Columbia University

Department of Economics
Discussion Paper Series

Can eliminating school fees in poor districts boost enrollment? Evidence from South Africa

Evan Borkum

Evan Borkum is a PhD student in the Department of Economics, Columbia University. Faculty membersLeigh Linden and Miguel Urquiola have recommended the inclusion of this paper in the Discussion Paper Series.

Discussion Paper No.: 0910-06

Department of Economics
Columbia University
New York, NY 10027

July 2009

Can eliminating school fees in poor districts boost enrollment? Evidence from South Africa

Evan Borkum, Columbia University

July, 2009

Abstract: The charging of school user fees is a much-debated policy issue in developing countries. In this paper, I evaluate the impact of a South African fee elimination program that was targeted at the poorest two quintiles of schools based on a community poverty score. Fixed effects estimates find that the program increased enrollment by almost 2% in treated secondary schools, an increase concentrated in earlier secondary grades. There is substantial heterogeneity in the estimated secondary school effect: it is driven entirely by an increase of around 3.5% in the poorer of the two treated quintiles. Regression discontinuity estimates confirm that the relatively wealthy schools near the treatment cutoff did not experience any effects on enrollment. Overall, the abolition of fees seems to have been reasonably effective in increasing secondary school enrollment in particularly poor communities. This is despite the fact that the eliminated fees were relatively low, comprising only around 1.5% of annual household income (per child) in these communities.

JEL No.: I28, O15

Keywords: User fees, Enrollment, South Africa

I would like to thank Bradley Page of the Eastern Cape Department of Education and Nosipho Zantsi of the South African Department of Education for assistance with the data. I am grateful to Leigh Linden and Miguel Urquiola for their guidance throughout this project. Thanks also to seminar participants at Columbia for other helpful comments. Contact email: etb2102@columbia.edu.

1. Introduction

Many developing countries continue to charge fees to students attending public schools. Considerable concern has been expressed that such fees may represent a barrier to school enrollment, lowering investment in human capital and reducing economic growth. School fees may be especially damaging for poor households, which are more likely to be credit constrained and to fall into an intergenerational poverty trap due to the high costs of schooling (see for example Barham *et al.*, 1995)[1].

Motivated by these concerns, some developing countries have recently adopted policies to eliminate school fees, predominantly in primary education[2]. However, the impact of these programs on enrollment has rarely been studied using anything more than a simple before-after comparison of total enrollment figures. In part this reflects that these programs have involved the simultaneous country-wide elimination of fees, making other identification strategies difficult to implement[3]. As a result, there is still a paucity of well-identified evidence on the quantitative impact of such large-scale fee elimination programs.

---

[1] On the other hand, others present arguments in favor of school fees, mainly that they can improve quality by increasing available resources and by providing incentives for parental monitoring. For a summary of this debate see Hillman and Jenkner (2002) and Reddy and Vandemoortele (1996).

[2] In Africa alone examples of countries that have eliminated primary school fees include Kenya, Malawi, Uganda, Mozambique, Tanzania, Ethiopia and Ghana. In general the initiatives in these low-income countries seem to have been accompanied by sudden and dramatic increases in enrollment, often at the cost of severe overcrowding and reduced quality (UNICEF, 2006).

[3] An exception is Fafchamps and Minten's (2007) study on Madagascar, in which a civil conflict led to arguably exogenous geographical variation in the implementation of the fee elimination.

In order to contribute such evidence, this paper investigates the impact on enrollment of a unique targeted national fee elimination initiative in South Africa, under which some 40% of public school students have benefited from the abolition of fees. Specifically, schools were first divided into national quintiles using a poverty score based on the poverty of the surrounding community. The poorest two quintiles of schools were then declared “no-fee”, a status which required them to eliminate school fees on a mandatory basis starting in the 2007 academic year.

I evaluate the impact of this program using two empirical approaches to analyze data on public school enrollment in South Africa. The first is a standard fixed effects (FE) estimator that takes advantage of the panel structure of the data to identify the treatment effect. This strategy produces little evidence of any impact on enrollment at the primary level. In contrast, it suggests that the program increased national secondary school enrollment by almost 2 percentage points, driven by an increase in the earlier secondary school grades. This may reflect that secondary enrollment started from a lower base than primary enrollment, which was already high in South Africa prior to this initiative. In addition, the initial median secondary fee was more than double the median primary fee, so that the fee elimination entailed a greater cost reduction in secondary schools. The program also does not appear to have caused substantial student migration from fee-charging to no-fee schools, so that the impact on secondary enrollment likely occurred through the enrollment of children who would otherwise have been out of school.

The evidence suggests that the average secondary school impact estimated by FE masks important heterogeneities in the treatment effect. Specifically, it is the schools in the poorer of the treated communities - where sensitivity to school fees is likely to be highest - that experience positive treatment effects. Of the two bottom quintiles of schools that were treated, the national secondary school effect is almost entirely concentrated in the poorest quintile where enrollment increased by almost 3.5%. In the other, wealthier, treated quintile, the effect is near zero.

To confirm the lack of a treatment effect in the less poor of the treated schools I use a second empirical approach, a regression discontinuity (RD) design, which is feasible because data on the exact poverty score awarded to each school were obtained for a single large province. This allows me to take advantage of the sharp discontinuity in treatment status around the treatment cutoff poverty score that separates schools in quintiles 2 and 3. If the treatment effect is concentrated in the poorer of the treated schools, RD estimates are likely to find zero program effects since they are only valid locally for the relatively wealthy schools close to the cutoff.

Indeed, the RD estimates suggest that the program had virtually no effect on enrollment in either primary or secondary schools. In contrast, applying the FE estimator to the same provincial sample finds that the program increased secondary enrollment by almost 6 percentage points in the poorer of the two treated quintiles. Consistent with the findings for the national sample, this suggests that enrollment increases were concentrated in the poorer communities far from the treatment cutoff score. Overall these results imply that the elimination of the relatively modest pre-existing school fees was effective in increasing secondary school enrollment, but only in the poorer of the treated communities.

The decision to target this program at schools serving poor communities is in contrast to the recent fee elimination programs in many lower-income countries that have involved scrapping fees across the board. The alternative of targeting such a program at the poor allows the government to allocate school funding progressively by relying on parents to continue to meet much of the costs of schooling in wealthier communities. In the context of conditional cash transfer programs, Morley and Coady (2003) suggest that a geographical method of targeting such as the one described here may be especially appropriate for middle income countries like South Africa in which inequality is high and poor communities are spatially distinct from wealthy ones, so that they can be easily identified for targeting.

Another alternative is to target fee reduction at the household level, an example of which is the *Gratuidad* intervention in Colombia evaluated by Barrera-Osorio, Linden and Urquiola (2007). This program provided various levels of fee reductions to families based on their poverty score from a household survey, lending itself to an RD design. Suggestive of the results here, different effects are found by poverty level and for primary and secondary enrollment. However, there are potential problems associated with targeting at the household level. These include the potentially high costs of identifying poor households and verifying eligibility, the existence of perverse incentives to gain or retain eligibility and potential intra-community conflict between those selected for and those excluded from treatment. In fact, the introduction of the geographically targeted no-fee program in South Africa was partly motivated by the problems experienced with the household means-based system of fee exemptions that was in place prior to the program.

The remainder of this paper is organized as follows. Section 2 describes the no-fee program, including some background on schooling in South Africa. Section 3 discusses the identification strategies for evaluating the program, while section 4 describes the data and presents some descriptive statistics. Section 5 presents the FE results and some evidence on heterogeneous effects for the national sample. Section 6 investigates the existence and implications of heterogeneous treatment effects for the provincial sample, including the results from the RD design. Section 7 concludes.

2. The no-fee program

a. Background: schooling and school fees in South Africa

Schooling in South Africa consists of 12 grades of which the first 7 constitute primary school. Education is officially compulsory until grade 9, although this has rarely been enforced. The vast majority of students attend public school: the private sector is growing but still accounts for under 3% of enrollments (Department of Education, 2006a). The national Gross Enrollment Rate (GER) in 2005 was 103% for primary school, 89% for secondary school and 82% for the post-compulsory phase of secondary school, grades 10-12 (Department of Education, 2006b)[4]. As one of the few middle income countries in Sub-Saharan Africa, it is not surprising that these enrollment rates are far higher than those typical of the region.

The fundamental piece of education-related legislation in the post-apartheid era is the South African Schools Act (SASA) of 1996. Its stated aim is to "provide for a uniform system for the organization, governance and financing of schools". Under the SASA, School Governing Bodies (SGB's) must be constituted in each school, consisting of the principal as well as elected representatives of teachers, students (in secondary schools) and parents. The exact composition of the SGB can vary, but parents must retain an absolute majority. The precise functions of the SGB also vary by school. Schools can apply for "section 21" status that gives the SGB full control over its budgetary allocation from the education department. However, all SGB's have

[4] The GER is simply the number of students enrolled as a fraction of the number of school-age children in the population. It can be over 100% if there are many underage or overage children enrolled. The Net Enrollment Ratio (NER) is the number of school-age children enrolled as a fraction of the number of school-age children in the population. The Department of Education does not publish NER estimates, but figures available from the World Bank suggest that the 2005 NER was 86% for primary school and 72% for secondary school (World Bank, 2009).

power over individual teacher appointments (subject to the number of posts granted by the education department), language policy, admissions policy and the right to set fees.

In fact, the charging of school fees where appropriate is encouraged by the act, which exhorts SGB's to take "all reasonable measures" to supplement their funding from the state in order to improve quality. However, the act is also explicit that fees may not be imposed on parents who cannot afford to pay them. The provision of exemptions for parents who cannot afford fees has been set out in subsequent government notices that mandate partial or full fee exemptions based on the relative size of household income and school fees. Despite these provisions, anecdotal evidence from media reports and speeches from Department of Education officials suggests that fee exemptions are not operating as intended. One of the main problems is that the government does not compensate schools for fee exemptions, so that any exemption granted effectively means a lower amount of per-student funding. There is therefore an incentive for SGB's to avoid fulfilling their obligation to inform parents of their right to an exemption and to exert significant pressure on parents to pay fees even when they are eligible for an exemption[5].

b. School poverty scores and quintiles

The Education Laws Amendment Act of 2005 laid the foundation for the quintile-based declaration of no-fee schools that is the subject of this study. Subsequent guidelines published by the Department of Education clarified the basic procedure to be followed in order to classify schools into poverty quintiles. According to the guidelines, schools are to be awarded a poverty

[5] The media has reported on several instances where SGB's have attempted to seize the assets of parents who have not paid fees (see for example Stolley, 2005, and Macfarlane, 2007). Pressure on parents to pay fees may also arise through the (illegal) withholding of report cards and examination results of students who have not paid.

score based solely on the poverty of the surrounding community. The exact determination of the score is to be performed by the nine individual provincial departments, using the following basic approach. First, schools are to be linked to a certain geographical area such as an electoral voting ward. Then, variables from the latest census are to be selected to reflect average income, unemployment and the level of education/literacy in that area. Finally, these variables are to be weighted in some way to arrive at a final score. All schools within a certain small geographical area are therefore assigned the same score that is meant to reflect the poverty of the community served by these schools. It should be emphasized that the national legislation only sets out this basic framework: the final determination of poverty scores is up to the individual provinces.

