due to reduced statistical power resulting from the smaller sample sizes for the cutoff neighborhood samples relative to the SLQAT samples.

Table 9 presents means and standard deviations for selected outcomes measured at baseline, again separately by phase and for the SLQAT and cutoff neighborhood samples. In the phase-3 SLQAT sample (Column 1), we find that schools on average have 253 students; 10 teachers, classrooms, and blackboards; and 3 toilets. Mean student-teacher and student-classrooms ratios are 25:1 and 28:1, respectively, which, to begin with, are below the stipulated maximums for program benefit maintenance. The differences across phases in the SLQAT samples (Columns 1–3) lose their statistical significance when we restrict our attention to the cutoff neighborhood samples (Columns 4–6). Again, some part of the explanation for the loss in significance is likely to be the lower statistical power of the cutoff neighborhood samples relative to the SLQAT samples. Notwithstanding, the sizes of the inter-phase differences in the cutoff neighborhood samples are generally substantially smaller than the corresponding differences in the SLQAT samples. In sum, the phase-3 and phase-4 cutoff neighborhood samples appear to be similarly composed with respect to mean outcomes measured at baseline.

6. Results

Before presenting our RD estimates of the impacts of the FAS program, we discuss results from a couple of model specification tests. First, we show that there is a discontinuous change in the conditional probability of treatment at the SLQAT pass rate cutoff, as the suitability of the RD model hinges on this feature of the data. Second, we examine whether the density function of the SLQAT pass rate exhibits local smoothness at the cutoff. This test is particularly useful if data are unavailable to directly test whether the conditional mean untreated outcomes at the cutoff exhibit local smoothness. However, in our case, we can directly test the identifying condition for the RD design by using our baseline data on outcomes. Thus, we perform the density test mainly to check whether the result of this test is consistent with the result of the test of the identifying condition.

Discontinuity in the probability of treatment at the cutoff

Figure 3 plots the local linear regression functions estimated separately above and below the cutoff, separately for the phase-3 and phase-4 cutoff neighborhood schools *with* follow-up data on, among other things, program participation status (the optimal bandwidth is six percentage points for both samples). The jump in the probability of treatment is clearly visible for both samples. For the phase-3 sample, as expected, the probability of treatment is one above the cutoff and positive but less than one below (with 42% of marginal failers in phase 3 gaining participation status in phase 4). Furthermore, we observe that the probability of treatment is increasing as we approach the cutoff from below. A plausible and likely explanation for this pattern is that phase-3 marginal failers closer to the cutoff have a greater chance of passing the phase-4 SLQAT and entering the program than marginal failers farther away. For the phase-4

sample, as expected, the probability of treatment jumps discontinuously from zero to one at the cutoff. These patterns in the conditional probabilities of treatment motivate our selections of partially-fuzzy and sharp RD designs for the phase-3 and phase-4 SLQAT samples, respectively.

Local smoothness in the density function of SLQAT pass rates

Rejection of local smoothness in the density function of the assignment variable at the cutoff may suggest manipulation of the assignment variable z (McCrary 2008). Manipulation is plausible in our case as the SLQAT pass rate cutoff for program participation was advance public knowledge, and schools and perhaps even program managers may have an interest in a particular treatment outcome. The pass rate that the school *actually* received on the SLQAT may be subject to partial manipulation, which is typically benign.²⁵ However, PEF staff exercise full control over the calculation of SLQAT pass rates, as well as their reporting, thus, introducing the risk of complete manipulation.²⁶ Given that we use SLQAT pass rates calculated by us based on the individual student test score data as our measure of the assignment variable, and these calculations nearly always match the pass rates calculated by PEF, manipulation would have to occur at the individual student level. Manipulation at this level is less straightforward; nonetheless, we cannot rule out the possibility.

Figures 4 and 5 present frequency histograms of schools with bins equal to the integer values of SLQAT pass rates for the SLQAT and cutoff neighborhood samples, respectively, separately by phase. A simple visual inspection of the SLQAT pass rates for both the phase-3

²⁵ The day the SLQAT is offered, the classes which are tested, and the specific SLQAT papers offered at the school are stochastic elements from the perspective of the school and outside its control.

²⁶ PEF staff may face incentives to manipulate the pass rates of marginal passers downwards (to just below the cutoff) if, for example, the program budget limits the number of schools that the program can be offered to. Given that PEF reports that they were flush with funds during this period, we can probably discount downward manipulation due to this reason. Alternatively, sympathetic PEF staff might manipulate the pass rate of marginal failers upwards (to just above the cutoff), so that these schools become eligible to participate in the program.

and phase-4 samples suggests the presence of discontinuities in their densities in the immediate neighborhood of the cutoff. This is confirmed by formally testing for discontinuities in the densities at the cutoff by applying a test proposed by McCrary (2008) involving separate kernel density estimations below and above the cutoff.²⁷ However, this pattern is not unique to the cutoff location; the histograms show multiple sharp frequency peaks and troughs (at times, at adjacent pass rates) across the SLQAT pass rate support.

Local smoothness in conditional mean outcomes at baseline

As a direct test of the identifying assumption, we examine whether mean outcomes measured at baseline satisfy local smoothness at the cutoff using two approaches: by (1) examining differences in simple means between marginal failers and passers for alternative cutoff neighborhood sizes and (2) estimating the discontinuity in the conditional means at the cutoff via local linear regressions at the selected alternative bandwidths. Table 10 presents the simple mean outcomes at baseline for marginal failers and differences from these means for marginal passers for two cutoff neighborhood sizes: ± 15 percentage points and ± 5 percentage points. The evidence suggests that, for both phase 3 and 4, by and large, there are no statistically-significant differences between marginal failers and passers in mean outcomes measured at baseline.

Table 11 presents sharp RD estimates of baseline mean outcomes at the cutoff using local linear regressions, separately by phase and for the selected alternative bandwidths (the optimal bandwidth is four and three percentage points for the phase-3 and phase-4 cutoff neighborhood samples, respectively). In addition, Figures 6a–6c depict the estimated local linear regression functions for the phase-3 and phase-4 cutoff neighborhood samples, using their respective optimal bandwidths. In general, the evidence suggests that we can reject local smoothness in

²⁷ Results from the tests are available from the authors upon request.

baseline mean outcomes at the cutoff more frequently than what our earlier simple comparisons of baseline mean outcomes between marginal failers and passers would have led us to state, as well as more frequently than we would have expected by random chance given standard significance levels.

Out of the 42 individually-estimated RD parameters across phases, outcomes, and selected bandwidths, seven estimates (17%) are statistically-significant. Virtually all the statistically-significant RD estimates are negative in sign, suggesting that program schools tend to be smaller and have lower levels of inputs at baseline. The inference results are however not robust to the selected alternative bandwidths. Given this sensitivity in inference, we consider the results to generally satisfy the identifying condition.

RD impact estimates

Table 12 presents our RD impact estimates based on local linear regressions, separately by phase and for the selected alternative bandwidths (the optimal bandwidth is six and three percentage points for the phase-3 and phase-4 cutoff neighborhood samples, respectively). Figures 7a–b plot the estimated local linear regression functions using the optimal bandwidth choice for the phase-4 cutoff neighborhood sample.

For the phase-3 cutoff neighborhood sample (Columns 1–3), whose data are subject to a partially-fuzzy RD design, we find no evidence of significant program impacts across the outcomes of interest. The magnitudes of the impact estimates also appear to be sensitive to the selected alternative bandwidths. In addition, the empirical standard errors associated with the impact estimates are inordinately large, suggesting weak identification.

For the phase-4 cutoff neighborhood sample (Columns 4–6), whose data are subject to a sharp RD design, we find evidence of significant positive impacts on the number of students, teachers, classrooms, and blackboards in marginal passers; furthermore, their significance is robust to the selected alternative bandwidths. The accompanying figures show a discernible structural change in the mean levels for these outcomes marginally above and below the cutoff. However, as with the partially-fuzzy RD estimates for the phase-3 cutoff neighborhood sample, the magnitudes of the impacts appear to be sensitive to bandwidth. In particular, the estimates drop fairly sharply when we increase the bandwidth from the optimal choice of 3 percentage points by 50% to 4.5 percentage points. The most conservative impact estimates at the cutoff on the number of students, teachers, classrooms, and blackboards are 85, 3.4, 4, and 2.8, respectively. Relative to the baseline means for these outcomes in the phase-4 cutoff neighborhood sample, the absolute impact estimates translate into percent impact estimates of roughly 37%, 37%, 47% and 27%, respectively. These impacts are substantial. They are particularly impressive given the short treatment exposure under investigation: the follow-up telephone interview data were collected only some ten months after phase-4 program schools received their first subsidy payment.

Notwithstanding, we do not find evidence of impacts at the cutoff on the number of toilets and student-teacher and student-classroom ratios. This finding is robust to the selected alternative bandwidths. The lack of evidence of a positive effect on the number of toilets is concerning given the large expansion in enrollment in marginal passers and its potential negative bearing on the use and maintenance of the facility. The finding may reflect that PEF does not encourage the adequate provision of this facility to the same degree as the provision of other types of physical infrastructure (such as classrooms). Given the expansion in enrollment in

marginal passers, the absence of effects on the ratios suggests that program schools have expanded the number of teachers and classrooms (which we find) in lock-step with this expansion. This behavior may be driven in large part by PEF's condition that program schools maintain these ratios below stipulated levels.

