How price sensitive is secondary school enrollment? Evidence from nationwide tuition fee reforms in South Africa

Robert Garlick†

October 21, 2012

PRELIMINARY AND INCOMPLETE. PLEASE DO NOT CITE WITHOUT CONSULTING THE AUTHOR.

Abstract

This paper advances the literature on education demand in developing countries by studying a policy of nationwide school fee eliminations in South Africa. I use a regression discontinuity design that compares schools just above a threshold poverty level, most of which eliminate fees, to schools just below the threshold, most of which continue to charge fees. I show that fee elimination has a relatively small effect, increasing enrollment by approximately 10 students, or 3% of baseline enrollment. These results are considerably smaller than those reported elsewhere in the literature and imply a relatively inelastic demand for enrollment in this setting. The effects are substantially larger in urban than rural areas and larger in secondary schools than in primary schools. I argue that this at least partially reflects more elastic demand for enrollment at the secondary level.

JEL codes: I25, O15

*I thank Manuela Angelucci, John Bound, John DiNardo, Brian Jacob, David Lam, Jeffrey Smith; seminar participants at the University of Michigan; and conference participants at ESSA and MEA for helpful suggestions. Justice Libago, Christo Lombaard, Erna Lubbe, Hersheela Narsee, Ralph Mehl, Siza Shongwe, and Hylton Visagie from South Africa’s Department of Basic Education provided invaluable assistance in obtaining the data used in this project. All errors are my own.

†Ph.D. candidate in Public Policy and Economics, University of Michigan; rgarlick@umich.edu.
1 Introduction

Increasing school enrollment is an important goal for policy makers in many developing and developed countries. This reflects both widespread political claims that education has desirable economic, political and social consequences and a large and growing body of academic research that supports this argument. Labor economists have amassed substantial evidence of a positive causal effect of educational attainment on individual earnings (Card, 1999; Heckman, Lochner, and Todd, 2006) and some research suggests that these pecuniary returns to education may be higher in developing than developed countries (Lam, Ardington, Branson, Goostrey, and Leibbrandt, 2010). A complementary literature in macroeconomics shows a positive association at the country level between education and economic growth (Hanushek and Kimko, 2000; Mankiw, Romer, and Weil, 1992), though it is not clear whether a causal interpretation should be assigned to these associations.

Given the many potential benefits of improving education at the individual and country level, policy-makers have experimented with several interventions aimed at increasing education participation in developing countries. In particular, more than ten African countries have eliminated tuition fees in all or selected primary schools during the past decade. These interventions have affected several million students and involve a large redistribution to households with enrolled children, who are no longer required to pay tuition fees, from the general fiscus, which must increase government grants to schools that are no longer permitted to charge fees. Despite the very large implications of such policies, there is relatively little credible evidence regarding the causal effect of eliminating school fees on enrollment and other student outcomes. Most existing studies of fee eliminations in developing countries have compared enrollment before and after fees were eliminated for the entire country (Deininger, 2003) or time trends in enrollment across districts that eliminated fees at different points in time (Fafchamps and Minten, 2007). Such comparisons may, however, generate biased estimates of the treatment effect of school fee elimination if other policy changes occur at the same time as the fee elimination or if districts that abolish fees at different times differ in other characteristics.

I contribute to this literature by providing some of the first clean evidence on the causal effects of school fee eliminations. I study a natural experiment in South Africa, in which schools were required to eliminate fees if and only if they were located in neighborhoods whose poverty rate surpassed a cutoff value. The design of this policy ensures that schools just above the cutoff are very similar to schools just below the cutoff in all respects other than their fee policy. Any differences in enrollment and other student outcomes between the two groups must then reflect the causal effect of the fee policies. This regression discontinuity design recovers the causal effect of eliminating fees under much weaker assumptions than those made by prior studies. I also provide
some of the first evidence regarding the effect of secondary school fee eliminations on enrollment and the first evidence regarding the effect of fee eliminations on student’s academic performance, measured by high school graduation test results.

I show that fee elimination has a relatively small effect, inducing approximately 10 additional students to enroll in the average school. This increases baseline enrollment by less than 3% and under plausible assumptions on the costs of education suggests an uncompensated elasticity of demand for enrollment of less than 0.1. These results are considerably smaller than those reported elsewhere in the literature and are insignificant in many specifications, suggesting that eliminating school fees in this setting had a modest effect on enrollment.\footnote{Preliminary results from a power analysis suggest that the insignificant results are driven more by high residual variance than by small treatment effects.} This effect is not driven by student transfers from control into treatment schools and appears to be a pure demand-side effect, with little evidence of correlated changes in school characteristics that might be interpreted as supply-side effects. I show that the effect is concentrated in urban areas and in secondary schools and report suggestive evidence that this reflects more elastic demand for enrollment at the secondary level.

The remainder of the paper is organized as follows. I develop a simple theoretical framework in section 2 that motivates some of the empirical content of the paper. I then discuss the South African education system and fee elimination policy in more detail in section 3. This section also shows how the design of the policy motivates the regression discontinuity design I use to estimate the effect of fee elimination on enrollment and presents evidence that the identification assumptions of the regression discontinuity design are plausible in this setting. Section 4 reports and discusses treatment effects for the full sample and selected subsamples. Section 5 presents evidence that there are no substantial student transfers from control to treatment schools, though I cannot yet rule out the possibility of other behavioral responses, such as changes in fees charged at control schools. I conclude in section 6 and discuss some alternative specifications of the regression discontinuity design in the appendix.

This paper uses a school-year panel of administrative data collected by the South African Department of Education, to which I have negotiated unprecedented access. Unfortunately, I do not yet have access to the entire dataset, so this version of the paper presents results for only five of the country’s nine provinces.

This paper directly advances a growing literature that studies the effect of tuition fee eliminations on student enrollment in primary and secondary schools. Studies in Madagascar (Fafchamps and Minten, 2007), Malawi (Al-Samarrai and Zaman, 2000), Kenya (Lucas and Mbiti, 2009), and Uganda (Deininger, 2003) have found very large increases in enrollment of up to 100% off relatively
low bases. These results are at least an order of magnitude larger than those I find, which may reflect more elastic education demand in these settings or limitations of these research designs. Most studies rely on simple comparisons of enrollment before and after a nationwide fee elimination. The implementation of other simultaneous policy changes may also have affected enrollment and resulted in upward biased estimates of the effect of fee elimination.