In practice, a poverty index ("deprivation index") at the electoral ward level devised by Noble *et al.* (2006) was made available to provincial education departments in order to assist with them with the assignment of poverty scores[6]. Wildeman (2008) interviewed education department officials in all nine provinces (anonymously), and reports that most provincial departments made some use of this deprivation index in their score assignments. However, some provinces also felt that the ward was too crude a geographical level for the purpose of assigning scores and/or that the indices for some wards were outdated since they were using data from 2001. They therefore tended to make many of their own adjustments before arriving at the final

[6] This "Provincial Index of Multiple Deprivation" for each electoral ward in the country was obtained by combining data on 13 variables from the 2001 census variables into 5 "domains" of deprivation covering income, employment, health, education and living environment. Within a province, wards were ranked by their score in each of these five domains. A transformation was then applied, using each ward's rank in a particular domain to arrive at a new domain score with a common support across domains. These scores were then averaged to obtain the final deprivation index for each ward. Since the transformations were carried out using the within province ranking, the final index is not comparable across provinces.

scores, using additional local knowledge of schools and locations to distinguish between schools within a ward. For example, Wildeman reports that one province used the deprivation index as a component in its own formula while another made adjustments to the score if there was a discrepancy with a pre-existing provincial poverty ranking. Fortunately, communication with officials from the Eastern Cape confirmed that this province assigned poverty scores directly using the provided deprivation index. This is important because the Eastern Cape is the province for which the actual poverty scores are used in the subsequent RD analysis.

Once a province has assigned poverty scores it is then to establish poverty score cutoffs to enable it to classify schools according to *national* quintiles. This classification is performed using a "poverty distribution table" that was drawn up by the national government based on census household income data (Table 1). For example, the Western Cape is a relatively wealthy province, with only 6.5% of households in the province falling into the bottom national income quintile. Therefore, after ranking schools by their poverty score, enrollment data should be used to set the first provincial cutoff in such a way that 6.5% of students in the province are placed in the bottom national quintile. On the other hand, the Eastern Cape is a relatively poor province, where the first cutoff should ensure that 34.8% of students fall into the lowest quintile.

c. Declaration of "no-fee" schools

In 2006, schools in the bottom national quintile were invited to adopt "no-fee" status[7] on a voluntary basis, with individual provinces free to decide on an appropriate allocation to make up for the loss in schools fees. However, it is unclear to what extent the policy was implemented in

[7] In South Africa public schools are only meant to charge a single flat fee – separate administrative or registration fees are prohibited by law. "No-fee" status therefore eliminates the requirement to pay any fee at all for attending school, although parents are still responsible for other costs such as uniforms.

2006, both because of its voluntary nature and because of initial doubts over whether the legislation would be passed in time. As a result of the latter, many provincial departments were not completely prepared for its implementation and struggled to identify eligible schools and obtain the required budgetary allocation until well into the academic year, if at all.

Implementation of the program changed for the 2007 academic year, by which time provincial departments were well prepared for it. For this year, the policy was made mandatory and was extended so that schools in the bottom two national quintiles were declared "no-fee"[8]. A minimum per-student budgetary allocation was also specified at the national level in order to compensate schools for lost fees, with the allocation varying by quintile so that poorer schools received a higher per-student non-personnel allocation. Evidence from provincial budgets suggests that this additional component to the education budget was indeed provided for in 2007[9]. Particularly for schools in the no-fee quintiles, this allocation represented a substantial increase over that in previous years and is considered by the government to be sufficient to sustain a school in the absence of any income from school fees[10].

---

[8] Since it contained so few schools in the bottom two quintiles, the Western Cape province decided to extend no-fee status to the bottom three quintiles.

[9] Provincial governments are prohibited from raising their own revenue and instead are funded by a bulk allocation from the national budget. However, provincial governments have ultimate responsibility for the administration of the education system, including drawing up and disbursing their own education budgets subject to certain constraints.

[10] There may be a concern that schools deliberately inflated their reported enrollment figures after the introduction of the program in order to benefit from this increased funding. However, the fact that I find that the program had little effect in primary schools argues against this, since this incentive was also in place for primary schools.

3. Identification strategies

If treatment with no-fee status had been randomly assigned, a simple comparison of enrollment in treated and untreated schools would yield an unbiased estimate of the causal effect of the treatment. Since this was not the case, treated and untreated schools are likely to be quite different from one another on average and a simple comparison of enrollment is likely to reflect many factors other than the impact of the no-fee program. Two main approaches will be used to mitigate this problem and identify the causal effect of the program on enrollment for different groups of schools.

First, since panel data are available for each school, one can eliminate the impact of any school characteristic that is fixed over time using a fixed effects (FE) estimator. At the same time, one can attempt to control as best as possible for other pre-existing trends in enrollment that may be confounded with the estimated treatment effect. The basic estimates obtained are an average treatment effect for the entire group of treated schools. The second approach is to note that treatment status was determined by a discontinuous function of the province-specific poverty score: schools beyond a certain score cutoff were declared no-fee. This suggests the use of a regression discontinuity (RD) design, which attributes any discrete change in the level of 2007 enrollment at the cutoff to the causal effect of the treatment. In contrast to the FE estimates, the RD estimates are only valid locally for the limited group of treated schools near the treatment cutoff between quintiles 2 and 3, which are the least poor of the treated schools. A more formal discussion of the two identification strategies is presented below.

a. Fixed Effects (FE)

The FE estimator effectively uses the group of untreated schools as the comparison group for treated schools and makes use of the panel structure of the data to eliminate unobserved fixed effects. Consider the following reduced form expression for log enrollment, Y_{it}, in school i in year t:

$$Y_{it} = \alpha + \gamma\mu_t + \delta treat_{it} + \phi Z_i + \upsilon_{it} \qquad (1)$$

Where μ_t is a vector of time fixed effects and Z_i is a vector of observed school fixed effects. The variable $treat_{it}$ is a dummy for experiencing a no-fee policy in year t: δ is the coefficient of interest. Since the paper focuses on the treatment effect in 2007, the first year of mandatory and uniform application of the program, we can replace $treat_{it}$ with the interaction $I_i * 2007$ where I_i is a dummy that is unity for schools that were treated in 2007 (those in quintiles 1 and 2)[11]. Effectively we are identifying the treatment effect as a shock to enrollment in treated schools in 2007.

The error term υ_{it} in equation (1) can be thought of as $\upsilon_{it} = c_i + \varepsilon_{it}$, where c_i are unobserved fixed effects and ε_{it} are unobserved time-varying effects. Estimating δ in equation

[11] Although all quintile 1 schools were meant to be eligible for treatment in 2006, the program was voluntary. In theory I could therefore also estimate a separate 2006 treatment effect and interpret it as an intention to treat (ITT) estimate, despite the fact that I do not have any information on which schools (even as a proportion of the total) were actually treated in 2006. These estimates suggest a small and marginally significant effect of around 1.3% in quintile 1 schools. However, the ITT interpretation is complicated by the fact that it is unclear whether provincial departments succeeded in identifying their quintile 1 schools on time and whether the necessary funding was made available in 2006. In that sense we cannot identify the schools that were actually made eligible for the program in its intended form. Since these 2006 estimates do not have a clear interpretation and are therefore of limited value from a policy perspective, the focus in the paper is entirely on the 2007 treatment.

(1) by OLS will be biased if there is a relationship between $I_i * 2007$ and either of the components of υ_{it}. The FE estimator deals with the potential correlation between c_i and $I_i * 2007$ by eliminating c_i through a "within" transformation of the data. Each variable is differenced by its mean for that particular school over the entire period of observation and OLS is then applied to these transformed variables. The estimate of δ is then consistent under the assumption that $\text{cov}(I_i * 2007, \varepsilon_{it}) = 0$ for all $t \in \{2003,...,2007\}$, the period covered by the data. This implies zero covariance between the transformed versions of` $I_i * 2007$ and ε_{it}, resulting in a consistent estimate of δ. With the FE estimator, δ is identified purely from within school variation in enrollment over time.

b. Regression discontinuity (RD) approach

The intuition behind the RD approach is that, while the no-fee schools are likely to be quite different from other schools on average, schools in the vicinity of the treatment-determining cutoff score should be similar to one another in all characteristics other than treatment status. If this assumption holds then schools just below the cutoff can be used as a comparison group for those just above. More formally, if baseline enrollment and all other characteristics affecting enrollment are continuous in the poverty score at the cutoff, then any discrete change in the level of 2007 enrollment at the cutoff can be attributed to the causal effect of the treatment. We can therefore write the treatment effect in province j, T_j, as:

$$T_j = E[Y_{ij,t=2007}(1) - Y_{ij,t=2007}(0) \mid X_{ij} = s_j] = \lim_{x \downarrow s_j} E[Y_{ij,t=2007} \mid X_{ij} = x] - \lim_{x \uparrow s_j} E[Y_{ij,t=2007} \mid X_{ij} = x] \quad (2)$$

where $Y_{ijt}(I_{ijt})$ denotes the log of enrollment in school i in province j in year t, I_{ijt} is a binary indicator for treatment, X_{ij} denotes the poverty score and s_j is the poverty score cutoff beyond which schools in the province are treated. The subscript j is introduced to emphasize that each province is responsible for the final assignment of poverty scores and uses its own cutoff for division into quintiles. For this study, the necessary poverty score data and cutoffs were only obtained for a single large province, the Eastern Cape.

In a regression framework, one also needs to account for the possibility that the poverty score may be correlated with enrollment aside from its role in determining treatment. An appropriate specification would be:

$$Y_{ij,t=2007} = \alpha + \beta I_{ij,t=2007} + f(X_{ij} - s_j) + \theta Z_{ij,t<2007} + \omega_{ij,t=2007} \qquad (3)$$

where $f(.)$ is a smooth function of the score and $Z_{ij,t<2007}$ is a vector of predetermined covariates such as lagged log enrollment ($Y_{ij,t=2006}$), the lagged student-teacher ratio and regional dummies amongst others. The covariates are important for reducing bias as one uses observations further away from the cutoff and for improving the precision of the estimates. The coefficient of interest is β, the difference at the discontinuity. A well known point is that this treatment effect is valid only for the sub-population in the vicinity of the cutoff. If treatment effects are heterogeneous and vary by poverty score, schools far from the cutoff could experience very different treatment effects. Estimates presented later suggest that this is indeed the case and is a serious issue here.

4. Data and descriptive statistics

Each of the nine provinces of South Africa has its own Education Management Information Systems (EMIS) unit that maintains a database on all educational institutions in that province. Data from the provincial units are then forwarded to the EMIS unit at the national Department of Education, which collates all the data into a single database. School information on the provincial EMIS databases is updated through two major annual surveys completed by school principals. The first is the Snap survey that is conducted on the 10th day of the academic year and collects basic information on students and teachers such as enrollment numbers. The information on enrollments from the Snap survey is used for allocating the budget to schools each year. The second survey is the Annual Survey of Schools that is conducted every year in March or April[12]. It collects far more detailed information on students and teachers than the Snap survey such as a subject-wise breakdown of student numbers and information on the SGB and school infrastructure. However, since this survey is relatively lengthy and complex there are often problems experienced in obtaining fully completed surveys from schools.