Falsification test: RD estimates at false cutoffs

Table 13 presents sharp RD estimates for our outcomes of interest based on local linear regressions for the phase-4 cutoff neighborhood sample at two arbitrarily-selected false cutoffs: 57% and 77%. The false cutoffs are equidistant from the true cutoff of 67%. We expect to find local smoothness in the conditional mean outcomes at these cutoffs. The subsample for the investigation at 57% is schools with SLQAT pass rates between 52% and 66%. Likewise, the subsample for the investigation at 77% is schools with SLQAT pass rates between 67% and 82%. The optimal bandwidths for the lower and upper samples are six and three percentage points, respectively. At both cutoffs, the evidence suggests that we cannot reject the null hypothesis of local smoothness in the conditional mean outcomes. The inference results are robust to the selected alternative bandwidths.

Cost-effectiveness estimates

We estimate the cost-effectiveness of the program in relation to enrollment gains using two alternative approaches. Using the conservative estimate of the sharp RD impact on enrollment for the phase-4 cutoff neighborhood sample (85 students), a first way that we estimate the program's cost-effectiveness is by deriving the annual rupee cost of one additional student in a program school *induced* by the program. Given a baseline mean school size of 232 students for

schools in the phase-4 cutoff neighborhood sample (which we treat as the number of children that would have attended the program school in the absence of the treatment) and an annual subsidy amount of 3,600 rupees (US\$42.4) per student,²⁸ it costs 13,426 rupees ($=3,600 \times (232+85)/85$) (US\$157.7) to induce an additional student per year. In comparison, this cost is roughly half of the cost of inducing an additional student per year through conditional cash transfers to female students in public secondary schools in Punjab (Andrabi et al. 2008a).

Following the approach by Evans and Ghosh (2008), a second way we calculate the program's cost-effectiveness is by deriving the annual per-student cost of increasing enrollment in program schools by 1%. Using an annual subsidy amount of 3,600 rupees per student and the sharp RD impact estimate on enrollment of 37%, we estimate a cost-effectiveness ratio of 97 rupees ($=3,600/37$) (US\$1.1). This estimate compares extremely favorably with the estimated cost-effective ratios of other evaluated education interventions across the developing world which generated enrollment gains as reported in Evans and Ghosh. In fact, the FAS program's ratio ranks among the very lowest.²⁹ What is more, our cost-effectiveness estimate surprisingly neighbors Evan and Ghosh's estimate of the cost-effectiveness of the per-student subsidy program in Balochistan, Pakistan evaluated by Kim et al. (1999) and noted in Section 1.

7. Conclusion

In this paper, we estimate the causal effects of the FAS program, a recently-instituted public subsidy program to low-cost private schools, in the province of Punjab, Pakistan on student enrollment and schooling inputs. Our findings on the impacts of the program differ by phase. For

²⁸ The annual subsidy amount per student is roughly equal to the annual *program amount* per student, as the per-student amounts for program administrative costs and the teacher and school bonuses add less than 1% to the amount. This is principally due to the large number of students currently covered under the program.

²⁹ This result remains qualitatively unaltered if we precisely follow the currency conversion and inflation adjustments steps taken by Evans and Ghosh to fix all ratios in 1997 US dollars.

phase-3 SLQAT takers, applying a partially-fuzzy RD design, we find no evidence of program impacts at the cutoff on our outcomes of interest. We posit that this is likely due to weak identification arising from a small jump in the probability of treatment at the cutoff. In contrast, for phase-4 SLQAT takers, applying a sharp RD design, we find robust evidence that the program significantly increased the number of students, teachers, classrooms, and blackboards in marginal passers. The impact estimates at the cutoff were sizeable: our conservative estimates indicate that the program expanded marginal passers by roughly 85 students, and 3–4 teachers, classrooms, and blackboards. These impact estimates are also large relative to the mean baseline values for these outcomes for phase-4 SLQAT takers near the cutoff. What is more, they are particularly impressive given that phase-4 program schools were exposed to only about ten months of treatment before the follow-up telephone interview data were gathered. Finally, our cost-effectiveness estimates suggest that the program is among the cheapest interventions in developing countries for inducing enrollment gains.

Two potential threats to the internal validity of our impact estimates are however present. These threats arise from specific design and implementation features of the program. The first threat is program spillovers to non-program schools (of which marginal failers are but a specific subset) that operate in the same local schooling markets as program schools. Design features of the program such as the free-schooling condition for program benefit maintenance can alter the terms of local market competition, providing program schools with a competitive edge vis-à-vis non-program schools. The altered terms could result in impact estimates being upwardly biased if, for example, they induce the flow of students (and accompanying them, teachers) from non-program to program schools leading to the shrinking (or, at an extreme, the shutdown) of non-program schools and/or they discourage investments in the quantity and quality of inputs. On the

other hand, the altered terms could result in impact estimates being downwardly biased if they induce non-program schools to adapt to more effectively compete in the new local market environment created by the program such as by ratcheting up their investments in the quantity and quality of inputs and/or altering their fee structures to retain their existing student bodies and attract new students. Both types of effects could be present simultaneously; consequently, the direction of the net effect is theoretically ambiguous.

The second threat arises from anticipation of future treatment. Program entry is not a one-off event. To date, there have five calls for applications over a three-year period. Given this pattern, it is conceivable that non-program schools interested in joining the program might alter their behavior in anticipation of a future call for applications and in an effort to increase the likelihood of program entry. These actions could, for example, take the form of non-program schools investing in more and better quality inputs and resources. Anticipation in this case would result in impact estimates being downwardly biased. In particular, behavioral changes due to anticipation might be most applicable to non-program schools who just failed to achieve the SLQAT pass rate cutoff in an earlier phase of entry (our marginal failers), as we would expect that the marginal costs of investments and efforts required for failers from a previous phase to gain program entry are likely to be decreasing as one approaches the cutoff from below.

While when the RD model is applied to the right data designs, it yields internally-valid estimates, the generalizability of these impact estimates is likely to be limited as they are in principle only valid for narrowly-defined subpopulations. In the case of the phase-4 SLQAT taker data, which are fitted using a sharp design, the RD impact estimates are valid for low-cost private schools that successfully applied to the program, cleared the physical inspection, and obtained pass rates near the SLQAT cutoff. Even if we generalize our estimated treatment effects

to be average treatment effects over the full range of SLQAT pass rates, the further generalizability of the estimated effects is limited by sample selection, as the SLQAT pass rate is only available for low-cost private schools in the seven main program districts that learned about the call for applications, applied to the program, passed the inspection screening, and took the SLQAT. This group of schools may not be representative of the population of low-cost private schools in the program districts, leave alone the population of low-cost private schools more widely. Given what we know about the steps that precede taking the SLQAT, the extent to which the population of SLQAT takers diverges from the broader population of low-cost private schools in the program districts is largely determined by how demanding PEF's physical inspection screening is. Evidently, this screening has bite: as mentioned before, only 75% and 61% of schools that were inspected by PEF cleared it and took the SLQAT in phases 3 and 4, respectively.

References

- Andrabi, Tahir, Jishnu Das, Asim Ijaz Khwaja, Tara Vishwanath, Tristan Zajonc and LEAPS team. 2008a. Pakistan Learning and Educational Achievements in Punjab Schools (LEAPS): Insights to inform the education policy debate. World Bank Working Paper No. 43750.
- Andrabi, Tahir, Jishnu Das, Asim Ijaz Khwaja. 2008b. A dime a day: The possibilities and limits of private schooling in Pakistan. *Comparative Education Review* 52(3):329–355.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91(434):444–455.
- Battistin, Erich, and Enrico Rettore. 2008. Ineligible and eligible non-participants as a double comparison group in regression-discontinuity designs. *Journal of Econometrics* 142:715–730.
- Bloom, Howard S. 1984. Accounting for no-shows in experimental evaluation designs. *Evaluation Review* 8(2):225–246.
- Cheng, Ming-Yen, Jianqing Fan, and James S. Marron. 1997. On automatic boundary corrections. *Annals of Statistics* 25(4):1127–1160.
- Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. Using randomization in development economics research: A toolkit. In *Handbook of Development Economics* eds. T. Paul Schultz and John Strauss, Vol 4, UK: Elsevier B. V.:3895–3962.
- Evans, David, and Arkadipta Ghosh. 2008. Prioritizing educational investments in children in the developing world. RAND Working Paper 587.
- Gauri, Varun, and Ayesha Vawda. 2003. Vouchers for basic education in developing countries: A principal-agent perspective. Policy Research Working Paper No. 3005. Washington, D.C.: World Bank.
- Government of Punjab, Pakistan (Planning and Development Department). 2004. *Punjab District-based Multiple Indicators Cluster Survey 2003–04*. Lahore: Government of Punjab.
- 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.
- Imbens, Guido W., and Karthik Kalyanaraman. 2009. Optimal bandwidth choice for the regression discontinuity estimator. National Bureau of Economic Research Working Paper No. 14726.