Barrera-Osorio, Linden, and Urquiola (2007) and Borkum (2011) are the only existing studies that use more credible research designs. The former paper examines a natural experiment in Bogota, Colombia, where the local government reduced school fees for households whose socio-economic status fell below a threshold level. The authors find that enrollment increased by 0-5 percentage points in households just below the threshold, relative to households just above the threshold who did not qualify for fee reductions. The latter paper also studies the South African fee reform in a single province using more limited data and finds a rise in enrollment of 0-2%.

The contrast between these two strands of the literature may reflect a number of differences. First, the interrupted time series and panel data used in the former literature may be subject to substantial upward biases due to correlated policy changes. Second, the regression discontinuity designs used in the latter literature and in my own work estimate valid treatment effects only in the neighborhood of the cutoff. If poorer households are more responsive to fee eliminations, treatment effects may be considerably larger well below the cutoff. Third, countries studied by the latter literature typically have much higher baseline enrollment rates so the treatment effects may be restricted by ceiling effects. I cannot directly separate these hypotheses but I briefly suggest a framework for beginning to do so in section 4.

This paper is also related to a substantial literature on the enrollment effects of conditional cash transfers. These transfers have typically been offered in countries without tuition fees and are designed to offset the lost contributions to home production and labor market earnings from enrolled children. These programs typically raise enrollment by less than 10%, though there is some heterogeneity across studies (Angelucci and di Giorgi, 2009; Glewwe and Kassouf, 2011; Schultz, 2004; Todd and Wolpin, 2006). This heterogeneity has been ascribed to differences in baseline enrollment rates across different countries and to differences in the extent to which conditions bind. For example, some cash transfers are only conditional on school enrollment, while others also impose conditions on attendance or grades. All else equal, the latter tend to have smaller effects, perhaps because the stricter conditions render more potential students inframarginal with respect to the intervention (Filmer and Schady, 2008; Kremer, Miguel, and Thornton, 2009). The effect sizes that I find from eliminating tuition fees are roughly comparable to those associated with conditional cash transfers in Brazil and Mexico, suggesting that the two policies may have relatively similar effects when applied in countries with comparable levels of baseline enrollment.
There is also a small literature that studies the enrollment effects of changes in the cost of education in the developed world. See Dynarski (2003), Dynarski, Gruber, and Li (2009), Kane (1994), and Seftor and Turner (2002) for examples and Neal (2002) for a discussion of the challenges faced by this research agenda. However, these research designs typically focus on margins different to those in the development literature: the choice between private and (free) public primary or secondary education and the choice between different types of postsecondary institutions with different costs. These are very different to the choice between fee-charging public and no primary or secondary education, which is perhaps the more relevant margin in much of the developing world. While these studies provide useful insights into the nature of education investment decisions, they do not reduce the need for better evidence on the effect of school fees on enrollment.

2 Theoretical Framework

This paper employs changes in school-level enrollment counts as the primary outcome of interest whereas most of the existing literature uses household survey data that records the enrollment status of each child or young adult in the household. I can transform the school-level treatment effects that I estimate into individual-level treatment effects under mild additional assumptions. This permits an easier comparison of my results with the existing literature.

To illustrate this transformation, consider a very simple reduced form model of the enrollment decision problem. Assume that there exists a countable population of agents \(i = 1, \ldots, N\), each of whom chooses whether to enroll in school in period \(t \in \{0, 1\}\). I refer to these agents as “potential students” and abstract away from the fact that they enter period \(t\) with different levels of prior schooling. I define \(U_{ijt}\) as the benefit that potential student \(i\) receives from enrolling at school \(j\) in period \(t\), and \(C_{ijt}\) as the analogous cost. I remain agnostic regarding the structure of the benefit and cost functions and at this stage impose no restrictions on the distribution of benefits and costs across the population. These functions are likely to vary systematically across grades and \(U_{ijt}\) implicitly includes the option value of enrollment in higher grades in future years. I interpret the cost function as including direct pecuniary costs (school fees), indirect pecuniary costs (transport, uniforms, etc.) and non-pecuniary costs (psychic costs, foregone leisure and employment opportunities, etc.) Using a potential outcomes framework, I define \(C_{ijt}(1)\) and \(C_{ijt}(0)\) as the costs of enrollment if the fee elimination treatment is and is not applied respectively. As indirect pecuniary and non-pecuniary costs remain even after fees are eliminated, \(C_{ijt}(1)\) is not in general equal to zero. The distribution of costs may vary across potential students and schools, so it is likely that that \(C_{ijt}(1) - C_{ijt}(0)\) varies across \(i\) and/or \(j\). The benefits of enrollment may depend on whether fees are charged or not, so \(U_{ijt}(1)\) and \(U_{ijt}(0)\) will not in general be equal.
Total enrollments with and without the fee elimination treatment at time $t$ are

$$\sum_{i=1}^{N} 1 \{U_{ijt}(1) - C_{ijt}(1) \geq 0\}$$

and

$$\sum_{i=1}^{N} 1 \{U_{ijt}(0) - C_{ijt}(0) \geq 0\}$$

respectively. The average treatment effect at time $t$ is simply the proportion of agents who enroll if and only if the policy is implemented:

$$\tau_{individual} = \frac{1}{N} \sum_{i=1}^{N} 1 \{U_{ijt}(1) - C_{ijt}(1) \geq 0 > U_{ijt}(0) - C_{ijt}(0)\}$$

$$- \frac{1}{N} \sum_{i=1}^{N} 1 \{U_{ijt}(1) - C_{ijt}(1) < 0 \leq U_{ijt}(0) - C_{ijt}(0)\}$$

(1)

If the benefit of enrollment is unaffected by the fee elimination policy, the second expression is equal to zero. However, if schools experience overcrowding or decreases in peer quality due to the policy, it may be the case that $U_{ijt}(0) > U_{ijt}(1)$ for some $i$ and $j$. This reduces the magnitude of the treatment effect relative to a pure demand effect that holds $U_{ijt}(0) = U_{ijt}(1)$. In section ?? I explore whether there is empirical evidence of such “school quality” effects. I suppress the second term to simplify the notation in the next paragraph.