Data on the poverty quintile for the universe of public schools in seven of the nine provinces were obtained from the national EMIS database, together with annual data on enrollments by grade and gender and the number of teachers from 2003 to 2007. Data on the level of school fees are also available, but were last collected in 2005. Detailed information on the location of each school was also obtained from the database, together with some information on school characteristics such as the apartheid-era racial classification of the school or whether it is urban or rural. Further details on the sample composition are available in section a of the appendix.

As mentioned, data on the precise poverty score awarded to each school could only be obtained for a single province, namely the Eastern Cape province. This is the third largest of the nine provinces by population and one of the poorest (see Table 1). It also has the largest number of public schools of any province in the country and the second largest number of enrolled students (Department of Education, 2006b). The poverty scores and cutoffs for allocation into national quintiles were obtained from a spreadsheet provided by the provincial Department of Education. There are 573 unique poverty scores for 5048 ordinary schools in the Eastern Cape

[12] The South African school year runs from the end of January to the beginning of December.

sample, with poorer schools receiving a higher score. For the purposes of this paper, the scores have been normalized so that they represent standard deviations from a no-fee cutoff of zero.

Selected descriptive statistics for the national sample of schools are provided in Table 2, presented separately for primary schools (panel A) and secondary schools (panel B). This distinction proves convenient and will be followed throughout the paper. Table 2 shows that schools in higher poverty quintiles – those located in more prosperous communities – tend to consist of a greater proportion of secondary schools than those in lower quintiles (panel B). This reflects the general neglect of secondary education for black South Africans during the apartheid era, particularly in the former tribal homelands. Schools in poorer quintiles are far less likely to be located in an urban area, while the formerly white schools are concentrated in the wealthiest quintile. The student-teacher ratio is fairly similar across quintiles with the exception of the wealthiest quintile where it is substantially lower. These wealthy schools typically raise enough funds from school fees to be able to hire extra teachers in addition to their allocated government teaching posts. Schools in higher quintiles tend to be larger in terms of their enrollment level per school, while secondary schools are larger on average than primary schools in the same quintile.

Fee data are presented for 2005, the last year in which they were collected[13]. The median annual school fee for primary schools is R50 (PPP US$19) per child. For secondary schools, the median fee increases to R130 (PPP US$49) per annum. Fees tend to be similar in magnitude across quintiles (with the exception of quintile 5), with secondary school fees being uniformly higher than the corresponding fees for primary schools. A crude calculation using data from the 2001 census indicates that, for the median household in each school targeted by the no-fee program, school fees amounted to 0.5% (primary schools) and 1.5% (secondary schools) of

[13] Fee data for the Western Cape were not available for 2005: the 2004 figures are used instead.

annual household income on average[14]. The figures are virtually identical if one calculates a separate figure for schools in each of the treated quintiles, quintile 1 and quintile 2, since fees and household income are proportionally higher.

While these amounts are not particularly large, several points should be emphasized. First, these are per-child amounts: for households with many children of school-going age, the cumulative effect could be large. Second, school fees are often due up front at the start of each semester. For poor rural households who are credit constrained this could potentially be problematic. Finally, the median fee for secondary schools is more than double that for primary schools. For schools in the poorest quintile, it is more than three times as high. In addition, SGB's in secondary schools are much more likely to have control over their allocated budget ("section 21 status") which would increase the incentive for them to enforce payment of fees. *A priori*, one might therefore expect the effect of the school fee elimination to be higher for secondary schools, especially given the fact that initial levels of secondary enrollment were much lower.

As a prelude to the more detailed analysis to follow, Figures 1a and 1b present the change in enrollment at an aggregated level. Each observation in Figure 1a represents one of the 60 education districts in the seven provinces in the sample that contained at least one public primary school in 2007. The change in total primary school log enrollment between 2006 and 2007 for each district is plotted against the proportion of primary schools in that district that were declared no-fee. A fitted line and 95% confidence interval for the simple binary regression are shown, in

[14] This calculation uses the 2001 census to assign to each treated school in the sample the median household income in the surrounding subplace (village or suburb), converted into 2005 rands. The 2005 school fee as a percentage of this income was then calculated for each school, and the mean taken over all treated schools.

which observations are weighted by the 2006 level of enrollment in each district. Figure 1b presents analogous results for secondary schools. The slope coefficient from this simple exercise is essentially equal to zero for primary schools. For secondary schools, the slope coefficient is positive and significant at the 1% level. The point estimate is 0.066 which implies that, on average, a district with 90% of its secondary schools treated would have experienced a 5.3% increase in enrollment relative to a district with 10% of its secondary schools treated. This first pass at the data suggests that, at a relatively high level of aggregation, secondary school enrollment did increase more in areas where the no-fee program was more widely implemented. Of course, these results are merely correlations and need not have any causal interpretation. The next two sections attempt to draw causal inferences on the program's effects using the identification strategies discussed above.

5. FE results

a. Basic results

Columns 1-3 of Table 3 present FE results for the full sample of primary schools (panel A) and secondary schools (panel B). The coefficients are those from applying the FE estimator to variants of equation (1), with $treat_{it}$ replaced by the interaction $I_i * 2007$. Column 1 includes only a vector of year dummies as controls. Province-year dummies are added in column 2 and a time trend specific to the treated group is added in column 3. The full specification in column 3 is therefore:

$$Y_{it} = \alpha + \gamma_1 \mu_t + \gamma_2 prov_i * \mu_t + \gamma_3 I_i * t + \delta I_i * 2007 + \upsilon_{it} \qquad (4)$$

where Y_{it} is the log of enrollment, μ_t is a vector of year fixed effects, $prov_i * \mu_t$ is a vector of province-year fixed effects and $I_i * t$ is a linear time trend specific to the treated group. The treatment group time trend is important to mitigate the possibility that we are identifying an effect that is merely a result of a pre-existing trend in enrollment that is unique to treated schools. The coefficient of interest is δ which, under the FE approach, is identified solely from the variation in enrollment within schools over time.

The addition of province-year fixed effects and the treatment group specific trend decreases the magnitude of the coefficient of interest for both primary and secondary schools. The treatment group time trend seems to be particularly important. Treated primary schools were experiencing a significant negative enrollment trend relative to untreated primary schools prior to the program while treated secondary schools were experiencing a significant positive trend. The opposite trends in primary and secondary schools could well be related if, for example, they are driven by a trend towards lower repetition rates in the upper grades of primary school in poor communities. Ignoring these pre-existing trends results in estimated treatment effects that are biased down for primary schools and biased up for secondary schools. With the full set of controls (column 3), the effect for primary schools in panel A of Table 3 is close to zero. On the other hand, the effect for secondary schools in panel B is a positive 1.8 percentage points in magnitude and is significant at the 1% level. This is in keeping with the fact that, as discussed earlier, enrollment in secondary schools was initially lower and the level of fees higher compared to primary schools.

A natural question is whether the secondary school effect identified here is a “true” increase in enrollment of otherwise out of school children or whether it simply reflects transfers from untreated schools. Of course, the estimation techniques used here only identify the change in enrollment relative to untreated schools and not the absolute change[15]. While direct data on school transfers is not available, several pieces of ancillary evidence suggest that transfers are not playing a major role here. First, since poor communities tend to be clustered together, treated and non-treated schools tend to be largely geographically isolated from one another. Any large-scale transfers between treated and non-treated schools are therefore likely to be costly. Second, data from the March 2007 Labor Force Survey (LFS) indicate that, of those currently enrolled in each secondary school grade, around 97% did not move households in the previous year. These are identical to the Figures in the pre-program March 2006 LFS and even slightly higher than those in the March 2005 LFS. A story about the migration of children to (for example) move in with relatives located close to treated schools therefore does not seem plausible here.

As a final check for migration, one can run similar regressions to the above to check whether enrollment in quintile 3 secondary schools experienced a different shock to quintile 4 and 5 schools in 2007. If students were transferring schools, one would expect most of the transfers to come from quintile 3 schools, which are on average geographically closer to treated schools and likely to be closer substitutes[16]. These estimates (not shown) are effectively zero: quintile 3 secondary schools did not experience a drop in enrollment relative to schools in higher quintiles. Taken together, this evidence suggests that enrollment in secondary schools likely increased due to the program's impact on the enrollment decision itself rather than through the impact on school choice for those who would have enrolled in any case.

---

[15] One cannot infer the absolute change in enrollment directly using the coefficient on the year 2007 time dummy. In fact it is very difficult to interpret this coefficient due to the multiple omitted categories necessitated by collinearity between the various time dummies, interactions and trends in the full specification.

[16] As mentioned, the private sector is still very small (3% of total enrollments in 2006): transfers from private to public schools are unlikely to play much of a role here.

Panels A-E of Table 4 presents FE estimates for individual secondary school grades (grades 8-12), in order to investigate which grades are driving the overall secondary school effect. Once again column 1 includes year dummies, column 2 adds the full set of province-year dummies and column 3 adds a linear trend specific to the treated group. In the full specification of column 3, the impact on secondary enrollment appears to be concentrated in the earlier secondary school grades of 8-10. This impact is around 2.5 percentage points for grades 8 and 9 (albeit only significant at the 10% level) and 3.0 percentage points for grade 10 (significant at the 1% level). In contrast, the inclusion of the treatment group time trend reduces the estimates for grades 11 and 12 to close to zero and statistically insignificant. The evidence using the national sample thus suggests that the secondary school effect is taking place in earlier secondary grades, particularly in grade 10 which is the first year of post-compulsory education[17].

It is interesting to note that administrative data suggest that the GER in grade 10 is unusually high compared to other secondary school grades (Department of Education, 2006b), which is typically taken as evidence of substantial grade retention at this level. This could be because the school curriculum tends to become noticeably more advanced in grade 10 and/or because of conscious "gate keeping" whereby schools block the progression of weak students at an early stage in order to avoid having them eventually take the matric exams in grade 12. These

[17] The lack of a positive grade 12 effect addresses an obvious concern, namely that the secondary school impact could reflect a school-leaving ("matric") exam shock. If these exams were unusually difficult in 2006 and the pass rate in poor schools is more sensitive to the difficulty level, more students than usual could have repeated in these schools in 2007. This would bias the estimated treatment effect up. Indeed, the national matric pass rate dropped from 68.3% in 2005 to 66.6% in 2006. However, the grade 12 effect is small, insignificant and negative in sign in specifications that use the full set of controls, suggesting that a matric exam shock is not driving the secondary school results.

academic difficulties in progressing past grade 10 could explain why enrollment in later secondary grades is less sensitive to school fees.

b. Robustness checks

The main concern with the panel data methods used here is that the treatment effect is identified simply as a shock to enrollment in treated schools in 2007. The analysis does control for differential *trends* in treated schools but any *shock* in 2007 that differentially affects treated schools will be part of the estimated treatment effect even if this is entirely unrelated to the no-fee policy. One way to address this is to retain the full treatment group but restrict the comparison group to schools in quintile 3. A similar intuition to that underlying the RD design suggests that this should result in a comparison group that is more similar to the treatment group and less likely to experience different shocks unrelated to the no-fee policy. The results for restricting the sample in this way are presented in column 4 of Tables 3 and 4. The estimated treatment effects for both primary and secondary schools (Table 3) are virtually identical to the original estimates. Turning now to the specific secondary grades (Table 4), the estimate for grade 8 increases substantially compared to the original estimate in column 3, while the estimates for grades 9 and 10 decrease in both magnitude and significance. The estimates for grades 11 and 12, however, remain close to zero and insignificant as before. The overall primary and secondary results are therefore quantitatively robust to the selection of a more appropriate comparison group, while those for individual grades are less so but retain the same overall pattern as before.