- Imbens, Guido W., and Thomas Lemieux. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142(2):615–635.
- Imbens, Guido W., and Joshua D. Angrist. 1994. Identification and estimation of local average treatment effects. *Econometrica* 62:467–476.
- Kim, Jooseop, Harold Alderman, and Peter F. Orazem. 1999. Can private school subsidies increase schooling for the poor? The Quetta urban fellowship program. *World Bank Economic Review* 13(3):443–466.
- Ludwig, Jean, and Douglas L. Miller. 2005. Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics* 122(1):159–208.
- Malik, Allah Bakhsh. 2007. *Freedom of choice: Affordable quality education in public private partnership*. Lahore, Pakistan: Maqbool Academy Press.
- McCrary, Justin. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142(2):698–714.
- McCrary, Justin, and Heather Royer. 2006. The effect of female education on fertility and infant health: Evidence from school entry policies using exact date and birth. National Bureau of Economic Research Working Paper No. 12329.
- Nichols, Austin. 2007. `rd`: Stata module for regression discontinuity estimation. <http://ideas.repec.org/c/boc/bocode/s456888.html>. Last accessed: January 2009.
- Social Policy and Development Centre. 2003. *Social Development in Pakistan: Annual Review 2002-03*. Karachi: Social Policy Development Centre.
- Todd, Petra. 2007. Evaluating Social Programs with Endogenous Program Placement and Selection of the Treated. In: *Handbook of Development Economics*, eds. T. Paul Schultz and John Strauss, Vol 4., UK: Elsevier B.V.:3847–3894.
- United Nations Educational, Scientific, and Cultural Organization (UNESCO-UIS). 2007. *Global Education Digest 2007: Comparing Education Statistics Across the World*. Montreal, Canada: UNESCO Institute for Statistics.
- van der Klaauw, Wilbert. 2008. Breaking the link between poverty and low student achievement: An evaluation of Title I. *Journal of Econometrics* 142(2):731–756.
- World Bank. 2009. *The role and impact of public-private partnerships in education*. Washington, D.C.: World Bank.

Table 1. Selected education statistics for Pakistan and Punjab province, 1998/99 and 2004/05

| Sample | 1998/99 | | 2004/05 | | % change, 1998/99–2004/05 | |
|--|-----------------|---------------|-----------------|---------------|------------------------------|---------------|
| | Pakistan (1) | Punjab (2) | Pakistan (3) | Punjab (4) | Pakistan (5) | Punjab (6) |
| <i>Panel 1: School participation rate, ages 6–15 years</i> | | | | | | |
| All | 51.6 | 54.2 | 60.9 | 65.7 | 18.1 | 21.2 |
| Rural | 45.3 | 48.9 | 54.6 | 60.9 | 20.5 | 24.5 |
| Female | 43.0 | 48.1 | 53.7 | 61.2 | 24.7 | 27.3 |
| Poorest quintile | 30.2 | 32.6 | 45.0 | 48.7 | 49.1 | 49.4 |
| <i>Panel 2: Public school participation rate, ages 6–15 years</i> | | | | | | |
| All | 38.8 | 39.0 | 43.7 | 45.5 | 12.5 | 16.9 |
| Rural | 38.7 | 40.0 | 44.7 | 47.4 | 15.5 | 18.5 |
| Female | 31.6 | 33.7 | 37.6 | 42.0 | 18.7 | 24.8 |
| Poorest quintile | 27.2 | 28.5 | 39.3 | 41.1 | 44.5 | 44.2 |
| <i>Panel 3: Private school participation rate, ages 6–15 years</i> | | | | | | |
| All | 11.2 | 13.5 | 15.8 | 18.5 | 40.7 | 37.4 |
| Rural | 5.2 | 7.4 | 8.7 | 12.1 | 66.9 | 64.8 |
| Female | 10.2 | 13.0 | 14.9 | 17.8 | 45.2 | 36.9 |
| Poorest quintile | 2.0 | 2.9 | 4.5 | 6.4 | 119.2 | 120.2 |

Notes: Statistics for Pakistan actually represent statistics for the four provinces of Pakistan (unit data for the territories were not available). All statistics are corrected for sampling weights. The selected age group for measuring participation (ages 6–15 years) is one year removed from the official ages for primary and secondary schooling (ages 5–14 years) as this range is more consistent with when children typically attend school. The private and public school participation rates do not necessarily added up to the school participation rate as school participation also includes religious and NGO/community school participation, with are treated as separate school types.

Data sources: 1998/99 Pakistan Integrated Household Survey (PIHS); 2004/05 Pakistan Social and Living Standards Measurement Survey (PSLM). Fieldwork for the 2004/05 PSLM was conducted between September 2004 and March 2005, preceding the FAS program start date.

Table 2. FAS program timeline

| | |
|--|------------------------------------|
| Phase 1 application deadline | October 2005 |
| Phase 1 school agreements signed | November 2005–January 2006 (52/54) |
| Phase 1 first monthly subsidy payment | January 2006 |
| QAT 1 | March–April 2006 |
| Phase 2 application deadline | June 2006 |
| Phase 2 school agreements signed | September–December 2006 (143/150) |
| Phase 2 first monthly subsidy payment | October 2006 |
| QAT 2 | October–November 2006 |
| First payment of annual teacher bonus | January 2007 |
| QAT 3 | March–April 2007 |
| Phase 3 application deadline | April 2007 |
| Phase 3 school agreements signed | July–August 2007 (473/482) |
| Phase 3 first monthly subsidy payment | August 2007 |
| Phase 4 application deadline | July 2007 |
| QAT 4 | October–November 2007 |
| Phase 4 school agreements signed | November 2007 (424/424) |
| Phase 4 first monthly subsidy payment | December 2007 |
| Second payment of annual teacher bonus | February 2008 |
| First payment of annual school bonus | February 2008 |
| QAT 5 | March–April 2008 |
| Phase 5 application deadline | April 2008 |

Notes: The statistics in parentheses represent the number of schools that signed agreements in those months relative to the number of schools that entered the program in that phase.

Table 3. District- and phase-wise frequency distribution of FAS program schools

| District | Adult literacy rate (10+ years) | Adult literacy rate rank (Low number = high rank) | Phase | | | | Total |
|----------------------------------|---------------------------------|--|-------|-----|-----|-----|-------|
| | | | 1 | 2 | 3 | 4 | |
| Rawalpindi | 78 | 1 | | | | | |
| Lahore | 70 ^a | 2 | 7 | 11 | | | 18 |
| Sialkot | 70 | 3 | 7 | 9 | | | 16 |
| Chakwal | 69 | 4 | 9 | 17 | | | 26 |
| Jhelum | 68 | 5 | | | | | |
| Gujranwala | 67 | 6 | | | | | |
| Gujrat | 65 | 7 | | 15 | | | 15 |
| Faisalabad | 60 | 8 | | | | | |
| Narowal | 60 | 9 | | 6 | | | 6 |
| Sargodha | 58 | 10 | | | | | |
| Toba Tek Singh | 58 | 11 | | | | | |
| Attock | 57 | 12 | | | | | |
| Mandi Bahauddin | 57 | 13 | | | | | |
| Mianwali | 56 | 14 | | 12 | | | 12 |
| Hafizabad | 55 | 15 | | | | | |
| Sahiwal | 54 | 16 | | | | | |
| Khushab | 52 | 17 | 10 | 19 | | | 29 |
| Sheikhupura | 50 | 18 | 2 | | | | 2 |
| Khanewal | 49 | 19 | | | | | |
| Jhang | 47 | 20 | | | 44 | 48 | 92 |
| Layyah | 46 | 21 | | | | | |
| Multan | 46 | 22 | | | 58 | 80 | 138 |
| Bahawalnagar | 44 | 23 | | 14 | 115 | 43 | 172 |
| Okara | 43 | 24 | | | | | |
| Vehari | 43 | 25 | | 1 | | | 1 |
| Kasur | 42 | 26 | | | | | |
| Pakpattan | 42 | 27 | | | | | |
| Rahim Yar Khan | 42 | 28 | | | | | |
| D. G. Khan | 40 | 29 | | | | | |
| Bahawalpur | 37 | 30 | 10 | 16 | 114 | 88 | 228 |
| Bhakkar | 37 | 31 | | 11 | | | 11 |
| Lodhran | 37 | 32 | | | 42 | 49 | 91 |
| Muzaffargarh | 36 | 33 | | 1 | 72 | 81 | 154 |
| Rajanpur | 34 | 34 | | | 35 | 35 | 70 |
| Nankana Sahib^b | -- | -- | | 1 | | | 1 |
| | | Number of districts | 6 | 13 | 7 | 7 | 18 |
| | | Number of program schools (as of June 2008) | 45 | 133 | 480 | 424 | 1082 |

High presence
 Low presence
Bold Program district

Notes: District-level adult literacy rates obtained from Government of Punjab (2004) based on data from the 2003-04 Multiple Indicators Cluster Survey (MICS). ^bThe literacy rate for Nankana Sahib was unavailable. ^aThe statistic for Lahore district is a simple mean of the literacy rates for towns and cantonments in Lahore.