The parameter defined in equation (1) cannot be directly estimated without individual-level data. I can, however, define total enrollments with and without the policy as

$$\sum_{j=1}^{J} \sum_{i=1}^{n_j} 1 \{U_{ijt}(1) - C_{ijt}(1) \geq 0\}$$

and

$$\sum_{j=1}^{J} \sum_{i=1}^{n_j} 1 \{U_{ijt}(0) - C_{ijt}(0) \geq 0\}$$

respectively, where $n_j$ is the number of potential students linked to school $j$ and $J$ is the total number of schools. The average treatment effect on the treated schools is the average change in

---

2This is technically an average treatment effect on the treated (ATT), rather than an average treatment effect (ATE). All the estimates I report in this paper are ATT estimates but this distinction is unlikely to be important in practice. The regression discontinuity design should (locally) balance students on observed and unobserved characteristics, in which case the two parameters are identical. See Heckman and Robb (1985) for more discussion on this point.
the number of students enrolled at each school:

\[
\tau_{school} = \frac{1}{J} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \mathbf{1} \{U_{ijt}(1) - C_{ijt}(1) \geq 0 > U_{ijt}(0) - C_{ijt}(0)\}.
\]

These parameters are related by the inverse of the average number of potential students linked to each school:

\[
\tau_{individual} = \frac{J}{\sum_{j=1}^{J} n_j^t} \times \tau_{school}
\]

This rescaling cannot be implemented using school-level data, which permit estimates of only \(J\) and \(\tau_{school}\). However, the denominator can in principle be identified using census data, as it equals the number of “school-age” children and adults in the population. Subsequent versions of the paper will attempt this exercise once data from the 2011 national census becomes publicly available. As an interim measure, I discuss with selected results the value of \(\tau_{individual}\) implied by the estimated \(\tau_{school}\) and plausible values of the baseline enrollment rate.

### 3 Policy Background, Design, and Implementation

#### 3.1 Background on South African Education

South Africa is a middle income country with a history of sharp economic, political, and social inequality. The public education system was racially segregated until the early 1990s, and per capita government expenditure on white schools was orders of magnitude larger than on black schools. Both enrollment and high school graduation rates differed sharply by race and there is some evidence of large differences by household socio-economic status (Fedderke, Luiz, and de Kadt, 2000; Seekings and Nattrass, 2005). A small number black students from low income households enrolled in private schools, mostly church-run. Despite recent growth in the number of small, for-profit private schools in low income communities, the best available data suggests that the private sector remains negligible relative to the public sector and considerably smaller than in South Asian countries (Centre for Development Enterprise, 2010).

State expenditure on black education rose substantially in the 1970s, 1980s, and 1990s. This was associated with rising enrollment rates and table 1 shows that primary enrollment surpassed 90% by the mid-1990s, although the secondary enrollment rate remained closer to 50%.

\footnote{These are net enrollment rates, defined as the total number of enrollments amongst age-eligible individuals, divided by the total number of age-eligible individuals. This measure is widely used but risks understating actual enrollment if there are large numbers of older individuals enrolled, due to grade repetition, late school entry, or re-enrollment after a period of drop-out.}

However, substantial evidence suggests that the quality of education remained very low in historically black
schools. Curricula at these schools had deliberately focussed on non-academic subjects until the 1990s, reflecting the apartheid government’s insistence on preparing black students for manual employment only. Few students completed secondary schooling, pass rates on high school graduation examinations were low, and even fewer students took mathematics or physical science as high school subjects (Fedderke, Luiz, and de Kadt, 2000).

Enrollment rates appeared to plateau shortly after 2000. Administrative records on enrollment and census data suggest that approximately 9 in 10 primary school-age individuals and 2 in 3 secondary school-age individuals were enrolled at this time (Department of Education, 2009). Survey data suggest that secondary school enrollment may be as high as 4 in 5 age-eligible individuals, though these data also point to a plateau at considerably less than full enrollment (Borkum, 2011). Survey evidence suggests that this is driven by high rates of drop-out during high school, rather than cohort-level differences in enrollment patterns, and that the remaining non-enrollment is concentrated amongst low-income households. This has led some policy-makers to advocate policy changes to reduce the pecuniary cost of schooling (Pampallis, 2008). Such policies are closely aligned with the ruling party’s long-standing rhetorical commitment to free education and perhaps motivated by widespread primary school fee eliminations in other African countries during the 1990s and 2000s.

Another school of thought argues that drop-out reflects the low quality of schools in low income communities and is a rational response to low returns to education in these schools. There are no direct measures of returns to education at different types of schools, so these arguments typically infer low returns to education from evidence of low school quality. While measuring “school quality” is a difficult process, it is true that South African students attending schools in low-income neighborhoods perform considerably worse on international literacy and numeracy assessments than poorer students from other African countries (van den Berg and Louw, 2007). South Africa has a system of nominal school choice, under which individuals from low income communities can in principle enroll in schools in high income neighborhoods but the limited data available on this phenomenon suggests that it is uncommon. This may reflect a combination of high commuting costs in cities that are still highly segregated by income and race, and social and cultural barriers that limit low income students’ ability to integrate into schools in high income neighborhoods.

Two policies were introduced in order to reduce the pecuniary cost of education and promote higher enrollment: means-tested individual-level tuition fee waivers and school-level tuition fee eliminations. The former policy was introduced in 1996 and required that schools grant partial tuition fee waivers to any household that either earned less than 10 times the per-student tuition fee or was eligible to receive a means-tested government child grant. The latter requirement meant that a large proportion of the country’s students were eligible for the waiver. However, the 2005
General Household Survey shows that approximately 2% of students benefitted from this policy and some media reports suggest that parents were discouraged by schools from requesting waivers (Hanes, 2006). This policy reduced the actual price of education below its nominal level and so risks attenuating any effect of the second policy on tuition fee eliminations. I argue below that any bias is likely to be small, due to the small scale of the fee waiver policy and the discontinuous implementation of the fee elimination policy.

3.2 The Fee Elimination Policy

The tuition fee elimination policy was announced in 2006 and implemented in 2007. Schools treated by this policy were required to eliminate all tuition fees, though the status of additional fees for extra-curricular activities was not regulated by the act. These “no fee” schools were chosen by a complex three-stage interaction between provincial and national governments, laid out in guidelines published by the national Department of Education.

In the first stage, provincial governments assigned each school in their province a “poverty score” based on characteristics of the electoral ward in which it was located. These scores ranked all schools within the province from least to most poor, with ties permitted. The national Department of Education provided each province with ward-level data on income, employment, education, health, and “living environment” from the 2001 census as a starting point for the assignment of poverty scores. Provinces were permitted choose their own weighting of these five data series and to make ad hoc adjustments to the resultant score based on within-ward heterogeneity. They were not permitted to use any data collected directly from schools, such as administrative data on school’s physical facilities or student-teacher ratios. Wildeman (2008) conducted anonymous interviews with provincial officials responsible for creating the poverty scores and reports that most of the ad hoc adjustments were made for schools near the boundaries of electoral wards, as the socio-economic characteristics of their students may have differed from those of the electoral ward. Wildemann’s interviewees reported no incidents of schools lobbying provincial officials to change their scores, although lobbying may have occurred in the third stage described below. The formulae used to determine the poverty scores were left to the discretion of the provinces and none have publicly announced these formulae or provided them to the author.