Another way to address the concern of other shocks is to augment equation (4) with the term $I_i * 2005$, the interaction between the treatment group dummy and a year 2005 dummy. Since the no-fee policy was not in effect in 2005, one would expect to see no effect on the placebo dummy. Of course, this does not get directly at the issue of whether there was some other unusual shock to treated schools in 2007, but rather gives an idea of how large such off-trend shocks are in a "typical" year absent the no-fee policy[18]. The results (not shown) indicate that the coefficients on $I_i * 2005$ for the primary and secondary samples are close to zero. The same is true if one runs the same estimates using quintile-year dummies instead. Other shocks unrelated to the no-fee program are therefore unlikely to be large enough to drive the effects identified here, although of course one cannot entirely rule out that such shocks may have occurred in 2007.

As a final robustness check of the overall results, one can use an entirely different identification strategy that takes advantage of the level of geographical targeting. As described in section 2b, the poverty scores determining treatment were assigned predominantly (or at least partially) using census data at the ward level, a geographical division that typically encompasses many different suburbs or villages ("subplaces")[19]. It is therefore possible that schools in two subplaces with similar observable characteristics would have been assigned a different treatment status since one subplace happened to be located in a ward that was less poor overall. Schools in the untreated subplace can therefore serve as the counterfactual for those in the treated subplace, allowing one to implement Heckman, Ichimura and Todd's (1997) difference-in-differences

[18] It is important to retain $I_i * 2007$ in equation (5) when running these falsification tests. Otherwise, if there is a positive effect on enrollment in 2007 and we ignore this term, the fitted treatment group time trend will be too steep. One could (and does) obtain significant negative estimates of $I_i * 2005$ even though enrollment is simply following the true trend in the year 2005.

[19] There are 21243 subplaces and 3754 wards in South Africa. While subplace level data from the census are publicly available, wards are used for electoral purposes and are not part of the geographical hierarchy in the census. Ward level data are thus not readily available although they were calculated by Noble *et al.* (2006) specifically for use in poverty targeting.

matching estimator (see section c of the appendix for details). Nearest neighbor matching estimates, performed using the method of Abadie *et al.* (2004), are presented in Table 5. The estimates are largely insensitive to the number of nearest neighbors used for matching or to a regression adjustment for differences in covariates to reduce bias. They suggest a treatment effect of near zero for primary schools (the estimates are actually slightly negative) and a significant effect of around 2 percentage points for secondary schools – very similar to the previous estimates.

c. Heterogeneous effects

Several empirical studies have presented evidence suggesting that, even within poor countries, poorer households are substantially more sensitive to schooling costs in their enrollment decisions[20]. In the current context, heterogeneous effects can be investigated for the national sample of schools at the poverty quintile level. Accordingly, Table 6 replaces the single treatment dummy in the basic FE estimates with a treatment dummy for each of the two treated quintiles, quintile 1 and quintile 2. Column 1 includes only the vector of year dummies as controls, column 2 adds province-year dummies and column 3 adds a time trend specific to each treated quintile[21]. The estimates for the primary sample in panel A of Table 6 confirm that there was little differential change in enrollment in either of the treated quintiles once the quintile specific time trends are properly accounted for. The secondary school estimates in panel B are more interesting. In the full specification of column 3, the estimates suggest that the impact of the no-fee program was almost entirely concentrated in the poorer schools classified into quintile

[20]See for example Mingat and Tan (1986), Gertler and Glewwe (1990), Birdsall and Orival (1996) and Glick and Sahn (2006).

[21] One could include a time trend for each of the non-treated quintiles too: this makes no difference to the results.

1. The impact on enrollment in these schools in these specifications is 3.4 percentage points (significant at the 1% level), compared to a near zero estimate for quintile 2. These results suggest that the earlier estimates of almost 2 percentage points for secondary schools were an average of a much higher 2007 effect in quintile 1 schools and a near-zero effect in quintile 2 schools.

d. Comparison to other programs

How does the estimated effect of around 2-3 percentage points for the South African no-fee program compared to other programs aimed at boosting enrollment in developing countries? Fiszbein and Schady (2009) provide a recent meta-analysis of a common type of program, namely conditional cash transfer (CCT) programs. Not surprisingly, the impact of these programs on enrollment is highly context specific. Impacts tend to be highest (up to 31 percentage points) for poorer countries, where baseline enrollment levels tend to be particularly low. Perhaps the more sensible comparison for South Africa is to several of the Latin American countries, which have more comparable levels of GDP per capita and baseline enrollment. The range of effects for CCT programs in these countries is between zero and 10 percentage points. Given that the size of the transfer in these settings (between 7 and 20 percent of per-capita expenditure) is much higher than the fee reduction in the South African program, the estimated impact of around 2-3 percentage points found here is quite respectable. The magnitudes are also in line with those for the Colombian fee reduction program evaluated in Barrera-Osorio *et al.* (2007), which has an impact of 3-6 percentage points on enrollment in a setting where baseline enrollment rates are similar to South Africa but the fee subsidy is somewhat higher.

6. Heterogeneous effects with RD: Eastern Cape sample

An alternative analysis of heterogeneous secondary school effects is possible for the sample of schools from the Eastern Cape province, since the exact poverty score awarded to each school was obtained for this sample. In particular, the sharp discontinuity in treatment status around the treatment cutoff can be used to implement an RD design in order to investigate the treatment effect for schools whose poverty score is in the vicinity of the cutoff.

Since the sample has changed, Table 7 first repeats some of the earlier FE estimates on the Eastern Cape sample in order to provide a baseline[22]. Columns 1 and 3 of Table 7 present FE estimates for primary and secondary schools respectively that are analogous to those using the full specification of Table 3, only restricted to the Eastern Cape sample. Columns 2 and 4 distinguish between the two treated quintiles and are analogous to the estimates for heterogeneous effects by quintile using the full specification of Table 6. In contrast to the national estimates, there is some evidence in favor of a small positive effect on primary enrollment in column 1 of Table 7, with the estimate 1.5 percentage points in magnitude and

[22] Around 6.5% of Eastern Cape schools are not classified as "ordinary" schools by the provincial education department. The bulk of these (5.5% of the total schools) are "farm schools". These schools are located on private land, typically a commercial farm, and were originally built during apartheid by white landowners to provide basic education to the children of their black laborers. Since the poverty score allocated to these schools likely reflected the status of the wealthy surrounding farming community rather than that of those actually attending the school, these schools were manually reassigned to quintile 2. Since their treatment status was not determined by the poverty score, the subsequent RD analysis drops these schools, together with the remaining 1% of schools not classified as "ordinary" by the provincial Department such as church schools. For comparability to the RD estimates, the estimates in Table 7 also omit these schools. For the earlier national analysis, the national department's definition of "ordinary" schools – which includes farm schools – was used.

significant at the 5% level. The estimates with heterogeneous effects in column 2 suggest that any impact on primary school enrollment is concentrated almost entirely in quintile 1 schools.

For the secondary school sample in column 3 of Table 7, the overall estimated impact of the program in the Eastern Cape is 3.2 percentage points (significant at the 5% level): far higher than the point estimate for the national sample in column 3 of Table 3. This makes sense given the fact that the Eastern Cape is one of the poorest provinces and had the lowest secondary and post-compulsory secondary enrollment ratios in the country in 2005 (72% and 65% respectively). For the estimates with heterogeneous effects in column 4, the coefficients suggest an increase in enrollment of 5.9 percentage points in quintile 1 secondary schools (significant at the 1% level) and no effect on enrollment in quintile 2 secondary schools. While the magnitude of the quintile 1 effect is much larger than for the national secondary school sample, the national pattern of heterogeneous effects seems to hold in the Eastern Cape in that the effects are concentrated entirely in the group of quintile 1 schools.

Preliminary visual evidence in favor of heterogeneous effects using the detailed poverty score data is presented in Figure 2. The figure plots a quintic fit of the change in log enrollment between 2007 and 2006 against the poverty score for each Eastern Cape school, allowing for a break at the cutoff. This figure is analogous to RD estimates for 2007 log enrollment that use 2006 log enrollment as a covariate and constrain its coefficient to unity. The advantage of presenting the figure in this way is that the change in enrollment in regions far from the treatment cutoff is made clear. The solid line in Figure 2 is the fit for the change in log enrollment in primary schools and is relatively flat across the domain of poverty scores. However, the slope does increase slightly towards the poorer schools in quintile 1, implying that there may have been a small impact on enrollment in this subset of treated primary schools as suggested by the previous FE estimates.

The dashed line in Figure 2 is the fit for secondary schools and suggests a large positive change in secondary enrollment, but again not in a homogeneous fashion. The change in secondary enrollment does not jump noticeably at the treatment cutoff, but instead begins to increase gradually a little past the cutoff. Towards the poorest quintile 1 schools, the change in secondary enrollment reaches a maximum and begins to decrease slightly. This could reflect the fact that, in the very poorest communities, the decrease in fees was not sufficient to overcome other barriers to enrollment such as a lack of access. In any case, Figure 2 is consistent with the FE estimates in that the change in secondary enrollment is small over the score range of quintile 2 schools but large for the poorer quintile 1 schools.

Figure 2 suggests that RD estimates of the jump at the discontinuity are likely to find that the program had small or zero effects on enrollment at any level. This is simply because these RD estimates are valid locally around the treatment cutoff between quintiles 2 and 3 where the least poor of the treated schools are located. In the presence of heterogeneous effects that are larger in poorer communities, these relatively wealthy schools could well experience no treatment effects. The remainder of this section focuses on the econometric evidence in support of this visual impression by checking the underlying assumptions behind the RD design and presenting RD estimates for enrollment.

a. First stage

Figure 3 shows how the normalized poverty score of each Eastern Cape school is related to its quintile assignment. The figure suggests that, except for a handful of schools, ordinary public schools were indeed placed into the quintile suggested by their poverty score. This motivates the use of the sharp RD design, in which having a normalized poverty score above zero (the cutoff between quintiles 2 and 3) determines treatment with no-fee status.