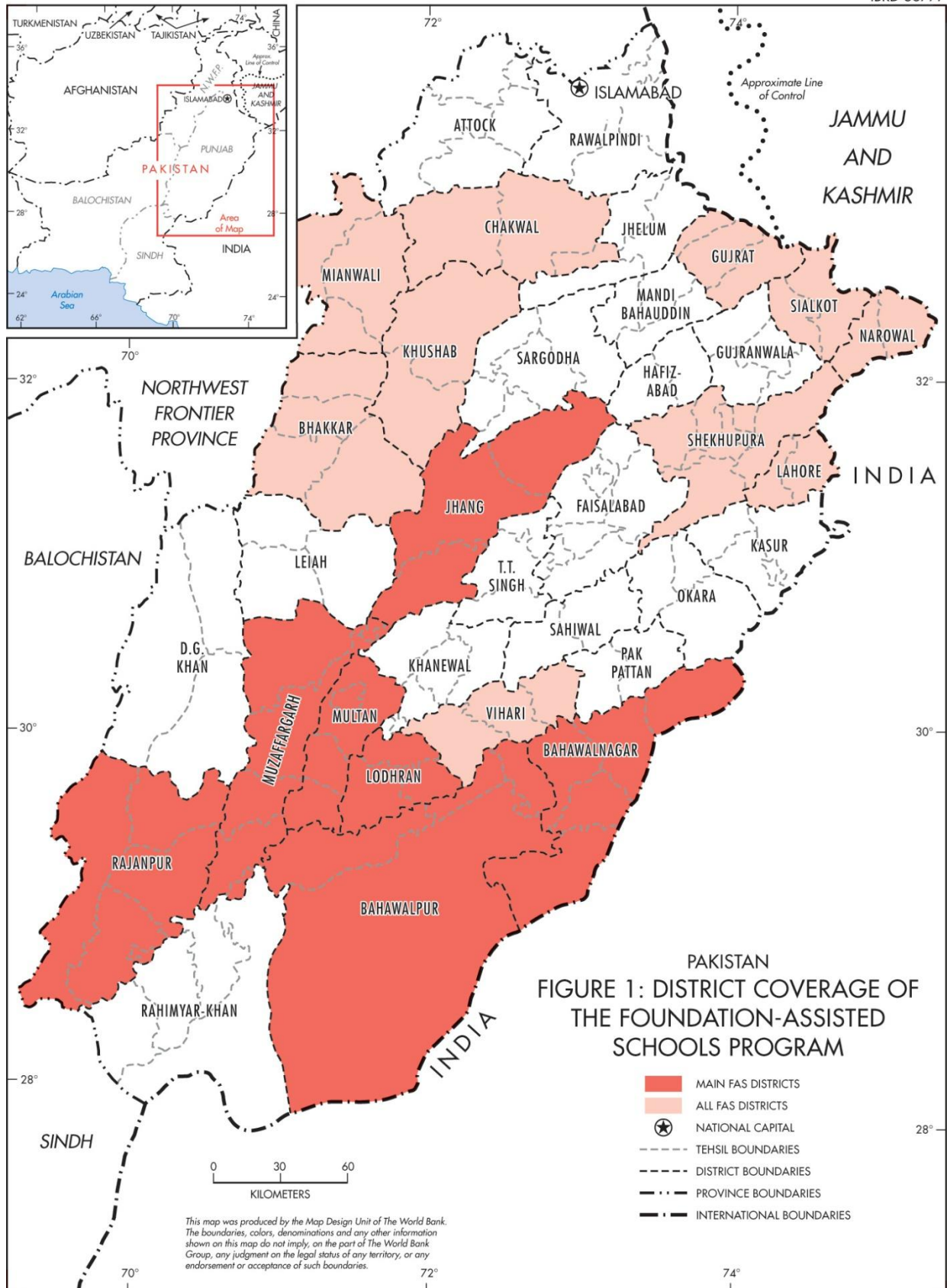


Table 4. FAS program coverage of private schools in selected program districts

| Program district | (1) Number of program schools | (2) Number of primary, middle and secondary private schools | (3) Program share |
|------------------|----------------------------------|--|----------------------|
| Bahawalnagar | 172 | 447 | 0.38 |
| Bahawalpur | 230 | 871 | 0.26 |
| Jhang | 93 | 686 | 0.14 |
| Lodhran | 93 | 284 | 0.33 |
| Multan | 139 | 1,411 | 0.10 |
| Muzaffargarh | 157 | 690 | 0.23 |
| Rajanpur | 71 | 226 | 0.31 |
| Total | 955 | 4,615 | 0.21 |

Notes: Numbers of private schools provided by Corinne Siaens using the 2005 National Education Census (NEC) data. Column (2) reports numbers of relevant schools strictly classified as private in the NEC.

Table 5. Mean characteristics of FAS program schools

| Characteristic | (1) Phase 1 | (2) Phase 2 | (3) Phase 3 | (4) Phase 4 | (5) All phases |
|----------------------------|----------------|----------------|----------------|----------------|-------------------|
| Total number of students | 561.40 | 547.42 | 373.83 | 241.66 | 351.18 |
| <i>Level</i> | | | | | |
| Primary | 0.02 | 0.05 | 0.05 | 0.11 | 0.07 |
| Middle | 0.24 | 0.31 | 0.60 | 0.69 | 0.59 |
| Secondary | 0.73 | 0.65 | 0.35 | 0.20 | 0.34 |
| <i>Gender type</i> | | | | | |
| Coeducational | 0.69 | 0.86 | 0.83 | 0.82 | 0.83 |
| Girls-only | 0.20 | 0.11 | 0.09 | 0.11 | 0.11 |
| Boys-only | 0.11 | 0.03 | 0.07 | 0.07 | 0.07 |
| <i>Registration status</i> | | | | | |
| Registered | 0.91 | 0.97 | 0.89 | 0.81 | 0.87 |
| Unregistered | 0.09 | 0.03 | 0.11 | 0.19 | 0.13 |
| <i>Location</i> | | | | | |
| Urban | 0.36 | 0.45 | 0.48 | 0.42 | 0.45 |
| Rural | 0.64 | 0.55 | 0.52 | 0.58 | 0.55 |
| <i>N</i> | 45 | 133 | 480 | 424 | 1,082 |

Notes: The statistics exclude the three higher secondary schools that are program schools. The statistics are constructed from administrative data from September 2008, one month before the follow-up data collection.

| | Met application requirements | Passed physical inspection | Passed SLQAT | Joined FAS program |
|----------|------------------------------|----------------------------|--------------|--------------------|
| Phase 3: | 1,070 | 799 | 514 | 482 |
| Phase 4: | 1,430 | 872 | 431 | 425 |

Figure 2. Numbers of schools which cleared the different FAS program entry steps

Table 6. FAS program school participation status, by phase

| Phase | (1) Entrants | (2) Disqualified, all reasons | (3) Disqualified, double QAT failure | (4) Current participation |
|-------|-----------------|-------------------------------------|--|---------------------------------|
| 1 | 54 | 9 | 7 | 45 |
| 2 | 150 | 16 | 13 | 133 |
| 3 | 482 | 2 | 0 | 480 |
| 4 | 425 | 1 | 0 | 424 |
| Total | 1,111 | 28 | 20 | 1,082 |

Notes: Disqualification also includes voluntary exits. Statistics reflect program school participation status at the time of the follow-up data collection in October 2008.