In the second stage, the national government specified the number of schools in each province to be treated. This number was chosen to satisfy two criteria: that 40% of the students nationwide would attend treated schools (based on 2006 enrollment data) and that the proportion of students in each province treated by the policy would reflect the relative poverty rates of the province. For

---

4South Africa’s academic year runs from January to December.

5The electoral ward is not an administrative unit in South Africa. Assignment took place at this level because it is the smallest geographic unit at which census data is publicly available.
example, 28% of schools in the relatively wealthy Western Cape province were treated and only 65% of schools in the relatively poor Eastern Cape province were treated. The choice of how many schools were to be treated in each province was based on province-level data from the 2001 census but the exact algorithm used for this decision is unclear. By determining the number of schools to be treated in each province, the national government implicitly specified a cutoff value of the poverty score above which all schools were the “intention to treat” and below which all schools were the “intention to control” group. National government officials report that the number of schools to be treated was chosen after the poverty scores had already been assigned, so it was not possible for poverty scores to be precisely manipulated in the neighborhood of the cutoff.

In the third and final stage, provincial governments decided which schools were to abolish fees, which created an “actual treatment” group of schools. The actual treatment assignments followed the intended treatment assignments relatively closely: 5% of schools below the cutoffs and 98% of schools above the cutoff are treated. The discrepancies may reflect lobbying by schools above the cutoffs who wished to continue charging fees or by schools below the cutoff who wished to eliminate them. The frequency of these discrepancies varies across provinces: 15% of schools have different intended and actual treatment statuses in the least compliant province (Northern Cape), while intended and actual treatment statuses are the same for all schools in one other province (Gauteng).

3.3 Research Design and Validation Tests

The design of the fee elimination policy makes it a natural candidate for analysis by a regression discontinuity design. I compare schools just below the cutoffs, the intended control group, with schools just above the cutoffs, the intended treatment group. If the poverty scores are “as good as randomly assigned” (Lee and Lemieux, 2010) in the neighbourhood of the cutoffs, these two groups differ only in their treatment status and so any differences in enrollment between the two groups may be interpreted as a causal effect of the fee elimination policy.

The regression discontinuity design I use differs from the standard application in two ways. The first difference arises because I have access to an eight year panel of school enrollment, beginning four years before the fee elimination policy was implemented. I therefore use the change in enrollment from before the policy to after the policy in each school as the primary outcome:

$$\Delta Y = \frac{1}{4} \sum_{t=2007}^{2010} Y_t - \frac{1}{4} \sum_{t=2003}^{2006} Y_t$$

This has two advantages relative to the standard regression discontinuity design based on cross-sectional data. First, using the change in enrollment from before to after the policy removes the
influence of any time-invariant unobserved school characteristics that are associated with both the treatment status and the outcome. This means that the only potential threat to the causal interpretation of the regression discontinuity design is if an unobserved time-varying school characteristic changes discontinuously at the cutoffs. Second, averaging enrollment in the four years before and the four years after reduces the influence of measurement error in the enrollment data and so reduces the risk of attenuation bias.\footnote{I also estimate pre-treatment discontinuities in the level of enrollment and test to see whether these are constant through time. If not, it is possible that there is a pre-treatment trend in enrollment in the neighborhood of the cutoffs and this trend should be explicitly modeled, rather than simply averaging all pre-treatment data. The test strongly rejects the presence of a pre-treatment trend: a test of zero time trend has a \( p \)-value of 0.93 and a test that the pre-treatment discontinuities are equal has a \( p \)-value of 0.26.} Furthermore, the differencing-and-averaging transformation I apply is statistically conservative. It emphasizes that the source of variation in my data is at the school, rather than school-year, level. This transformation collapses my data to the school level and so reduces the scope for type I errors due to inappropriately narrow confidence intervals.\footnote{I could instead use one of the cluster-robust inference procedures available in the literature, such as constructing bootstrap standard errors or confidence intervals and resampling at the school level. Simulation results from linear models suggest that these methods lead to higher rates of type I errors than simply collapsing the data. Note that because I study school-level enrollments, rather than individual-level enrollment decisions, these data are already aggregated and the tests are likely to be conservative.}

The second difference arises because the scale of the running variable, poverty scores, varies across the five provinces in my sample. Two provinces, for example, report poverty scores between zero and one, one reports poverty scores between plus and minus fifteen. These do not necessarily reflect differences in the underlying poverty levels of the provinces; they instead reflect differences in the way in which provincial officials choose to report the scores. Based on the interviews conducted by Wildeman (2008), I believe that the scores in each province are best interpreted as an approximately monotonic transformation of a common latent poverty measure.\footnote{The scores could alternatively be interpreted as a strictly monotone transformation with some degree of measurement error applied after the transformation.} I therefore normalize the scores so that within each province, the cutoff separating treatment and control schools equals zero and the scores have unit variance. I discuss some alternative transformations in appendix A.

The validity of the regression discontinuity relies on the assumptions that (i) schools just above the cutoff have a significantly higher probability of treatment than schools just below the cutoff and (ii) there are no other systematic differences between schools just above and just below the cutoff (Imbens and Lemieux, 2008; Lee and Lemieux, 2010). Under these assumptions, any differences in enrollment between schools just above and just below the cutoff are interpreted as causal effects of the fee elimination policy. Each of these assumptions is at least partially testable.

I begin by estimating the difference in the probability of treatment between schools just above and below the cutoff.\footnote{I implement the regression discontinuity using a local linear regression estimated separately on either side of the cutoff with an edge kernel and a bandwidth chosen to minimize the approximate mean squared error of the treatment effect (Porter, 2003; Imbens and Wooldridge, 2009). I use the same approach throughout the paper. This estimation strategy also yields a simple analytical formula for a heteroscedasticity-robust variance-covariance matrix.} The first panel of figure 1 and the first row of table 3 show that the
probability of treatment jumps by 81.2 percentage points (standard error 1.3 percentage points) at
the cutoff. As intended and actual treatment status are not identical for all schools, I use intended
treatment status, \( 1\{\text{pov. score} > 0\} \), as an instrument for actual treatment status. The jump in
the probability of treatment at the cutoff is the first stage of the instrumental variables estimation.
The first stage is strong and the first condition for the validity of the fuzzy regression discontinuity
design is clearly satisfied.