In any RD design there may be a concern that agents can select in or out of treatment by manipulating the running variable (the poverty score in this case). If this is the case, the key underlying assumption of the RD design – namely that all characteristics but treatment status are continuous around the cutoff - may not hold[23]. The nature of the no-fee program suggests some financial incentive for poor schools to claim no-fee status, since such schools obtain an increase in the per-student allocation from the government that is supposed to be larger than the lost income from school fees. For example, one could imagine quintile 3 schools run by a hard-working principal pressuring education department officials to change their score so as to become eligible for treatment. Although the level of discretion given to provinces in the final assignment of poverty scores may lead to serious concerns about score manipulation, this is less of an issue in the Eastern Cape where officials made direct use of the ward-level deprivation index computed independently by Noble *et al.* (2006). Moreover, since identification in the RD context focuses on schools near the cutoff, the concern is more specifically related to schools that may have manipulated their scores so as to become "just" eligible for the program. Figure 4 displays a histogram of the number of schools observed in small bins of the poverty score. There is little visual evidence of the stacking that might be expected to the right of the cutoff if this was the case.

b. Continuity checks

In evaluating the plausibility of the underlying assumptions in the RD design, we would ideally verify that all characteristics determined prior to treatment are continuous at the cutoff. While it

[23] See, for example, McCrary (2008).

is obviously not feasible to verify this for unobservable characteristics, it is essential do so for observable ones. For illustrative purposes consider the log of enrollment in 2006. Following Imbens and Lemieux (2008), the circles in Figure 5 represent bin averages for 2006 log enrollment. The open circles are for the sample of primary schools and the solid circles for secondary schools. The solid line plots fitted values of a regression of 2006 log enrollment on a quintic in the poverty score and a dummy for having a score greater than the cutoff for the sample of primary schools. This dummy allows for a break at the cutoff: any sharp changes in baseline 2006 enrollment could be a concern in the RD context. The dashed lines are fitted values for the same regression for the sample of secondary schools, around 15% of the total sample. Baseline enrollment displays a U-shaped relationship with the poverty score for primary schools and is generally lower than that of secondary schools with the same poverty score. With regards to identifying any breaks at the cutoff, the curves both appear to be smooth around the cutoff, supporting the RD approach.

Table 8 presents corresponding statistical evidence for log enrollment and a variety of other school characteristics measured in 2006: the student-teacher ratio, the gender ratio and a binary variable for connection to the electricity grid. The table shows the results of regressions of the variable of interest on a polynomial in the poverty score and a dummy for being beyond the no-fee cutoff. The coefficients shown are for the dummy, which again represents a possible break at the cutoff. Columns 1-4 show the results for primary schools while columns 5-8 are for secondary schools. The polynomials in the normalized score range from cubic to quintic as indicated in the table. The quintic specification in columns 3 and 7 corresponds to the fitted curves in Figure 5. Some specifications (columns 4 and 8) also include an interaction between the treatment dummy and the poverty score to allow for a different slope either side of the cutoff.

For the primary school sample, the difference in the 2006 enrollment level across the cutoff is close to zero (albeit quite imprecisely estimated) once the control polynomial is quintic. With regards to the other baseline characteristics, the estimates of the differences in student-teacher ratio and probability of connection to the electricity grid become weakly statistically significant with the quintic specification, but remain relatively small in economic terms. Allowing for a change in slope across the cutoff yields very similar results; the straightforward quintic specification is therefore the preferred one. For the sample of secondary schools, the estimates for the jump in 2006 enrollment are similarly small and close to zero once higher order polynomials are used. The differences in the other secondary school baseline characteristics are also generally small in magnitude and statistically insignificant. Overall, the point estimates using the preferred quintic specification suggest that the underlying assumption of similarity in baseline characteristics is plausible in the current setting.

c. Results for enrollment

Figure 6 illustrates fitted values from a regression of 2007 log enrollment in primary schools (solid line) and secondary schools (dashed line) on a quintic in the poverty score and a dummy for being beyond the no-fee cutoff. As in Figure 2 there appears to be little visual evidence of a jump around the cutoff, suggesting that the program did not substantially boost enrollment in treated schools of either level in the vicinity of the cutoff. The left hand panel of Table 9 presents corresponding statistical evidence for primary schools using a variety of specifications that correspond to versions of equation (3) using the preferred quintic polynomial to control for the poverty score. Column 1 presents the results of the simple regression of primary 2007 log enrollment on this quintic in the score and a treatment dummy. Column 2 adds the lagged level of enrollment as a control, while column 3 includes a variety of other baseline characteristics: the gender ratio, student-teacher ratio and a dummy for an electricity connection. Finally, column 4 adds six district dummies to the set of controls. Columns 5-8 repeat these regressions for the sample of secondary schools. Since the poverty score is not a continuous variable, standard errors are clustered by the poverty score following Lee and Card (2008), although in practice this makes little difference to the results here.

The coefficient of interest, namely that on the treatment dummy, is always insignificant and slightly negative across most specifications for primary schools. Once the full set of controls is included, the point estimate is below 1 percentage point in absolute value and still negative in sign. There does not appear to be any evidence that the program had a positive effect on enrollment in the vicinity of the cutoff. The estimates for secondary schools are presented in the right hand panel of Table 9. The point estimates of interest are less precisely estimated than those for primary schools but are close to zero in all specifications, providing little evidence of a strong positive effect on enrollment.

These estimates – together with the FE estimates broken down by treatment quintile – thus support the notion that school fees represented a barrier to secondary school enrollment, but only in poorer communities. Schools in quintile 2, particularly those relatively wealthy treated communities close to the treatment cutoff that are the focus of the RD estimates, experience virtually no treatment effect. On the other hand, the poorer quintile 1 schools experience relatively large and significant effects.

7. Conclusion

The estimates presented here have provided evidence on the impact of South Africa's no-fee program on school enrollment. Compared to the across the board fee elimination programs implemented elsewhere, the geographical targeting of the South African program at poor communities is an alternative that may prove attractive to countries aiming to reduce school fees in a manner compatible with their limited budgets. The national FE estimates suggest that the program had little effect on primary school enrollment, suggesting that school fees were not playing a major role in household schooling decisions at this level. The lack of a large primary school response is not surprising given that the initial level of fees was relatively low and enrollment was high.

On the other hand, the FE estimates suggest that the program did boost overall secondary school enrollment. For the national sample, the best estimate of this increase is just under 2 percentage points. The evidence for individual grade enrollments, while weaker, suggests that these increases are concentrated in the earlier grades of secondary school. There is strong evidence in favor of substantial heterogeneity by community poverty so that enrollment is affected more in poorer communities. In fact, the overall national secondary effect consists of an increase of almost 3.5 percentage points in the poorer of the two treated quintiles diluted by a near-zero effect in the other. The presence of heterogeneous effects is also evident in the estimates for a provincial sample, for which RD estimates find a zero effect on secondary school enrollment for the relatively wealthy group of treated schools near the cutoff. On the other hand, FE estimates for the same sample suggest an increase of almost 6 percentage points due to the program in the poorer of the two treated quintiles located far from the cutoff score.

The overall estimates are comparable to those for other subsidy-type programs in middle income countries, particularly CCT programs. The evidence thus indicates that secondary school enrollment in South Africa is somewhat sensitive to even modest schooling costs in particularly poor communities and that other interventions aimed at reducing costs (such as transport subsidies) might be effective if targeted at these communities using a similar method. However, the estimates do suggest that more careful attention should be paid to the setting of the treatment cutoff in these initiatives in order to target the poorest communities. If the goal of the no-fee program was simply to increase enrollment for example, the cutoff for no-fee status could have been set at a far lower level and had a similar effect. Of course, one cannot conclude that fees were not burdensome to households simply because their enrollment decisions were unaffected by the program. In that sense, the more generous cutoff and accompanying increase in government funding may have had desirable redistributive effects.

However, the results also suggest that efforts to substantially improve secondary school enrollment across a wider range of poor communities in South Africa – particularly in upper secondary grades where enrollment is most in need of improvement – should focus on other initiatives besides a simple fee reduction. While much progress has been made in equalizing and improving school resources since the end of apartheid, student performance on the matric exams and international tests such as TIMMS indicate that average school quality remains poor[24]. An obvious possibility is therefore that children are choosing not to enroll in secondary school since poor school quality offers low returns to the investment in schooling or that they are simply unable to attain the required academic standard to progress to higher secondary grades.

A recent paper by Lam, Ardington and Leibbrandt (2008) uses a panel dataset from a large South African city to confirm that repetition and failure rates in formerly black urban schools are very high. At least in the urban context, it does indeed appear that many students in poor communities experience severe academic difficulties in secondary school grade

[24] Almost 35% of candidates failed the national matric (grade 12) exams in 2006, while only 17% of candidates attained the basic standard necessary for university admission ("matric exemption"). South Africa also ranked bottom of 45 countries in math and science in the 2003 TIMMS test, below other African nations such as Botswana and Ghana.

progression. One would therefore expect a reduction in schooling costs to play only a limited role in increasing secondary school enrollment in this setting, particularly in higher grades. The no-fee program may have another role to play here, since it entails an effective increase in funding for poor schools. Whether or not this will eventually translate into an improvement in quality that could feed back into increased enrollment is an interesting issue for future research.

References

Abadie, A., Drukker, D., Herr, J.L. and Imbens, G.W. (2004). "Implementing Matching Estimators for Average Treatment Effects in Stata." *The Stata Journal*. vol. 4, pp. 290-311.

Barham, V., Boadway, R., Marchand, M. and Pestieau, P. (1995). "Education and the Poverty Trap." *European Economic Review*. vol. 39, pp. 1257-1275.

Barrera-Osorio, F., Linden, L.L. and Urquiola, M. (2007). "The Effects of User Fee Reductions on Enrollment: Evidence from a Quasi-Experiment." Columbia University: Working Paper.

Birdsall, N. and Orival, F. (1996). "Demand for Primary Schooling in Rural Mali: Should User Fees Be Increased?" *Education Economics*. vol. 4, pp. 279-296.

Department of Education (2006a). *School Realities 2006*. Republic of South Africa.

Department of Education. (2006b). *Education Statistics in South Africa at a glance 2005*. Republic of South Africa.

Fafchamps, M. and Minten, B. (2007). "Public Service Provision, User Fees and Political Turmoil." *Journal of African Economies*. vol. 16, pp. 485-518.

Fiszbein, A. and Norbert, S. (2009). *Conditional Cash Transfers: Reducing Present and Future Poverty*. Washington D.C: World Bank.

Gertler, P. and Glewwe, P. (1990). "The Willingness to Pay for Education in Developing Countries: Evidence from Rural Peru." *Journal of Public Economics*. vol. 42, pp. 251-275.

Glick, P. and Sahn, D.E. (2006). "The Demand for Primary Schooling in Madagascar: Price, Quality, and the Choice Between Public and Private Providers." *Journal of Development Economics*. vol. 79, pp. 118-145.

Heckman, J.J., Ichimura, H. and Todd, P.E. (1997). "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies*. vol. 64, pp. 605-654.