Table 7. Correlates of nonresponse in follow-up telephone interview data

| | Phase-3 cutoff neighborhood sample | | | Phase-4 cutoff neighborhood sample | | |
|---|------------------------------------|---------------------|-----------------|------------------------------------|---------------------|-------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Marginal failers | Marginal passers | All | Marginal failers | Marginal passers | All |
| <i>Panel A: Nonresponse counts and rates</i> | | | | | | |
| Number of nonresponses | 29 | 47 | 76 | 26 | 45 | 71 |
| Nonresponse rate | 0.27 | 0.29 | 0.28 | 0.20 | 0.24 | 0.22 |
| <i>Panel B: Bivariate correlation estimates</i> | | | | | | |
| Above cutoff | -- | -- | 0.03 (0.06) | -- | -- | 0.04 (0.05) |
| <i>Baseline characteristics</i> | | | | | | |
| Registered | -0.14 (0.16) | 0.01 (0.11) | -0.04 (0.09) | -0.06 (0.12) | 0.00 (0.07) | -0.01 (0.06) |
| Coeducational | 0.06 (0.12) | -0.03 (0.11) | 0.01 (0.08) | -0.03 (0.14) | -0.10 (0.08) | -0.06 (0.07) |
| Girls-only | -0.10 (0.13) | 0.19 (0.13) | 0.07 (0.10) | 0.03 (0.20) | 0.05 (0.10) | 0.02 (0.09) |
| Boys-only | 0.07 (0.27) | -0.31*** (0.04) | -0.19 (0.10) | 0.03 (0.20) | 0.15 (0.14) | 0.10 (0.11) |
| Primary | -0.02 (0.16) | 0.27* (0.16) | 0.15 (0.12) | 0.26* (0.15) | 0.10 (0.10) | 0.15* (0.08) |
| Middle | -0.12 (0.10) | 0.00 (0.08) | -0.05 (0.06) | -0.21** (0.10) | -0.10 (0.07) | -0.14** (0.06) |
| Secondary | 0.16 (0.11) | -0.10 (0.08) | 0.00 (0.07) | 0.12 (0.11) | 0.07 (0.08) | 0.09 (0.07) |
| Urban | 0.01 (0.09) | -0.12* (0.07) | -0.07 (0.05) | 0.07 (0.08) | -0.05 (0.06) | 0.00 (0.05) |
| Rural | -0.01 (0.09) | 0.12* (0.07) | 0.07 (0.05) | -0.07 (0.08) | 0.05 (0.06) | 0.00 (0.05) |
| Bahawalnagar | -0.01 (0.11) | -0.07 (0.08) | -0.05 (0.06) | 0.09 (0.14) | -0.08 (0.08) | -0.01 (0.08) |
| Bahawalpur | -0.12 (0.09) | 0.01 (0.08) | -0.05 (0.06) | 0.18* (0.11) | -0.02 (0.07) | 0.06 (0.06) |
| Jhang | -0.21* (0.13) | 0.05 (0.11) | 0.13 (0.09) | 0.06 (0.14) | 0.00 (0.08) | 0.01 (0.07) |
| Lodhran | 0.07 (0.20) | -0.17* (0.09) | -0.09 (0.09) | 0.00 (0.14) | -0.14** (0.07) | -0.07 (0.07) |
| Multan | -0.14 (0.11) | 0.11 (0.12) | 0.00 (0.08) | -0.18** (0.09) | 0.02 (0.07) | -0.06 (0.06) |
| Muzaffargarh | 0.03 (0.12) | 0.11 (0.11) | 0.08 (0.08) | -0.18 (0.09)** | 0.05 (0.08) | -0.05 (0.06) |
| Rajanpur | 0.07 (0.20) | -0.13 (0.11) | -0.05 (0.11) | 0.16 (0.16) | 0.14 (0.11) | 0.14 (0.09) |
| <i>N</i> | 108 | 160 | 268 | 128 | 191 | 319 |

Notes: * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level. Heteroskedasticity-robust standard errors are reported in parentheses.

Table 8. Distribution of schools by selected characteristics at baseline, SLQAT and cutoff neighborhood samples

| Characteristic | SLQAT sample $z \in [0,100]$ | | | Cutoff neighborhood sample $z \in [52,82]$ | | |
|----------------------------|---------------------------------|-------------------------|-------------------------|---|-------------------------|-------------------------|
| | (1) Phase 3 Share | (2) Phase 4 Share | (3) Diff. (1)–(2) | (4) Phase 3 Share | (5) Phase 4 Share | (6) Diff. (4)–(5) |
| <i>Level</i> | | | | | | |
| Primary | 0.09 | 0.12 | –0.02 | 0.07 | 0.12 | –0.05 |
| Middle | 0.63 | 0.72 | –0.09*** | 0.71 | 0.70 | 0.01 |
| Secondary | 0.28 | 0.16 | 0.12*** | 0.22 | 0.18 | 0.04 |
| <i>Gender type</i> | | | | | | |
| Coeducational | 0.82 | 0.87 | –0.05*** | 0.87 | 0.86 | 0.01 |
| Girls-only | 0.12 | 0.08 | 0.03** | 0.10 | 0.09 | 0.01 |
| Boys-only | 0.07 | 0.05 | 0.02 | 0.04 | 0.06 | –0.02 |
| <i>Registration status</i> | | | | | | |
| Registered | 0.88 | 0.81 | 0.07*** | 0.88 | 0.83 | 0.05* |
| Unregistered | 0.12 | 0.19 | –0.07*** | 0.12 | 0.17 | –0.05* |
| <i>Location type</i> | | | | | | |
| Urban | 0.45 | 0.41 | 0.04* | 0.47 | 0.41 | 0.06 |
| Rural | 0.54 | 0.59 | –0.04* | 0.53 | 0.59 | –0.06 |
| <i>District</i> | | | | | | |
| Bahawalnagar | 0.21 | 0.11 | 0.10*** | 0.20 | 0.09 | 0.11*** |
| Bahawalpur | 0.24 | 0.20 | 0.04** | 0.25 | 0.21 | 0.04 |
| Jhang | 0.13 | 0.11 | 0.02 | 0.14 | 0.14 | 0.00 |
| Lodhran | 0.08 | 0.12 | –0.04*** | 0.07 | 0.09 | –0.02 |
| Multan | 0.13 | 0.16 | –0.03* | 0.13 | 0.20 | –0.07** |
| Muzaffargarh | 0.14 | 0.20 | –0.05*** | 0.14 | 0.17 | –0.03 |
| Rajanpur | 0.07 | 0.10 | –0.03* | 0.06 | 0.09 | –0.03 |
| <i>N</i> | 747 | 830 | -- | 268 | 319 | -- |

Notes: * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

Table 9. Summary statistics of outcome measures at baseline, SLQAT and cutoff neighborhood samples

| Outcome measure | SLQAT sample $z \in [0,100]$ | | | Cutoff neighborhood sample $z \in [52,82]$ | | |
|-------------------------|---------------------------------|--------------------------------|--------------------------------------|---|--------------------------------|--------------------------------------|
| | (1) Phase 3 Mean (SD) | (2) Phase 4 Mean (SD) | (3) (2)–(1) Difference (SE) | (4) Phase 3 Mean (SD) | (5) Phase 4 Mean (SD) | (6) (5)–(4) Difference (SE) |
| Number of students | 252.85 (155.70) | 222.63 (106.68) | –30.22*** (6.79) | 240.82 (134.67) | 232.27 (108.05) | –8.56 (10.21) |
| Number of teachers | 10.22 (5.03) | 9.00 (3.69) | –1.22*** (0.22) | 9.82 (4.30) | 9.32 (3.77) | –0.50 (0.34) |
| Number of classrooms | 9.56 (4.71) | 8.43 (3.82) | –1.13*** (0.22) | 8.90 (3.78) | 8.59 (3.76) | –0.31 (0.31) |
| Number of blackboards | 9.93 (5.03) | 9.03 (4.00) | –0.90*** (0.23) | 9.39 (4.11) | 9.27 (3.93) | –0.11 (0.33) |
| Number of toilets | 3.23 (1.80) | 2.95 (1.94) | –0.28*** (0.09) | 3.09 (1.59) | 2.97 (1.75) | –0.12 (0.14) |
| Student-teacher ratio | 25.06 (9.27) | 25.67 (9.51) | 0.60 (0.47) | 24.98 (9.59) | 25.59 (8.26) | 0.60 (0.75) |
| Student-classroom ratio | 27.51 (12.63) | 28.13 (12.13) | 0.62 (0.63) | 28.30 (13.99) | 28.35 (11.47) | 0.05 (1.07) |

Notes: * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level. SD denotes standard deviation; SE standard error; and z the treatment assignment variable, the SLQAT pass rate.

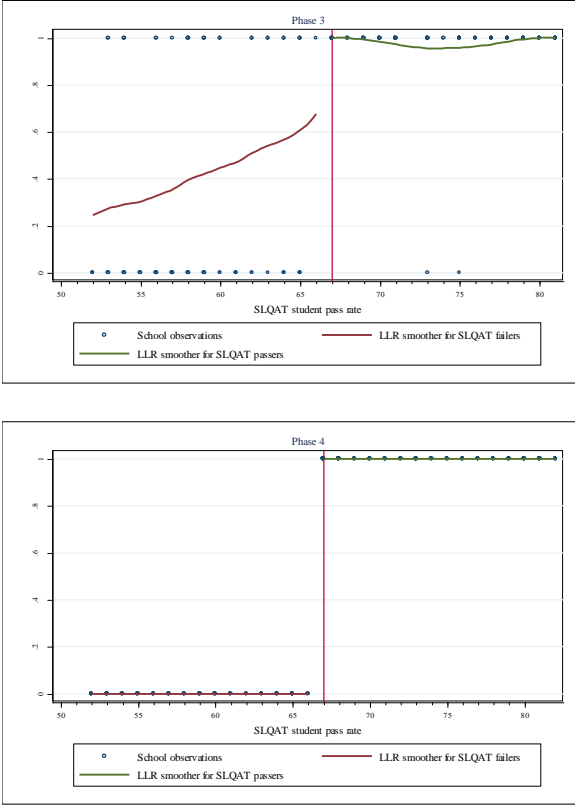


Figure 3. Probability of treatment, cutoff neighborhood sample, by phase

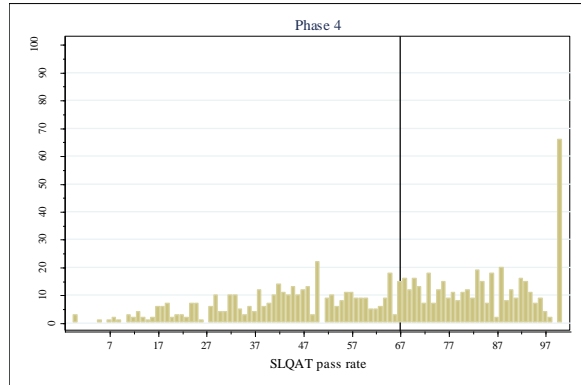
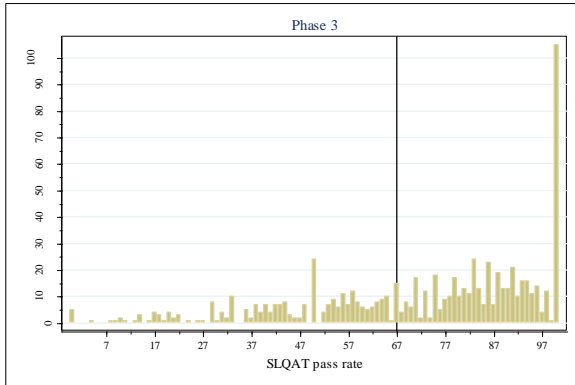