I assess the plausibility of the second condition by testing for discontinuities in any pre-treatment
school characteristics. Table 2 shows that there are no significant differences between schools
just above and just below the cutoff in pre-treatment enrollment, school structure or governance,
location, student-teacher ratio, or socio-economic status of the student body. The proportion of
historically white schools is slightly higher below the cutoff than above, but these schools make up
a negligible proportion of all the schools near the cutoff and there are no significant differences in
the proportion of any other historical racial classification. This suggests that the schools on either
side of the cutoff are relatively similar on baseline observed characteristics, although it is slightly
concerning that some of the insignificant differences are still substantively large. This concern is,
however, offset by the fact that I am using a stricter test than necessary: the test shows that there is
little evidence of differences in levels of covariates at the cutoff, whereas my identification strategy
merely requires that there are no differences in changes in these covariates from before to after the
treatment. Estimating an analogue of table 2 using first differences shows no evidence of changes.

The second condition of the regression discontinuity design may be violated if schools are able
to precisely manipulate their poverty score and hence intended treatment status. The balance tests
reported in the previous paragraph provide some reassurance that this is not occurring. I also
test directly for evidence of such manipulation by examining whether the density of poverty scores
jumps at the cutoff, following McCrary (2008). Figure 2 shows that there is no evidence of such a
discontinuity and the difference at the cutoff of -0.002 is insignificant. Given the design of the policy,
this is unsurprising. Neither provincial governments nor individual schools knew where the cutoffs
would be located when poverty scores were assigned, making it impossible for them to precisely
manipulate these scores. If schools or provincial governments wished to subvert the system, it would
have been considerably easier for them to do so by changing schools’ actual treatment status after
their intended treatment status was determined. Such discrepancies do not pose a problem for the
validity of the regression discontinuity design. I simply estimate intention to treat effects based on
intended treatment status and rescale them by the jump in the probability of treatment to obtain
instrumental variables estimates of the treatment on the treated effect in the neighborhood of the
cutoff.

The balance tests in table 2 and figure 2 provide support to the idea that the only school-
level characteristic that changes discontinuously at the cutoff is the probability of treatment. In particular, panel E of table 2 shows that the proportion of students living in households that are eligible for means-tested government child grants is approximately equal on both sides of the cutoff. This provides a noisy measure of the proportion of households eligible for individual-level partial tuition fee waivers before the policy, and suggests that this proportion did not change discontinuously at the cutoff. The existence of the tuition fee waiver policy will thus not bias the estimated enrollment effects in the regression discontinuity design but will affect their interpretation. Specifically, the treatment effects should be interpreted as the effect of removing fees \textit{conditional on the fact that not all potential students would need to pay these fees}. This means that the price elasticity implied by a given enrollment change should be rescaled by the inverse of the proportion of students who were required to pay paid fees. Given that take-up of the fee waivers was so low (see the discussion above), this effect is likely to be small.

4 Effects on Enrollment

Figure 3 and table 3 show the effects of the tuition fee elimination policy on enrollment. The second column of table 3 reports the intention to treat estimate: this shows that the change in enrollment from 2003–2006 to 2007–2010 was 9 students higher for schools just above the cutoff than just below the cutoff. Mean baseline enrollment in schools just above the cutoff was approximately 378 students, so this represents a 2.4% increase in enrollment. The instrumental variables estimate of the treated on the treats reports a slightly larger effect of 11 students or 2.9% of baseline enrollment. The treatment effects are significant but not very precisely estimated: the IV estimator’s 95% confidence interval ranges from 0 to 22 students.

4.1 Benchmarking the Size of the Enrollment Effect

This effect size cannot be directly compared to other studies that use individual-level data on enrollment decisions. However, I can make some suggestive comparisons under additional assumptions. Recall from section 2 that the individual-level treatment effect of fee elimination can be written as

\[
\sum_{j=1}^{J} \frac{n_j \times \tau_{school}}{\sum_{j=1}^{J} n_j} \times \tau_{school},
\]

where \(J\) is the total number of schools, \(n_j\) is the number of “potential students” in school \(j\)’s catchment area, and \(\tau_{school}\) is the school-level treatment effect. I can therefore recover the individual treatment effect from the school treatment effect under assumptions on \(\sum_{j=1}^{J} n_j\). The survey and administrative data discussed in section 3 suggest that the baseline enrollment rate amongst age-eligible individuals over the entire country is between 75 and 85%. If the baseline enrollment for schools just above the cutoff was 75%, then

\[
\frac{\sum_{j=1}^{J} n_j \times \tau_{school}}{\sum_{j=1}^{J} n_j} = \frac{14197}{14197 \times 378/0.75} \times 11 = 0.022,
\]

so the eliminating fees increased enrollment from 75% to approximately 77%. This conversion is
very sensitive to the choice of the baseline enrollment rate. Perhaps the most useful insight from this conversion is that the estimated treatment effect using school-level enrollment counts imposes an upper bound on the treatment effect on the enrollment rate of three percentage points. This is a very small effect relative to prior studies of fee eliminations and is smaller than those reported by many studies of conditional cash transfers in other middle income countries.

The effect size can also be converted into a (Walrasian) price elasticity under assumptions on the baseline cost of schooling. The simplest possible conversion calculates the elasticity as \(
\frac{\% \Delta \text{enrollment}}{\% \Delta \text{cost of enrollment}} \)
using the effect size as the numerator and considering different values of the denominator. If tuition fees are the only pecuniary cost of education and all students paid fees before the policy, the denominator is one and the the effect size implies an elasticity of approximately 0.03. More generally, the elasticity is approximately \( \frac{0.03}{F/C} \), where \( F \) and \( C \) are respectively the expenditure on tuition fees and expenditure on all education costs, averaged over all students in all schools. I am still negotiating access to baseline fee data from the South African Department of Education but for all plausible values of \( F \) and \( C \), this implies an elasticity of less than 0.1.

4.2 Heterogeneous Effects on Enrollment

Tables 4 and 5 present evidence of two important dimensions of heterogeneity in the effect of fee elimination. The former table shows that fee elimination increases enrollment in urban schools (relative to the trend for schools that continue to charge fees) by 23 students, or 4.7% of baseline enrollment. In rural schools, the policy increases enrollment by a statistically insignificant 2.2 students, or 0.6% of baseline enrollment. The direction of this difference is perhaps surprising, given that survey data point to lower baseline enrollment rates in rural areas and hence more scope for the policy to have an effect.

There are at least two other differences between urban and rural schools that may have outweighed this “ceiling effect” and resulted in a larger treatment effect in rural areas. First, the fee elimination policy may have had a smaller effect on the cost of education in rural schools if baseline fees were lower in rural schools and/or transport costs were a larger percentage of the pecuniary cost of education. Second, rural schools may be of lower quality than urban schools, so that the gains to enrollment (broadly defined) are lower and hence the willingness to pay for enrollment is lower.