Hillman, A.L. and Jenkner, E. (2002). "User Payments for Basic Education in Low Income Countries." IMF Working Paper WP/02/182.

Imbens, G.W. and Lemieux, T. (2008). "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*. vol. 142, pp. 613-635.

Lam, D., Ardington, C. and Leibbrandt, M. (2008). "Schooling as a Lottery: Racial Differences in School Advancement in Urban South Africa." University of Michigan: PSC Research Report No. 08-632.

Lee, D.S. and Card, D. (2008). "Regression Discontinuity Inference with Specification Error." *Journal of Econometrics*. vol. 142, pp. 655-674.

Macfarlane, D. (2007). "Govt to Act on School Fees." Mail & Guardian Online: *http://www.mg.co.za/printformat/single/2007-06-22-govt-to-act-on-school-fee*

McCrary, J. (2008). "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*. vol. 142, pp. 698-714.

Mingat, A. and Tan, J.P. (1986) "Expanding Education through User Charges: What Can Be Achieved in Malawi and Other LDC's?" *Economics of Education Review*. vol. 5, pp. 273-286.

Morley, S. and Coady, D. (2003). *From Social Assistance to Social Development: Targeted*

Education Subsidies in Developing Countries. Washington D.C: IFPRI Center for Global Development.

Noble, M., Babita, M., Barnes, H., Dibben, C., Magasela, W., Noble, S., Ntshongwana, P., Phillips, H., Rama, S., Roberts, B., Wright, G. and Zungu, S. (2006). *The Provincial Indices of Multiple Deprivation for South Africa 2001*. Oxford University, U.K.

Reddy, S. and Vandemoortele, J. (1996). “User Financing of Basic Social Services”. UNICEF Staff Working Papers: Evaluation, Policy and Planning Series.

Stolley, G. (2005). “Unpaid School Fees a `Serious Problem' in KZN.” Daily News (July 1, 2005)

UNICEF (2007). “School Fees for Africa: Coming to Grips with an Elusive Promise.” *http://www.unicef.org/media/media_33184.html*

Wildeman, R.A. (2008). “Reviewing eight years of the school norm funding reforms, 2000 to 2008.” IDASA research paper.

World Bank (2009). *EdStats Database*.

Table 1: Poverty distribution table for allocation of schools to poverty quintiles

Province	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
Eastern Cape	34.8	21.6	21.0	11.6	10.9	100.0
Free State	30.8	14.9	20.1	18.8	15.4	100.0
KwaZulu-Natal	24.2	18.8	25.6	17.3	14.1	100.0
Gauteng	10.5	11.4	27.4	27.2	23.6	100.0
Limpopo	34.0	22.3	24.9	11.6	7.2	100.0
Mpumalanga	16.7	20.2	29.8	19.9	13.5	100.0
North West	22.7	15.2	30.5	20.5	11.0	100.0
Northern Cape	26.3	17.7	21.6	14.8	19.6	100.0
Western Cape	6.5	8.0	23.1	27.7	34.6	100.0
South Africa	*20.0*	*20.0*	*20.0*	*20.0*	*20.0*	*100.0*

Notes: Q1-Q5 refers to national poverty quintile. Entries are the percentage of students in each province that fall in each national poverty quintile.

Table 2: Descriptive statistics for ordinary public schools

	Panel A: Primary schools					
	Q1	Q2	Q3	Q4	Q5	All
N	4469	2821	2686	1265	1046	12287
2006 enrollment	286.2	322.2	430.9	587.0	608.0	384.7
	(254.1)	(280.3)	(303.7)	(339.1)	(305.7)	(308.0)
2006 student-teacher ratio	28.5	28.6	29.9	30.4	24.9	28.7
	(9.7)	(8.2)	(7.3)	(6.9)	(7.2)	(8.5)
2006 gender ratio	1.09	1.10	1.09	1.06	1.02	1.08
	(0.26)	(0.22)	(0.17)	(0.15)	(0.14)	(0.21)
Urban	0.07	0.11	0.30	0.73	0.80	0.26
Formerly white	0.00	0.01	0.03	0.10	0.63	0.08
Section 21 status	0.52	0.58	0.60	0.71	0.81	0.60
2005 fees (Rand)						
25%	20	20	30	50	220	25
50%	30	40	50	80	1900	50
75%	50	60	75	160	3000	70
	Panel B: Secondary schools					
	Q1	Q2	Q3	Q4	Q5	All
N	855	727	972	494	497	3545
As proportion of total schools in qunitile	0.16	0.20	0.27	0.28	0.32	0.22
2006 enrollment	508.3	575.3	718.9	988.6	925.8	707.3
	(307.4)	(360.6)	(391.2)	(387.1)	(346.6)	(400.0)
2006 student-teacher ratio	31.2	30.7	30.8	31.0	24.7	30.0
	(7.2)	(6.7)	(6.5)	(5.6)	(6.2)	(6.9)
2006 gender ratio	0.90	0.94	0.95	0.93	0.88	0.92
	(0.19)	(0.18)	(0.20)	(0.17)	(0.17)	(0.19)
Urban	0.14	0.17	0.32	0.78	0.85	0.39
Formerly white	0.01	0.01	0.01	0.08	0.63	0.12
Section 21 status	0.96	0.94	0.76	0.82	0.92	0.87
2005 fees (Rand)						
25%	75	100	100	100	450	100
50%	100	120	120	190	2800	130
75%	130	150	160	350	4200	225

Notes: Standard deviations are in parentheses. Q1-Q5 refers to national poverty quintile. Upper and lower 1% of observations excluded as outliers for 2006 enrollment, 2006 student-teacher ratio and 2006 gender ratio.

Table 3: FE estimates for log enrollment – national sample.

	Full sample			Omit Q4 and Q5
	(1)	(2)	(3)	(4)
	Panel A: Primary schools			
I*2007	-0.039***	-0.017***	0.005	0.007
	(0.004)	(0.004)	(0.004)	(0.004)
I*t			-0.009***	-0.002
			(0.002)	(0.002)
year dummies	Y	Y	Y	Y
province-year dummies	N	Y	Y	Y
N	57818	57818	57818	46845
	Panel B: Secondary schools			
I*2007	0.052***	0.042***	0.018***	0.018***
	(0.006)	(0.007)	(0.005)	(0.007)
I*t			0.010***	0.011***
			(0.003)	(0.004)
year dummies	Y	Y	Y	Y
province-year dummies	N	Y	Y	Y
N	16382	16382	16382	11683

Notes: Coefficients are from FE estimation with school log enrollment in each year from 2003-2007 as the dependent variable. Independent variables are as indicated in the table. $I*2007$ is the interaction between a dummy for no-fee status and a 2007 dummy (the year of implementation). $I*t$ is a linear time trend specific to the group of no-fee schools. Schools with the 2006/2007 change in enrollment below the 1st percentile or above the 99th percentile for each sub sample were omitted as outliers. Robust standard errors are in parentheses: *significant at 10%, **significant at 5%, ***significant at 1% levels.

Table 4: FE estimates for log enrollment by grade – national secondary school sample.

	Full secondary school sample			Omit Q4 and Q5
	(1)	(2)	(3)	(4)
	Panel A: Grade 8			
I*2007	0.053***	0.014	0.024*	0.040**
	(0.011)	(0.013)	(0.014)	(0.017)
I*t			-0.004	0.001
			(0.005)	(0.006)
N	13789	13789	13789	9343
	Panel B: Grade 9			
I*2007	0.022**	0.011	0.025*	0.017
	(0.010)	(0.012)	(0.014)	(0.017)
I*t			-0.006	-0.001
			(0.005)	(0.006)
N	13972	13972	13972	9458
	Panel C: Grade 10			
I*2007	0.036***	0.040***	0.030***	0.024*
	(0.009)	(0.010)	(0.010)	(0.013)
I*t			0.004	0.007
			(0.004)	(0.005)
N	16626	16626	16626	11871
	Panel D: Grade 11			
I*2007	0.035***	0.035***	0.004	0.006
	(0.010)	(0.011)	(0.012)	(0.015)
I*t			0.012***	0.011**
			(0.005)	(0.006)
N	16548	16548	16548	11804
	Panel E: Grade 12			
I*2007	0.091***	0.046***	-0.011	0.003
	(0.013)	(0.014)	(0.016)	(0.019)
I*t			0.023***	0.015**
			(0.006)	(0.007)
N	16450	16450	16450	11732
year dummies	Y	Y	Y	Y
province-year dummies	N	Y	Y	Y

Notes: Coefficients are from FE estimation with school-grade log enrollment in each year from 2003-2007 as the dependent variable. Independent variables are as indicated in the table. $I*2007$ is the interaction between a dummy for no-fee status and a 2007 dummy (the year of implementation). $I*t$ is a linear time trend specific to the group of no-fee schools. Schools with the 2006/2007 change in enrollment below the 1st percentile or above the 99th percentile for each grade were omitted as outliers. Robust standard errors are in parentheses: *significant at 10%, **significant at 5%, ***significant at 1% levels.

Table 5: Matching estimates – national sample

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Primary schools					
Treatment effect	0	0.004	0.003	-0.001	0.004	0.003
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Nearest neighbors	1	3	5	1	3	5
Bias adjustment	N	N	N	Y	Y	Y
Obs.	11237	11237	11237	11237	11237	11237
	Panel B: Secondary schools					
Treatment effect	0.020**	0.022***	0.021***	0.017**	0.020***	0.019***
	(0.008)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)
Nearest neighbors	1	3	5	1	3	5
Bias adjustment	N	N	N	Y	Y	Y
Obs.	3246	3246	3246	3246	3246	3246

Notes: Nearest neighbor estimates of average treatment effect on the treated performed using the "nnmatch" command in Stata (Abadie *et. al.*, 2004). Matching covariates (subplace-level) are: log average income, unemployment rate, employment rate, proportion without any secondary education, proportion with a secondary school diploma (matric), proportion with a toilet, proportion with access to piped water and proportion with electricity in the home. Bias adjustment refers to the option for regression adjustment of estimates based on difference in covariates (the matching covariates were used here). Schools with the 2006/2007 change in enrollment below the 1st percentile or above the 99th percentile for each sub sample were omitted as outliers. Robust standard errors are in parentheses: *significant at 10%, **significant at 5%, ***significant at 1% levels.

Table 6: Heterogeneous effects by quintile - national sample.

	Full sample (1)	(2)	(3)
	Panel A: Primary schools		
Q1*2007	-0.047***	-0.014***	0.006
	(0.005)	(0.005)	(0.004)
Q2*2007	-0.028***	-0.020***	0.001
	(0.005)	(0.006)	(0.005)
Q1*t and Q2*t	N	N	Y
year dummies	Y	Y	Y
province-year dummies	N	Y	Y
N	57818	57818	57818
	Panel B: Secondary schools		
Q1*2007	0.072***	0.061***	0.034***
	(0.008)	(0.009)	(0.007)
Q2*2007	0.036***	0.028***	-0.002
	(0.009)	(0.009)	(0.007)
Q1*t and Q2*t	N	N	Y
year dummies	Y	Y	Y
province-year dummies	N	Y	Y
N	16382	16382	16382

Notes: Coefficients are from FE estimation with school log enrollment as the dependent variable. Independent variables are as indicated in the table. $Q1*2007$ is the interaction between a quintile 1 dummy and a 2007 dummy (the year of implementation). $Q1*t$ is a linear time trend specific to quintile 1 schools. Identical terms for quintile 2 schools are denoted by $Q2$. Schools with the 2006/2007 change in enrollment below the 1st percentile or above the 99th percentile for each sub sample were omitted as outliers. Robust standard errors are in parentheses: *significant at 10%, **significant at 5%, ***significant at 1%.