Figure 4. Frequency histogram of SLQAT pass rates, SLQAT sample, by phase

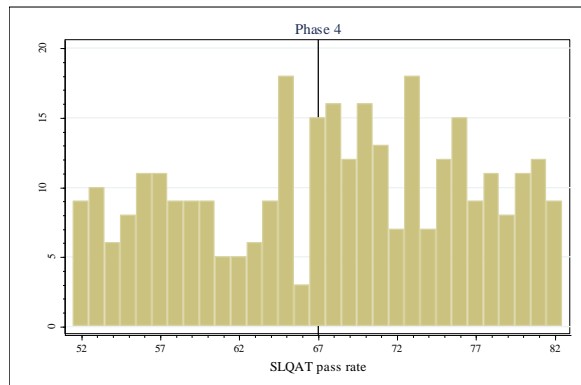
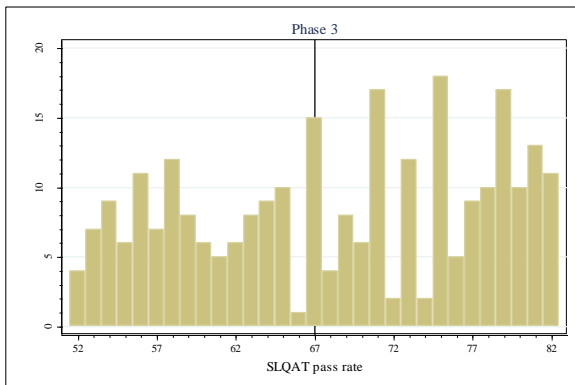


Figure 5. Frequency histogram of SLQAT pass rates, cutoff neighborhood sample, by phase

Table 10. Differences in mean outcomes at baseline between SLQAT marginal passers and marginal failers

| Outcome measure | Phase-3 cutoff neighborhood sample | | | | Phase-4 cutoff neighborhood sample | | | |
|-------------------------|--|---|--|---|--|---|--|---|
| | $z \in [52, 82]$ | | $z \in [62, 72]$ | | $z \in [52, 82]$ | | $z \in [62, 72]$ | |
| | (1) Marginal failers Mean (SD) | (2) Marginal passers Diff. (SE) | (3) Marginal failers Mean (SD) | (4) Marginal passers Diff. (SE) | (5) Marginal failers Mean (SD) | (6) Marginal passers Diff. (SE) | (7) Marginal failers Mean (SD) | (8) Marginal passers Diff. (SE) |
| Number of students | 242.32 (138.80) | -2.52 (16.78) | 268.50 (188.03) | -48.17 (34.82) | 229.68 (99.26) | 4.33 (12.36) | 254.83 (114.57) | -28.01 (20.60) |
| Number of teachers | 9.91 (4.22) | -0.15 (0.54) | 9.82 (5.41) | -0.79 (1.08) | 9.23 (3.36) | 0.14 (0.43) | 9.80 (4.03) | -0.63 (0.71) |
| Number of classrooms | 8.88 (3.83) | 0.04 (0.47) | 8.41 (4.10) | 0.01 (0.80) | 8.53 (3.83) | 0.10 (0.43) | 9.61 (4.12) | -1.22* (0.68) |
| Number of blackboards | 9.29 (4.10) | 0.16 (0.51) | 8.74 (4.59) | -0.02 (0.93) | 9.17 (3.72) | 0.18 (0.45) | 10.22 (4.34) | -1.16 (0.73) |
| Number of toilets | 3.08 (1.66) | 0.01 (0.20) | 2.74 (1.83) | 0.23 (0.36) | 2.98 (1.74) | -0.03 (0.20) | 3.37 (2.44) | -0.32 (0.43) |
| Student-teacher ratio | 24.67 (9.28) | 0.53 (1.19) | 25.95 (8.10) | -0.53 (2.33) | 25.48 (7.90) | 0.18 (0.94) | 26.57 (7.32) | -1.10 (1.52) |
| Student-classroom ratio | 28.69 (13.58) | -0.66 (1.74) | 31.97 (15.61) | -4.87 (3.26) | 28.18 (10.69) | 0.28 (1.32) | 27.44 (8.77) | 1.55 (2.32) |

Notes: * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level. z denotes the treatment assignment variable, the SLQAT pass rate; SD standard deviation; and SE standard error.

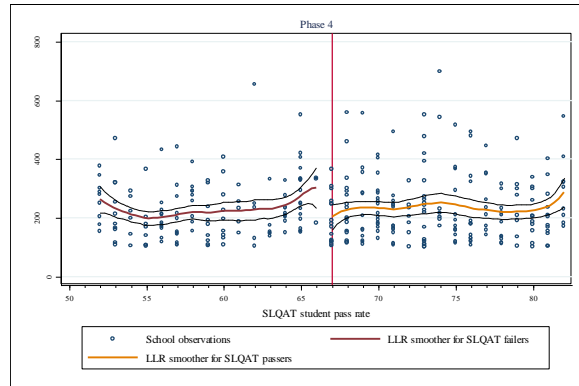
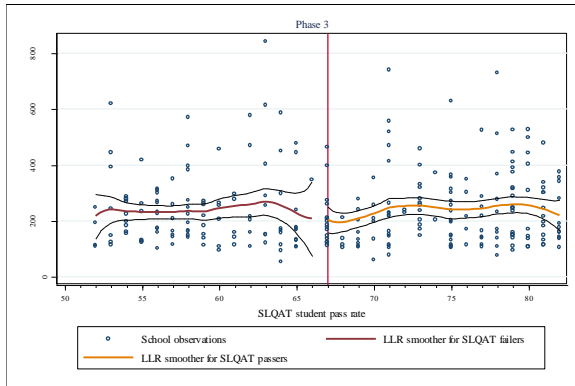
Table 11. Local smoothness in conditional mean outcomes at baseline

Local linear regression with triangular kernel and bandwidth h

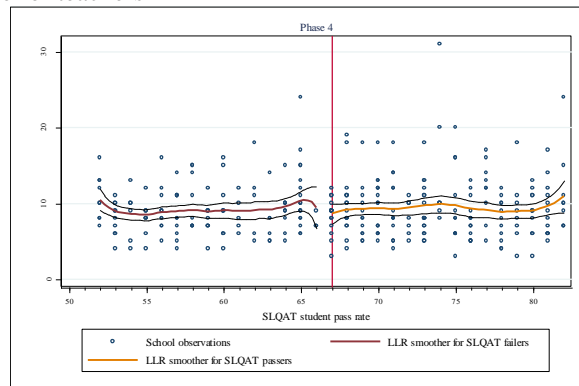
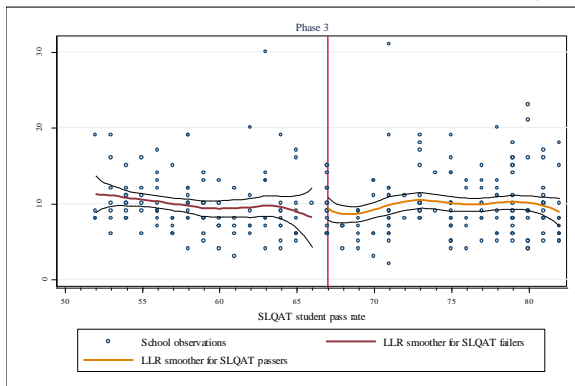
| Outcome measure | Phase-3 cutoff neighborhood sample | | | Phase-4 cutoff neighborhood sample | | |
|-------------------------|------------------------------------|----------------------|----------------------|------------------------------------|------------------------|----------------------|
| | (1) $h = 4\%$ pts | (2) $h = 6\%$ pts | (3) $h = 8\%$ pts | (4) $h = 3\%$ pts | (5) $h = 4.5\%$ pts | (6) $h = 6\%$ pts |
| Number of students | -97.05 (140.67) | 14.32 (88.06) | -39.12 (60.96) | -78.43 (92.74) | -132.88** (65.79) | -107.76** (51.50) |
| Number of teachers | -0.72 (3.56) | 1.35 (2.57) | -0.59 (1.75) | 4.37* (2.51) | -1.41 (3.26) | -1.78 (2.08) |
| Number of classrooms | 0.31 (3.24) | 1.78 (2.23) | 0.61 (1.57) | 2.43 (2.04) | -2.17 (2.72) | -2.44 (1.77) |
| Number of blackboards | -0.74 (3.24) | 1.71 (2.13) | 0.69 (1.56) | 0.67 (1.73) | -2.98 (2.30) | -2.95* (1.73) |
| Number of toilets | 1.36 (1.24) | 1.45* (0.74) | 0.93 (0.61) | -1.32 (1.65) | -2.13 (1.56) | -2.19* (1.25) |
| Student-teacher ratio | -8.53 (8.52) | -1.37 (4.15) | -0.15 (3.39) | -22.37* (11.98) | -11.91 (8.20) | -8.70 (5.46) |
| Student-classroom ratio | -14.37 (12.67) | -3.80 (7.14) | -4.19 (5.69) | -12.26 (9.49) | -5.73 (7.28) | -3.12 (5.85) |

Notes: * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level. Bootstrapped standard errors based on 500 replications are reported in parentheses.