Table 5 reports the results of regression discontinuity designs estimated separately for primary, secondary, and combined schools. The effect of fee elimination is small and insignificant in primary

---

10\( \hat{\mu} \) calculate this bound by finding the value of the enrollment rate \( R \) that solves \( \frac{J}{\mu \hat{\mu}} R \times \tau_{\text{school}} + R = 1 \), where \( \mu \) is the average school size.

11Urban schools are those in cities and suburbs, while rural schools are those in small towns and farming areas.

12In South Africa, primary school covers kindergarten or grade 1 up to grade 7 and secondary school covers grades
and combined schools (1.3 and 0.6% of baseline enrollment respectively). Secondary schools see a rise in enrollment of 29 students or 5.7% of baseline enrollment but this effect is relatively imprecisely estimated, which in part reflects the relatively small number of high schools. Figure 4 reports treatment effects for each grade and shows that the effects are largest in grades 10, 11, and 12, although the grade-specific effects are again very imprecisely estimated. This pattern may reflect some combination of (i) higher baseline enrollment rates in primary schools that impose a ceiling on the treatment effect, (ii) a different relationship between fee and non-fee costs of education at primary and secondary schools, or (iii) more elastic demand for enrollment in secondary schools.

I cannot directly test the relative importance of these three explanations but I can perform some back–of–the–envelope calculations to explore whether the difference may be driven entirely by the first two explanations. Given the relatively imprecise estimates, these calculations should be interpreted with a high degree of caution. Survey data mentioned in section 3 suggest that the ratio of the primary and secondary enrollment rates for the entire country is at most 1.5 (although these data are for the entire country, not only schools located near the cutoff). Applying this ratio and the estimated treatment effects to the relationship $\tau_{individual} = \frac{1}{\sum_{j=1}^{J} n_j} \times \tau_{school}$ suggests that the individual treatment effect on enrollment rates in secondary schools is approximately three times higher than in primary schools.

Using the same crude elasticity formula from the previous subsection, I can also establish that the primary and secondary school elasticities implied by the treatment effects are equal only if

$$\frac{F_{primary}}{C_{primary}} = \frac{1}{4} \frac{F_{secondary}}{C_{secondary}},$$

i.e. if fees constitute four times as large a percentage of the total cost of education in secondary schools as in primary schools. Both these calculations suggest that ceiling effects and differences in the relationship between fees and non–fee costs of education do not fully explain the difference between the primary and secondary school treatment effects. I interpret this as weak evidence that demand for enrollment at schools near the cutoff is more elastic at the secondary than the primary level.

5 Spillover Effects on Control Schools

My estimation strategy assumes that the school fee elimination policy has no effect on enrollment levels at the control schools that continue to charge fees. This assumption may be violated if students who attended control schools before the policy change transferred to treatment schools after the policy change to take advantage of their new no fee status. Such behavior would result in an upward bias in the estimated treatment effect of the fee elimination policy.

I do not observe student–level data on transfers that would permit a direct test of this hypothesis.

8 to 12. Intermediate or middle schools are very rare and make up only 2.5% of my sample.
I therefore implement an indirect test that examines whether control schools that are geographically
closer to treatment schools experience falls in enrollment from 2006 to 2007 relative to farther away
control schools and relative to their own enrollment change in previous years. Figure 5 shows a
local linear regression of the change in grade-level enrollment from 2006 to 2007 at control schools
against the distance from the nearest treatment school offering the same grade.\textsuperscript{13} Control schools
nearer to treatment schools actually experience small gains in enrollment relative to control schools
farther away. I cannot reject that this pattern of changes is identical to that observed between 2005
and 2006, before the fee elimination policy was implemented. The result is robust to restricting the
sample to control schools within one half standard deviation of the cutoff. I interpret this as strong
evidence against the spillover hypothesis.

I also estimate a linear regression of change in enrollment by grade at control schools from
2006 to 2007 on the same measure at the nearest treatment school. If the treatment effect is driven
entirely by transfers from control to treatment schools, the slope coefficient should be approximately
equal to one. Instead, it equals 0.045 (standard error 0.016). This is not significantly different to the
coefficient in the equivalent regression using changes from 2005 to 2006 (0.037, with standard error
0.08). This result is robust to weighting the regression by the inverse distance between treatment
and control schools, to restricting the sample to control schools within one half standard deviation of
the cutoff, and to excluding control schools that are more than 10 miles from the nearest treatment
school. These results strongly suggest that the treatment effects are not driven by transfers from
treatment to control schools.

6 Conclusion

This paper advances the literature on education demand in developing countries by studying a
policy of nationwide school fee eliminations in South Africa. I use a regression discontinuity design
that compares schools just above a threshold poverty level, most of which eliminate fees, to schools
just below the threshold, most of which continue to charge fees. I show that fee elimination has
a relatively small effect, increasing enrollment by approximately 10 students, or 3% of baseline
enrollment. These results are considerably smaller than those reported elsewhere in the literature
and imply a relatively inelastic demand for enrollment in this setting. The effects are substantially
larger in urban than rural areas and larger in secondary schools than in primary schools. I argue
that this at least partially reflects more elastic demand for enrollment at the secondary level.

The small effect sizes and inelastic demand that they imply suggest that there are relatively
low marginal returns to additional years of enrollment at schools in the neighborhood of the cutoff.

\textsuperscript{13}I construct this distance measure using GIS codes for every school in the sample.
This is perhaps surprising in view of the general consensus that returns to education in South Africa are high in absolute terms and relative to other countries. One way to rationalize this finding is to note that the Mincer returns to additional years of high school education estimated using cross-sectional data are relatively low but that there is a large earnings premium associated with high school graduation. If the probability of high school graduation conditional on enrollment up to grade 12 is low at schools in the neighborhood of the cutoff, particularly for the marginal enrollees who may come from particularly low income households, then the expected return to enrolling may be relatively low. The availability of high school graduation examination results in future work will facilitate exploration of this hypothesis.
References


Table 1: Net enrollment rates through time

<table>
<thead>
<tr>
<th></th>
<th>Primary education</th>
<th>Secondary education</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>92</td>
<td>56</td>
</tr>
<tr>
<td>1998</td>
<td>93</td>
<td>59</td>
</tr>
<tr>
<td>1999</td>
<td>93</td>
<td>60</td>
</tr>
<tr>
<td>2000</td>
<td>90</td>
<td>60</td>
</tr>
<tr>
<td>2001</td>
<td>89</td>
<td>67</td>
</tr>
<tr>
<td>2002</td>
<td>89</td>
<td>64</td>
</tr>
<tr>
<td>2003</td>
<td>89</td>
<td>66</td>
</tr>
<tr>
<td>2004</td>
<td>87</td>
<td>67</td>
</tr>
<tr>
<td>2005</td>
<td>85</td>
<td>67</td>
</tr>
</tbody>
</table>