Table 7: FE estimates - Eastern Cape ordinary schools.

	Eastern Cape sample: **Primary** (1)	**Primary** (2)	**Secondary** (3)	**Secondary** (4)
I*2007	0.015**		0.032**	
	(0.006)		(0.014)	
Q1*2007		0.021***		0.059***
		(0.007)		(0.019)
Q2*2007		0.007		0.003
		(0.008)		(0.016)
I*t	Y	N	Y	N
Q1*t and Q2*t	N	Y	N	Y
year dummies	Y	Y	Y	Y
N	20915	20915	3809	3809

Notes: Coefficients are from FE estimation with school log enrollment as the dependent variable. Independent variables are as indicated in the table. $I*2007$ is the interaction between a dummy for no-fee status and a 2007 dummy (the year of implementation). $I*t$ is a linear time trend specific to the group of no-fee schools. $Q1*2007$ is the interaction between a quintile 1 dummy and a year 2007 dummy. $Q1*t$ is a linear time trend specific to quintile 1 schools. Identical terms for quintile 2 schools are denoted by $Q2$. Schools with the 2006/2007 change in enrollment below the 1st percentile or above the 99th percentile were omitted as outliers. Robust standard errors are in parentheses: *significant at 10%, **significant at 5%, ***significant at 1%.

Table 8: Continuity checks for baseline variables in the RD design (Eastern Cape ordinary schools)

	Primary schools				Secondary schools			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Score polynomial	Cubic	Quartic	Quintic	Quintic	Cubic	Quartic	Quintic	Quintic
2006 log enrollment	0.071	0.07	0.009	0.002	0.051	0.003	0.005	0.006
	(0.059)	(0.058)	(0.071)	(0.070)	(0.107)	(0.111)	(0.125)	(0.125)
2006 gender ratio	0.01	0.01	0.007	0.008	0.048	0.067	0.03	0.049
	(0.016)	(0.016)	(0.019)	(0.019)	(0.042)	(0.043)	(0.051)	(0.049)
2006 teacher ratio	-0.526	-0.56	-1.634**	-1.678**	1.477	0.787	-0.293	-0.306
	(0.630)	(0.627)	(0.756)	(0.750)	(1.458)	(1.558)	(1.694)	(1.734)
2006 electricity	0.005	0.006	0.067*	0.072*	-0.018	0.003	0.03	0.052
	(0.035)	(0.035)	(0.041)	(0.041)	(0.058)	(0.063)	(0.066)	(0.066)
I*score	N	N	N	Y	N	N	N	Y
N (2006 log enrol.)	4187	4187	4187	4187	762	762	762	762

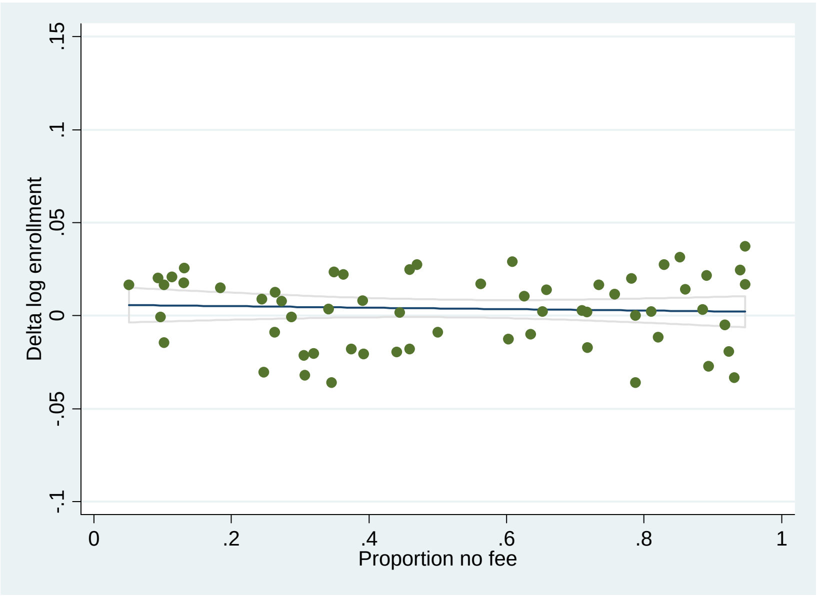
Notes: Coefficients are from OLS regressions of each baseline characteristic on I , a dummy for having a poverty score above the no-fee cutoff of zero and a polynomial in the poverty score (of the order indicated). Columns 4 and 8 include the interaction between I and the poverty score. For each baseline characteristic, observations with the dependent variable below the 1st percentile or above the 99th percentile were omitted as outliers. Standard errors are in parentheses: *significant at 10%, **significant at 5%, ***significant at 1% levels.

Table 9: RD estimates of effect of the program on 2007 enrollment

	Primary schools				Secondary schools			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Score polynomial	Quintic	Quintic	Quintic	Quintic	Quintic	Quintic	Quintic	Quintic
I	0.007	-0.012	-0.011	-0.004	-0.001	-0.004	-0.006	0.001
	(0.157)	(0.016)	(0.016)	(0.014)	(0.172)	(0.024)	(0.024)	(0.021)
2006 log enrollment	N	Y	Y	Y	N	Y	Y	Y
2006 gender ratio	N	N	Y	Y	N	N	Y	Y
2006 teacher ratio	N	N	Y	Y	N	N	Y	Y
2006 electricity	N	N	Y	Y	N	N	Y	Y
District dummies	N	N	N	Y	N	N	N	Y
N	4189	4188	4179	4179	764	763	761	761

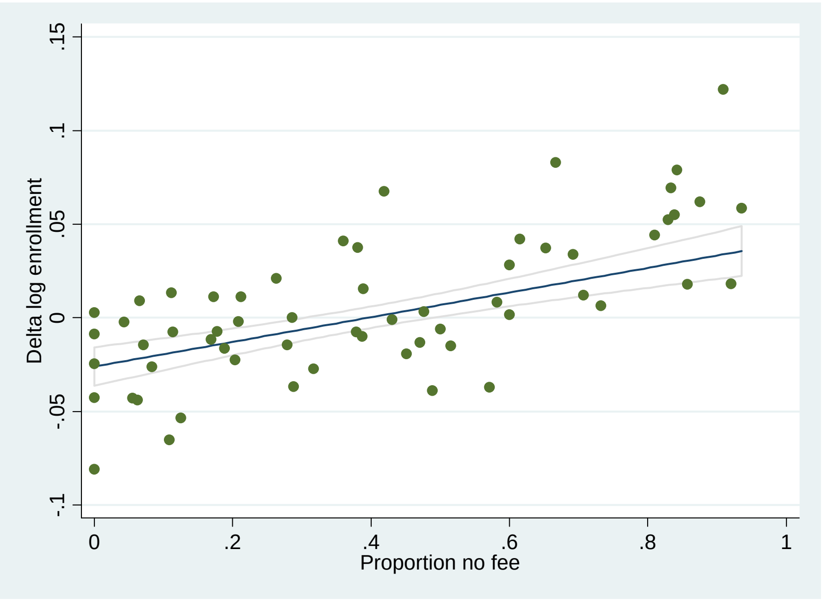
Notes: Coefficients are from OLS regressions of school log enrollment in 2007 on I , a dummy for having a poverty score above the no-fee cutoff of zero; a quintic polynomial in the poverty score; and various baseline characteristics measured in 2006 as indicated. Columns 4 and 8 also include the full set of education district dummies. Observations with 2007 enrollment below the 1st percentile or above the 99th percentile for each sub sample were omitted as outliers. Standard errors are in parentheses and are clustered by poverty score: *significant at 10%, ** significant at 5%, ***significant at 1% levels.

Figure 1a: Change in log enrollment and proportion of no-fee schools by education district (primary schools)



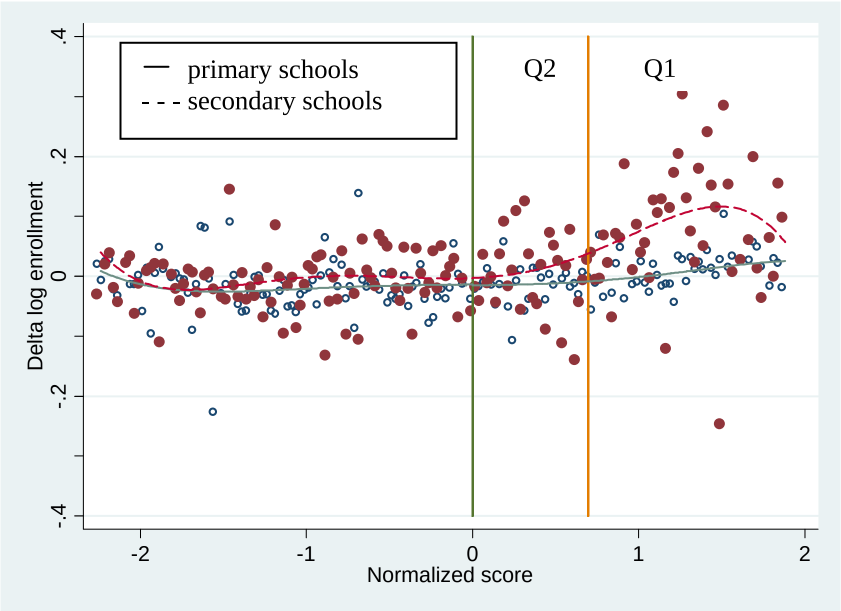
Notes: Each dot represents an education district. The lines are for a simple binary regression and a 95% confidence interval. Observations are weighted by the level of enrollment in 2006 using the "aweight" command in Stata. The slope coefficient is -0.004 with a standard error of 0.008.