1. Number of students



2. Number of teachers



3. Number of classrooms

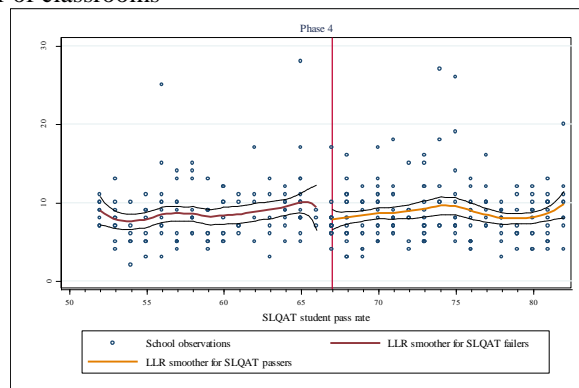
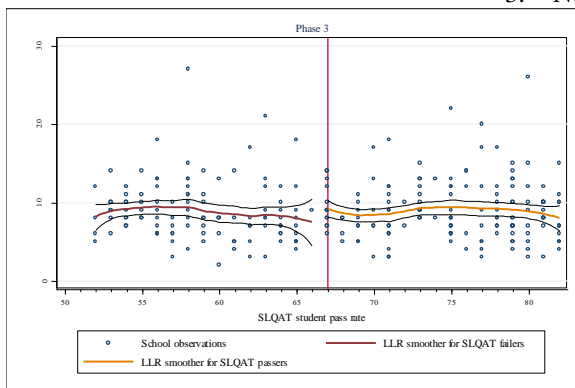
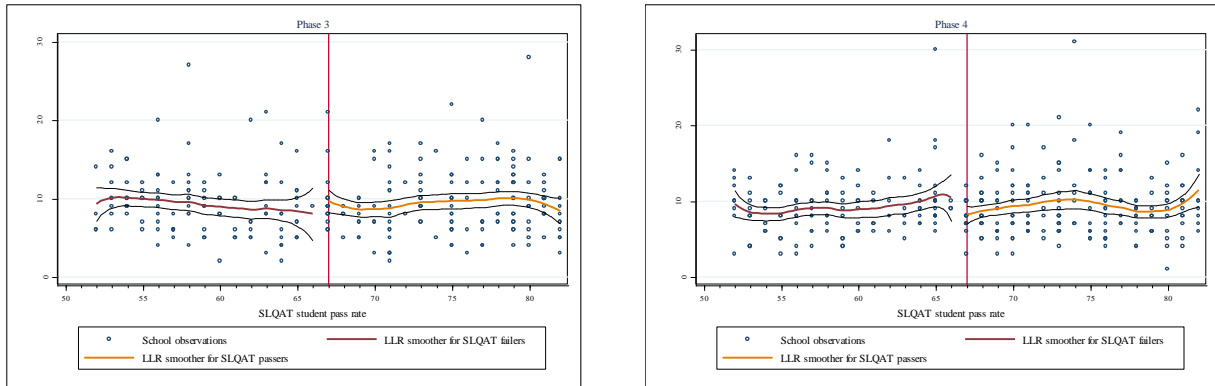
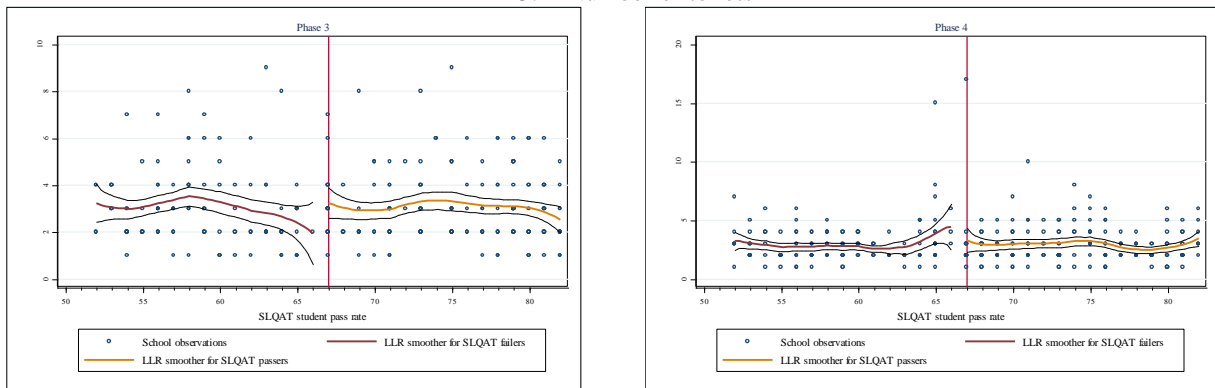


Figure 6a. Local smoothness in conditional mean outcomes at baseline, by phase

4. Number of blackboards



5. Number of toilets



6. Student-teacher ratio

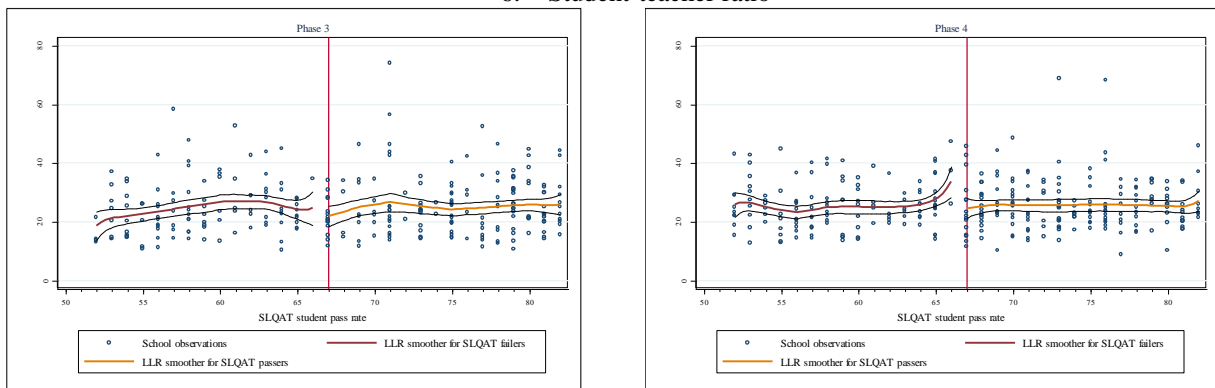


Figure 6b. Local smoothness in conditional mean outcomes at baseline, by phase

7. Student-classroom ratio

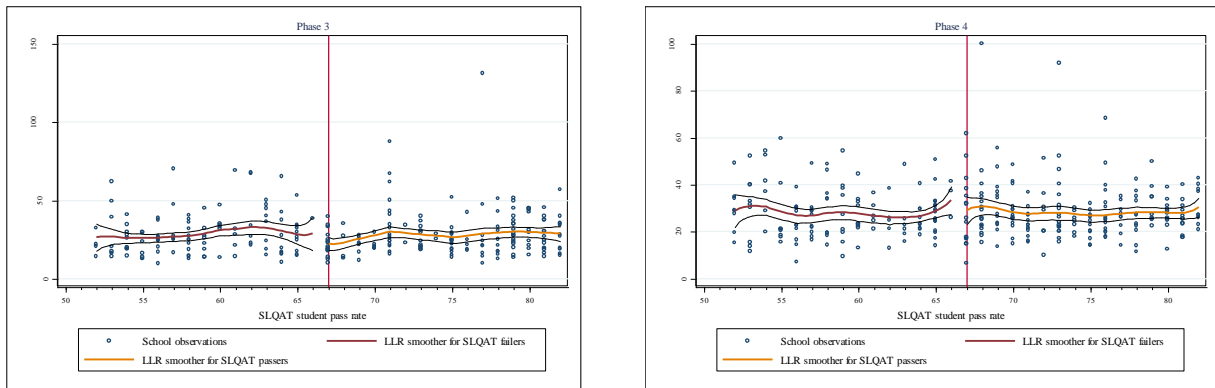


Figure 6c. Local smoothness in conditional mean outcomes at baseline, by phase

Table 12. Discontinuity estimates of conditional mean outcomes
Local linear regression with triangular kernel and bandwidth h