Notes: Adapted from Department of Education (2009). Population totals are mid-year imputations from census data. Net enrollment rate is defined as the number of enrolled age-eligible individuals divided by the number of age-eligible individuals. Primary education covers grades 1-7; secondary covers grades 8-12.
<table>
<thead>
<tr>
<th>Panel A: Enrollment</th>
<th>Control group mean at cutoff</th>
<th>Discontinuity</th>
</tr>
</thead>
<tbody>
<tr>
<td>In 2003</td>
<td>558.8</td>
<td>-91.4 (84.2)</td>
</tr>
<tr>
<td>In 2004</td>
<td>556.7</td>
<td>-85.4 (90.3)</td>
</tr>
<tr>
<td>In 2005</td>
<td>565.1</td>
<td>-93.9 (94.7)</td>
</tr>
<tr>
<td>In 2006</td>
<td>554.5</td>
<td>-90.9 (98.5)</td>
</tr>
<tr>
<td>Average 2003–2006</td>
<td>560.2</td>
<td>-90.1 (92.6)</td>
</tr>
</tbody>
</table>

| Panel B: School structure | % primary | .672 | .058 (0.088) |
| % intermediate          | .003 | -0.003 (0.005) |
| % secondary             | .222 | .005 (0.038) |
| % combined              | .057 | -0.039 (0.082) |
| # grades offered        | 6.89 | .12 (0.26) |
| % partially self-governing | .79 | .027 (0.103) |

| Panel C: Historical racial classification | % “homeland” schools | .738 | .052 (0.032) |
| % historically black     | .323 | -0.014 (0.027) |
| % historically white     | .023 | -0.014 (0.008) |
| % historically Indian    | .016 | -0.007 (0.033) |
| % historically mixed race | .098 | .002 (0.022) |
| % founded after desegregation | .052 | .007 (0.029) |

| Panel D: School location | % urban | .233 | -.015 (0.035) |
| % rural                 | .651 | .075 (0.072) |

| Panel E: Socio-economic status of students | % of students who are orphans | .151 | -0.011 (0.013) |
| % of students who receive child grants | .27 | .031 (0.024) |

| Panel F: Teachers | Student-teacher ratio | 37.925 | -569 (2.027) |
|                  | % of teachers privately hired | .027 | -.007 (0.01) |
Table 3: Regression discontinuity estimates using standardized poverty scores

<table>
<thead>
<tr>
<th>Estimator</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr(Treatment)</td>
<td>.812***</td>
<td>(\cdot 14)</td>
<td></td>
</tr>
<tr>
<td>∆ in enrollment</td>
<td>8.98**</td>
<td>11.06**</td>
<td>(4.58) (5.64)</td>
</tr>
<tr>
<td>Pre-treatment mean</td>
<td>378</td>
<td>378</td>
<td>378</td>
</tr>
<tr>
<td>Implied %∆ in enrollment</td>
<td>2.38**</td>
<td>2.94**</td>
<td>(1.22) (1.48)</td>
</tr>
<tr>
<td>Sample size</td>
<td>14197</td>
<td>14197</td>
<td>14197</td>
</tr>
</tbody>
</table>

Table 4: Regression discontinuity estimates for urban and rural schools using standardized poverty scores

<table>
<thead>
<tr>
<th>Sample Estimator</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban schools</td>
<td>ITT</td>
<td>IV</td>
<td>ITT</td>
<td>IV</td>
<td>ITT</td>
<td>IV</td>
</tr>
<tr>
<td>Pr(Treatment)</td>
<td>.935***</td>
<td>(\cdot 14)</td>
<td>.751***</td>
<td>(\cdot 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆ in enrollment</td>
<td>21.86**</td>
<td>23.39**</td>
<td>1.65</td>
<td>2.19</td>
<td>(9.47)</td>
<td>(10.15)</td>
</tr>
<tr>
<td>Pre-treatment mean</td>
<td>494</td>
<td>494</td>
<td>357</td>
<td>357</td>
<td>(1.92)</td>
<td>(2.05)</td>
</tr>
<tr>
<td>Implied %∆ in enrollment</td>
<td>4.43**</td>
<td>4.73**</td>
<td>.46</td>
<td>.61</td>
<td>(1.92)</td>
<td>(2.05)</td>
</tr>
<tr>
<td>Sample size</td>
<td>8346</td>
<td>8346</td>
<td>8346</td>
<td>5703</td>
<td>5703</td>
<td>5703</td>
</tr>
</tbody>
</table>

Table 5: Regression discontinuity estimates for primary, secondary, and combined schools using standardized poverty scores

<table>
<thead>
<tr>
<th>Sample Estimator</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary schools</td>
<td>ITT</td>
<td>IV</td>
<td>Combined schools</td>
<td>ITT</td>
<td>IV</td>
<td>Secondary schools</td>
<td>ITT</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Pr(Treatment)</td>
<td>.778***</td>
<td>(\cdot 19)</td>
<td>.923***</td>
<td>(\cdot 21)</td>
<td>.806***</td>
<td>(\cdot 35)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>∆ in enrollment</td>
<td>3.23</td>
<td>4.16</td>
<td>2.13</td>
<td>2.31</td>
<td>23.37</td>
<td>28.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-treatment mean</td>
<td>332</td>
<td>332</td>
<td>388</td>
<td>388</td>
<td>509</td>
<td>509</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implied %∆ in enrollment</td>
<td>.97</td>
<td>1.25</td>
<td>.55</td>
<td>.6</td>
<td>4.59</td>
<td>5.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>7959</td>
<td>7959</td>
<td>7959</td>
<td>2940</td>
<td>2940</td>
<td>2942</td>
<td>2942</td>
<td>2942</td>
<td>2942</td>
</tr>
</tbody>
</table>