Figure 1b: Change in log enrollment and proportion of no-fee schools by education district (secondary schools)



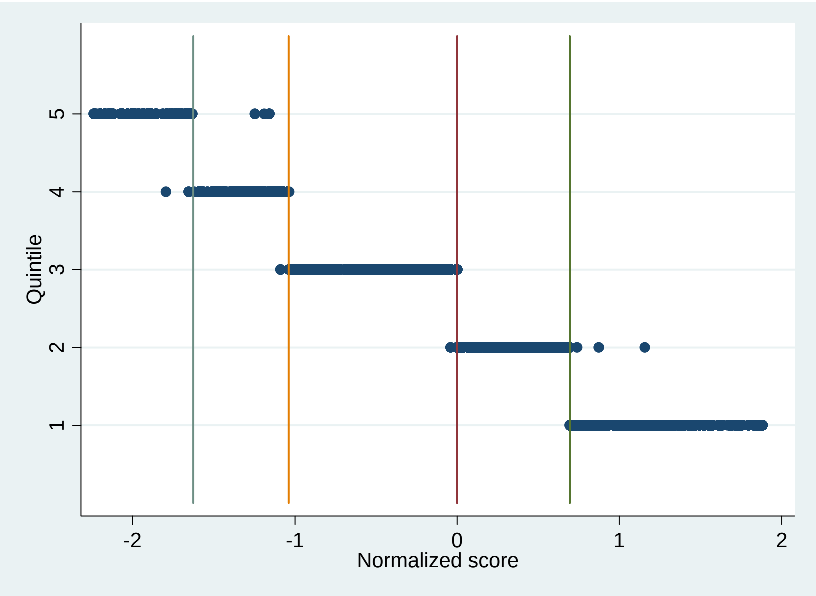
Notes: Each dot represents an education district. The lines are for a simple binary regression and a 95% confidence interval. Observations are weighted by the level of enrollment in 2006 using the "aweight" command in Stata. The slope coefficient is 0.066 with a standard error of 0.011.

Figure 2: 2007/2006 change in log enrollment versus normalized poverty score for Eastern Cape ordinary schools



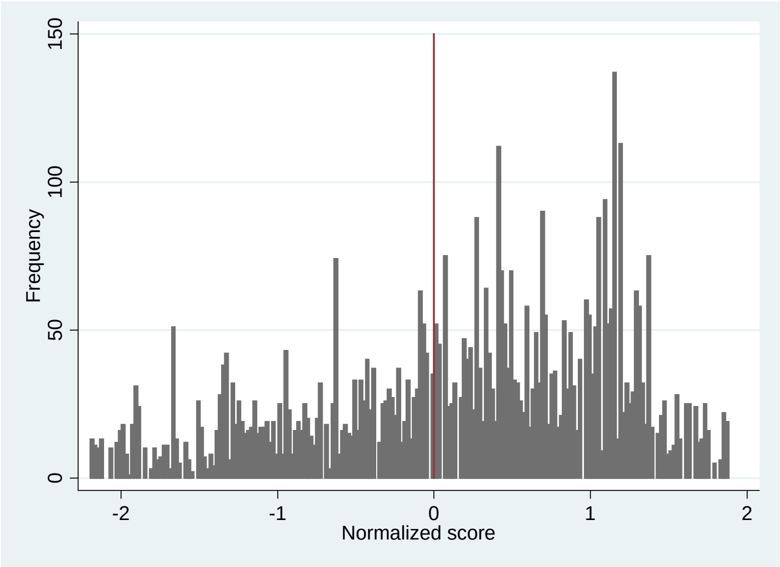
Notes: The open circles are bin averages for primary schools, the solid circles for secondary schools. The solid line is the regression of 2007/2006 change in log enrollment for primary schools on a quintic in the score and a dummy for being beyond the cutoff using individual school observations. The dashed line is the regression for secondary schools. Schools with 2007/2006 change in enrollments below the 1st percentile and above the 99th percentile were dropped as outliers. Q1 and Q2 demarcate the region of quintile 1 and quintile 2 schools respectively.

Figure 3: School quintile assignment and normalized poverty score for Eastern Cape ordinary schools



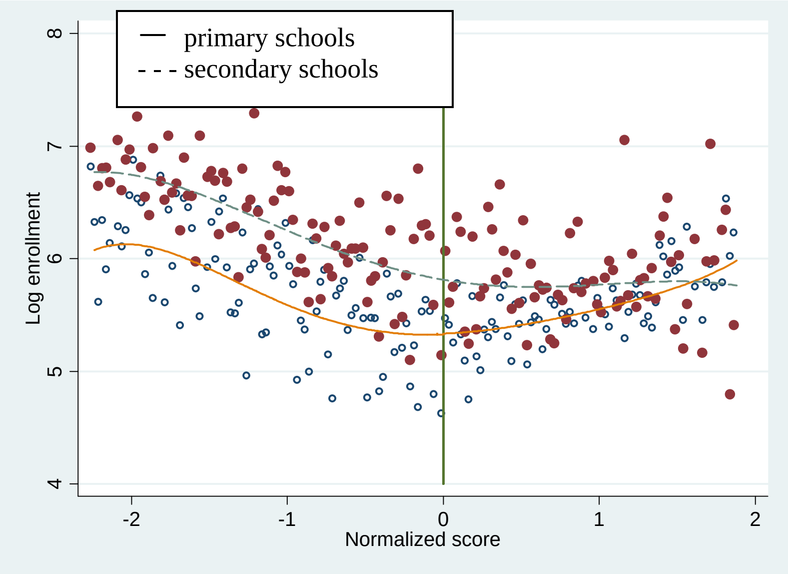
Notes: Each vertical line represents the poverty score cutoff between quintiles. Each dot represents a school.

Figure 4: Histogram of the normalized poverty score for Eastern Cape ordinary schools



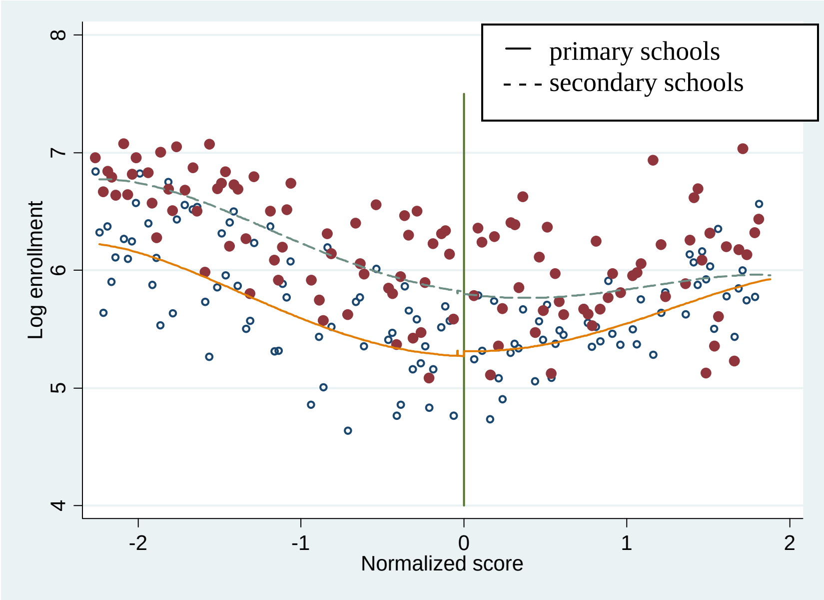
Notes: Schools are in bins of size 0.02. The vertical line at zero represents the cutoff between quintiles 2 and 3: schools to the right are no-fee.

Figure 5: 2006 log enrollment versus normalized poverty score for Eastern Cape ordinary schools



Notes: The open circles are bin averages for primary schools, the solid circles for secondary schools. The solid line is the regression of 2006 primary school log enrollment on a quintic in the score and a dummy for being beyond the cutoff of zero using individual school observations. The dashed line is the regression for secondary schools. Schools with 2006 enrollment below the 1st percentile and above the 99th percentile were dropped as outliers.

Figure 6: 2007 log enrollment versus normalized poverty score for Eastern Cape ordinary schools.



Notes: The circles are bin averages for primary schools, the solid circles for secondary schools. The solid line is the regression of 2007 primary school log enrollment on a quintic in the score and a dummy for being beyond the cutoff of zero using individual school observations. The dashed line is the regression for secondary schools. Schools with 2007 enrollments below the 1st percentile and above the 99th percentile were dropped as outliers.

Appendix

a. Sample composition

The data provided by the national Department of Education contained information on the universe of public schools in South Africa. Of the 29171 public schools on the system, 25129 reported positive total enrollment in 2007. The remaining 4042 may well be schools that closed but remained on the system, since many of these had no enrollment information in the previous few years either. In addition, schools could only be used as observations in the regressions if their quintile (and hence treatment status) could be determined. No quintile information was available for schools in two out of the nine provinces: KwaZulu-Natal (around 23% of the total sample) and North West (around 8% of the total sample). In the other provinces, only a few schools – at most 22 per province - did not have this information available. As discussed in section b below, some schools were also omitted from the final sample since it was difficult to classify them as primary or secondary schools. The final sample therefore consists of 15832 schools in seven of the nine provinces which reported positive enrollment in 2007, had quintile information available on the system and could be classified as primary or secondary schools.

b. Classification of schools

The standard classification of public schools in South Africa has primary schools offering grades 1-7 and secondary schools offering grades 8-12. There is a third category known as "combined" schools which typically offer grades 1-9 (i.e. the grades comprising compulsory education) and are concentrated predominantly in the Eastern Cape province. However, around 31% of schools do not follow this neat classification scheme. For example, some schools only offer the post-compulsory grades while others are particularly small and offer only a few primary grades. For the purposes of the analysis in this paper, primary schools are classified as those terminating at or before grade 7. Combined schools are counted as primary schools throughout the paper, since their grades largely overlap with primary school grades[25]. Secondary schools are those commencing at grade 8 or above and terminating at grade 12. This classification scheme leaves around 7% of the schools as unclassified, the bulk of which are schools offering all grades 1-12. Since the difference in effects between primary and secondary schools is a key feature of the analysis and these schools cannot be readily classified into either category, they have been dropped from the sample. For 2007, this approach yields a sample consisting of 12287 "primary" schools (of which 22% are combined schools) and 3545 secondary schools.

c. Matching estimator

A basic version of the difference-in-differences matching estimator introduced by Heckman, Ichimura and Todd (1997) gives the average treatment effect on the treated as:

$$\hat{ATT} = \frac{1}{N}\sum_{i=1}^{N}\left\{[Y_{i,t=2007}(1) - Y_{i,t=2006}(0)] - \frac{1}{n_i}\sum_{k=1}^{n_i}[Y_{k,t=2007}(0) - Y_{k,t=2006}(0)]\right\} \tag{A7}$$

where $Y_{it}(treat_{it})$ is the log enrollment in school i in year t with treatment status $treat_{it}$, N is the number of treated schools and n_i is the number of untreated schools (indexed by k) matched to treatment school i. Effectively this estimator uses the average change in enrollment for the set of untreated schools matched to each treated school as the relevant counterfactual. Traditional cross-sectional matching estimators that focus only on ex-post outcomes assume that, conditioning on a set of observables, mean outcomes are conditionally mean independent of treatment. The advantage of using a difference-in-differences type matching estimator instead is that it only requires conditional mean independence for the *change* in outcomes. This allows for time invariant unobserved differences by treatment status. For implementation here, each school

[25] Omitting combined schools makes little difference to the results for the primary school sample.

was first linked to a subplace using its GPS coordinates. Nearest neighbor matches were then selected for each treated school based on a vector of subplace-level covariates and matching estimates obtained as described in Abadie *et al.* (2004). These are the estimates that are presented in Table 5 and discussed in section 5b.