| Outcome measure | Phase-3 cutoff neighborhood sample <i>Partially-fuzzy RD estimates</i> | | | Phase-4 cutoff neighborhood sample <i>Sharp RD estimates</i> | | |
|-------------------------|---|----------------------|---------------------|---|--------------------|--------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | $h = 6\%$ pts | $h = 9\%$ pts | $h = 12\%$ pts | $h = 3\%$ pts | $h = 4.5\%$ pts | $h = 6\%$ pts |
| Number of students | 180.3 (8.09e+14) | -29.03 (18002.95) | -28.03 (6230.31) | 122.66*** (45.75) | 85.20** (42.77) | 88.84** (41.32) |
| Number of teachers | 6.46 (3.96e+13) | 3.34 (331.89) | 3.09 (42.19) | 5.07** (2.35) | 3.39* (1.92) | 3.46* (1.82) |
| Number of classrooms | 2.63 (1.77e+13) | -0.75 (121.08) | 1.43 (95.78) | 9.85** (3.89) | 4.55* (2.38) | 4.00** (2.02) |
| Number of blackboards | -- | -- | -- | 6.56** (2.71) | 3.14* (1.85) | 2.83* (1.72) |
| Number of toilets | -- | -- | -- | 0.07 (1.28) | -0.37 (0.95) | -0.14 (0.85) |
| Student-teacher ratio | 0.39 (97.98) | -6.17 (86.51) | -5.07 (114.35) | 1.23 (3.97) | -0.85 (2.77) | -1.04 (2.54) |
| Student-classroom ratio | 11.48 (88.45) | 1.45 (131.10) | -4.77 (59.86) | -29.18 (22.29) | -11.39 (9.89) | -8.12 (7.24) |

Notes: * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level. Bootstrapped standard errors based on 500 replications are reported in parentheses.

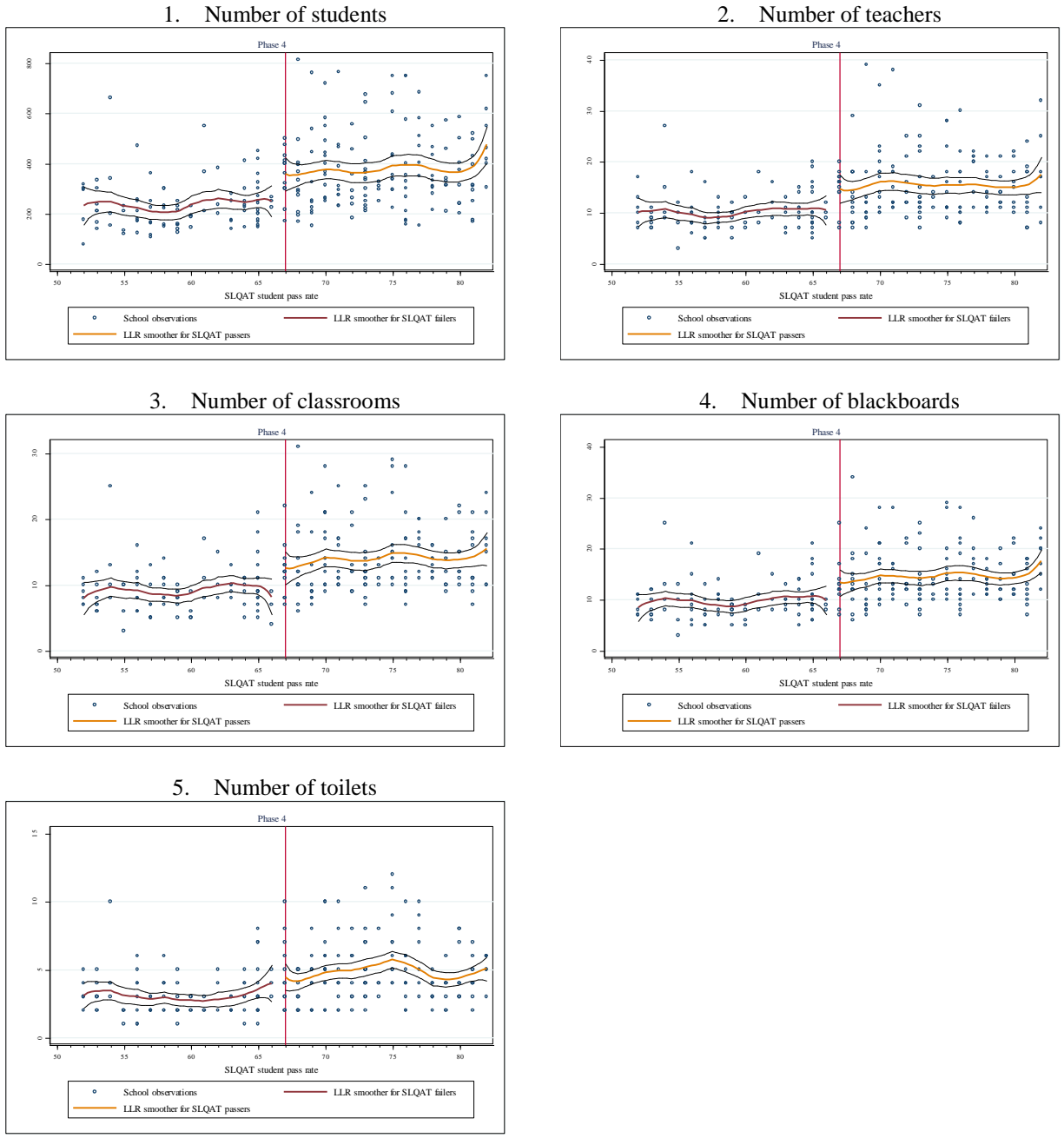
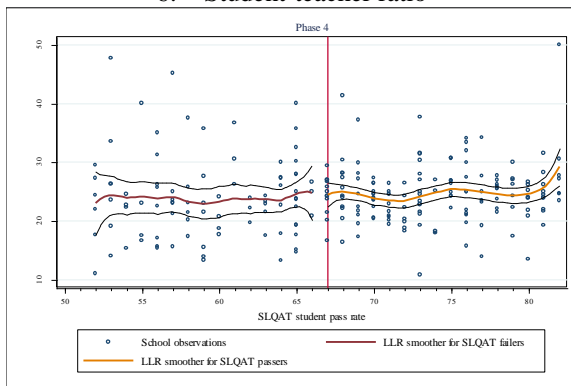


Figure 7a. Local discontinuities in conditional mean outcomes, phase 4

6. Student-teacher ratio



7. Student-classroom ratio

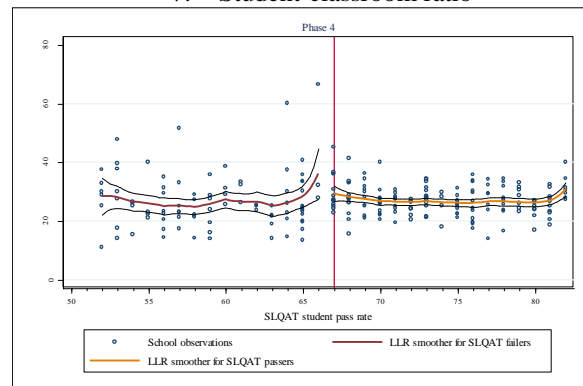


Figure 7b. Local discontinuities in conditional mean outcomes, phase 4

Table 13. Discontinuity estimates of conditional mean outcomes at false cutoffs, phase-4 cutoff neighborhood sample
Local linear regression with triangular kernel and bandwidth h

| Outcome measure | Cutoff = 57% | | | Cutoff = 77% | | |
|-------------------------|----------------------|-------------------------|-----------------------|----------------------|----------------------|----------------------|
| | (1) $h = 7\%$ pts | (2) $h = 10.5\%$ pts | (3) $h = 14\%$ pts | (4) $h = 4\%$ pts | (5) $h = 6\%$ pts | (6) $h = 8\%$ pts |
| Number of students | -21.38 (48.84) | -18.77 (47.11) | -18.59 (47.28) | 57.99 (122.09) | 8.42 (85.87) | 9.86 (72.00) |
| Number of teachers | -1.21 (1.79) | -1.02 (1.65) | -0.99 (1.65) | 3.80 (3.70) | 2.11 (2.69) | 2.00 (2.35) |
| Number of classrooms | -1.74 (1.84) | -1.69 (1.61) | -1.67 (1.58) | 2.39 (3.57) | -0.30 (2.58) | -0.05 (2.10) |
| Number of blackboards | 0.95 (3.70) | 0.29 (3.52) | 0.14 (3.50) | -1.81 (3.23) | -2.64 (2.23) | -2.85 (1.92) |
| Number of toilets | 2.91 (4.46) | 2.38 (4.08) | 2.12 (4.00) | 0.56 (3.73) | 1.38 (2.65) | 0.78 (2.38) |
| Student-teacher ratio | -2.46 (2.28) | -2.49 (2.13) | -2.50 (2.11) | 2.21 (4.10) | 0.06 (2.78) | 0.21 (2.30) |
| Student-classroom ratio | -0.21 (0.79) | -0.32 (0.76) | -0.36 (0.76) | 1.51 (1.76) | -0.40 (1.40) | -0.64 (1.14) |

Notes: * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level. Bootstrapped standard errors based on 500 replications are reported in parentheses. The subsample for the investigation at 57% is schools with pass rates between 52% and 66%. Similarly, the subsample for the investigation at 77% is schools with pass rates between 67% and 82%.