23
Table 6: Tests for changes in the “quality” of schools at the cutoff

<table>
<thead>
<tr>
<th>Panel A: Socio-economic status of the enrolled students</th>
<th>Pre-treatment level</th>
<th>Change from pre- to post-treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control mean</td>
<td>Discontinuity</td>
</tr>
<tr>
<td># of orphans</td>
<td>81.3</td>
<td>-8.3</td>
</tr>
<tr>
<td></td>
<td>(7.8)</td>
<td>(3.4)</td>
</tr>
<tr>
<td>% of students who are orphans</td>
<td>.163</td>
<td>-.014</td>
</tr>
<tr>
<td></td>
<td>(.01)</td>
<td>(.007)</td>
</tr>
<tr>
<td># of students receiving child grants</td>
<td>132.6</td>
<td>-1.3</td>
</tr>
<tr>
<td></td>
<td>(11.3)</td>
<td>(8.7)</td>
</tr>
<tr>
<td>% of students receiving child grants</td>
<td>.332</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>(.021)</td>
<td>(.016)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Teacher characteristics</th>
<th>Pre-treatment level</th>
<th>Change from pre- to post-treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control mean</td>
<td>Discontinuity</td>
</tr>
<tr>
<td># teachers</td>
<td>12.7</td>
<td>-.6</td>
</tr>
<tr>
<td></td>
<td>(.7)</td>
<td>(.2)</td>
</tr>
<tr>
<td>Student–teacher ratio</td>
<td>36.4</td>
<td>-.2</td>
</tr>
<tr>
<td></td>
<td>(1.2)</td>
<td>(1.2)</td>
</tr>
<tr>
<td>% of teachers paid for by parents</td>
<td>.024</td>
<td>-.002</td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td>(.005)</td>
</tr>
</tbody>
</table>
Figure 1: Test for a discontinuity in the probability of treatment
Figure 2: Falsification test for a discontinuity in the density of standardized poverty scores
Figure 3: Test for a discontinuity in the change in enrollment

-20 0 20 40 60
Change in enrollment 2003–06 to 2007–10
-3 -2 -1 0 1 2
Standardized poverty score
Fitted change in enrollment 95% confidence interval
Enrollment change in bin
Figure 4: Test for a grade-specific discontinuities in the probability of treatment (first panel) and change in enrollment (second panel)
Figure 5: Test whether control schools’ distance from treatment schools predicts their change in enrollment (0.1 degrees ≈ 7 miles)
A Constructing the Running Variable

As noted in section 3, my application differs from a textbook regression discontinuity design because the running variable is scaled in different units in each province. Two provinces, for example, report poverty scores between zero and one, one reports poverty scores between plus and minus fifteen. These do not necessarily reflect differences in the underlying poverty levels of the provinces; they instead reflect differences in the way in which provincial officials choose to report the scores. Based on the interviews conducted by Wildeman (2008), I believe that the scores in each province are best interpreted as an approximately monotonic transformation of a common latent poverty measure.

I use five strategies to aggregate the results across the provinces. First, I estimate five province-specific regression discontinuities and take the average of the five treatment effects. Second, I estimate five province-specific regression discontinuities and average them using weighted least squares with weights equal to the inverse of their variances. This is similar in spirit to the two-step generalized method of moments estimator sometimes for repeated cross-section data with clustered errors. Third, I estimate a single regression discontinuity using province-specific ranks as the running variable (the first school above the cutoff is assigned a value of one, the first school below the cutoff a value of minus one, etc.). Fourth, I estimate a single regression discontinuity using province-specific percentiles as the running variable. Substantively, this embodies the assumption that the difference in socio-economic status between between the poorest and the wealthiest wards in each province are roughly equal across provinces. Fifth, I estimate a single regression discontinuity using province-specific poverty scores normalized to have unit variance. Substantively, this embodies the assumption that the variance of ward-level socio-economic status is roughly equal across provinces. In all cases I normalize the value of the cutoff in each province to zero.

Table 7 reports the results from all five estimation strategies. The strong first stage is clearly evident, with the probability of treatment jumping by at least 77 percentage points in all designs. However, the estimated the change in enrollment is not robust across the five designs. The weighted least squares design, which does not impose any common structure on the poverty scores across the different provinces, imposes the fewest restrictions on the data and suggests an effect on enrollment of approximately 11 students. The standardized variance design yields a very similar estimate but considerably tighter confidence intervals, as there are efficiency gains to aggregating data rather than estimating separate models and then combining them. I use the standardized variance design as my preferred estimation strategy because for the main effect and most subgroup effects it produces estimates similar to the very unrestrictive weighted least squares design but with smaller standard errors.

The first and fourth designs generate slightly larger and smaller estimates respectively. The
clear outlier is the third, rank-based design, which estimates a treatment effect of more than two hundred students. Figure 6 shows that this is driven by the presence of several outliers near the cutoff.
<table>
<thead>
<tr>
<th>Running variable</th>
<th>Separation strategy</th>
<th>Province-specific poverty score</th>
<th>Rank</th>
<th>Combined RDs</th>
<th>Percentile</th>
<th>Standardized score</th>
</tr>
</thead>
<tbody>
<tr>
<td>First stage</td>
<td>OLS</td>
<td>.835***</td>
<td>.741***</td>
<td>.838***</td>
<td>.812***</td>
<td>(.069) (.059) (.058) (.01) (.014)</td>
</tr>
<tr>
<td></td>
<td>WLS</td>
<td>.93***</td>
<td>.774***</td>
<td>.838***</td>
<td>.812***</td>
<td>(.16) (.059) (.058) (.01) (.014)</td>
</tr>
<tr>
<td>Second stage</td>
<td>OLS</td>
<td>16.72</td>
<td>183.5***</td>
<td>5.45</td>
<td>8.98**</td>
<td>(16.43) (14.23) (66.8) (3.8) (4.58)</td>
</tr>
<tr>
<td></td>
<td>WLS</td>
<td>6.17</td>
<td>183.5***</td>
<td>5.45</td>
<td>8.98**</td>
<td>(16.43) (14.23) (66.8) (3.8) (4.58)</td>
</tr>
<tr>
<td>IV estimator</td>
<td>OLS</td>
<td>20.41</td>
<td>237.06***</td>
<td>6.51</td>
<td>11.06**</td>
<td>(18.17) (17.78) (86.83) (4.54) (5.64)</td>
</tr>
<tr>
<td></td>
<td>WLS</td>
<td>11.35</td>
<td>237.06***</td>
<td>6.51</td>
<td>11.06**</td>
<td>(18.17) (17.78) (86.83) (4.54) (5.64)</td>
</tr>
</tbody>
</table>
Figure 6: Test for a discontinuity in the probability of treatment (left panel) and change in enrollment (right panel), using different running variables and normalizing the cutoff to equal zero within each province.

Panel A: Standardizing poverty scores to have unit variance in each province

Panel B: Using within-province percentiles as the running variable (i.e. imposing unit range)

Panel C: Using within-province ranks as the running variable
Figure 7: Test for a discontinuity in the change in enrollment, using province-specific poverty scores

Panel A: Eastern Cape

Panel B: Gauteng

Panel A: KwaZulu-Natal

Panel D: Northern Cape

Panel E: Western Cape